

Perspektiven auf gute instruktionale Erklärungen
im Biologieunterricht
Komplexität als didaktische Herausforderung

Dissertation zur Erlangung des akademischen Grades
einer Doktorin
der Didaktik der Naturwissenschaften „Dr. phil. nat“ (doctor philosophiae naturalis)
im Promotionsfach Didaktik der Biologie
der Fakultät für Biologie und Vorklinische Medizin
der Universität Regensburg

vorgelegt von
Christina Ehras

aus
Lauf a. d. Pegnitz

im Jahr
2023

Erstgutachter: Prof. Dr. Arne Dittmer
Zweitgutachter: Prof. Dr. Stefan Krauss
Drittgutachter: Prof. Dr. Sven Hilbert

Zusammenfassung

Die Studie zielt darauf ab, unter Berücksichtigung verschiedener Perspektiven eine Konzeptualisierung guten instruktionalen Erklärens im Biologieunterricht zu erarbeiten sowie fächerübergreifende und fachspezifische Kriterien guten Erklärens zu untersuchen. Im Hinblick auf den Biologieunterricht und dessen Bedeutung bei der Vermittlung eines angemessenen Verständnisses biologischer Phänomene besteht insbesondere ein Interesse daran, Unterschiede bzw. Schwierigkeiten beim Erklären komplexer Phänomene im Vergleich zu Erklärungen von linearen Ursache-Wirkungs-Beziehungen herauszuarbeiten.

Für die fachspezifische Begriffsbestimmung des instruktionalen Erklärens im Biologieunterricht werden sprachwissenschaftliche, lerntheoretische bzw. didaktische sowie wissenschaftstheoretische und fachliche Perspektiven aufeinander bezogen. Mit der fachspezifischen Begriffsbestimmung instruktionalen Erklärens wird unter anderem der Forderung nachgekommen, über Unterrichtsqualität nicht ausschließlich fachübergreifend nachzudenken, sondern auch fachspezifische bzw. fachdidaktische Besonderheiten zu berücksichtigen.

Für die empirische Bestimmung von Kriterien guten instruktionalen Erklärens wurde ein computerbasierter Fragebogen eingesetzt, der sechs Videovignetten instruktionaler Erklärungen enthält. Die Erklärungen greifen verschiedene biologische Phänomene auf, wobei diese im Grad der dargestellten Komplexität systematisch variieren. In drei der Erklärungen wird das zu erklärende Phänomen auf eine eindeutige, lineare Ursache-Wirkungs-Beziehung zurückgeführt. In drei weiteren Erklärungen wird ein komplexes Ursache-Wirkungs-Gefüge dargestellt, wobei die Lehrkraft explizit auf Nichtwissen und unsicheres Wissen eingeht.

Die Erklärungen werden von vier verschiedenen Statusgruppen (Schüler:innen, Studierende, Lehrkräfte, Didaktiker:innen) hinsichtlich ihrer Güte eingeschätzt. Die Perspektiven der Statusgruppen werden im Sinne einer Perspektiventriangulation für ein umfassendes Verständnis guten instruktionalen Erklärens zusammengeführt. Im Verlauf des Fragebogens geben die Teilnehmenden zunächst zu jeder Erklärung ein holistisches Globalurteil in Form einer Schulnote ab und können ihre Bewertung danach in einem offenen Textfeld begründen. Anschließend werden die Videos erneut präsentiert und anhand geschlossener Items zu den Aspekten der Adressatenorientierung, der Strukturiertheit, des Sprech- und Körperausdrucks, der sprachlichen Verständlichkeit, der Wahrnehmung des dargestellten Komplexitätsgrades sowie der Persönlichkeitswirkung der erklärenden Person eingeschätzt. Die Daten des nicht-standardisierten und des standardisierten Zugangs werden im Zuge einer Methodentriangulation aufeinander bezogen ausgewertet.

Die Ergebnisse geben Einblick in das breite Spektrum an Kriterien, das die Teilnehmenden nutzen. Die Kriterien lassen sich den Aspekten Adressatenorientierung, Strukturiertheit, Repräsentationen, Sprech- und Körperausdruck, Verständlichkeit der Sprache und dem fachlichen Aspekt zuordnen. Auch Merkmale der Lehrperson selbst sowie die Merkbarkeit und Lernbarkeit des Gesagten werden in die Bewertung miteinbezogen. Dabei differieren unter anderem die Häufigkeiten zwischen den Statusgruppen, mit denen auf die unterschiedlichen Aspekte Bezug genommen wird. So äußern sich beispielsweise Studierende und Lehrkräfte am häufigsten zu den Repräsentationen und zur Adressatenorientierung, wohingegen Didaktiker:innen neben den Repräsentationen vor allem auf fachliche Kriterien bei der Bewertung zurückgreifen.

Bei den geschlossenen Items zeigen sich insbesondere Bewertungsunterschiede zwischen Schüler:innen und den verschiedenen Gruppen der Metaperspektive (Studierende, Lehrkräfte, Didaktiker:innen). Die Häufigkeit und Ausprägung der Unterschiede variieren dabei in Abhängigkeit von dem zu bewertenden Aspekt – besonders heterogene Bewertungen liegen beispielsweise für den Aspekt der Strukturiertheit vor, die Adressatenorientierung wird hingegen nicht signifikant unterschiedlich bewertet.

Maßgebliche Einflussfaktoren auf die Güte der Erklärung sind die Adressatenorientierung und die Strukturiertheit, die in der Literatur auch als zentrale Aspekte effektiven Unterrichts diskutiert werden. In der statusgruppenspezifischen Analyse zeigt sich für Schüler:innen und die Gruppe der Metaperspektive jeweils noch ein weiterer Aspekt als relevanter Faktor, der die Unterschiedlichkeit der Perspektiven markiert: für Schüler:innen der Sprech- und Körperausdruck und für die Gruppe der Metaperspektive der fachspezifische Aspekt zur Wahrnehmung des dargestellten Komplexitätsgrades.

In der Zusammenschau sprechen die empirischen Befunde für eine gute Konstruktvalidität. Damit liegt erstmals ein Vorschlag für eine validierte Konzeptualisierung guten instruktionalen Erklärens vor, der die Spezifika des Biologieunterrichts bzw. biologischer Phänomene berücksichtigt.

Beim Vergleich der Bewertung von komplexen und linearen Erklärungen ergibt sich folgendes Bild: Schüler:innen bewerten die komplexen Varianten schlechter, Studierende und Lehrkräfte bewerten sie überwiegend, aber nicht in allen Fällen schlechter und Didaktiker:innen bewerten konträr dazu eher lineare Erklärungen schlechter. Anhand der Begründungen aus dem offenen Textfeld lassen sich dafür folgende Gründe rekonstruieren: Über alle Statusgruppen hinweg wird bei komplexen Erklärungen eine entstehende Vagheit negativ bewertet, wobei insgesamt keine Verunsicherung durch das Aufzeigen von Wissenslücken zu entstehen scheint. Während Studierende und Didaktiker:innen die explizite Erwähnung von Nichtwissen und unsicherem Wissen unter Bezugnahme auf die Ziele eines wissenschaftspropädeutischen Biologieunterrichts (Nature of Science) positiv bewerten, scheint aus der Perspektive von Lehrkräften die Darstellung von Wissenslücken dem Ziel der Wissensvermittlung zu widersprechen. Für die Gruppe der Adressat:innen zeichnet sich ab, dass die Visualisierungen bei komplexen Erklärungen vermehrt in den Fokus rücken und für ein Verständnis an Bedeutung gewinnen.

Die Ergebnisse implizieren unter anderem, dass Schüler:innen ein Umgang mit Nichtwissen und unsicherem Wissen zugetraut werden kann. Um ein adäquates Verständnis komplexer biologischer Phänomene zu fördern, sollten Lehrkräfte **Nichtwissen und unsicheres Wissen als ‚normale‘** Bestandteile biologischen Wissens bzw. des Wissens über die biologischen Phänomene in den Unterricht implementieren. Vor dem Hintergrund der theoretischen Überlegungen zu Nichtwissen und unsicherem Wissen sowie der empirischen Befunde sollte aus fachdidaktischer Sicht über die Auffassung von fachlicher Korrektheit und fachlicher Vollständigkeit selbstkritisch reflektiert werden.

Die Forschungsarbeit entstand im Rahmen des inter- und transdisziplinären Projekts FALKE (Fachspezifische LehrkräfteKompetenzen im Erklären) als eine Maßnahme von KOLEG (KOoperative LEhrkräftebildung Gestalten) an der Universität Regensburg.¹

¹ Das Vorhaben KOLEG wurde im Rahmen der gemeinsamen „Qualitätsoffensive Lehrerbildung“ von Bund und Ländern mit Mitteln des Bundesministeriums für Bildung und Forschung unter dem Förderkennzeichen FKZ01JA1512 gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei der Autorin.

Inhalt

1	Einleitung: Gute instruktionale Erklärungen im Biologieunterricht	10
1.1	Gute instruktionale Erklärungen im Biologieunterricht als Forschungsgegenstand.....	10
1.2	Forschungsinteresse und Verortung des Forschungsgegenstands	12
1.3	Aufbau der Arbeit.....	15
2	Instruktionale Erklärungen im Biologieunterricht	18
2.1	Erklären aus sprachwissenschaftlicher Perspektive: konversationsanalytische Sicht und funktionale Pragmatik	20
2.1.1	Inhalte und Typen instruktionaler Erklärungen aus sprachwissenschaftlicher Perspektive	21
2.1.2	Ziele und Funktionen instruktionaler Erklärungen aus sprachwissenschaftlicher Perspektive	23
2.2	Lerntheoretische und didaktische Perspektive auf instruktionales Erklären	23
2.2.1	Ziele und Funktionen instruktionaler Erklärungen im Biologieunterricht aus didaktischer Perspektive	23
2.2.2	Inhalte und Typen instruktionaler Erklärungen aus didaktischer Perspektive.....	24
2.3	Wissenschaftstheoretische und fachliche Perspektiven auf Erklärungen	26
2.3.1	Wissenschaftliche Erklärungsformen	27
2.3.2	<i>Exkurs:</i> Wissen über wissenschaftliche Erklärungsformen als Element einer wissenschaftspropädeutischen Lehrkräftebildung, der Unterrichtsgestaltung und des fachlichen Lernens	34
2.4	Synthese – instruktionales Erklären als fachlicher Kommunikationsprozess	37
3	Komplexität, Nichtwissen und unsicheres Wissen als charakteristische Eigenschaften biologischer Phänomene und Herausforderung für Erklärende und Rezipient:innen.....	41
3.1	Komplexität als Eigenschaft von biologischen Systemen.....	43
3.2	Wissenschaftliche Erklärungen von und in komplexen Systemen	46
3.3	Nichtwissen und unsicheres Wissen als Teil von Erkenntnis	48
3.3.1	Betroffene Subjekte – Wer weiß etwas (nicht)?	50
3.3.2	Wissensstati – Wie viel wissen wir über die Welt? Wie viel können wir über die Welt wissen?.....	51
3.3.3	Auswirkungen – Kommunizieren, Bewerten und Handeln	55
3.4	Herausforderungen bei der Vermittlung und dem Verständnis komplexer Phänomene	58
3.5	Hinweise zum Stellenwert sowie zur Bedeutung eines Verständnisses von Komplexität und Nichtwissen im Biologieunterricht	61
3.6	Zusammenfassung und Fazit: instruktionale Erklärungen komplexer biologischer Phänomene	66
4	Kriterien guten Erklärens	68
4.1	Adressatenorientierung	70

4.2	Strukturiertheit	75
4.3	Einsatz von Repräsentationen	77
4.4	Verständliche Sprache	84
4.5	Sprech- und Körperausdruck.....	89
4.6	Fachliche Kriterien	92
4.7	Fazit – Anforderungen an eine erklärende Lehrkraft	95
4.8	Gutes instruktionales Erklären – eine Frage der Perspektive	100
5	Zielsetzung und Konzeption der empirischen Untersuchung zum instruktionalen Erklären im Biologieunterricht.....	103
5.1	Zielsetzung und Fragestellungen	103
5.2	Forschungszugang und Untersuchungsdesign.....	106
5.2.1	Aufbau und Ablauf des Fragebogens	107
5.2.2	Videovignetten als Stimulus	110
5.2.2.1	Auswahl der Erklärgegenstände	111
5.2.2.2	Konzeptualisierung und Operationalisierung der unabhängigen Variable ‚Grad der dargestellten Komplexität‘ in den Videovignetten	112
5.2.2.3	Gestaltungsprämissen – technische und methodische Umsetzung	113
5.2.2.4	Entstehungsprozess der Videovignetten und Qualitätssicherung.....	116
5.2.2.5	Beschreibung der Videovignetten.....	117
5.2.3	Triangulation als Möglichkeit, komplexen Forschungsgegenständen zu begegnen....	135
5.2.3.1	Methodentriangulation – Verschränkung qualitativer und quantitativer Methoden.....	137
5.2.3.2	Datentriangulation – Erfassung des Messgegenstands aus unterschiedlichen Perspektiven	137
5.2.4	Konzeptualisierung und Operationalisierung des Messgegenstandes ‚Gutes instruktionales Erklären im Biologieunterricht‘	138
5.2.4.1	Adressatenorientierung.....	141
5.2.4.2	Strukturiertheit	142
5.2.4.3	Sprachliche Verständlichkeit.....	144
5.2.4.4	Sprech- und Körperausdruck.....	144
5.2.4.5	Wahrnehmung des dargestellten Komplexitätsgrades	145
5.2.5	Kontrolle möglicher Störvariablen innerhalb des Fragebogens Kontrollvariablen	147
5.2.6	Pilotierung des Messinstruments mittels kognitiver Interviews	150
5.2.7	Stichprobe und Durchführung der Erhebung	153
5.3	Datenaufbereitungs- und Auswertungsmethoden	157
5.3.1	Datenaufbereitung und Vorbereitungen für Auswertungen der geschlossenen Items	157
5.3.2	Auswertungen der Begründungen des Globalurteils	158

5.3.2.1	Die inhaltlich-strukturierende qualitative Inhaltsanalyse	159
5.3.2.2	Umsetzung der inhaltlich-strukturierenden qualitativen Inhaltsanalyse.....	160
5.3.2.3	Umsetzung der Erweiterung zur evaluativen qualitativen Inhaltsanalyse	165
5.3.2.4	Kategoriensystem	166
6	Untersuchungsergebnisse.....	168
6.1	Gütekriterien und konfundierende Einflussgrößen.....	168
6.1.1	Kontrollvariablen und mögliche Konfundierungen.....	168
6.1.2	Gütekriterien der quantitativen Erhebung.....	174
6.1.3	Gütekriterien der qualitativen Erhebung	180
6.2	Ergebnisse.....	182
6.2.1	Ergebnisse zur Nutzung von Kriterien für die Bewertung instruktionaler Erklärungen	183
6.2.1.1	Ergebnisse zur statusgruppenübergreifenden Nutzung von Kriterien für die Bewertung instruktionaler Erklärungen	183
6.2.1.2	Überblick und Einordnung der Ergebnisse zur statusgruppenübergreifenden Nutzung von Kriterien für die Bewertung instruktionaler Erklärungen.....	203
6.2.1.3	Ergebnisse zur statusgruppenspezifischen Nutzung von Kriterien für die Bewertung instruktionaler Erklärungen	206
6.2.1.4	Exkurs: Personenspezifische Nutzung von Kriterien für die Bewertung instruktionaler Erklärungen	228
6.2.1.5	Überblick und Einordnung der Ergebnisse zur statusgruppenspezifischen Nutzung von Kriterien zur Bewertung instruktionaler Erklärungen	233
6.2.2	Ergebnisse zu Bewertungsunterschieden zwischen verschiedenen Statusgruppen in Hinblick auf vorgegebene Kriterien	238
6.2.2.1	Ergebnisse zur Bewertungsunterschieden zwischen Schüler:innen, Studierenden, Lehrkräften und Didaktiker:innen in Hinblick auf vorgegebene Kriterien	238
6.2.2.2	Überblick und Einordnung der Ergebnisse zu Bewertungsunterschieden zwischen verschiedenen Statusgruppen in Hinblick auf vorgegebene Kriterien	251
6.2.3	Ergebnisse zu maßgeblichen Einflussfaktoren auf die globale Bewertung instruktionaler Erklärungen	254
6.2.3.1	Ergebnisse zu statusgruppenübergreifend maßgeblichen Einflussfaktoren auf die globale Bewertung instruktionaler Erklärungen	254
6.2.3.2	Ergebnisse zu maßgeblichen Einflussfaktoren auf die globale Bewertung instruktionaler Erklärungen für Lernende und Lehrende	257
6.2.3.3	Überblick und Einordnung der Ergebnisse zu maßgeblichen Einflussfaktoren auf die globale Bewertung instruktionaler Erklärungen	260
6.2.4	Ergebnisse zu den Bewertungs- und Begründungsunterschieden zwischen linearen und komplexen instruktionalen Erklärungen in Abhängigkeit von der Statusgruppe.....	261
6.2.4.1	Ergebnisse zu den Bewertungsunterschieden zwischen linearen und komplexen instruktionalen Erklärungen in Abhängigkeit von der Statusgruppe	261

6.2.4.2	Überblick und Einordnung der Ergebnisse zu Unterschieden der globalen Bewertung von linearen und komplexen instruktionalen Erklärungen in Abhängigkeit von der Statusgruppe	267
6.2.4.3	Ergebnisse zu den Begründungsunterschieden zwischen linearen und komplexen instruktionalen Erklärungen in Abhängigkeit von der Statusgruppe	268
6.2.4.4	Überblick und Einordnung der Ergebnisse zu Begründungsunterschieden zwischen linearen und komplexen instruktionalen Erklärungen in Abhängigkeit von der Statusgruppe	279
7	Zusammenfassung und Diskussion	282
7.1	Zusammenfassung und Einordnung zentraler Befunde	282
7.2	Reflexion des Vorgehens, Limitationen und Ausblick	290
7.3	Fazit und Implikationen	294
	Abbildungsverzeichnis	297
	Tabellenverzeichnis.....	298
	Literaturverzeichnis	301
	Anhang.....	318

1 Einleitung: Gute instruktionale Erklärungen im Biologieunterricht

1.1 Gute instruktionale Erklärungen im Biologieunterricht als Forschungsgegenstand

Erklärungen kommen im Unterricht häufig (Dagher & Cossman, 1992) und in vielfältigen Formen vor (Findeisen, 2017; Geelan, 2012). Schüler:innen geben in Befragungen an, gut erklären zu können, sei eine wichtige Fähigkeit von guten Lehrkräften (Merzyn, 2015). Auch in Modellen zum Professionswissen von Lehrkräften wird das Wissen über Erklären und Repräsentieren als Facette im Bereich des fachdidaktischen Wissens berücksichtigt (Krauss et al., 2017; Shulman, 1986). Entsprechend wird es auch im Core-Practice-Ansatz (McDonald, Kazemi & Kavanagh, 2013) als Kernpraktik des Unterrichtens angesehen und teilweise werden hierzu auch Lernangebote im universitären Ausbildungsabschnitt für angehende Lehrkräfte geschaffen (Ehras, Asen-Molz, Frei, Schilcher & Krauss, 2021).

Die Zusammenstellung an Schlaglichtern gibt einen Einblick, dass dem Erklären von Lehrkräften aus verschiedenen Perspektiven – aus der Unterrichtspraxis und folglich auch aus der Lehrkräftebildung und Professionsforschung – eine zentrale Bedeutung beigemessen wird. Der mögliche Schluss, dass Erklärungen im Fachunterricht deshalb breit diskutiert werden und fachspezifische Konzeptionen vorliegen, trifft jedoch nicht zu. Das instruktionale Erklären im Biologieunterricht – also in der Regel kürzere Erklärsequenzen von Biologielehrkräften – stellt sich vielmehr als spannungsreicher Forschungsgegenstand dar:

Zum einen werden instruktionale Erklärungen als lehrkräftezentrierte Lehr-Lern-Form teilweise als Gegenpol zu einem konstruktivistischen Lernverständnis betrachtet und daher abgelehnt. Scheinbar hält sich unter einigen Lehrkräften, angehenden Lehrkräften und auch Lehrkräftebildner:innen die Meinung – zumindest wird sie ihnen in einigen Texten, die sich mit instruktionalem Erklären beschäftigen, zugeordnet –, dass das instruktionale Erklären durch Lehrkräfte eher vermieden werden sollte (Gaier, 2022; Hattie, 2015; Leisen, 2007). Lehr-Lehr-Formen mit – mitunter nur augenscheinlich – höherer Aktivität von Schüler:innen werden dabei insgesamt als geeigneter und wirksamer erachtet als solche, bei denen Lehrkräfte einen großen Redeanteil besitzen und aktiver erscheinen als ihre Schüler:innen. Dabei wird die Aktivität der Schüler:innen mit der Leistung der Wissenskonstruktion gleichgesetzt – instruktionale Erklärungen scheinen folglich einem konstruktivistischen Lernkonzept entgegenzustehen. Übersehen werden hier zwei Aspekte: Auch während der Rezeption instruktionaler Erklärungen sind Schüler:innen kognitiv aktiv (Brüning & Saum, 2019). Informationen müssen mit dem eigenen Vorwissen abgeglichen, Anknüpfungspunkte gesucht, Widersprüche identifiziert und mentale Bilder oder Modelle aufgebaut werden. Das Denken in den skizzierten Polen – überspitzt formuliert ein Frontalunterricht auf der einen Seite und vollkommen autonome, selbstgesteuerte Lerneinheiten auf der anderen Seite – versperrt auch den Blick darauf, dass es nicht die eine Methode gibt, die zum Ziel führt (Hattie, 2015; Helmke, 2007).

Denn instruktionale Erklärungen stehen nicht isoliert im Unterricht, sondern sind bzw. sollten stets in weitere Unterrichtsaktivitäten eingebettet sein, die zur Verarbeitung des Gesagten anregen und ein Anwenden ermöglichen, um ihr Potenzial für lernwirksamen Unterricht zu entfalten (Hattie, 2015; Webb, Troper & Fall, 1995). Den Einsatz von instruktionalen Erklärungen gilt es deshalb abzuwägen (Renkl, 2001). In einigen Unterrichtssituationen sind sie besonders geeignet, um den Lernprozess von Schüler:innen anzustoßen bzw. zu unterstützen, beispielsweise, wenn das

Vorwissen von Schüler:innen nicht ausreicht, um selbst Erklärungen zu konstruieren, oder korrespondierend dazu, sofern der Sachverhalt eine zu hohe Komplexität aufweist, sodass Schüler:innen die Zusammenhänge mit hoher Wahrscheinlichkeit nicht selbst rekonstruieren können (Renkl et al., 2007). Mit einer instruktionalen Erklärung können Lehrkräfte auch sicherstellen, dass Schüler:innen korrekte Formulierungen, Fachbegriffe und Zusammenhänge hören, auf die sie zurückgreifen können (Renkl, 2001). Sie können darüber hinaus aufzeigen, welche Fragen in der Disziplin der Biologie adressiert werden und wie Biolog:innen Phänomene erklären. Die unterrichtlichen Erklärungen besitzen dabei eine Modellierungsfunktion für Fragen, die mit Erkenntnismethoden der Disziplin beantwortet werden können, sowie fachlich angemessene Erklärungsformen (Larreamendy-Joerns & Muñoz, 2010). Instruktionale Erklärungen sind weiterhin zu bevorzugen, wenn es um das Auflösen von bereits vorhandenen oder antizipierten Verständnisschwierigkeiten bei den Rezipient:innen geht (Renkl, 2001). Insgesamt stehen instruktionale Erklärungen damit nicht im Widerspruch zu einer konstruktivistischen Auffassung von Lernen. Aus dem „Schattendasein“ (Schilcher, Krauss, Rincke & Hilbert, 2017, S. 444) scheint das Erklären auch durch die Metastudie von Hattie (2015) entkommen zu sein, die die Lernwirksamkeit der Methode der direkten Instruktion, innerhalb derer auch instruktionalen Erklärungen eine prominente Rolle zukommt, aufzeigt. Insgesamt werden nicht mehr nur Schüler:innen und deren Lernprozessen fokussiert, sondern vermehrt Lehrkräfte und Lehrprozesse in den Blick genommen (Geelan, 2013; Schilcher et al., 2017; Trautmann & Wischer, 2016). Auch durch Diskussionen um Erklärvideos auf Plattformen wie YouTube, die Schüler:innen rege zum eigenständigen Lernen zu Hause nutzen (Medienpädagogischer Forschungsverbund Südwest [mpfs], 2020), wird die Lehr-Lern-Form des Erklärens theoretisch und empirisch untersucht (Frei, Asen-Molz, Hilbert, Schilcher & Krauss, 2020; Kulgemeyer, 2018).

Zum anderen zeigt sich in der bildungswissenschaftlichen Diskussion zur Frage, ob instruktionale Erklärungen effektiv sind, also zu einem Lernwachstum bei Schüler:innen führen, ein Spannungsverhältnis. Während instruktionale Erklärungen auf der einen Seite als bewährte und allgegenwärtige Elemente unterrichtlicher Praxis beschrieben werden, sprechen auf der anderen Seite Befunde einiger Untersuchungen eher für eine geringe Wirksamkeit in Bezug auf den Lernzuwachs von Schüler:innen (Kuhn, 2007; Wittwer & Renkl, 2008). Auch hier können die Ergebnisse der Metastudie Hatties (2015) zur Relativierung der geringen Wirksamkeit von instruktionalen Erklärungen herangezogen werden, denn die Methode der direkten Instruktion ist mit einer mittleren Effektstärke vergleichsweise lernwirksam. Die Autoren Wittwer und Renkl (2008) analysieren die teils fehlende Wirksamkeit: Sie führen an, dass Faktoren aufseiten der Lehrkräfte dazu beitragen – beispielsweise die Ausprägung diagnostischer Fähigkeiten, um fachlich nicht angemessene Vorstellungen zu erfassen und angepasste Erklärungen zu konstruieren –, aber auch Faktoren aufseiten der Schüler:innen zu finden sind – wie die Motivation und Interesse am Erklärgegenstand als erste Hürde das Gesagte zu verarbeiten und zu verstehen. Mitunter muss dabei konstatiert werden, dass die isolierte Messung der Wirksamkeit bzw. Effektivität von instruktionalen Erklärungen eine große Herausforderung darstellt, da sie Teil eines komplexen Lehr-Lern-Prozesses mit zahlreichen Variablen sind (Brüning & Saum, 2019; Leinhardt, 1987; Webb et al., 1995). Einen Schritt zurückzutreten und zunächst Kriterien guter Erklärungen aus einer theoretisch fundierten, normativen Sicht zu erfassen und zu diskutieren, ist daher lohnend, um weitere Ansätze bzw. Variablen für eine umfangreiche Untersuchung zu Wirkeffekten zu ermitteln.

Ein letzter interessanter Reibungspunkt ergibt sich hinsichtlich der Charakteristika biologischer Phänomene – das heißt in Bezug auf Eigenschaften möglicher Inhalte instruktionaler Erklärungen im Biologieunterricht – und der Ziele von instruktionalen Erklärungen: Biologische Phänomene

weisen häufig einen hohen Grad an Komplexität auf, das heißt, dass viele Faktoren zur Phänomenentstehung beitragen, die wiederum über Ursache-Wirkungs-Beziehungen miteinander in Verbindung stehen (Mainzer, 2008; Mitchell, 2008; Potochnik, 2013). Ändert sich die Ausprägung eines Faktors bzw. einer Anfangsbedingung in diesem Ursache-Wirkungs-Netz, wirkt sich das auf das zu erklärende Phänomen aus. Die hohe Anzahl von Faktoren, unterschiedlichen Ausprägungsvarianten der Faktoren und Wechselwirkungen zwischen den Faktoren erschweren Wissenschaftler:innen erheblich bzw. machen es teils unmöglich, das Phänomen vollständig zu erklären oder Vorhersagen über eine weitere Entwicklung zu treffen (Kuhlmann, 2007; Mainzer, 2008; Mitchell, 2008). Dies führt dazu, dass jedes Phänomen einer individuellen Betrachtung bedarf. Dabei verlieren allgemeine Prinzipien, Gesetzmäßigkeiten oder Regeln zwar nicht ihre Gültigkeit, können das reale Phänomen jedoch nur in Ansätzen erklären (Mitchell, 2008; Potochnik, 2013). Insgesamt bleibt in Erklärungen komplexer Phänomene aufgrund epistemologischer und methodischer Grenzen häufig ein Anteil an Nichtwissen und unsicherem Wissen enthalten (Kuhlmann, 2007). Dem steht das allgemeine Ziel unterrichtlicher Erklärungen gegenüber, Wissen zu vermitteln (Ehlich, 2009; Hohenstein, 2006; Treagust & Harrison, 2000). Dabei entsteht auch eine weiterführende diskussionswürdige Frage, die über die nach der guten Umsetzung der Unterrichtsmethode der instruktionalen Erklärungen hinaus geht und sich auf das vermittelte Wesen von biologischem Wissen in Bezug auf die Bedeutung von Komplexität, Individualität, Nichtwissen und unsicherem Wissen allgemein bezieht (Dittmer & Ehras, 2024).

Die Frage nach gutem, instruktionalem Erklären im Biologieunterricht kann als spannungsreich und damit einhergehend als interessant beschrieben werden. Sie ist gleichzeitig höchst relevant, da sie Grundlage für allgemeine und fachspezifische Forschung zum Lernen von Schüler:innen mit instruktionalen Erklärungen ist sowie die Basis für die Ausbildung von angehenden Lehrkräften und Weiterbildung von Lehrkräften zum Thema des Erklärens im Unterricht schafft.

1.2 Forschungsinteresse und Verortung des Forschungsgegenstands

Geleitet von den im vorherigen Abschnitt aufgeführten Punkten bezieht sich das Forschungsinteresse der vorliegenden Arbeit allgemein auf eine Konzeptualisierung instruktionalen Erklärens im Biologieunterricht sowie die empirische Untersuchung von fachspezifischen und fächerübergreifenden Kriterien guten Erklärens aus verschiedenen Perspektiven. Im Spezifischen sollen dabei Unterschiede in der Einschätzung der Güte instruktionaler Erklärungen, die lineare Ursache-Wirkungs-Beziehungen aufzeigen, und jener, die komplexe Ursache-Wirkungs-Beziehungen rekonstruieren, betrachtet werden.

Grundlegend kann das Forschungsvorhaben in der Unterrichtsforschung verortet werden. Dabei wird nach relevanten Merkmalen des Unterrichts gesucht bzw. werden deren Wirkung und ihr Zusammenspiel untersucht. Berücksichtigung finden Komponenten, die eine Wirkung auf die Bildung bzw. auf den Erwerb von Fähigkeit, Fertigkeiten und Wissen von Schüler:innen haben. Dabei können zwei Sichtweisen bzw. Paradigmen der pädagogisch-psychologisch orientierten Unterrichtsforschung unterschieden werden: zum einen das Strukturparadigma, innerhalb dessen Theorien des Lehrens als Basis für Modellierungen von Unterrichtskomponenten herangezogen werden und Wirkungen auf Lernergebnisse als Zieldimension untersucht werden, zum anderen das Prozessparadigma, innerhalb dessen Theorien des Lernens als Grundlage für Modellierungen für Unterrichtskomponenten herangezogen werden und Effekte auf spezifische Prozesse des Lernens und nicht auf Ergebnisse des Lernens im Fokus stehen (Seidel, 2014). In Angebots-Nutzungs-

Modellen werden beide Paradigmen berücksichtigt (Seidel, 2014). Sie spiegeln auch die „Komplexität des Zusammenspiels unterschiedlichster Einflussfaktoren“ (Seidel, 2014, S. 857) wider. So sind beispielsweise im Angebots-Nutzungs-Modell von Helmke (2007) sechs Elemente enthalten, die Wirkung entfalten können: Merkmale der Lehrperson, Klassen- und Fachkontext, der Unterricht mit seinen Merkmalen als Angebot, individuelle Eingangsvoraussetzungen von Schüler:innen, Mediationsprozesse auf Schüler:innenseite sowie Lernaktivitäten der Schüler:innen als Nutzung des Angebots. Instruktionale Erklärungen sind dabei als Lernangebot von Lehrkräften und als Teil des Unterrichts zu verstehen.

Ein gut untersuchtes Merkmal von Lehrpersonen ist deren Professionswissen, das heißt dasjenige Wissen, auf das Lehrkräfte bei der Bewältigung von Aufgaben innerhalb ihres Berufsfeldes zurückgreifen und dem eine zentrale Bedeutung innerhalb der professionellen Lehrkräftekompetenz zukommt (Shulman, 1986; Baumert & Kunter, 2011). Das Professionswissen lässt sich unter anderem in drei weitere Wissensarten differenzieren: Fachwissen (content knowledge; CK), fachdidaktisches Wissen (paedagogical knowledge; PCK) und pädagogisches Wissen (paedagogical knowledge; PK). Innerhalb der Wissensart des fachdidaktischen Wissens ist wiederum die Facette ‚Wissen über Erklären und Repräsentieren‘ enthalten (Shulman, 1986), sodass das Erklären in der Professionsforschung Berücksichtigung findet (Krauss et al., 2017; Kunter et al., 2011). Insgesamt stellt sich empirisch das fachdidaktische Wissen als ein relevanter Prädiktor für Unterrichtserfolg dar (Kunter et al., 2013), sodass auch das Wissen über Erklären und Repräsentieren ein wichtiges Element für erfolgreiches Unterrichten bzw. den Lernerfolg von Schüler:innen ist.²

Während bei Professionswissenstests die Facette ‚Wissen über Erklären und Repräsentieren‘ in der Regel die Verfügbarkeit von verschiedenen Erklärungsansätzen erfasst und damit einen Eindruck vermittelt, wie flexibel Lehrkräfte auf verschiedene Vorstellungen und Probleme von Schüler:innen beim Verständnis mit unterschiedlichen Ansätzen reagieren können, um schließlich Aussagen über das fachdidaktische Wissen von Lehrkräften zu treffen (Krauss et al., 2017), stehen im vorliegenden Forschungsvorhaben die Güte von instruktionalen Erklärungen als Teil der Unterrichts und damit die Bestimmung von Kriterien guten Erklärens im Fokus.

Den Rahmen für die normative Erfassung von Kriterien guten Erklärens bildet die Unterscheidung zwischen ‚gut‘ und ‚effektiv‘, die Berliner (2005) aufgreift:

Good teaching occurs when the standards of the field are upheld. If you are a physician or a waiter, good practice includes washing your hands frequently. If you are a teacher, good practice may include greeting students warmly at the classroom door. Good is normative. It is what is expected of people in a position. In contrast, effective teaching is about reaching achievement goals. It is about students learning what they are supposed to in a particular class, grade, or subject. A high-quality teacher shows evidence of both good and effective teaching. (Berliner, 2005, S. 207)

Eine Erfassung von Kriterien *guten* Erklärens scheint nicht nur sinnvoll, weil die empirische Erfassung der Wirksamkeit von instruktionalen Erklärungen aufgrund vieler Einflussfaktoren schwierig ist (siehe Kap. 1.1), sondern auch, da bisher eine fachspezifische Betrachtung aussteht, die durch die Berücksichtigung verschiedener Perspektiven und einer normativen Sicht realisiert werden kann.

Nicht allein im Hinblick auf die Untersuchung instruktionalen Erklärens, sondern insgesamt auf die Unterrichtsforschung bezogen wird häufig die Berücksichtigung des fachspezifischen Kontextes bzw. eine explizit fachdidaktische Perspektive gefordert. So formuliert Helmke (2007) mit Bezug zur

² Rückschlüsse auf relevantes Wissen bzw. wichtige Fähigkeiten von Lehrkräften, um gut erklären zu können, stehen nicht im Fokus dieser Arbeit, finden aber in Kapitel 4.7 Eingang.

Unterrichtsqualität, dass „fachbezogene Aspekte der Unterrichtsqualität von großer Wichtigkeit sind“ (Helmke, 2007, S. 13). In diesem Kontext wird häufig darauf Bezug genommen, dass viele Untersuchungen zur Unterrichtsqualität im Mathematikunterricht durchgeführt wurden. So kommen auch Schilcher und Kolleg:innen (2021) zu der Schlussfolgerung, dass bei der Übertragung von in spezifischen Unterrichtsfächern gewonnenen Erkenntnissen Vorsicht geboten ist:

And one might well doubt whether [...] generalized statements really could apply to all school subjects. We rather need to ask ourselves – across disciplinary boundaries – which evidence and statements we might be able to generalize from one subject to others. Conversely, we need to consider when we should only look at teaching and learning through the lens of the specific subject, taking into account its highly complex content and domainspecific learning processes. (Schilcher et al., 2021, S. 2)

Schilcher und Kolleg:innen (2021) führen dazu domänenspezifische Lernprozesse und die komplexe Beschaffenheit der Inhalte als Begründung für fachspezifische Unterrichtsforschung an. Baumert und Kunter (2006) beziehen sich bei ihren Begründungen nicht auf die Prozesse des Lernens, sondern gehen auf die Charakteristika der Fachinhalte ein, die die Struktur des Unterrichts und des Unterrichtens formen:

Das Fach bestimmt bis in Einzelheiten hinein die Textur des Unterrichts. Dies beginnt mit Vorstellungen über die Sequenzialität und Anordnung von inhaltlichen Komponenten und reicht bis hin zum Modus der Repräsentationen und Erklärungen. (Baumert & Kunter, 2006, S. 492)

Damit führen Baumert und Kunter (2006) Erklärungen sogar als unmittelbares Beispiel für Fachspezifika des Unterrichtens an. Folgende Beispiele können verdeutlichen, dass Erklärungen verschiedene Modi annehmen: Während die physikalischen Phänomene des Physikunterrichts überwiegend durch Rückbezug auf Gesetze – also universelle, ausnahmslos geltende Zusammenhänge – erklärt werden können, spielen Gesetze in der Biologie und damit auch für Erklärgegenstände des Biologieunterrichts kaum eine Rolle (Bartelborth, 2007; Lange, 2013; siehe auch Kap. 2.3). Im Bereich der Evolutionsbiologie ähnelt die Herangehensweise sogar eher demjenigen von Historiker:innen: In beiden Disziplinen werden multikausale Ursachenkonstellationen durch existierende Artefakte rekonstruiert, um singuläre Ergebnisse zu erklären, die im Labor weitestgehend nicht wiederholt hergestellt werden können. Schließlich muss durch Plausibilitätsüberlegungen auf die wahrscheinlichste bzw. beste Erklärung geschlossen werden (Bartelborth, 2007; Potochnik, 2013). Die Beispiele geben Einblicke, dass Erklärungen der Fächer in besonderer Weise in Beziehung zur Erkenntnistheorie der jeweiligen Fächer stehen. Die Art und Weise, wie Wissen gewonnen wird, was als Wissen angesehen wird und welche Vorstellung über die Wissensstruktur vorherrscht, spiegelt sich dabei in den Erklärungen der Wissenschaftler:innen eines Faches wider. Das Erkenntnisinteresse der Fächer und das Vorgehen bei der Erkenntnisgewinnung prägen schließlich – wie im Zitat von Baumert und Kunter (2006) verkürzt enthalten – den Modus von fachlichen Erklärungen und die Gestalt des Unterrichts. Somit hat die Frage nach gutem instruktionalen Erklären im Fachunterricht auch eine wissenschaftspropädeutische Dimension und Bezüge zum naturwissenschaftsdidaktischen Nature-of-Science-Diskurs (Heering & Kremer, 2018).

Eine fachspezifische Betrachtung von instruktionalen Erklärungen bzw. von Kriterien guten Erklärens im Biologieunterricht ist daher neben einer fachübergreifenden Betrachtung notwendig. Dabei können – wie das bereits im vorherigen Abschnitt geschehen ist – wissenschaftstheoretische Überlegungen herangezogen werden (zu einer Reflexion über die Nützlichkeit wissenschaftstheoretischer Überlegungen für die Lehrkräftebildung siehe auch Kap. 2.3.2). Das Argument spiegelt auch das der Arbeit zugrunde liegende Verständnis zum Verhältnis von

Wissenschaftstheorie, wissenschaftlichen Erklärungen und Gestaltung instruktionaler Erklärungen im Fachunterricht wider: Wissenschaftstheoretiker:innen setzen sich aus einer analytischen Perspektive heraus mit der Struktur von wissenschaftlichen Aussagen bzw. Erklärungen des Faches auseinander. Sie beschreiben, was gültige und ungültige Erklärungen unterscheidet und somit als fachlich angemessene Erklärung in der wissenschaftlichen Disziplin gilt (Bartelborth, 2007; Dittmer, 2010; Schlosser & Weingarten, 2002). Da im Unterricht Wissen über Phänomene der Disziplin vermittelt werden soll und gleichzeitig die fachliche Korrektheit gewahrt bleiben muss, spiegeln instruktionale Erklärungen immer auch Formen fachlicher Erklärungen wider (siehe Kap. 2.3.1).

Die Diskussionen der Wissenschaftstheoretiker:innen können den Blick für Fachspezifika schärfen sowie Unterschiede zu anderen wissenschaftlichen Disziplinen markieren und damit eine fundierte fachspezifische Konzeptualisierung instruktionalen Erklärens ermöglichen. Wissenschaftstheoretische Überlegungen bilden für das Vorhaben der fachspezifischen Konzeptualisierung instruktionaler Erklärungen sowie der Bestimmung fachspezifischer Kriterien guten Erklärens im Fall der Biologie daher einen wichtigen Rahmen.

In Hinblick auf aktuelle Problemlagen wie den Klimawandel oder die Coronapandemie haben dabei insbesondere Überlegungen zur Komplexität von biologischen Phänomenen und zu dem Umgang mit Nichtwissen sowie unsicherem Wissen an Bedeutung gewonnen (Dittmer & Ehras, 2024; Grotzer, 2012; Mehren, Mehren, Ohl & Resenberger, 2015). Die wachsende Relevanz für ein Verständnis von Komplexität sowie der angemessenen Kommunikation von Nichtwissen und unsicherem Wissen steckt den inhaltlichen Rahmen dieser Arbeit.

1.3 Aufbau der Arbeit

Die hier vorgestellte Forschungsarbeit ist im Rahmen des interdisziplinären Forschungsprojekts FALKE³ (FACHspezifische LEhrkräfteKompetenzen im Erklären), das wiederum Teil des Projekts KOLEG (KOoperative LEhrkräftebildung Gestalten) ist, an der Universität Regensburg entstanden. Das Projekt KOLEG wurde 2015 bis 2018 im Rahmen der Qualitätsoffensive Lehrerbildung vom Bundesministerium für Bildung und Forschung gefördert (Förderkennzeichen FKZ 01JA1512). Innerhalb des FALKE-Projekts kooperierten elf Fachdidaktiken (Biologie, Chemie, Deutsch, Englisch, Geschichte, Mathematik, Physik, Grundschulpädagogik, evangelische Religionspädagogik, Musikpädagogik, Bildende Kunst und Ästhetische Erziehung) sowie die deutsche Sprachwissenschaft, die Sprecherziehung und die Professur für Methoden der empirischen Bildungsforschung der Universität Regensburg. Da die intensive, diskursive Zusammenarbeit und das Ziel eines fachübergreifend analogen Studiendesigns zu projektübergreifenden Entscheidungen geführt haben, wird in der Arbeit an einigen Stellen auf das Projekt FALKE verwiesen und damit werden die gemeinsamen Entscheidungen respektive Ergebnisse kenntlich gemacht.⁴

³ Das Forschungsprojekt FALKE (Fachspezifische Lehrerkompetenzen im Erklären) wurde 2022 in FALKE-q (Fachspezifische Lehrkräftekompetenzen im Erklären – Quality) umbenannt, sodass die inhaltliche Schwerpunktsetzung der mittlerweile drei FALKE-Maßnahmen in den Akronymen deutlich wird.

⁴ Neben dem gemeinsamen Studiendesign bzw. der gemeinsam durchgeführten interdisziplinären Studie wurden in der Projektlaufzeit auch sogenannte Tandemseminare konzipiert und realisiert (Asen-Molz, Ehras & Krauss, im Druck; Ehras, Asen-Molz, Frei, Schilcher & Krauss, 2021). Dabei kooperierten jeweils mindestens zwei am Projekt beteiligte Fächer für ein Semester. Die Seminare für Lehramtsstudierende fokussierten dabei auf den Aufbau von Erklärkompetenz durch praktische Übungen zum Erklären. Diese wurden videografiert und anschließend durch Feedback und Reflexionsschleifen nach theoretischen Inputs in mehreren Schritten überarbeitet. Dabei ermöglichten die Kooperationen zum einen den Vergleich von instruktionalen Erklärungen aus verschiedenen Unterrichtsfächern, sodass fachspezifische Unterschiede im Kontrast sichtbar werden konnten, sowie zum anderen das Feedback von Personen, die weniger Fachwissen aufweisen,

Um Kriterien guten instruktionalen Erklärens im Biologieunterricht unter Berücksichtigung verschiedener Perspektiven zu bestimmen ist die vorliegende Arbeit wie folgt aufgebaut:

Im ersten Teil der Arbeit (Kapitel 2) werden konzeptionelle Überlegungen zum Begriff des instruktionalen Erklärens vorgestellt und damit der Frage nachgegangen, was instruktionales Erklären grundlegend kennzeichnet. Dazu wird eine diskursorientierte Struktur verfolgt, sodass zunächst eine Betrachtung aus sprachtheoretischer Perspektive erfolgt, die allgemein die Stellung der beteiligten Personen sowie den Ablauf der Sprachhandlung des Erklärens in verschiedenen Kontexten untersucht (Kap. 2.1). Darauffolgend wird eine didaktische bzw. lehr-lerntheoretische Sicht eingenommen (Kap. 2.2), die das Erklären insbesondere in Hinblick auf den damit intendierten Lernprozess und das Gelingen dieses Vermittlungsprozesses beschreibt. Für eine biologiedidaktische Spezifizierung des bis dorthin allgemein betrachteten instruktionalen Erklärens werden im Anschluss wissenschaftstheoretische Überlegungen zu inhaltlich differenzierbaren Formen von Erklärungen in der Fachwissenschaft Biologie vorgestellt (Kap. 2.3). In diesem Abschnitt wird auch dargestellt, dass die wissenschaftstheoretische Sicht auf Erklärungen für ein Verständnis des Erklärens im Biologieunterricht nützlich bzw. sogar notwendig ist. Kapitel 2 schließt mit einer Synthese der drei Perspektiven (sprachwissenschaftliche Perspektive, lerntheoretische bzw. didaktische Perspektive und wissenschaftstheoretische Perspektive), die das grundlegende Verständnis des instruktionalen Erklärens für diese Arbeit bildet: das instruktionale Erklären als ein vielgestaltiger, komplexer Kommunikationsprozess in dem Ursache-Wirkungs-Beziehungen, Mechanismen, Funktionen, naturhistorische Entwicklungen, Begriffe, Handlungen oder Zwecke von Inhalten für das Lernen thematisiert werden können, wobei das Ziel verfolgt wird, dass Rezipient:innen Wissen erwerben oder ein Verstehensprozess bzw. ein Fähigkeitserwerb initiiert wird.

Mit dem Ziel einer fachspezifischen Ausrichtung der Konzeptualisierung des instruktionalen Erklärens widmet sich Kapitel 3 einem Merkmal biologischer Phänomene, das die Erklärungen im Biologieunterricht formt und Schwierigkeit für Erklärende und Rezipient:innen bedingen kann: Biologische Phänomene und damit Erklärgegenstände des Biologieunterrichts sind häufig komplex. Das bedeutet unter anderem, dass für die Erklärung des komplexen Phänomens keine einfache Ursache-Wirkungs-Beziehung angeführt werden kann, sondern ein multikausales Ursache-Wirkungs-Gefüge rekonstruiert werden muss. Bei dem Versuch, komplexe Ursache-Wirkungs-Gefüge aufzuklären, bleibt häufig ein Anteil an Nichtwissen und unsicherem Wissen enthalten. Auch hierfür wird zunächst auf wissenschaftstheoretische Literatur Bezug genommen, um den Begriff der Komplexität und die Phänomene ‚Nichtwissen und unsicheres Wissen‘ genauer zu bestimmen. Anschließend werden ab Kapitel 3.4 die Erkenntnisse mit Blick auf den Unterricht reflektiert. Dazu werden zunächst in Kapitel 3.4 die Herausforderungen bei der Vermittlung komplexer Phänomene – also aufseiten der Lehrkräfte – und beim Verstehen komplexer Phänomene – also aufseiten der Lernenden – betrachtet. Daraufhin wird der Blick auf das instruktionale Erklären geweitet und der Versuch unternommen, den Stellenwert der Aspekte ‚Komplexität‘ und ‚Nichtwissen‘ bzw. ‚unsicheres Wissen‘ im Biologieunterricht anhand der Bildungsstandards für das Fach Biologie abzuschätzen sowie die Bedeutung eines Verständnisses für die Bildung von Schüler:innen im Sinne einer Bildung zu Bürger:innen, die an öffentlichen Diskussionen zu naturwissenschaftlichen Fragen mündig teilhaben können, zu entfalten. Kapitel 3.6. enthält schließlich die Zusammenfassung der Erkenntnisse der vorhergehenden Abschnitte.

sodass durch die instruktionale Erklärung tatsächlich neues Wissen vermittelt werden konnte, wobei die fachfremden Rezipient:innen eher in der Lage sind, zum Verständnis des Inhalts Rückmeldung zu geben.

Kapitel 4 stellt bisherige Ergebnisse und Befunde zu Kriterien guten Erklärens zusammen. Dabei sind die Abschnitte nach sechs Aspekten guten Erklärens gegliedert, die sich aus der Literatur kategorisieren lassen: der Adressatenorientierung, der Strukturiertheit einer Erklärung, dem Einsatz von Repräsentationen in einer Erklärung, der Verständlichkeit der Sprache, der Gestaltung des Sprech- und Körperausdrucks sowie den fachlichen Aspekten bzw. der fachlichen Angemessenheit der Erklärung (Kap. 4.1 bis 4.6). Im darauffolgenden Unterkapitel (Kap. 4.7) werden schließlich die Ergebnisse der Literaturrecherche zusammengefasst und in einer tabellarischen Übersicht dargestellt. Im Spiegel der Literatursynopse wird abschließend der Blick von den Kriterien guten Erklärens auf die Lehrkräfte, die eben solche guten, instruktionalen Erklärungen produzieren sollen, sowie den Prozess des Erklärens im Unterricht gerichtet. Im letzten Unterkapitel (Kap. 4.8) wird der Frage nachgegangen, inwiefern die Aspekte guten Erklärens von der befragten Personengruppe abhängig sind, um Vorannahmen über erwartbare Tendenzen der Perspektiven zu erhalten.

Kapitel 5 leitet zum empirischen Teil der Arbeit über: Vor dem theoretischen Hintergrund werden vier Fragestellungen abgeleitet (Kap. 5.1): Welche Kriterien werden zur Bewertung instruktionaler Erklärung herangezogen? Wie unterscheiden sich die Bewertungen instruktionaler Erklärungen durch verschiedene Statusgruppen (Schüler:innen, Studierende, Lehrkräfte, Didaktiker:innen) in Hinblick auf vorgegebene Kriterien? Welche Kriterien sind für die Bewertung von instruktionalen Erklärungen große Einflussfaktoren? Wie unterscheiden sich die Bewertungen von komplexen und linearen instruktionalen Erklärungen in Abhängigkeit von verschiedenen Statusgruppen? Anschließend (Kap. 5.2) wird das Untersuchungsdesign vorgestellt: Mit einem computerbasierten Fragebogen, der sechs Videovignetten von instruktionalen Erklärungen einer Biologielehrkraft enthält, wurde eine Stichprobe bestehend aus Schüler:innen, Studierende, Lehrkräfte und Didaktiker:innen nach der Güte der Erklärung befragt. Der Fragebogen ist so aufgebaut, dass zunächst eine globale Einschätzung der Güte der Erklärung ohne Vorgabe von Kriterien anhand einer sechsstufigen Skala erfolgt sowie eine wörtliche Rückmeldung in einem offenen Textfeld, die diese Einschätzung begründet, abgegeben werden kann. Erst in einem nächsten Schritt werden geschlossene Items zur Einschätzung ausgewählter Aspekte bzw. Kriterien auf einer sechsstufigen Skala vorgelegt. Neben der Beschreibung des Fragebogens (Kap. 5.2.1) und der Videovignetten (Kap. 5.2.2) finden sich hier auch die Darstellung der Triangulation (Methodentriangulation und Datentriangulation; Kap. 5.2.3), der Konzeptualisierung und der Operationalisierung der im Fragebogen repräsentierten Merkmale guten Erklärens (Kap. 5.2.4), des Umgangs mit Störvariablen (Reihenfolge der Videovignetten, Wirkung der erklärenden Person, Vorwissen; Kap. 5.2.5), der Pilotierung mittels kognitiver Interviews (Kap. 5.2.6) sowie die Stichprobenbeschreibung (Kap. 5.2.7). Das Kapitel 5.3 macht das Vorgehen mit dem gewonnenen Datensatz transparent: Zum einen werden die notwendigen Aufbereitungen der Daten für die späteren quantitativen Analysen beschrieben (z. B. Rekodierungen, Vorgehen zur Bildung von Skalenmittelwerten, Umstrukturierung des Datensatzes). Zum anderen werden die zur Auswertung der Begründungen zur Vergabe des Globalurteils aus dem offenen Textfeld gewählten Methoden beschrieben und begründet (inhaltlich-strukturierende und evaluative qualitative Inhaltsanalyse; Kuckartz, 2018). Das Methodenkapitel wird von der Darstellung der erarbeiteten Kategoriensystems für die Auswertung der wörtlichen Rückmeldungen abgeschlossen.

Kapitel 6 enthält die empirischen Ergebnisse der Arbeit: Im ersten Unterkapitel (6.1) werden zunächst die Gütekriterien der quantitativen (Kap. 6.1.2) und qualitativen (in Anlehnung an Steinke (2012) und Kuckartz (2018); Kap. 6.1.3) Forschungszugänge reflektiert. Anschließend finden sich die Ergebnisse der konfundierenden Einflussgrößen und Folgerungen für die Auswertung der Daten

(Kap. 6.1.1). Im darauffolgenden Unterkapitel (Kap. 6.2) werden die Studienergebnisse dargestellt. Diese sind nach den Forschungsfragen strukturiert (Kap. 6.2.2 bis 6.2.3).

Kapitel 7.1 dient der Zusammenfassung zentraler Befunde und deren Diskussion. Anschließend werden Limitationen in Bezug auf die Forschungsmethodik sowie Grenzen in Bezug auf mögliche Aussagen anhand der Daten reflektiert und es wird ein Ausblick auf weiteren Forschungsbedarf gegeben (Kap. 7.2). Im letzten Kapitel (Kap. 7.3) werden unter Rückgriff auf den theoretischen Hintergrund und die Befunde dieser Studie die Relevanz für die schulische Praxis sowie die Aus- und Weiterbildung von Lehrkräften skizziert.

2 Instruktionale Erklärungen im Biologieunterricht

Wie bereits in Kapitel 1 beschrieben, stehen in dieser Arbeit instruktionale Erklärungen einer Lehrkraft – also Erklären im Sinne von kürzeren Darlegungen, die im unterrichtlichen Kontext von Lehrkräften für Schüler:innen gegeben werden – im Fokus. Um zu klären, was allgemein unter instruktionalem Erklären im Unterricht und im Speziellen unter instruktionalem Erklären im Biologieunterricht verstanden werden kann, werden drei Disziplinen als Bezugspunkte herangezogen: erstens die Sprachwissenschaft, die Erklärsituationen beispielsweise in Hinblick auf die sprachliche Gestaltung des Kommunikationsprozesses, den Ablauf der Interaktionen der Kommunikationspartner:innen oder die Abgrenzung des Erklärens von anderen Sprachhandlungen wie dem Beschreiben, Begründen und Definieren untersucht (z. B. Ehlich, 2009; Hohenstein, 2006; Klein, 2009), zweitens die Didaktik, die sich mit dem Erklären in Hinblick auf eine gelingende Gestaltung von Lern- und Bildungsprozessen beschäftigt und die Bedeutung des Erklärens für fachliche Lernprozesse in den Blick nimmt (z. B. Kiel, 1999; Larreamendy-Joerns & Muñoz, 2010; Wittwer & Renkl, 2008), und drittens Überlegungen aus der Wissenschaftstheorie der Biologie. Sie beschäftigt sich mit der Analyse fachspezifischer Formen⁵ wissenschaftlicher Erklärungen sowie mit der Frage nach der Gültigkeit von Erklärungen in der Disziplin der Biologie⁶ (z. B. Hempel & Oppenheim, 1948; Krohs, 2004; Mayr, 2002).

Dabei wird bereits deutlich, dass sich die sprachwissenschaftliche und die didaktische Perspektive auf den kommunikativen Prozess des Erklärens beziehen, der ein verständlich machen, Verstehen oder einen Wissensaufbau bei Adressat:innen zum Ziel hat (z. B. Hohenstein, 2006; Kiel, 1999). Im Gegensatz dazu fokussiert die wissenschaftstheoretische Perspektive das Produkt eines Forschungsprozesses, durch den neues Wissen generiert wurde – eine wissenschaftliche Erklärung (Schilcher et al., 2017; Schlosser & Weingarten, 2002; Strube, 1986). Schilcher und Kollegen (2017) gehen auf die Unterscheidung zwischen wissenschaftlichen Erklärungen und unterrichtlichem Erklären ein und machen dies an insgesamt acht Kategorien, unter anderem dem dafür einsetzbaren Synonym, deutlich: Während unterrichtliches Erklären mit „vermitteln“ oder „verständlich machen“ (Schilcher et al. 2017, S. 442) gleichgesetzt werden kann, sind wissenschaftliche Erklärungen Schlussfolgerungen in Form logischer Argumente. Strube (1986), der sich mit dieser Unterscheidung

⁵ Mit ‚Form einer fachlichen Erklärung‘ oder ‚Form einer instruktionalen Erklärung‘ wird im Folgenden stets eine Einteilung bezeichnet, deren Differenzierungsmerkmal sich auf den Inhalt bezieht: zum Beispiel ob ein Mechanismus, eine Ursache-Wirkungs-Beziehung oder eine Funktion erklärt wird.

⁶ Die Gültigkeit von Erklärungen wird daran gemessen, ob das Aussagensystem logisch und widerspruchsfrei ist.

in einem Text mit dem Titel „explanation and explaining in science education“ beschäftigt, spricht von einer Struktur im Gegensatz zu einem Akt bzw. einer Handlung:

[...] an explanation here is taken to be an argument crafted to represent the current best view of the question under consideration. For example, if someone makes the request – “Why does an iron rod expand upon heating? – the explanation will consist of statements from physics representing the best current views from physical theory. [...] Therefore an explanation is seen to be a structure, if you like, and the statements and their connections can be written and displayed independently of any person providing the explanation. [...] Unlike an explanation, explaining is an act – it is the act of getting someone else to understand something. This activity is not necessarily discipline-bound, but is primarily an educational activity. (Strube, 1986, S. 135–136)

Das Erklären im Unterricht kann gewissermaßen als Überschneidungsbereich der drei disziplinären Perspektiven (Sprachwissenschaft, Didaktik, Wissenschaftstheorie) betrachtet werden, da es sich als *sprachlich basierter Kommunikationsakt* zwischen Lehrkraft und Schüler:innen *in einer Lernsituation* über einen *fachlichen Gegenstand* oder eine fachliche Fragestellung darstellt.⁷ Die wissenschaftstheoretische Perspektive – auch wenn sie sich nicht mit der Vermittlung von Inhalten beschäftigt, sondern Erklärungen von Biolog:innen in den Blick nimmt – kann dabei für die Begriffsbestimmung des instruktionalen Erklärens im Biologieunterricht Ansatzpunkte für die fachliche Spezifizierung des allgemeinen Begriffs insbesondere in Hinblick auf inhaltlich differenzierbare Formen liefern.⁸ Die folgende Tabelle (Tab. 1) greift die relevanten Unterscheidungen der disziplinären Perspektiven in Hinblick auf die jeweiligen Forschungsschwerpunkte im Kontext von (instruktionalen) Erklärungen, den Forschungsgegenständen, die die empirischen Daten dafür liefern, sowie das grundlegende Verständnis des Begriffs ‚Erklärung‘ in Bezug auf die Unterscheidung von Strube (1986) und Schilcher et al. (2017) noch einmal auf.

Tabelle 1: Relevante Unterschiede der drei Disziplinen Sprachwissenschaft, Didaktik und Wissenschaftstheorie der Biologie in Hinblick auf (instruktionale) Erklärungen

Disziplin	Forschungsgegenstände	Forschungsschwerpunkte	Verständnis des Begriffs ‚Erklärung‘
Sprachwissenschaft	Erklärungen und Erklärprozesse aus dem Alltag und aus Lehr-Lern-Kontexten	Analyse der Sprache und der Kommunikationssituation	Erklärungen zum Zweck der Weitergabe von bestehendem Wissen oder dem Erzeugen von
Didaktik	Erklärungen und Erklärprozesse aus Lehr-Lern-Kontexten	Analyse von Gelingensbedingungen und der Relevanz von Erklärungen und Erklärprozessen für das Lernen	Verstehen bei den Rezipient:innen (Erklärung als Produkt von Erklärprozessen; Erklären als Akt)
Wissenschaftstheorie der Biologie	Erklärungen von biologischen Phänomenen	Analyse formaler Strukturen und der Gültigkeit von Erklärungen	Erklärungen als Resultat von Forschung (Erklärungen als Struktur)

⁷ In der unterrichtlichen Realisation von Erklären und für die Frage, was gutes Erklären kennzeichnet, ist auch die sprachwissenschaftliche Perspektive von Bedeutung. Für ein Begriffsverständnis respektive eine Konzeptualisierung tritt sie jedoch in den Hintergrund und wird hier deshalb nicht aufgegriffen.

⁸ Die Überlegungen aus der Wissenschaftstheorie zu gültigen Formen von Erklärungen der Disziplin werden darüber hinaus in Kapitel 4.6 auch zur Formulierung von Kriterien im Sinne der fachlichen Angemessenheit von instruktionalen Erklärungen genutzt.

Die drei disziplinären Perspektiven mit ihren unterschiedlichen Fokussen werden in den folgenden Kapiteln aufgegriffen und in Hinblick auf ihre Ergebnisse zur Begriffsbestimmung und Beschreibung des instruktionalen Erklärens im Biologieunterricht herangezogen. Vor allem folgende Fragen leiten die Literatursynopse zum instruktionalen Erklären im Biologieunterricht an:

- 1) Welche Merkmale charakterisieren instruktionale Erklärungen und wie können instruktionale Erklärungen von anderen vermittelnden Tätigkeiten von Lehrkräften abgegrenzt werden?
- 2) Welche inhaltlich differenzierbare Formen instruktionaler Erklärungen und Formen wissenschaftlicher Erklärungen können unterschieden werden?
- 3) Welche Funktionen werden instruktionalen Erklärungen zugeschrieben und welche Ziele werden damit verfolgt?

Die disziplinspezifischen Darstellungen in den Kapiteln 2.1 bis 2.3 werden in Kapitel 2.4 zu einem für den Biologieunterricht sowie die biologiedidaktische Forschung nützlichen Begriff zusammengeführt.

2.1 Erklären aus sprachwissenschaftlicher Perspektive: konversationsanalytische Sicht und funktionale Pragmatik

Grundlegende Aspekte zur Begriffsbestimmung des instruktionalen Erklärens sind zunächst in der Linguistik und der Diskursanalyse (als Teildisziplinen der Sprachwissenschaft) zu finden. Aus linguistischer Sicht wird das Verb ‚erklären‘ in der Argumentationsstruktur ‚jemand erklärt jemandem etwas‘ verwendet. Hohenstein (2006) expliziert, dass dies auf eine handelnde Person (im Satzgefüge als Nominativ) und eine empfangende Person (im Satzgefüge als Dativ) verweist sowie einen Handlungsgegenstand benötigt (im Satzgefüge als Akkusativobjekt; Hohenstein, 2006). Weiterhin stellt die Autorin fest, dass es keinen expliziten Verweis darauf gibt, dass das Erklären mit sprachlichen Mitteln realisiert werden muss, es jedoch wohl unmöglich ist, „**allein durch Zeigen**, durch Verbildlichung, graphische Darstellung, Gestik oder nonverbale Kommunikation eine **Erklärung durchzuführen**“ (Hohenstein, 2006, S. 99).

Die propositionale Struktur⁹ umfasst beim Erklären mindestens zwei Propositionen: eine, die sich auf den zu erklärenden Sachverhalt (Explanandum, das zu Erklärende) bezieht, und eine, die erklärend wirkt (Explanans, das Erklärende; Klein, 2009).¹⁰ Ein Beispiel kann diese Struktur verdeutlichen: Dass die Blätter des Ahorns im Sommer grün erscheinen (Explanandum), kann man durch das Absorptionsspektrum des Chlorophylls im Blatt (Explanans) erklären. Viele weitere Sprachhandlungen (z. B. Versprechen, Behaupten) beruhen hingegen auf einer einzigen Proposition.

Aus diskursanalytischer Sicht ist Erklären allgemein ein komplexer Kommunikationsprozess, der zwischen mindestens zwei Gesprächspartner:innen stattfindet (Ehlich, 2009; Hohenstein, 2006). Er ist weiterhin durch eine Divergenz in der mentalen Repräsentation zwischen den beteiligten Kommunikationspartner:innen (Ehlich, 2009) bzw. durch eine asymmetrische Wissensverteilung

⁹ Proposition wird als Begriff unter anderem in der Satzsemantik verwendet und bezeichnet hier eine Äußerung bzw. einen abgrenzbaren Inhalt. Dieser kann als wahr oder falsch beurteilt werden.

¹⁰ Die Begriffe Explanandum und Explanans wurden von Hempel und Oppenheim (1948) für die Beschreibung des deduktiv-nomologischen Erklärungsmodells eingeführt (siehe Kapitel 3.3.1). Die aus der Wissenschaftstheorie stammenden Begriffe werden jedoch auch in anderen Disziplinen gebraucht (siehe z. B. Klein, 2009; Morek, 2012; Wörn, 2014)

gekennzeichnet (Hohenstein, 2006). Eine Offenlegung des Wissensdefizits erfolgt durch die Zuhörer:innen selbst oder durch eine Annahme des Redners oder der Rednerin über ein Wissensdefizit bei den Adressat:innen (Hohenstein, 2006; Kotthoff, 2009). Dadurch ergibt sich wiederum eine typische Rollenverteilung zwischen nicht wissender und erklärender Person (Kotthoff, 2009).¹¹

2.1.1 Inhalte und Typen instruktionaler Erklärungen aus sprachwissenschaftlicher Perspektive

Zum Inhalt einer Erklärung wird von Becker-Mrotzek (2004) **Folgendes beschrieben**: „Ein zentrales gemeinsames Merkmal der erklärten Aspekte ist ihre Nichtwahrnehmbarkeit an der Oberfläche. Der kommunikative Zweck des Erklärens besteht darin, diese nicht sichtbaren Aspekte des Sachverhalts **sprachlich darzustellen**“ (Becker-Mrotzek, 2004, S. 183).

Theoretische Überlegungen zur Abgrenzung von anderen, teilweise recht ähnlichen Diskurs- und Handlungsmustern stammen vor allem aus Forschungsarbeiten, die mit dem Ansatz der funktionalen Pragmatik arbeiten. Das Erklären wird dabei – wie bereits in den obigen Abschnitten beschrieben – als Sprachhandlung aufgefasst, bei der ein Wissensdefizit zwischen Sprecher:in und Hörer:innen vorliegt, das beseitigt werden soll, wobei etwas thematisiert wird, das nicht unmittelbar wahrnehmbar ist (z. B. Zusammenhänge funktionaler oder kausaler Art), weshalb ein analytisches Durchringen sowie Abstraktion vom Sichtbaren notwendig sind (Becker-Mrotzek, 2004). Das Beschreiben wird hingegen als „ein „erfahrungsnahes“ sprachliches Muster“ (Morek, 2012, S. 30) charakterisiert, das dazu dient, den Hörer:innen etwas mental vorstellbar zu machen. Typischerweise kommt es aufgrund dieser Funktion oft eingeschoben in Erklärungen und Instruktionen vor (Becker-Mrotzek, 2004). Für das Erläutern, das in der Alltagssprache häufig synonym mit ‚Erklären‘ gebraucht wird, bestehen auch in der theoretischen Abgrenzung nur graduelle Unterschiede (Morek, 2012). Beim Erläutern wird Wissen versprachlicht, das den Hörer:innen helfen soll, eine Lücke zwischen bereits verbalisiertem Wissen zu schließen, beispielsweise in Form einer Begriffserläuterung (Morek, 2012). Es dient demnach nicht einem systematischen Wissensaufbau. Weiterhin werden weniger abstrakte Zusammenhänge aufgegriffen, sondern eher Erfahrungswissen verbalisiert (Becker-Mrotzek, 2004). Zuletzt sei das Begründen erwähnt: Hier werden bei den Kommunikationspartner:innen vorhandene Wissens Elemente umstrukturiert, um Verstehen zu erzeugen (Hohenstein, 2006) oder die Hörer:innen zur Ausführung einer Handlung bewegen sollen (Becker-Mrotzek, 2004). Dabei werden Erkenntnisgründe – im Gegensatz zu Realgründen bzw. Ursachen beim Erklären – angeführt (Hohenstein, 2006).

Erklären tritt jedoch in Kommunikationssituationen häufig im Verbund mit diesen theoretisch differenzierbaren Sprachhandlungen auf, wodurch die Abgrenzung des Erklärens von den beschriebenen Sprachhandlungen in der Praxis meist schwerfällt. Beispielsweise analysiert Hohenstein (2006) an Beispielen des wissenschaftlichen Vortrags, dass das sprachliche Handlungsmuster des Erklärens nicht isoliert genutzt wird, sondern in seiner Realisation bzw. im

¹¹ Die Voraussetzung der asymmetrischen Wissensverteilung ist im Kontext des Unterrichts mitunter nicht in allen Erklärsituationen gegeben: Werden beispielsweise Schüler:innen von der Lehrkraft aufgefordert, einen Sachverhalt aus der vergangenen Stunde in der Funktion einer Zusammenfassung oder einer Wiederholung erneut zu erklären, oder sollen Schüler:innen im naturwissenschaftlichen Unterricht im Verlauf des angeleiteten Experimentierens Erklärungen zum vermuteten Ausgang des Experiments geben, tritt die erklärende Person als bereits oder noch nicht bzw. mehr oder weniger wissend auf. Die Person, an die die Erklärung gerichtet ist, ist hingegen stets wissend, sodass die charakteristische Rollenverteilung nicht gegeben ist. Die weitere formale Struktur einer Erklärung wird dabei beibehalten.

Verfolgen seines (institutionalisierten) Zwecks als erklärendes Handeln – also als funktional-pragmatisches Handlungsmuster – im Verbund mit weiteren Sprechhandlungen auftritt. Sie zieht das Fazit, es handele sich häufig „um komplexe, d. h. mehrfach zusammengesetzte Formen sprachlichen Handelns, bei denen Zusammenschlüsse von erklärenden mit beschreibenden, begründenden, erläuternden und illustrierenden (Teil-)Handlungen vorliegen“ (Hohenstein, 2006, S. 113). Diese (Teil-)Handlungen dienen der Unterstützung der Rezipient:innen beim Aufbau von Wissen, dem Verstehen bzw. der Sinnkonstruktion. Die Verbindung von grundsätzlich **eigenständigen Sprechhandlungen zum erklärenden Handeln nennt sie einen „Illokutionsverbund“** (Hohenstein, 2006, S. 106). Die Ergebnisse bezüglich der Charakterisierung des Erklärens als Sprechhandlung werden von weiteren Autor:innen aufgegriffen und in Datenanalysen angewandt (z. B. Kotthoff, 2009; Morek, 2012).

Stukenbrock beschreibt darüber hinaus sogenannte familienähnliche Aktivitäten (Stukenbrock, 2009). Die Autorin fasst darunter verbale und visuelle Ausdrucksformen, die vom Sprecher bzw. der Sprecherin miteinander verschränkt werden, um Verstehen zu erzeugen (Stukenbrock, 2009). Die empirische Datenbasis der Studie legt nahe, „**dass Erklären, Zeigen und Demonstrieren keine diskreten Aktivitätskategorien mit scharfen Grenzen bilden, sondern vielmehr durch Familienähnlichkeiten miteinander verbunden sind und in der emergierenden Interaktion fließend ineinander übergehen können**“ (Stukenbrock, 2009, S. 173).

Ein allgemeines charakteristisches Aufbauprinzip von instruktionalen Erklärungen lässt sich nicht beschreiben (Hohenstein, 2006; Klein, 2009; Morek, 2012). Dies könne „**angesichts der Vielfalt möglicher Erklärgegenstände und der unter Umständen großen inhaltlichen Komplexität eines Explanandums kaum überraschen**“, konstatiert Morek (2012, S. 28 f.). Auch die Verbindung mit weiteren Sprechhandlungen trägt zur Komplexität und damit zu einer nicht verallgemeinerbaren Beschreibung des Aufbaus bei (Klein, 2009).

Unter Bezug auf die Sprechakttheorie¹² erarbeitet Klein (2009) eine Typologie, die neben der inhaltlichen Ebene der Erklärung auch die syntaktische sowie die pragmatische Ebene von Sprache mit einbezieht und damit eine typologische Unterscheidung der Sprechakte des Erklärens leisten kann. Mit dem vereinenden Merkmal, **dem pragmatischen Aspekt „Fraglichkeit zu beseitigen“** (Klein, 2009, S. 26), können auf syntaktischer Ebene drei Typen unterschieden werden, die durch unterschiedliche Fragesatzstrukturen erkennbar werden:

- ERKLÄREN-WAS:¹³ Was/Wer ist x?/Was bedeutet x?/...
- ERKLÄREN-WIE: Wie funktioniert y?/Wie führt man y durch?/...
- ERKLÄREN-WARUM: Warum ist z der Fall?/Warum hat C h getan?/...

Die semantischen Unterschiede beschreibt Klein (2009) **folgendermaßen**: „Die Was-Frage – und ähnlich die Wer-Frage – ist auf die konstitutiven und/oder sinngebenden Eigenschaften des zu

¹² Die Sprechakttheorie entstand als eine Antwort auf die Ende der 1960er Jahre geforderte Orientierung der linguistischen Methoden und Wissenschaften an relevanten Gesellschaftsthemen und weniger an der bloßen Linguistik (=Pragmatische Wende). „Die Ergebnisse der Forschung sollten helfen, ihre Beziehungen zur Gesellschaft und die Funktionsweisen in der gesellschaftlichen Ordnung besser zu verstehen“ (Ernst, 2011, S.238). Dabei wird davon ausgegangen, dass Sprache nicht nur eine Abbildfunktion erfüllt, sondern auch eine sogenannte Sprachhandlungsfunktion aufweist. Eine Sprachhandlung bezeichnet dabei eine Handlung, die allein durch Sprechen hergestellt werden kann (z. B. „Ich taufe dich auf den Namen ... [..]Ich vermache dir mein ...“; Ernst, 2011, S. 242). Der Sprechakt gliedert sich dabei in den lokutionären Akt (lautliche Produktion von Sprache), den illokutionären Akt (Handlungsabsicht einer Äußerung) und den perlokutionären Akt (die Folgewirkung bzw. der Zweck der Äußerung).

¹³ Die Großschreibung ist eine Konvention der Linguistik, um die Begriffe der linguistisch definierten Sprachhandlungen von der Nutzung der identischen Worte als Verben in der Alltagssprache abgrenzen zu können (Klein, 2009).

Erklärenden gerichtet, die *Wie*-Frage auf die Modalität von Prozessen, insbesondere Handlungen, und die *Warum*-Frage auf kausale Faktoren (Ursachen, Gründe)" (Klein, 2009, S. 26).

2.1.2 Ziele und Funktionen instruktionaler Erklärungen aus sprachwissenschaftlicher Perspektive

Über das Ziel einer Erklärung besteht in der Literatur Uneinigkeit bzw. es werden unterschiedliche Ziele genannt. Hohenstein (2006) stellt den strukturierten Wissensaufbau und Wissensausbau als zentrales Ziel des Erklärens dar. Noch stärker auf die aktive Rolle der Rezipient:innen verweisend postuliert Ehlich (2009), dass das Verstehen Ziel des Erklärens ist:

[...] das Verstehen als eine mentale Aktivität des Hörers, ist nun aber der zentrale Zielpunkt des Erklärens. Dieses Verstehen in einem substanziellen Sinn ist zugleich eine zentrale mentale Aktivität, deren Ergebnisse in H [(= Hörer:in)] charakteristische Veränderungen von dessen mentalem Bereich bedeuten. (Ehlich, 2009, S. 16)

Dabei soll die Abweichung der mentalen Repräsentation der Wirklichkeit zwischen Sprecher:in und Hörer:in durch sprachlichen Austausch überwunden werden. Auch Klein (2009) bezieht sich bei der Zielformulierung des Erklärens auf das Verstehen – in Paraphrasierungen an anderer Stelle auch auf die Beseitigung von Fraglichkeit und das Schaffen von Klarheit über Zusammenhänge im zu erklärenden Sachverhalt. Ehlich (2009) führt gleichzeitig mit dem Verstehen die Sinnübertragung von Hörer:innen auf Zuhörer:innen bzw. die Sinnkonstruktion aufseiten der Zuhörer:innen als Ziel einer Erklärung ein. Den Begriff ‚Sinn‘ definiert er als eine Rekonstruktion der Strukturiertheit der Welt.

Mit der aus der Sprechakttheorie stammenden Unterscheidung von Illokution (Absicht der Äußerung) und Perlokution (tatsächliche Wirkung der Äußerung) kann die Sprechhandlung des Erklärens weiter beschrieben werden: Das Erklären zielt zwar auf eine Wirkung – auf Verstehen, Sinnkonstruktion oder Wissensaufbau – und damit auf einen perlokutionären Effekt, wird jedoch im Deutschen als illokutives Verb gebraucht und verstanden (Klein, 2009). Klein formuliert hierzu prägnant: „Die Semantik von ERKLÄREN legt zwar eine Bringschuld auf Seiten des Erklärenden nahe, beinhaltet aber bei Misserfolg kein sprachliches Verwendungsverbot“ (Klein, 2009, S. 27). Die perlokutive Wirkung zu erreichen wird von ihm weiterhin als komplex beschrieben – beispielsweise im Gegensatz zum Beleidigen –, da sie nicht allein durch die Befolgung und Kenntnis sprachlicher Regeln und kommunikativer Normen erreichbar ist (Klein, 2009).

2.2 Lerntheoretische und didaktische Perspektive auf instruktionales Erklären

2.2.1 Ziele und Funktionen instruktionaler Erklärungen im Biologieunterricht aus didaktischer Perspektive

Die in der didaktischen Literatur aufgegriffenen Ziele, die mit unterrichtlichem Erklären verbunden werden, sind ebenso vielfältig, wie dies in der sprachwissenschaftlichen Betrachtung herausgearbeitet wurde. Ganz allgemein wird ihnen eine Funktion beim Lernen zugeschrieben (Leinhardt, 1990; Wittwer & Renkl, 2008). Leinhardt (1990) sieht im Erklären eine Möglichkeit des pädagogischen Handelns, mit dem der Versuch unternommen werden kann, Fragen der Schüler:innen zu klären. Diese Fragen können sowohl explizit von den Lernenden geäußerte als auch von der Lehrkraft antizipierte Fragen sein (Wagner & Wörn, 2011). Sie dienen damit dem Lernen

der Schüler:innen (Leinhardt, 1990). Kiel (2009) betrachtet es in seiner umfassenden theoretischen Beschäftigung mit dem Erklären als Möglichkeit, kognitive Konflikte bei Kindern zu beseitigen. Kognitive Konflikte beschreibt der Autor dabei als Inkonsistenzen innerhalb der Kognitionen, also zwischen „Wissen, Glauben, Einstellungen“ (Kiel, 1999, S. 30) oder „Kognitionen und der Umweltsituation“ (Kiel, 1999, S. 30). Die erfahrenen Inkonsistenzen können sich sowohl auf „Aussagen, Theorien, Begriffe“ (Kiel, 1999, S. 29) als auch auf „Handlungen“ (Kiel, 1999, S. 29) beziehen. Die Definition des Erklärens und die darin enthaltenen Ziele des Erklärens der Autorinnen Wagner und Wörn (2011) verweisen hingegen auf einen Wissensaufbau und die Förderung von Verständnis: Das Erklären „hat ein Erwachsen von Wissen und/oder Verständnis durch Umorganisation und Intensivierung von bereits vorhandenem Wissen und/oder Verständnis zum Ziel“ (Wagner & Wörn, 2011, S.20). Renkl et al. (2007) sehen für Erklärungen auch das Potenzial, beim Aufbau kognitiver Fertigkeiten zu unterstützen. Am häufigsten wird insgesamt das Verstehen als Zieldimension von Erklären angegeben (Leinhardt, 1990; Leisen, 2013; Renkl et al., 2007; Treagust & Harrison, 2000; Wittwer & Renkl, 2008). Dabei formulieren Treagust und Harrison (2000): „[explanations] provide an avenue to enhance students’ understanding of scientific phenomena“ (Treagust & Harrison, 2000, S. 1157). Darauf verweisen auch Wagner und Wörn (2011), indem sie das Erklären als eine „auf Erfolg ausgerichtete [...] Handlung“ (Wagner & Wörn, 2011, S. 20) definieren, bei der der Wissensaufbau sowie das Verständnis der erklärten Inhalte aber nicht automatisch eintreten muss. Ähnlich wie aus sprachwissenschaftlicher Sicht mit der Unterscheidung zwischen Illokution und Perlokution wird damit auch aus didaktischer Sicht darauf verwiesen, dass ein Verstehen oder Wissensaufbau lediglich angebahnt werden kann. Grundlegende Annahme dafür ist, dass Wissen, Verständnis oder Sinn eines Lerngegenstandes nicht übertragen werden kann, sondern eigentätig konstruiert werden muss.

Der vorgestellte Kanon an Zielen, die auf einen kognitiven Lernfortschritt in Bezug auf einen konkreten Sachverhalt bzw. ein Phänomen ausgerichtet sind, wird von Larreamendy-Joerns und Muñoz (2010) erweitert: „Instructional explanations not only communicate content but also convey a sense of what disciplinary fields are, of how they are organized, and of what it takes to be a legitimate member in disciplinary communities“ (Larreamendy-Joerns & Muñoz, 2010, S. 23). Sie richten damit die Aufmerksamkeit auf die Möglichkeit, mit lehrerzentriertem Erklären implizit auch eine Vorstellung über das Wesen des Faches vermitteln zu können, da die Lehrkräfte modellieren, welche Fragen von Wissenschaftler:innen des Feldes gestellt werden und wie in der jeweiligen Disziplin gültige Erklärungen aussehen. Damit leisten fachliche Erklärungen von Lehrkräften auch einen Beitrag für die wissenschaftspropädeutische Bildung der Lernenden. Diese Funktion von Erklärungen ist insbesondere im Kontext der Diskussion um die Angemessenheit von instruktionalen Erklärungen interessant: Erklären sich die Lernenden Sachverhalte selbst, so ist diese Lernmöglichkeit zu adäquaten Erklärungen des Faches nicht mehr gegeben (Renkl, 2001) und Schüler:innen erklären sich die Phänomene mitunter anhand der von ihnen eher präferierten, alltäglichen bzw. fachlich weniger angemessenen Formen (z. B. Trommler, Gresch & Hammann, 2017). So werden von Schüler:innen häufig teleologische Erklärungsformen bevorzugt, obwohl eine Erklärung, die die Kausalität berücksichtigt, fachlich angemessen wäre (Trommler et al., 2017).

2.2.2 Inhalte und Typen instruktionaler Erklärungen aus didaktischer Perspektive

Auf Grundlage von Unterrichtsbeobachtungen kommen – ebenso wie Sprachwissenschaftler:innen – auch Autor:innen didaktischer Arbeiten zum Ergebnis, dass das instruktionale Erklären nicht mit einem uniformen detaillierten Schema beschrieben werden kann (Dagher & Cossman, 1992; Kiel,

1999; Leinhardt, 1990; Treagust & Harrison, 2000). Beispielsweise beeinflussen aufseiten der Lehrkräfte deren Wissen über den Erklärinhalt sowie Vorstellungen über die wissenschaftliche Disziplin, individuelle Präferenzen für Unterrichtsstile oder das persönliche Verhältnis zu den Schüler:innen der Klasse die Ausgestaltung der Erklärung (Treagust & Harrison, 2000). Aufseiten der Schüler:innen prägen das Vorwissen, Gruppendynamiken, kulturelle Einflüsse sowie Motivation und Interesse den Erklärprozess (Treagust & Harrison, 2000). Neben der konkreten Ausgangsfrage für die Erklärung spielen außerdem Kontextfaktoren wie verfügbare Zeit oder Veranstaltungsformat eine Rolle (Treagust & Harrison, 2000). Durch diese und weitere Faktoren wird die Konstruktion der Erklärung maßgeblich bestimmt. So kann das vorhandene Vorwissen beispielsweise Einfluss darauf nehmen, wie viele Einschübe in Form von Beschreibungen und Definitionen die Erklärsequenz aufweisen muss, um verständlich zu sein, oder die jeweiligen Interessen können mitbestimmen, welche Beispiele gewählt werden, um die Erklärung für die Zuhörer:innen interessant und motivierend zu gestalten. Dabei wird für die Bestimmung einer Erklärung als solche kein fester Katalog abgrenzbarer, formaler Kriterien herangezogen, sondern die Bestimmung zielorientiert – am Wissenserwerb oder Verstehen – vorgenommen. Die Zielorientierung als Kriterium der Abgrenzung einer unterrichtlichen Erklärung zu anderen Sprachhandlungen wird auch bei Larreamendy-Joerns und Muñoz (2010) deutlich: **„Instructional explanations can adopt the form of expositions, conversations, demonstrations, and narratives and make use of a variety of representations to capture critical aspects of subject matter knowledge“** (Larreamendy-Joerns & Muñoz, 2010, S. 23).

Für die Beschreibung von Typen des Erklärens in der didaktischen Literatur zu unterrichtlichen Erklärungen wird die in Kapitel 2.1 beschriebene Einteilung von Klein (2009) in ‚Was‘-, ‚Wie‘- und ‚Warum‘-Erklärungen, die er aufgrund sprachlicher Analysen differenziert, häufig aufgegriffen (z. B. Findeisen, 2017; Harren, 2009; Schopf & Zwischenbrugger, 2015; Wagner & Wörn, 2011). Eine formal ähnliche Einteilung stammt von Brown und Atkins (1986), die jedoch bei der Analyse nicht auf die Sprechakttheorie zurückgreifen, sondern die Erklärungen aus verschiedenen Kontexten und Professionen in Hinblick auf deren inhaltliche Ausrichtung kategorisieren. Sie beschreiben die Erklärungstypen des ‚reason-giving‘, die Warum-Fragen beantworten, ‚interpretive‘, die Was-Fragen beantworten, sowie ‚descriptive‘, die Wie-Fragen beantworten (Brown & Atkins, 1986). Die Dreiteilung ist hinreichend allgemein, sodass sie für alle Unterrichtsfächer eine Orientierung zur Beschreibung von Erklärungstypen bieten kann.

Aus Sicht der Wirtschaftspädagogik bzw. -didaktik formulieren Schopf und Zwischenbrugger (2015) sowie aus Sicht der Biologiedidaktik Dittmer und Langlet (im Druck), dass auch Wozu-Fragen und damit Wozu-Erklärungen als wichtiger Typus im Unterricht vorkommen sollten. Die Hintergründe der Autor:innen unterscheiden sich jedoch in ihrem Kern: Schopf und Zwischenbrugger (2015) bzw. die von ihnen in Expert:inneninterviews befragten Wirtschaftslehrenden der Universität Wien plädieren für ein Aufgreifen von **Wozu-Fragen und Erklärungen**, **„um die Bedeutung bzw. den Sinn und Zweck des neuen Inhalts aufzuzeigen“** (Schopf & Zwischenbrugger, 2015, S. 13). Damit kann ein Verständnis für den Anwendungsbereich des neuen Inhalts bzw. des erklärten Grundprinzips erzeugt werden. Die Darlegung des ‚Wozu‘ wird von den befragten Expert:innen als selten beschrieben (Schopf & Zwischenbrugger, 2015). Hingegen beziehen sich Dittmer und Langlet (im Druck) auf die Zweckmäßigkeit von Strukturen oder Verhaltensweisen, also die Fragen wozu es in der Natur etwas gibt oder wozu es entstanden ist. Damit richten sich diese Fragen auf Funktionen, die jedoch ohne die Annahme von Zielgerichtetheit beantwortet werden müssen (siehe auch Kap. 2.3.1 zu funktionalen bzw. teleologischen Erklärungen).

Von einigen Autor:innen wird die Formenvielfalt von Erklärungen und damit der Begriff des Erklärens noch etwas weiter gefasst. So formulieren beispielsweise Schilcher et al. (2017), „**dass Fachinhalte selbst als Erklärungen fungieren können** [...]“. So ist beispielsweise eine Definition in der Mathematik (‚ein Viereck mit zwei parallelen Seiten heißt Parallelogramm‘) immer auch ein ‚Erklären-WAS‘, ein Rechenalgorithmus immer ein ‚Erklären-WIE‘ und ein mathematischer Beweis immer ein ‚Erklären-WARUM“ (Schilcher et al., 2017, S. 443). Mit diesem weiten Verständnis werden die Unterscheidungen zwischen den Sprachhandlungen (z. B. Becker-Mrotzek, 2004; Hohenstein, 2006; Morek, 2012; siehe Kap. 2.1) – die von der KMK auch als Operatoren definiert wurden – weitestgehend aufgelöst und das Ziel des verständlich Machens tritt in den Vordergrund.

2.3 Wissenschaftstheoretische und fachliche Perspektiven auf Erklärungen

Erklärungen für beobachtete Phänomene zu finden, ist „**zentrales Ziel wissenschaftlicher Forschung**“ (Bartelborth, 2007, S. 5), denn dadurch wird neues Wissen generiert. Bis eine wissenschaftliche Erklärung für ein Phänomen gefunden ist, werden im Erkenntnisprozess viele Zwischenschritte wie das Aufstellen von Vermutungen, die Planung und Durchführung von Beobachtungen oder Experimenten sowie die Auswertung der gesammelten empirischen Daten geleistet. Dabei wird die Suche nach Erklärungen zum einen vom Interesse der Wissenschaftler:innen angetrieben, Phänomene zu verstehen und Zusammenhänge aufzudecken. Zum anderen kann die Möglichkeit zum Eingreifen in Entwicklungen (z. B. die (Um-)Gestaltung der Umwelt zur Erhöhung der Artenvielfalt; die Milderung des Symptomverlauf von Krankheiten) eine Motivation für die Suche nach Erklärungen sein (Bartelborth, 2007).

Welche Erklärungen von anderen Wissenschaftler:innen als gut – im Sinne von logisch und widerspruchsfrei – bewertet werden, hängt dabei unter anderem von formalen Kriterien ab. Die Wissenschaftstheorie der Biologie als Teildisziplin der Wissenschaftsphilosophie¹⁴ analysiert solche wissenschaftlichen Aussagesysteme wie Erklärungen in Hinblick auf ihre formale Beschaffenheiten und beschäftigt sich mit deren Genese, Etablierung, Reichweite und Gültigkeit (Dittmer, 2010).

Über Formen biologischer Erklärungen wird unter Wissenschaftstheoretiker:innen viel diskutiert und das Verständnis darüber unterliegt einem historischen Wandel. Im Folgenden sollen Ansätze zur Beschreibung und Kategorisierung biologischer Erklärungen skizziert werden, mit denen der Begriff des instruktionalen bzw. allgemein des unterrichtlichen Erklärens mit der fachspezifischen Perspektive erweitert werden kann. Dabei sollen die Darstellungen nicht dazu dienen, den aushandelnden, kommunikativen Prozess des unterrichtlichen Erklärens an die formale Struktur wissenschaftlicher Erklärungen anzugleichen. Da jedoch im Unterricht die in der Wissenschaft untersuchten Phänomene und Fragen thematisiert werden, wird dabei unumgänglich auch auf entsprechende Erklärungsformen der Disziplin zurückgegriffen. Ausgehend davon können die wissenschaftstheoretischen Überlegungen für das Vorhaben der Beschreibung instruktionaler Erklärungen im Biologieunterricht nützlich sein, indem sie einen Ausgangspunkt für eine Einteilung von inhaltlich differenzierbaren Formen bieten. Die systematisierten, wissenschaftstheoretischen Beschreibungen der zugrunde liegenden Erklärungsformen können deshalb auch als Anhaltspunkte

¹⁴ Die Wissenschaftsphilosophie stellt sich allgemein als ein heterogenes Feld verschiedener Forschungsbereiche und Disziplinen dar, die sich vor allem mit der Reflexion über das grundsätzliche Wesen, angewandte und entwickelte Konzepte sowie die Bedeutung der jeweiligen Disziplin beschäftigen (Dittmer, 2010).

für die Beurteilung fachlicher Angemessenheit von Erklärungen im Biologieunterricht dienen bzw. dafür sogar notwendig sein.¹⁵

2.3.1 Wissenschaftliche Erklärungsformen

Es gibt keinen einheitlichen Ansatz bzw. kein einheitliches Modell für die Beschreibung der Form aller wissenschaftlicher Erklärungen (Schlosser & Weingarten, 2002; Potochnik, 2013). Hingegen ist ein historisch gewachsener Diskurs zu finden, der einerseits ausgehend von einer Vorstellung über Wissenschaft, die maßgeblich auf die Physik als klassische Naturwissenschaft fokussierte, geprägt ist und andererseits durch darauf aufbauende, disziplinspezifische Überlegungen eine Pluralität an differenzierbaren Erklärungsformen hervorgebracht hat (Mayr, 2002; Potochnik, 2013; Schlosser & Weingarten, 2002; Wouters, 2013).

Ein Ausschnitt aus diesem wissenschaftstheoretischen Diskurs soll in den folgenden Abschnitten skizziert werden. Dabei werden Erklärungsformen und -ansätze, die für die Biologie besonders charakteristisch oder herausfordernd sind, beschrieben, wobei an vielen Stellen lediglich auf weitere Diskussionen, konkurrierende Positionen und strittige Punkte verwiesen werden kann. Die jeweiligen Erklärungsformen werden an Beispielen illustriert.

Deduktiv-nomologische Erklärungen

Einen ersten Vorschlag zur formalen Überprüfung der Gültigkeit und damit zu einem Modell für wissenschaftliche Erklärungen liefern Hempel und Oppenheim (1948) in Form des deduktiv-nomologischen Erklärungsmodells. Für ein Verständnis der Diskussion um diesen Ansatz, gültige wissenschaftliche Erklärungen zu beschreiben, ist der historische Kontext der Entstehung des Ansatzes von Bedeutung: Beide Autoren gehören zu den Mitgliedern des sogenannten Wiener Kreises, der sich 1924 gründete und unter Mitwirkung zahlreicher Personen aus der Philosophie, Mathematik, Natur-, Geistes- und Sozialwissenschaft damit beschäftigte, eine wissenschaftliche Weltauffassung zu entwickeln, woraus die philosophische Position des logischen Empirismus hervorging (Stadler, 2015). Ziel war unter anderem, die seinerzeit als erfolgreich geltenden Methoden der Erkenntnisgewinnung der Naturwissenschaften – insbesondere der Physik als klassische Naturwissenschaft – und der Mathematik zu rekonstruieren und auf die unter Anhänger:innen des Wiener Kreises als weniger erfolgreich geltenden geisteswissenschaftliche Erkenntnismethoden zu übertragen. Eine grundlegende Annahme dabei ist, dass alle Wissenschaften einem einheitlichen Erkenntnisprozess und Erklärungsschemata folgen bzw. folgen sollten (Schurz, 2004; Stadler, 2015). Sie stehen damit dem aus heutiger Sicht formulierten einleitenden Satz dieses Kapitels entgegen, dass es kein einheitliches Modell wissenschaftlicher Erklärungen gibt.

Das deduktiv-nomologische Modell für wissenschaftliche Erklärungen sieht die logische Ableitung aus Aussagen über Anfangsbedingungen (sogenannten Antezedenzbedingungen) und universellen Gesetzesaussagen vor. Die Antezedenzbedingungen und die Gesetze werden als Explanans (= das Erklärende) bezeichnet; auf deren Basis erfolgt die logische Ableitung und lässt sich das sogenannte Explanandum (= das zu Erklärende) erklären. Aufgrund des logischen Schließens wird das Schema auch deduktiv-nomologisches Schema (kurz: DN-Schema) genannt (Bartelborth, 2007; Hempel &

¹⁵ Eine Reflexion zum Nutzen des Wissens über biologische Erklärungsformen für Lehrkräfte und die Gestaltung von Unterricht findet sich in Kapitel 2.3.2.

Oppenheim, 1948; Schurz, 2004). Ein einfaches Beispiel dafür aus dem Bereich der Physik ist: Dieser Draht leitet Strom (Explanandum), da er aus Metall ist (Antezedenzbedingung) und alle Metalle Strom leiten (Gesetzesaussage; ähnlich in: Schurz, 2004).

Der Versuch, alle Formen von Erklärungen, die in der wissenschaftlichen Praxis Verwendung finden, unter dieses Schema zu subsumieren – womit daraus folgend auch einhergeht, dass alle Erklärungen, die nicht diesem Schema folgen, nicht wissenschaftlich bzw. nicht gültig sind –, wurde über die Zeit mit viel Kritik bedacht (Bartelborth, 2007). So wurde auf formaler Ebene eine Asymmetrie der Erklärungsbeziehung als Mängel beschrieben: In einigen Fällen können auf diese Weise Teile des Explanans mit Teilen des Explanandums vertauscht werden (Bartelborth, 2007). Exemplarisch wird dies häufig an einem mathematisch-physikalischen Beispiel zur Erklärung der Länge eines Schattens über die Höhe des schattenwerfenden Turmes unter Bezug auf den Sonnenstand gezeigt (Bartelborth, 2007; Potochnik, 2013). Hier kann auch die Höhe des Turmes berechnet werden, indem die Länge des Schattens sowie der Sonnenstand gemessen werden. Beide Anfangsgrößen können als Explanans oder Explanandum dienen, sodass der logische Schluss formal in beide Richtungen funktioniert. Die letztere Variante – die Berechnung der Schattengröße über die Höhe des Turms und den Sonnenstand – führt jedoch zu keiner Aussage, die man intuitiv, also aus pragmatischer Sicht, auch als Erklärung bezeichnen würde, da lediglich feste Größen berechnet werden. Die formale Struktur des Erklärungsmodells, die Hempel und Oppenheim vorschlagen, ist für diesen Fall demnach nicht hinreichend genau beschrieben.¹⁶

Ein für die Disziplin der Biologie besonders relevant erscheinender Kritikpunkt bezieht sich auf einen inhaltlichen Aspekt des Modells, die Gesetzesaussagen, die für eine wissenschaftliche Erklärung nach dem DN-Schema notwendig sind (Bartelborth, 2007). Dabei sind Gesetze universelle, ausnahmslos geltende Zusammenhänge, mit denen das Auftreten eines Phänomens garantiert vorhergesagt werden kann (Bartelborth, 2007). So wird ein Gegenstand unter Normalbedingungen auf der Erde stets zu Boden fallen und sich die Geschwindigkeit über das Gravitationsgesetz relativ genau bestimmen lassen. Zwar wurde unter den Empiristen begrüßt, dass sich durch den Fokus auf Gesetzmäßigkeiten Spekulationen aus der Wissenschaft heraushalten lassen (Bartelborth, 2007), neben der Biologie stellt sich jedoch auch für weitere wissenschaftliche Disziplinen (z. B. Psychologie, Soziologie, Geschichte) die Frage, ob derartige ausnahmslose Gesetze überhaupt für deren jeweilige Gegenstände vorhanden bzw. wie Generalisierungen, die an Stelle von Gesetzen treten, gestaltet sind (Elgin, 2006; Lange, 2013). Die Regelmäßigkeiten, die Grundlage für biologische Fragestellungen sind, lassen jedoch meist Ausnahmen zu (z. B. treten Genmutationen nur mit einer bestimmten Wahrscheinlichkeit auf, spielen in der Evolution zufällige Ereignisse eine große Rolle, entwickeln sich zwei ähnliche Ökosysteme nie identisch, sondern in sensibler Abhängigkeit von komplexen Anfangsbedingungen).¹⁷ Diese Regelmäßigkeiten können daher nicht ohne Einschränkungen als Äquivalente zu den physikalischen Naturgesetzen definiert werden. Mit diesem vermeintlichen Defizit der Biologie wird verschiedentlich umgegangen und nach Lösungswegen für die Definition von Gesetzen in der Biologie gesucht (vgl. z. B. Bartelborth, 2007; Elgin, 2006; Lange, 2013).

¹⁶ Für weitere Ausführungen zur Kritik am DN-Schema von Hempel und Oppenheim empfiehlt sich das Buch von Bartelborth (2007).

¹⁷ Dittmer und Langlet (2021) verweisen darauf, dass vor allem in der Physiologie Mechanismen mit einer hohen Regelmäßigkeit ablaufen und daher als gesetzesähnlich bezeichnet werden können.

Induktiv-statistische Erklärung

Hempel und Oppenheim reagierten auf die Schwierigkeit mit Gesetzaussagen in vielen wissenschaftlichen Disziplinen mit dem Modell der induktiv-statistischen Erklärung (Bartelborth, 2007; Schurz, 2004). Hierfür werden eben keine strikten und ausnahmslosen Gesetze angeführt, sondern das zu Erklärende wird aus einem statistischen Zusammenhang induktiv abgeleitet. Eine essenzielle Bedingung ist jedoch, dass die Wahrscheinlichkeit als hoch bezeichnet werden kann (Bartelborth, 2007; Schurz, 2004). Beispielsweise können mit der dritten Mendelschen Regel, der Uniformitätsregel, die Häufigkeiten einiger phänotypischer Merkmale annähernd erklärt werden (Potochnik, 2013). Unregelmäßigkeiten in den Häufigkeiten trotz der Lokalisation der merkmalsstragenden Gene auf zwei unterschiedlichen Chromosomen können beispielsweise durch Crossing-Over entstehen und damit das statistische Ergebnis unvorhersehbar verändern.

Kritisiert wurde dieser Ansatz unter anderem, da nicht die Höhe der Wahrscheinlichkeit, die von Hempel und Oppenheim als Bedingung für die Gültigkeit der Anwendung des Schemas formuliert wurde, sondern die Relevanz des Zusammenhangs, ausschlaggebend scheint. Dadurch sind Erklärungen ohne erklärenden Wert wie Folgende formal möglich (Bartelborth, 2007, S. 31; in Anlehnung an Salmon, 1984):

- (G) Menschen, die erkältet sind, erholen sich mit hoher Wahrscheinlichkeit innerhalb von 14 Tagen, wenn sie Vitamin C nehmen [(statistische Gesetzaussage)].
- (A) Franz war erkältet und nahm Vitamin C [(Antezedenzbedingung)].
- (E) Franz erholte sich von seiner Erkältung in 14 Tagen [(Explanandum)].

Dabei muss in keiner Weise enthalten sein, ob dieser Zusammenhang tatsächlich besteht (oder ob die Erkältung nur kuriert war, weil die meisten Erkältungen nach dieser Zeit abklingen).¹⁸

Kausale Erklärungen

Während für Hempel und Oppenheim eine wissenschaftliche Erklärung durch die Subsumtion unter allgemeine Gesetzmäßigkeiten gekennzeichnet ist, sehen Vertreter:innen kausaler Erklärungstheorien die Angabe von Ursachen als Charakteristikum bzw. Bedingung an (Bartelborth, 2007). Dabei gehen die Ursachen zeitlich gesehen ihren jeweiligen Wirkungen stets voraus (Bartelborth, 2007; Schurz, 2004). So kann die Aufnahme von Influenzaviren in den Körper als Ursache für das Auftreten von Grippe-symptomen gelten. Hier wird auch deutlich, dass allein die Angabe der Ursachen nicht zwingend Verständnis erzeugt und damit keinen wesentlich erklärenden Charakter aufweist (Bartelborth, 2007; Potochnik, 2013). Zusätzlich müssen weitere Informationen gegeben werden (z. B. Eigenschaften des Phänomens und der Ursache) sowie der angenommene bzw. bestehende Zusammenhang zwischen Ursache und Wirkung aufgezeigt werden (Bartelborth, 2007; Potochnik, 2013). Für das Beispiel zur Entwicklung von Grippe-symptomen wären das beispielsweise die Eigenschaften von Influenzaviren und deren Vermehrung im Körper sowie das Immunsystem in seinen relevanten Teilen und Eigenschaften in Zusammenhang mit der Immunreaktion auf die Influenzaviren und Körperreaktionen, die als Symptome auftreten, zum Beispiel Fieber.

Bei kausalen Erklärungen ist auf zwei weitere Punkte zu verweisen, die eng miteinander verbunden sind und für die Gegenstände der Biologie von hoher Relevanz sein können. Zum einen werden

¹⁸ Auch hier führt Bartelborth (2007) weitere Kritikpunkte ausführlicher aus.

allgemein bei kausalen Erklärungen häufig keine absoluten, sondern wahrscheinlichkeitsabhängige Kausalitäten angeführt. Das bedeutet, dass nicht eine Ursache ein Phänomen verursacht, sondern eine Ursache die Auftretenswahrscheinlichkeit eines Phänomens begünstigt (Bartelborth, 2007).

Zum anderen sind einfache Vorstellungen von Kausalitäten, die das Phänomen auf eine Ursache zurückführen, häufig nicht zielführend, da es sich bei biologischen Phänomenen meist um komplexe, dynamische Systeme handelt, die sich aus vielen Faktoren und Wechselwirkungen zusammensetzen (Potochnik, 2013; Wouters, 2013; Ausführlicher wird auf Komplexität als Charakteristikum biologischer Gegenstände in Kapitel 3 eingegangen). Da auch die Wechselwirkungen zwischen den Faktoren Ursachen sein und so zur Phänomenentstehung beitragen können, muss diese systemische Perspektive bei der Erklärung miteinbezogen werden:

Yet research in biology amply demonstrates that most biological phenomena result from complex causal processes, with many factors combining and interacting at each step in the process. This renders impractical a simple causal approach to explanation, whereby to explain you simply cite all the causes. (Potochnik, 2013, S. 53)

Das Beispiel fortführend bedeutet dies, dass die Aufnahme von Influenzaviren in Abhängigkeit von deren Anzahl sowie dem gesundheitlichen Zustand der betreffenden Person und vorausgehenden Infektionen mit ähnlichen Erregern die Auftretenswahrscheinlichkeit von Grippe-symptomen erhöht oder verringert. Eine vollständige Erklärung würde dann auch die Zusammenhänge zwischen den Reaktionen des Immunsystems, den Mechanismen der Virenverbreitung und der individuellen körperlichen Verfassung der Person involvieren.

Zudem ergibt sich eine interessante Unterscheidung zwischen ‚how-actually‘- und ‚how-possibly‘-Erklärungen, die von Potochnik (2013) getroffen wird: „As the terminology suggests, a how-actually explanation tracks the actual causal process that brought about a phenomenon, whereas a how-possibly explanation outlines a process that *could have* (but may not in fact have) brought about a phenomenon“ (Potochnik, 2013, S. 52). Dabei kann durch die Form der ‚how-possibly‘-Erklärungen auch über vorläufige Erklärungen für Phänomene gesprochen werden, wodurch insbesondere der Vorläufigkeit und Veränderbarkeit des Wissens Rechnung getragen wird.¹⁹

Mit Blick auf biologisch relevante Fragestellungen bzw. untersuchte Phänomene werden weitere Formen von Erklärung differenzierbar. So fragen bzw. fragten sich Biolog:innen beispielsweise, wie die DNS in Proteine ‚übersetzt‘ wird, wozu das fetale Herz eine offene Verbindung zwischen den Herzvorhöfen (Foramen ovale) besitzt oder wie sich in pflanzlichen Zellen die Fähigkeit zur Photosynthese entwickelte. Sie stellen damit Fragen nach den zugrunde liegenden Mechanismen, der Funktion von Teilen eines Systems und der Naturgeschichte bzw. der evolutionären Entwicklung von Merkmalen. Auf diese Fragen wird jeweils mit differenzierbaren Formen von Erklärungen geantwortet, in denen das Kausalitätsprinzip im ontologischen Sinne nicht abgelehnt wird, die Zurückführung auf Kausalursachen – wie im deduktiv-nomologischen Erklärungsmodell oder bei kausalen Erklärungsansätzen – jedoch keine zentrale Rolle einnimmt. Wouters (2013) fasst zusammen, insbesondere dieses disziplinspezifische Charakteristikum sei als Kennzeichen angeführt

¹⁹ Neben den hier bisher angeführten Ansätzen zur Beschreibung wissenschaftlicher Erklärungen kann noch auf weitere verwiesen werden: den ‚unification account‘, der unter anderem durch Kitcher (1981) vertreten wird und bei dem ein Phänomen als erklärt gilt, wenn es mit Prinzipien verbunden wird, die auch weitere Phänomene erklären und dadurch in einem Zusammenhang eingebettet wird, den ‚kairitic account‘, begründet von Strevens (2004), der kausale Ansätze mit den vereinheitlichenden Ideen von Kitcher verbindet, den pragmatischen Ansatz, den unter anderem van Fraassen (1980) vertritt, indem er die Relevanz des Kontextes in das Zentrum stellt, sodass eine Erklärung als solche gilt, wenn die Ausgangsfrage der involvierten Personen damit für diese befriedigend beantwortet wird, die nomischen Muster, die von Bartelborth (2007) vorgeschlagen werden und die mit Gesetzen verbundenen Schwierigkeiten umgehen sollen.

worden, dass die Biologie lange als keine eigenständige wissenschaftliche Disziplin galt: „**To some scientists this is a sign of the immaturity of biology** [...]“ (Wouters, 2013, S. 706). Die erwähnten Erklärungsformen werden in den folgenden Abschnitten ebenfalls kurz skizziert und an einem Beispiel verdeutlicht.

Mechanistische Erklärungen

Um zu erklären, auf welche Art und Weise etwas vor sich geht, werden mechanistische Erklärungen genutzt. Dabei wird aufgezeigt, wie ein Phänomen durch die involvierten Komponenten des Systems und deren Interaktionen miteinander hervorgebracht wird (Bechtel, 2011; Kampourakis, 2013; Wouters, 2013;). Die Maschinerie – das System mit seinen interagierenden Komponenten – muss dafür in Teilprozesse zerlegt und diese müssen ihrerseits erklärt werden. Diese Beschreibung lässt zunächst die Annahme zu, dass damit eine reduktionistische Denkweise²⁰ der Phänomene befördert wird. Neben den Komponenten und ihren Interaktionen ist jedoch auch deren Organisation im System ein wichtiger Aspekt (Wouters, 2013).²¹ Grund dafür ist, dass die Komponenten niedrigerer Organisationsebenen nicht alle Informationen, die das Phänomen erklären können, enthalten: „**But even more fundamentally, the lower level does not provide all the information needed to account for the phenomenon. The working parts only produce the phenomenon when they are organized and their operations are appropriately orchestrated**“ (Bechtel, 2013, S. 488). Einhergehend damit ist eine reduktionistische Denkweise unangemessen.

So ist die Frage nach dem Ablauf der Synthese von Proteinen aus dem genetischen Code der DNS (= Mechanismus der Proteinbiosynthese) mit der Nennung der Involvierten Komponenten (unter anderem mRNA-Strang, t-RNA, Aminosäuren, Aminoacyl-tRNA-Synthetase, Ribosom aus zwei Untereinheiten), deren Interaktionen (z. B. Bindung eines Basentriplets sowie eines t-RNA Stücks nach dem Schlüsselschlossprinzip in der spezifischen Aminoacyl-tRNA-Synthetase, die die kovalente **Bindung des Aminoacylrests mit der freien 3'-OH-Gruppe des tRNA-Moleküls katalysiert**) und der Organisation (z. B. die örtliche Trennung von Transkription (Nukleus) und Translation (Cytoplasma)) zu beantworten (Campbell & Reece, 2009).

Funktionale Erklärungen

Krohs (2007) **bezeichnet die Beschäftigung mit Funktionen als „eine[n] der größten Unterschiede der Biologie im Vergleich mit Physik und Chemie“** (Krohs, 2007, S. 287), denn zu den relevanten Fragen kommt die nach Funktionen hinzu: die Funktion eines Organs im physiologischen System, eines Organismus im Ökosystem oder eines Mechanismus wie die Photosynthese für den Organismus einer Pflanze. Dafür wird auf bekannte Mechanismen zurückgegriffen, um die Funktion einer Struktur oder eines Prozesses zu erklären (Krohs, 2004; Wouters, 2013).

²⁰ Der Begriff ‚Reduktionismus‘ bezeichnet im Kern die Annahme, dass sich alle Phänomene erklären lassen, indem sie in die kleinsten differenzierbaren Teile zerlegt werden. Mayr (2002) führt hierzu ein historisches Beispiel an, dass die **Paradoxie der Annahme deutlich machen kann: „Darwins Freund Thomas Henry Huxley fragte die Reduktionisten, warum die Reduktion von Wasser (H₂O) in die Gase Wasserstoff und Sauerstoff nicht die Natur des Wassers erklärt?“** (Mayr, 2002, S. 25).

²¹ Wouters (2013) geht dabei sogar einen Schritt weiter und postuliert, dass zur Erklärung eines Mechanismus auch aufgegriffen werden muss, warum die realisierte Organisation das Problem löst, wohingegen es eine andere nicht täte: „**So in order to understand a mechanism one must not only explain how its organization results in a certain behavior (as is done in mechanistic explanations), but also why that mechanism’s organization solves the problem whereas other forms of organization (often called ‘designs’) do not**“ (Wouters, 2013, S. 708).

Krohs (2007) beschreibt als eine weitere Besonderheit funktionaler Erklärungen bzw. des Funktionsbegriffs, dass „mit ihm Vorstellungen von „richtig funktionieren“ und „falsch funktionieren“ verbunden [sind]: Die Verwendung von Funktionszuschreibungen impliziert, dass der Ausfall einer Funktion als Fehlfunktion zu klassifizieren ist“ (Krohs, 2007, S. 288). Damit umfassen funktionale Erklärungen auch immer ein normatives Moment.

Auch der Funktionsbegriff wird vielseitig zu definieren versucht. So lassen sich beispielsweise systemanalytische oder ätiologische (historisch-evolutionäre) Ansätze differenzieren (Krohs, 2007; McLaughlin, 2020). McLaughlin (2020) beschreibt die zwei Ansätze prägnant:

Die Zuschreibung einer Funktion zu einem Merkmal (Organ, Verhalten) kann lediglich eine Beschreibung seines Beitrags zu den normalen Leistungen des Organismus sein oder sie kann als Erklärung seiner Entstehung aufgefasst werden. Somit kann eine Funktionszuschreibung entweder erklären, was der Funktionsträger macht, oder sie kann darüber hinaus auch erklären, warum der Funktionsträger da ist. (McLaughlin, 2020, S. 24)

Demnach kann bei funktionalen Erklärungen entweder die Wirkung oder die Ursache der Existenz in Hinblick auf die Bedeutung einer Systemkomponente thematisiert werden.

Als Beispiel kann die Frage nach der Funktion des Foramen ovale, eine Verbindung der beiden Herzvorhöfe, im fetalen Herzen für den Organismus dienen. Grundlegende Mechanismen, um die Funktion der Struktur zu erklären, sind beispielsweise die Anreicherung von Blut mit Sauerstoff sowie die Kontraktionen des Herzens, durch die sauerstoffreiches Blut aus der rechten Herzkammer zu sauerstoffverbrauchenden Organen im ganzen Körper transportiert wird. Beim Kreislaufsystem des Fötus kann die Sauerstoffanreicherung nicht über den Lungenkreislauf erfolgen. Das Blut wird in der Plazenta mit Sauerstoff angereichert und gelangt von dort in die linke Herzkammer (beim erwachsenen Menschen gelangt das sauerstoffreiche Blut nach dem Lungenkreislauf hingegen direkt in die rechte Herzkammer). Durch das Foramen ovale besteht eine Verbindung zwischen dem linken und dem rechten Atrium, durch das sauerstoffreiches Blut hindurchgelangt und von dort durch Kontraktionen im Körper verteilt wird. In seltenen Fällen ist der Verschluss dieser Parallelschaltung in den ersten Lebenswochen nicht vollständig und es kommt zu einer Fehlfunktion des Herzens, die meist operativ reguliert werden kann (Gaehtgens, 1994). Damit ist mit McLaughlins (2020) Unterscheidung die Funktion durch Bezug auf das, was der Funktionsträger im Organismus macht, erklärt. Abgeleitet von der Funktion kann entsprechend eine Dysfunktion beschrieben werden (Krohs, 2007).

Teleologische Erklärungen

Als problematisch bei der Formulierung von biologischen Erklärungen, die Funktionen thematisieren, gilt es, wenn dabei auf zwecksetzende Instanzen verwiesen oder über die Zweckmäßigkeit der Phänomene oder Strukturen gesprochen wird (Schlosser & Weingarten, 2002; Krohs, 2007). Sie werden als teleologische Erklärungen bezeichnet, wobei ‚telos‘ aus dem Griechischen stammt und für ‚Ziel‘ oder ‚Zweck‘ steht (Krohs, 2007). Dabei werden Funktionen eines Prozesses oder einer Struktur für die Ursache gehalten (Lennox & Kampourakis, 2013; Schlosser & Weingarten, 2002).²² Es kann unterschieden werden zwischen Erklärungen mit direktem Verweis auf eine schöpferisch tätige Instanz und solche mit der Annahme, dass Prozesse natürlicherweise zielorientiert angelegt

²² Für eine Zusammenschau zur vielseitigen, historischen Entwicklung des Teleologiebegriffs empfiehlt sich der Artikel von Lennox und Kampourakis (2013) sowie für eine vertiefende Auseinandersetzung, auf mit verschiedenen aktuellen Teleologiebegriffen das Buch von Toepfer (2004).

sind und lebende Organismen deshalb Merkmale besitzen, um ihr Leben zu ermöglichen oder zu verbessern (Lennox & Kampourakis, 2013). Vor allem die zweite Variante wird häufig beleuchtet. Krohs (2007) verdeutlicht den Bezug auf eine allgemeine Zielorientierung an folgendem Beispiel: „Das Herz ist dazu da, Blut zu pumpen, um so die peripheren Gewebe mit Nährstoffen zu versorgen“ (Krohs, 2007, S. 287). Dabei wird eine Formulierung verwendet, die nahelegt, dass die Funktion des Herzens zweckhaft für den Organismus ist, sodass dieser als willentlich handelnde Entität aufgefasst werden kann. Ein weiteres Beispiel aus dem Bereich der Evolutionsbiologie ist die Aussage, dass Fische Flossen haben, *um* schwimmen zu können. Dabei wird die Finalursache als erklärendes Element (=Explanans) herangezogen (Schlosser & Weingarten, 2002).

Auch in der Biologiedidaktik wird gerade diese Form von Erklärungen vielseitig diskutiert, da sie als anschaulich und Ausgangspunkt für emotionale Beziehungen mit dem Gegenstand gilt, fachlich jedoch nicht mit wissenschaftlichen Erklärungen der Phänomene vereinbar ist (z. B. Kattmann, 2005; Trommler et al., 2017). In der Diskussion wird auch die Frage aufgeworfen, ob es sich bei teleologischen Formulierungen lediglich um sprachliche Muster handelt oder diese tatsächlich auch auf derartiges Denken schließen lassen (Trommler et al., 2017).

Naturhistorische Erklärungen

Organismen haben neben einer individuellen Entwicklungsgeschichte auch immer eine Geschichte ihrer Art, der sie angehören. Die Fragen und Erklärungen beziehen sich sodann auf die Evolution ihrer Merkmale. Aufgrund der Umstände, dass für die Gegenstände der Chemie oder Physik die Herkunft bzw. die Entwicklungsgeschichte meist keine besondere Bedeutung zukommt,²³ wird die naturhistorische Dimension der Biologie als ein Alleinstellungsmerkmal der Biologie als Wissenschaft diskutiert:

Unlike physical objects, organisms have the character they have because ancestral variants with those characteristics were favored over variants with other characteristics. For that reason in biology, but not in physics and chemistry, the answer to the question how a certain characteristic is produced must involve an answer to the question why (e.g. for what effects) that characteristic was favored in the selection process. (Wouters, 2013, S. 707)

Die allgemeinen Mechanismen, wie Merkmale in Populationen und schließlich Arten entstehen oder verschwinden, werden von der synthetischen Evolutionsbiologie beschrieben. Soll ein einzelnes Phänomen oder singuläres Ereignis erklärt werden (z. B. das Aussterben der Dinosaurier, die Entstehung der Landpflanzen, die Entwicklung des Covid-19 Erregers aus der Gruppe der Corona Viren), müssen Biolog:innen methodisch eher wie Historiker:innen vorgehen (Mayr, 2002). Dabei müssen anhand von Relikten – z. B. geologische Veränderungen im Gestein, physiologische oder genetische Vergleiche lebender Organismen oder Fossilien, etc. – und auf Basis der bestehenden Überlegungen zu Mechanismen der Evolution mögliche Entwicklungsszenarien rekonstruiert werden (Mayr, 2002). Auf Basis der zu diesem Zeitpunkt vorliegenden Indizien werden Plausibilitätsüberlegungen angestellt und unter den möglichen Szenarien auf das wahrscheinlichste geschlossen. Mit neuen Funden, die mit dem zu erklärenden Ereignis in Verbindung gebracht werden kann, kann diese Erklärung entweder weiter gestützt werden oder, falls der Fund der bisherigen Erklärung widerspricht, abgeändert werden (Mayr, 2002). Die Methodik unterscheidet

²³ In Hinblick auf den Aspekt, inwieweit die Disziplinen der Physik und der Chemie auch Fragen mit einer historischen Dimension nachgehen, könnten die Phänomene des Urknalls und der Entstehung der Elemente interessant sein. Die Limitation in Hinblick auf die Durchführung von Experimenten ergibt sich hier mitunter in ähnlicher Weise wie für Fragestellungen zur Evolution.

sich demnach von den weiteren Teildisziplinen der Biologie, da kaum Experimente und keine zeitgleichen Beobachtungen durchgeführt werden können, sondern das Geschehen rückwirkend rekonstruiert wird (historische Rekonstruktion; (Mayr, 2002). Hinzu kommt, dass zufällige Ereignisse die Anzahl möglicher Ursachen prinzipiell unabschätzbar groß werden lassen und die Rekonstruktion des tatsächlichen Hergangs damit erschwert wird: „In den **physikalischen** Wissenschaften führen die Naturgesetze normalerweise zu stark deterministischen Ergebnissen. Weder die natürliche noch die geschlechtliche Selektion gewährleisten einen solchen Determinismus. Tatsächlich ist das Ergebnis eines evolutionären Prozesses gewöhnlich die Folge von Wechselwirkungen zahlreicher **Zufallsfaktoren**“ (Mayr, 2002, S. 28). Potochnik (2013) macht die hohe Anzahl an Faktoren, die bei einer Rekonstruktion eines evolutionären Prozesses involviert sind und damit berücksichtigt werden müssen, an einem Beispiel deutlich:

Consider, for example, the causal processes involved in bringing about the long necks of giraffes. **In no particular order, these include, at least, features of ancestral giraffes’ environment**, including the presence of nutritious leaf matter high up in tall trees; various genetic influences on giraffe morphology; developmental processes, including additional regulator genes, involved in giraffe neck-development; certain genetic mutations arising; competition for resources such that giraffes with a greater reach enjoyed increased rates of survival; and changes in developmental processes resulting in longer necks. (Potochnik, 2013, S. 54)

Insgesamt lässt sich an der einführenden Übersicht erkennen, dass ein reger Diskurs über Erklärungsansätze und -formen in der Wissenschaftstheorie und so eine Pluralität an Theorien über wissenschaftliche Erklärungen besteht.

Braaten und Windschitl (2011) resümieren aus den unterschiedlichen Ansätzen, dass Erklärungen mehr als Beschreibungen sind und mit theoretischen Bezügen aufzeigen, wie Phänomene zustande **kommen**: „Many philosophers of science broadly conceptualize scientific explanations as attempts to move beyond descriptions of observable natural phenomena into theoretical accounts of how phenomena unfold the way they do“ (Braaten & Windschitl, 2011, S. 641). Potochnik (2013) macht – wie weitere Autor:innen (z. B. Bartelborth, 2007; Kampourakis & Zogza, 2009) das vereinende, zentrale Merkmal in der Annahme von **Kausalität fest**: „A causal understanding of explanation, in one version or another, seems to have gained dominance in philosophy of science, especially in philosophy of biology“ (Potochnik, 2013, S. 53). Kausalität ist dabei eine Grundannahme, wobei die kausalen Zusammenhänge in den Erklärungsformen unterschiedlich stark in den Vordergrund gerückt und explizit gemacht werden.

2.3.2 *Exkurs*: Wissen über wissenschaftliche Erklärungsformen als Element einer wissenschaftspropädeutischen Lehrkräftebildung, der Unterrichtsgestaltung und des fachlichen Lernens

Bisher wurde das Erklären und insbesondere die Vielfalt an Formen biologischer Erklärungen wenig in der Literatur, die zur Ausbildung von Lehrkräften dienen soll, berücksichtigt. Beispielsweise werden in einführenden Werken zur Biologiedidaktik wissenschaftliche Erklärungsformen meist nur kurz besprochen. Killermann, Hering und Starosta (2009) greifen den Begriff der Erklärung im Rahmen des Stichwortes ‚Kausalität‘ auf und grenzen ihn dabei von teleologischen Erklärungen biologischer Phänomene ab. Sie gehen damit auf eine häufig bei Schüler:innen vorkommende Vorstellung ein (Kattmann, 2015; Trommler et al., 2017). Dittmer und Langlet (im Druck) gehen in der Neuauflage der ‚Fachdidaktik Biologie‘ ausführlicher auf Formen biologischer Erklärungen ein. Anhand von vier für die Disziplin der Biologie relevanten Fragen unterscheiden sie vier

Erklärungsformen (mechanistische Erklärungen, historische Erklärung, funktionale Erklärung, teleologische Erklärung).

Aufgrund der wenig umfangreichen Thematisierung von wissenschaftstheoretischen Überlegungen zu Erklärungen in der Biologie kann angenommen werden, dass angehende sowie praktizierende Lehrkräfte nur geringes explizites Wissen über Erklärungsformen in der Biologie erwerben. Hingegen haben sie mit dem eigenen Lernen biologischer Inhalte ein implizites Wissen erworben, welche Fragen Biolog:innen stellen, welche Arten von Erklärungen daraufhin adäquate Antworten geben oder welchen Mustern die Erklärungen folgen. Ebenso erklären sie in ihrem Unterricht unter anderem Funktionen, Mechanismen, Kausalzusammenhänge und evolutionäre Entwicklungen und nutzen dabei intuitiv die entsprechenden Formen von Erklärungen. Ein explizites Wissen über biologische Erklärungsformen wird als Teil einer wissenschaftspropädeutischen Aus- und Weiterbildung von Lehrkräften gefordert (Dittmer & Zabel, 2019; Langlet, 2001) und kann in mehrerer Hinsicht für die Gestaltung von Lehr-Lern-Prozessen hilfreich sein.²⁴

Das Wissen um Formen biologischer Erklärungen bietet zum einen Potenzial für die Strukturierung des Biologieunterrichts, wenn die Erklärungsformen als Perspektiven auf ein Phänomen betrachtet werden (z. B. die mechanistische, funktionale, naturgeschichtliche Perspektive). Beispielsweise können das Phänomen der Photosynthese hinsichtlich der Struktur der Photosysteme und des zugrunde liegenden Mechanismus oder auch dessen Funktion sowie die evolutionäre Entstehungsgeschichte beleuchtet werden. Die systematische Betrachtung eines Themenkomplexes mithilfe der wissenschaftstheoretisch unterscheidbaren Erklärungsformen kann damit ein Mittel für Lehrkräfte sein, Phänomene vielseitig zu betrachten. Dabei können die Formen biologischer Erklärungen gezielt zur Reflexion genutzt werden, welche Perspektive auf das Phänomen bereits eingenommen wurden oder noch eingenommen werden sollten. Zudem können die Erklärungsformen auch dazu anregen, Fragen zu stellen, die weniger üblich bzw. in Lehrbüchern weniger häufig zu finden sind. So kann die Veränderung der Blattfärbung im Herbst, die meist recht knapp kausal über den Abbau und die anschließende Einlagerung der Chlorophyllbestandteile erklärt wird, zu Fragen über die evolutionäre Entstehung des Abbauvorgangs in fast allen bestehenden Laubbaumarten, zur Funktion der Carotinoide und Xantophylle als zurückbleibende Blattfarbstoffe oder zu Ursachen von Blattverfärbungen im Sommer führen.

Braaten und Windschitl (2011) resümieren aus einer empirischen Studienreihe zur Verbesserung der Performanz von Naturwissenschaftslehrkräften mit verschiedenen, selbstentwickelten Werkzeugen (z. B. Thompson, Braaten & Windschitl, 2009), dass die Bekanntheit der Formen biologischer Erklärungen Lehrkräfte und deren Unterrichtsgestaltung positiv beeinflusst: „We have found that familiarity with philosophical models for scientific explanation helps science teachers develop deeper understanding of science content informing their decision making about what to teach and how to **teach in science classes**“ (Braaten & Windschitl, 2011, S. 641).

Weiterhin kann durch die Bekanntheit der biologischen Erklärungsformen und die gezielte Reflexion über deren Einsatz verhindert werden, dass einzelne Erklärungsformen zu häufig und andere nur selten genutzt werden. Potochnik (2013) sowie Braaten und Windschitl (2011) verweisen beispielsweise darauf, dass im Biologie- bzw. Naturwissenschaftsunterricht häufig auf Regeln und Gesetzmäßigkeiten fokussiert wird. Für die Disziplin der Biologie sind Regeln und insbesondere

²⁴ Ein Wissen über Formen biologischer Erklärungen und die Grenzen der möglichen Aussagen wird von Shulman (1986) als Teil des Fachwissens von Lehrkräften beschrieben: „the teacher need not only to understand that something is so, the teacher must further understand why it is so, on what grounds its warrant can be asserted, and under what circumstances our belief in its justification can be weakened and even denied“ (Shulman, 1986, S. 9).

Gesetzmäßigkeiten jedoch nicht erschöpfend (Potochnik, 2013), da die Gegenstände durch ein hohes Maß an Komplexität und damit einhergehend einem hohen Maß an Individualität gekennzeichnet sind (siehe Kap. 3). So ist beispielsweise ein konkreter Verlauf einer Sukzession in einem Ökosystem nicht hinreichend genau mit Regeln und Gesetzmäßigkeiten zu beschreiben, da die vorherrschenden Umweltfaktoren stets individuelle Bedingungen schaffen, die den Verlauf entscheidend prägen. Mit einem einengenden Fokus auf Regeln und Gesetzmäßigkeiten kann dadurch ein falsches Bild der biowissenschaftlichen Forschung bzw. biologischer Phänomene vermittelt werden.

Weiterhin kann Wissen über Erklärungsformen Lehrkräften helfen, alltäglich formulierte Fragen von Schüler:innen zu verstehen. So beschreibt Langlet (2008) am Beispiel der Frage „Warum hat der Eisbär ein weißes Fell“ (Langlet, 2008, S. 78), dass damit drei wissenschaftliche Fragestellungen angesprochen werden. Diese sind: „Wozu besitzt der Eisbär ein weißes Fell? Wie kommt es, dass das Fell des Eisbären weiß ist? Warum hat der Eisbär ein weißes Fell bekommen?“ (Langlet, 2008, S. 78). Sie referieren in genannter Reihenfolge auf die Funktion des weißen Fells, auf die proximalen Wirkursachen sowie auf die evolutionäre Entstehung und damit auf funktionale, kausale und naturhistorische Erklärungsformen. Mit der Kenntnis über die Unterscheidungen von Erklärungsformen können Lehrkräfte diese auch in Fragen von Schüler:innen erkennen, gegebenenfalls zur Intention der Frage rückfragen und damit auch eine Sensibilisierung bei Schüler:innen für Formulierungen sowie Ausrichtungen von Fragestellungen erzielen sowie zielgerichtet passende erklärende Antworten formulieren. Hinzu kommt, dass auch Lehrkräfte bei der Formulierung ihrer Fragen, die sie an Schüler:innen richten, präzisieren können (Abrams, Southerland & Cummins, 2001).

Auch die explizite Thematisierung der Erklärungsformen im Biologieunterricht wird an verschiedenen Stellen gefordert, um Schüler:innen an adäquate wissenschaftliche Erklärungen heranzuführen und wissenschaftspropädeutisch auszubilden (Abrams et al., 2001; Braaten & Windschitl, 2011; Potochnik, 2013; Zabel, 2009). Dabei stehen häufig naturhistorische Erklärungen im Fokus, da Schüler:innen in vielen Fällen intuitiv auf die teleologische Erklärungsform zurückgreifen. So nutzt Zabel (2009) die Form von Geschichten als Antworten auf Fragen zur Entstehung von Arten, um mit Schüler:innen im weiteren Verlauf des Unterrichts über wissenschaftliche und nichtwissenschaftliche Erklärungsmuster kontrastierend sprechen zu können:

Die Schülerinnen und Schüler lernen bei der Arbeit mit ihren eigenen Texten und denen ihrer Mitschüler, Erklärungen von bloßen Beschreibungen eines Phänomens sowie physische Ursachen (*causes*) von psychischen Beweggründen (*reasons*) zu unterscheiden. Ursachen sind für naturwissenschaftliche Erklärungen konstitutiv, Beweggründe dagegen für Geschichten: Hinter jeder Handlung steht dort ein Motiv. (Dittmer & Zabel, 2019, S. 105)

Darüber hinaus stellen Abrams et al. (2001) in ihrer Studie fest, dass Schüler:innen zu einem großen Anteil nicht mit der in der Frage intendierten Erklärform antworten: „If asked a how question, as they were in every interview instance, 32 % [of] the students answered with a “why” response“ (Abrams et al., 2001, S. 1271). Auf dieser Basis fordern sie, dass Lehrkräfte mit Schüler:innen über den Unterschied zwischen funktionalen und kausalen Erklärungsformen sprechen. Für die Umsetzung dieser und ähnlicher Vorschläge wird ein explizites Wissen über biologische Erklärungsformen für Lehrkräfte notwendig.

2.4 Synthese – instruktionales Erklären als fachlicher Kommunikationsprozess

Die drei aufgezeigten disziplinären Perspektiven ergeben bei ihrer Zusammenführung sowie durch ihre Kontrastierung ein umfassenderes Bild von instruktionalem Erklären im Biologieunterricht. Ein Heranziehen der wissenschaftstheoretischen Überlegungen kann dabei insbesondere für die Beschreibung von differenzierbaren Formen instruktionaler Erklärungen des Biologieunterrichts als gewinnbringend bezeichnet werden. Die Einteilung auf Grundlage der sprachwissenschaftlichen Analysen und didaktischen Arbeiten beschreibt das Erklären von Begriffen (Erklären-Was), Prozessen (Erklären-Wie), Zusammenhängen (Erklären-Warum) oder des Zwecks (Erklären-Wozu; z. B. Brown & Atkins, 1986; Klein, 2009; Schopf & Zwischenbrugger, 2015). Die Kategorisierung aus der Wissenschaftstheorie unterscheidet folgende Formen: Neben Ursache-Wirkungs-Zusammenhängen (kausale Erklärungen), die häufig ein hohes Maß an Komplexität aufweisen (komplexe kausale Erklärungen) und als statistische Zusammenhänge aufgefasst werden müssen (wahrscheinlichkeitsabhängige kausale Erklärungen), werden auch Funktionen (funktionale Erklärungen), Mechanismen (mechanistische Erklärungen) und die Naturgeschichte bzw. die Evolution von Merkmalen (naturhistorische Erklärungen) erklärt (z. B. Mayr, 2002; Wouters, 2013). Die sprachwissenschaftliche und die wissenschaftstheoretische Kategorisierung weisen zwar teilweise Überlappungen auf, unterscheiden sich jedoch in anderen Teilen maßgeblich. Das Erklären von Begriffen (Erklären-Was) wird ausschließlich in der sprachwissenschaftlichen und didaktischen Kategorisierung aufgenommen (z. B. z. B. Brown & Atkins, 1986; Harren, 2009; Klein, 2009), wohl auch, weil die Form der Begriffserklärungen nicht als disziplinspezifisch gelten kann und sich die Erklärungen von Begriffen nicht an einem Gültigkeitsmaßstab messen lassen, da sie als Bezeichnungen von Strukturen, Phänomenen oder Konzepten von Wissenschaftler:innen gesetzt werden. Demnach besteht in keiner Weise die Notwendigkeit oder Möglichkeit der wissenschaftlichen Ableitbarkeit aus Aussagesystemen und Theorien, der standgehalten werden müsste. Instruktionales Erklären ist hingegen auf ein Verstehen durch die Rezipient:innen ausgerichtet, sodass Begriffe erklärbedürftig sind, die unbekannte Strukturen, Phänomene und Konzepte bezeichnen.

Das Erklären von Prozessen (Erklären-Wie) wird ebenfalls von Sprachwissenschaft und Didaktik aufgegriffen (z. B. Brown & Atkins, 1986; Klein, 2009), muss zur Klärung des Verhältnisses zu den Erklärungsformen der Wissenschaftstheorie der Biologie jedoch etwas näher betrachtet werden. Klein (2009, S. 26) paraphrasiert das Erklären-Wie mit „Wie funktioniert y?/Wie führt man y durch?“. Während unter ‚Wie funktioniert y?‘ ein Mechanismus verstanden werden kann – z. B. ‚Wie funktioniert die Photosynthese?‘ oder ‚Wie funktioniert die Reizweiterleitung?‘ –, ist eine Erklärung auf die Frage ‚Wie führt man y durch?‘ eher eine Handlungsanleitung wie beispielsweise ‚Wie führt man eine Gelelektrophorese durch?‘ oder ‚Wie führt man das Einfangen von Insekten in ein Becherglas artgerecht durch?‘. Das Erklären-Wie weist demnach Überschneidungen mit der Form der mechanistischen Erklärungen der Wissenschaftstheorie der Biologie auf, erschöpft sich aber nicht in dieser. Vielmehr wird eine zweite Form evident, die aus inhaltlicher Sicht davon differenziert werden kann und die Durchführung einer Handlungsabfolge thematisiert.

Die dritte Form, die aus sprachwissenschaftlichen Analysen hervorging und auch in der didaktischen Auseinandersetzung Anwendung findet, ist das Erklären-Warum (z. B. Brown & Atkins, 1986; Klein, 2009). Sie nimmt Zusammenhänge in den Blick. Auch hier muss aus biologiedidaktischer und fachwissenschaftlicher Sicht die Erklärungsform von Klein (2009) ausdifferenziert werden: Bei der Beschreibung der Erklärungsform nennt der Autor stellvertretend die Fragen „Warum ist z der Fall?“ und „Warum hat C h getan?“ (Klein, 2009, S. 26). Damit werden neben Ursache-Wirkungs-

Zusammenhängen, die in der Wissenschaftstheorie der Biologie mit kausalen Erklärungsformen bezeichnet werden, auch Gründe für Handlungen von Personen unter Erklären-Warum gefasst. Für die Biologie sind Handlungs- oder Beweggründe kein genuiner Fach- und Forschungsgegenstand, sodass sie auch im Biologieunterricht eine sehr untergeordnete Rolle spielen dürften. Eine Differenzierung zwischen der Erklärung von Ursache-Wirkungs-Zusammenhängen und Handlungs- und Beweggründen von Personen ist jedoch in jedem Fall notwendig: Die Erklärung von menschlichen Handlungs- und Beweggründen rekurriert auf ein willentliches, intentionales Handeln der Personen unter Angabe ihrer Motive oder Ziele als Gründe (teleologische Erklärung (Toepfer, 2004)). Eine Zielorientierung gibt es bei kausalen Erklärungen nicht.

Die von Schopf und Zwischenbrugger (2015) sowie Dittmer und Langlet (im Druck) aufgegriffenen Formen des Erklären-Wozu wurden bereits in Kapitel 2.3.2 gegeneinander abgegrenzt: Während die Wirtschaftspädagoginnen Schopf und Zwischenbrugger (2015) ein Erklären von Sinn und Zweck des neuen Inhalts bzw. des Grundprinzips fordern, um dessen Anwendungsbereich zu vermitteln, meinen die Biologiedidaktiker Dittmer und Langlet (im Druck) die Zweckmäßigkeit von Strukturen oder Verhaltensweisen, also die Funktion dieser in dem biologischen System. Letztere decken sich demnach mit der funktionalen Erklärungsform, die wissenschaftstheoretisch beschrieben ist. Das Erklären des Zwecks eines neu erlernten Inhalts oder Grundprinzips ist davon zu unterscheiden. Im Biologieunterricht könnte hierzu beispielsweise erklärt werden, warum das Wissen über die Bedingungen und der Verlauf des Umkippen eines Sees einen Zweck für die Schüler:innen oder die Gesellschaft allgemein hat oder das Lernen am historischen Fallbeispiel von Semmelweis und den von ihm entdeckten Leichenstoffen (später Bakterien) einen Zweck für das Lernen über die Disziplin hat. Der Zweck des Wissens über das Gegenstromprinzip als Grundprinzip kann durch das Aufgreifen von Beispielen aus der Natur (z. B. in der Niere von Säugetieren, in den Beinen bestimmter Vögel als Wärmetauscher) verdeutlicht werden. Sie stellt eine eigenständige Form des Erklärens dar, bei der über den eigentlichen Erklärgegenstand hinaus Verbindungen zu weiteren Inhalten der Disziplin aufgezeigt werden.

Die Form der naturhistorischen Erklärungen, die in der Wissenschaftstheorie der Biologie viel diskutiert ist (z. B. Mayr, 2002; Wouters, 2013), wird in den allgemeinen Beschreibungen aus Sicht der Sprachwissenschaft und Didaktik nicht explizit differenziert. Diese Form muss daher bei einer Zusammenschau der Erklärungsformen des Biologieunterrichts ergänzt werden.

Damit ergibt sich in der Synthese ein erweitertes Spektrum für inhaltlich differenzierbare Erklärungsformen im Biologieunterricht. Im Biologieunterricht können Fragen aufgeworfen bzw. Erklärgegenstände thematisiert werden, die kausale, mechanistische, funktionale sowie naturhistorische Erklärungen verlangen. Im Unterricht können darüber hinaus – unabhängig vom Fach – Begriffserklärungen, Handlungserklärungen und Erklärungen des Zwecks eines Inhalts auftreten. Tabelle (Tab. 2) stellt die fachlich angemessenen, differenzierbaren Erklärungsformen mit den jeweils verallgemeinerten, exemplarischen Fragestellungen, die dahinterstehen, zusammen.

Dabei treten die hier differenzierten Formen im Unterricht häufig nicht isoliert auf, sodass eine instruktionale Erklärung zum Beispiel ausschließlich eine Funktion oder einen Begriff thematisiert. Vielmehr greifen Lehrkräfte auf die Inhalte und damit Formen zurück, die einen Sachverhalt für Schüler:innen verständlich machen und greifen dabei mitunter auch mehrere erklärbedürftige Gegenstände auf. Beispielsweise müssen beim Erklären eines Zusammenhangs im Verlauf auch unbekannte Begriffe erklärt werden oder beim Erklären einer Funktion mitunter auch die zugrunde liegenden Mechanismen aufgegriffen werden, wenn diese nicht bekannt sind.

Tabelle 2: Übersicht zu den differenzierbaren Erklärungsformen im Unterricht und verallgemeinert exemplarische Fragestellungen zu den jeweiligen Erklärungsformen

Form	Zentrale Frage
Kausale Erklärung	Durch welche Ursache kann das Phänomen erklärt werden?
Komplexe kausale Erklärung	Durch welche Ursachen und deren Wechselwirkungen kann das Phänomen erklärt werden?
Wahrscheinlichkeitsabhängige kausale Erklärung	Durch welche Ursachen wird die Wahrscheinlichkeit für das Auftreten des Phänomens beeinflusst?
Mechanistische Erklärung	Wie interagieren die Bestandteile des Systems, um das Phänomen in dieser Weise hervorzubringen?
Funktionale Erklärung	Welche Funktion hat ein Bestandteil im System?
Naturhistorische Erklärung	Wie ist etwas entstanden?
Begriffserklärung	Was bedeutet der Begriff?
Handlungserklärung	Wie führt man die Handlung durch?
Erklärung zum Zweck	Welchen Zweck hat der Erklärinhalt für ein Verständnis weiterer Inhalte oder der Disziplin selbst?

Gemein ist allen Inhalten von instruktionalen Erklärungen dabei stets, dass sie sich der direkten Wahrnehmung entziehen (z. B. Bartelborth, 2007; Becker-Mrotzek, 2004; Braaten & Windschitl, 2011; Potochnik, 2013). So sind weder Kausalitäten oder Zusammenhänge noch Funktionen oder Entwicklungen (im Sinne der Naturgeschichte oder Evolution) direkt beobachtbar. Auch die Bedeutung von Begriffen oder der Zweck eines Inhalts offenbaren sich den Schüler:innen direkt, sondern müssen von der Lehrkraft expliziert und damit erklärt werden. Im Falle von Handlungserklärungen sind zwar die Handlungsschritte meist direkt beobachtbar, jedoch ist ein zentrales Element von Handlungserklärungen die Begründung der Reihenfolge von Teilschritten oder die Auswahl von Hilfs- und Arbeitsmitteln, um ein Verständnis für die Handlungsabfolge bzw. -ausführung zu stiften. Diese sind wiederum nicht direkt beobachtbar, sondern müssen expliziert werden.

Als weiterer Aspekt der Begriffsbestimmung wurde eingangs bereits die asymmetrische Wissensverteilung, die in sprachwissenschaftlichen Untersuchungen rekonstruiert werden konnte, aufgegriffen. Das instruktionale Erklären als Kommunikationsprozess ist durch eine solche asymmetrische Wissensverteilung zwischen einer Person, die bereits Wissen über den zu erklärenden Gegenstand besitzt und dieses Wissen im Erklärprozess unwissenden Personen verständlich machen kann, gekennzeichnet (Hohenstein, 2006). Im schulischen Kontext manifestiert sich dieser Wissensunterschied häufig – und beim instruktionalen Erklären stets – zwischen den Rollen der Lehrkraft als erklärende Personen mit einem Wissensvorsprung und den Schüler:innen als Rezipient:innen mit weniger Wissen zum Erklärgegenstand. Dabei kann der Ausgangspunkt für eine instruktionale Erklärung eine explizit verbalisierte Frage der Schüler:innen oder eine von der Lehrkraft antizipierte Frage sein, die sich durch die Unterrichtsplanung bzw. die Problemstellung der Unterrichtsstunde an sich ergibt.

Als Ziel des instruktionalen Erklärens wird mehrheitlich sowohl aus Sicht der Sprachwissenschaft als auch in lerntheoretischen bzw. didaktischen Überlegungen genannt, Schüler:innen Wissen zu

vermitteln oder einen Verstehensprozess zu initiieren (z. B. Ehlich, 2009; Hohenstein, 2006; Treagust & Harrison, 2000; Wagner & Wörn, 2011). Bei einem Erklärgegenstand, der einen Prozess beschreibt, beispielsweise die Einstellung eines Binokulars oder die Anfertigung eines Diagramms, kann das Erklären auch auf das Erlernen von Fähigkeiten ausgerichtet sein. Insgesamt müssen die Konstruktionsleistung und Behaltensleistung von Schüler:innen selbstständig und aktiv erbracht werden. Um erfolgreiches Lernen auch anhand von Erklärungen im Unterricht zu erreichen, ist deshalb eine Orientierung am Wissensstand der Adressat:innen eine grundlegende Voraussetzung. Die Sprachwissenschaft bringt dies zum Ausdruck, indem das Verb ‚erklären‘ als illokutives Verb eingeordnet wird: Die verfehlte Intention, zum Beispiel das Nichtverstehen durch die Rezipient:innen, führt nicht dazu, dass die Sprachhandlung nicht als Erklären bezeichnet werden kann (Klein, 2009). Aus fachdidaktischer Sicht können auch ein implizites Lernen über die Disziplin, deren Fragen und adäquate Erklärungsformen durch instruktionale Erklärungen – wobei die fachliche Erklärung der Lehrkraft hier als Modell dient – angestoßen werden (Larreamendy-Joerns & Muñoz, 2010).

In allen Diskursen wird weiterhin allgemein festgehalten, dass das Erklären vielgestaltig sein kann (z. B. Hohenstein, 2006; Leinhardt, 1990; Morek, 2012; Schlosser & Weingarten, 2002; Treagust & Harrison, 2000). Zum einen bedingt der fachliche Erklärgegenstand die Form der Erklärung. Darüber hinaus wurde bereits beschrieben, dass instruktionale Erklärungen nicht ausschließlich einen Erklärgegenstand thematisieren, sondern mitunter mehrere erklärbedürftige Gegenstände aufgegriffen werden, sodass mehrere inhaltlich differenzierbare Formen innerhalb einer instruktionalen Erklärung auftreten. Zum anderen werden je nach Rezipient:innen weitere Sprachhandlungen eingebettet – beispielsweise Definitionen von neuen Begriffen, Beschreibungen von unbekanntem Objekten oder Begründungen für Vorgehensweisen –, wodurch sich vielfältige Abläufe und Gestaltungen je nach Situation ergeben. Hinzu kommt die Verwendung von familienähnlichen Aktivitäten, also visuellen Darbietungsformen wie Zeigen und Demonstrieren, die in Verbindung mit Erklären eingesetzt werden können, um Verstehen zu erzeugen. Insgesamt wird der Kommunikationsprozess des instruktionalen Erklärens so von vielen Variablen beeinflusst.

Verdichtet kann Folgendes formuliert werden: Instruktionale Erklärungen im Biologieunterricht sind vielgestaltige Sprachhandlungen von Lehrkräften, die an Lernende gerichtet sind und die Initiierung einer Wissensvermittlung, eines Verstehensprozesses oder eines Fähigkeitserwerbs zum Ziel haben, wobei sich die Erklärgegenstände stets der direkten Wahrnehmung entziehen. Bei der Erklärung biologischer Phänomene orientieren sich Lehrkräfte an wissenschaftlichen Erklärungsformen der Biologie und modellieren damit implizit, welche Erklärungen aus Sicht des Faches angemessen sind.

3 Komplexität, Nichtwissen und unsicheres Wissen als charakteristische Eigenschaften biologischer Phänomene und Herausforderung für Erklärende und Rezipient:innen

Der These folgend, dass instruktionale Erklärungen im Biologieunterricht nur als *gut* gelten können, wenn das Phänomen fachlich korrekt bzw. adäquat dargestellt wird (siehe Kap. 3.6), wird im Folgenden auf ein Charakteristikum biologischer Phänomene näher eingegangen, das von Mitchell (2008) als ein überspannendes Prinzip für die Disziplin der Biologie beschrieben wird, dem jedoch im Kontext von instruktionalen Erklärungen bislang keine gesonderte Aufmerksamkeit geschenkt wurde: die Komplexität biologischer Systeme.²⁵

Viele Fragen, die sich Biolog:innen stellen und die auch im Unterricht so oder so ähnlich thematisiert werden, verlangen häufig nach Antworten, die Systeme mit vielen wechselwirkenden Elementen rekonstruieren und sich auf verschiedene Systemebenen beziehen, zum Beispiel: Welche Auswirkungen hat die Einschleppung einer neuen Art in ein Ökosystem? Wie entstehen Gedanken im menschlichen Gehirn? Wie kann das Immunsystem zwischen eigenen und fremden Zellen unterscheiden? Warum verlaufen Infektionen bei verschiedenen Menschen unterschiedlich, obwohl der gleiche Virus die Infektion hervorgerufen hat?

Antworten bzw. Erklärungen zu Fragen dieser Art sind (für Wissenschaftler:innen) nicht einfach zu finden und für eine Vermittlung (durch Lehrkräfte) mitunter schwer zu strukturieren: Es existiert nicht ein zentrales Element, das auf ein weiteres, isoliertes Element wirkt, also keine einzelne Ursache von der eine einzelne Wirkung ausgeht. Vielmehr ist eine Vielzahl von Elementen an Prozessen der Phänomenentstehung beteiligt. Diese Elemente sind innerhalb des Systems über Ursache-Wirkungs-Beziehungen miteinander vernetzt, sodass die Veränderung eines Elements Ursache für Veränderungen weiterer Elemente sein und über Rückkopplungsschleifen zurückwirken kann, wodurch die Differenzierung zwischen Ursache und Wirkung erschwert wird. Charakteristisch sind dadurch auch zahlreiche, gleichzeitig stattfindende Prozesse bzw. Veränderungen. Die Veränderungen sind dabei kaum kalkulierbar, da es sich meist nicht um lineare Kausalbeziehungen handelt, sondern kleine Ursachen (bzw. kleine Effekte) auch große Wirkungen nach sich ziehen können (nichtlineare Ursache-Wirkungs-Beziehung). Das Phänomen kann also nicht erklärt werden, indem nacheinander Ursache und Wirkung genannt und der Wirkmechanismus beschrieben wird. Regelmäßigkeiten im Verhalten des Systems sind dabei nicht einfach zu erkennen, Gesetzmäßigkeiten oder Gesetze spielen eine untergeordnete Rolle (Kuhlmann, 2007; Mainzer, 2008; Mitchell, 2008). Wissenschaftler:innen sind bei der Rekonstruktion komplexer Systeme zudem mit Nichtwissen und unsicherem Wissen konfrontiert. So sind die Grenzen der Systeme nicht eindeutig und Elemente des Systems werden mitunter nicht als solche erkannt (Kampourakis & McCain, 2020; Mitchell, 2008). Nichtwissen und unsicheres Wissen existieren bzw. entstehen zwar nicht ausschließlich in Zusammenhang mit komplexen Systemen, jedoch treten sie hier in besonders deutlicher Weise hervor.

Es wird auf wissenschaftstheoretische Überlegungen zurückgegriffen, die sich mit der interdisziplinären Komplexitätstheorie befassen. Die Überlegungen werden genutzt, um zu erörtern,

²⁵ Wie die im weiteren Verlauf aufgeworfenen fachlichen Fragen implizit zeigen, wird hier die fachliche Komplexität behandelt. Im didaktischen Diskurs – insbesondere dann, wenn Handlungen bzw. das Abwägen von Handlungsoptionen ins Spiel kommen und es um die Bewertungskompetenz von Schüler:innen geht – wird auch die ethische Komplexität relevant. Diese wird hier nicht näher betrachtet, da die ethische Komplexität eher nach einem diskursiven Erörtern im Unterricht verlangt, wofür instruktionale Erklärungen nicht geeignet sind.

wie die Gegenstände der Biologie beschaffen sind, die im Unterricht vermittelt werden sollen, und um darauf aufbauend eine Konzeptualisierung guten Erklärens im Biologieunterricht zu entwickeln. Die Gegenstände der Komplexitätsforschung sind nicht ausschließlich biologische Systeme. Allgemein komplexe Systeme werden untersucht und hinsichtlich ihrer Gemeinsamkeiten und Unterschieden analysiert, die sich zum Beispiel in der Physik (Bar-Yam, 1997; Kuhlmann, 2007), den Sozialwissenschaften (Mainzer, 2008), der Wirtschaft (Mainzer, 2008), der Biologie (Mainzer, 2008; Mitchell, 2008) und der Informatik (Mainzer, 1997) finden lassen.

Komplexitätsforschung beschäftigt sich fachübergreifend mit der Frage, wie durch Wechselwirkungen vieler Elemente eines komplexen Systems (z. B. Moleküle in Materialien, Zellen in Organismen oder Menschen in Märkten und Organisationen) Ordnungen und Strukturen entstehen können, aber auch Chaos und Zusammenbrüche. (Mainzer, 2008, S. 10)

Das Thema der Komplexität bzw. die Komplexitätsforschung gewinnt dabei zunehmend an Beachtung und Bedeutung, da sich aktuelle Probleme und bisher kaum fassbare Phänomene damit bearbeiten lassen. Mainzer (2008) formuliert dazu:

Komplexität bestimmt die Wissenschaft des 21. Jahrhunderts. Die Expansion des Universums, die Evolution des Lebens und die Globalisierung von Wirtschaft, Gesellschaft und Kulturen führen zu Phasenübergängen komplexer, dynamischer Systeme. Die sich abzeichnenden Schlüsselthemen dieses Jahrhunderts haben mit Komplexität zu tun. Globale Klimaveränderungen, Erdbeben und Tsunamis werden in Computermodellen komplexer dynamischer Systeme untersucht. Die Gentechnologie analysiert DNS-Information, die komplexe zelluläre Organismen wachsen lässt. Die Life Sciences beschäftigen sich mit der Komplexität des Lebens. Artificial Life simuliert die komplexe Selbstorganisation des Lebens in geeigneten Computermodellen. Komplexität bestimmt auch die moderne Medizin. Heutige Krankheitsgeißeln der Menschheit wie z. B. Krebs, Herz-Kreislauf- und Gefäßerkrankungen hängen von hochkomplexen Wechselwirkungen ab. Viren mutieren und schaukeln sich zu globalen Infektionen auf, die sich wie Wellen über den Erdball ausbreiten. (Mainzer, 2008, 10 f.)

Für Wissenschaftler:innen stellen komplexe Phänomene eine Herausforderung dar, die nach neuen (mathematischen) Konzepten zur Erforschung und Darstellung der Phänomene und Strukturen sowie neuen wissenschaftlichen Methoden bzw. einer neuen Erkenntnistheorie verlangt (Mitchell, 2008).

Auch bei der Vermittlung – also für Erklärende und Rezipient:innen unter anderem im Schulunterricht – stellen komplexe Phänomene eine Herausforderung dar: Gleichzeitig stattfindende Veränderungen und Prozesse müssen in eine nacheinander erzählbare Reihenfolge gebracht werden, verwobenen Wirkungsgefüge müssen strukturiert und visualisiert werden, nichtlineare Wechselwirkungen sind schwieriger verständlich als lineare Wirkungszusammenhänge und der Umgang bzw. die Kommunikation von Nichtwissen und unsicherem Wissen ist in Hinblick auf eine Orientierung an Wissen und Fakten eher ungewohnt (Dittmer, 2010; Jacobson, 2000; Jacobson & Wilensky, 2006; Krüger, Ruhrig & Höttecke, 2013; Mambrey, Timm, Landskron & Schmiemann, 2020).

Gleichzeitig sollte der Darstellung komplexer Phänomene im Unterricht Raum gegeben werden, da die Anbahnung eines Verständnisses komplexer, dynamischer Systeme für eine adäquate Deutung aktueller Problemlagen wie Klimawandel oder Covid-19-Pandemie sinnvoll erscheint (Grotzer, 2012; Jacobson & Wilensky, 2006). Auch die Akzeptanz von Nichtwissen und unsicherem Wissen als ‚normale‘ Wissensstati ist für das Verständnis der Phänomene und der Diskurse über diese Phänomene hilfreich.

Die folgenden Kapitel geben vertieften Einblick in die angesprochenen Facetten: In Kapitel 3.1 werden zunächst Eigenschaften komplexer Systeme vorgestellt sowie daran anschließend in Kapitel 3.2 die Charakteristika komplexer Systeme in Bezug zu wissenschaftlichen Erklärungen von komplexen Phänomenen gesetzt. Anschließend werden in Kapitel 3.3 die Phänomene ‚Nichtwissen und Unsicherheit‘ näher betrachtet. Die Kapitel beziehen sich dabei auf wissenschaftstheoretische Überlegungen. Sie geben Anhaltspunkte für die Kriterien guten Erklärens biologischer Phänomene, also der fachspezifischen Kriterien instruktionaler Erklärungen (siehe Kapitel 4.6). In Kapitel 3.4 wird der Fokus auf die didaktische Perspektive gerichtet: Es werden Herausforderungen für Lernenden und Lehrende skizziert, die sich bei der Darstellung von Komplexität in Erklärungen bzw. dem Verstehen von komplexen Phänomenen ergeben. In Kapitel 3.5 wird der Versuch unternommen, den Stellenwert und die Bedeutung einer Thematisierung bzw. eines Anbahnens eines Verständnisses komplexer Phänomene und damit einhergehend auch des Nichtwissens und des unsicheren Wissens zu umreißen. Abschließend wird in Kapitel 3.6 auf Grundlage der vorhergehenden Kapitel ein Fazit für Erklärungen im Biologieunterricht gezogen.

3.1 Komplexität als Eigenschaft von biologischen Systemen

Im alltäglichen Sprachgebrauch wird das Adjektiv ‚komplex‘ häufig synonym zum Adjektiv ‚kompliziert‘ verwendet. Dabei verschwimmen jedoch zwei Konzepte, die sich grundlegend unterscheiden: Komplizierte Systeme enthalten eine große Anzahl an Elementen, die meist in linearen Ursache-Wirkungs-Beziehungen stehen. Die Kompliziertheit ist dabei stets vom betrachtenden Subjekt abhängig. Sie ergibt sich aus einer Relation zwischen dem Wissensstand sowie den kognitiven Fähigkeiten einer Person auf der einen Seite und dem Gegenstand auf der anderen Seite (Richter & Rost, 2014). Während beispielsweise die Wahlvorgänge für das Europäische Parlament für einige kompliziert erscheinen, empfinden andere diese vor dem Hintergrund ihrer Erfahrungen und ihres erworbenen Wissens als einfach. Die Kompliziertheit als solche kann durch Wissenszuwachs bzw. Lernen überwunden werden, sodass komplizierte Systeme für Personen zu einfachen Systemen werden können (Richter & Rost, 2014).

Komplexität ist hingegen eine feststehende Eigenschaft eines Systems und damit durch den Betrachtenden nicht veränderbar (Richter & Rost, 2014). Beispiele für komplexe Phänomene bzw. komplexe Systeme – in der Biologie und nahestehenden Wissenschaften – sind unter anderem die Funktion des menschlichen Gehirns (Mainzer, 2008), die Evolution (Mainzer, 2008), Klimaphänomene bzw. der Klimawandel (Rahmstorf & Schellnhuber, 2019) sowie Ökosysteme und ihre Entwicklungen (Bar-Yam, 1997).

Sie bezieht sich nicht auf die bloße Größe eines Systems, sondern ihr sind spezifische Charakteristika inhärent, die im Folgenden in Anlehnung an die Wissenschaftsphilosoph:innen bzw. -theoretiker:innen Kuhlmann (2007), Mainzer (2008) und Mitchell (2008) sowie den Systemforscher Bar-Yam (1997) skizziert werden sollen:

1. Komplexe Systeme sind in der Regel offene Systeme, das heißt, sie stehen in stetigem Austausch mit ihrer Umgebung.
2. Komplexe Systeme weisen eine sehr hohe Anzahl an Elementen (auch Bestandteile oder Freiheitsgrade genannt) auf.²⁶

²⁶ Es gibt jedoch auch einzelne Beispiele komplexer Systeme, bei denen nicht die Anzahl der Elemente ausschlaggebend für die Beschreibung als komplexes System ist, sondern die Komplexität der Dynamik der Entwicklung. Als Beispiel führt

3. Einzelne Elemente können Netzwerke bilden, die wiederum Elemente unter- oder übergeordneter Systemebenen sind. Insgesamt ergibt sich häufig ein System von Systemen.
4. Diese Elemente wechselwirken in vielfältiger Weise miteinander, sodass zwischen den Elementen nicht immer nur eine einzelne Ursache-Wirkungs-Beziehung besteht, sondern es zahlreiche Interaktionen gibt, bei denen ein Element von mehreren Elementen beeinflusst werden kann, gleichzeitig auf einige in Rückkopplungsschleifen zurückwirkt oder Ursache für weitere Elemente ist. Wirkungen können dabei ihrerseits wieder zu Ursachen werden. Durch den hohen Grad an Vernetzung der Elemente können sich einzelne mitunter auch sehr kleine Veränderungen in großem Ausmaß global auswirken.
5. Die Ursache-Wirkungs-Beziehungen müssen dabei nichtlinear verlaufen, wobei eine kleine Ursache auch zu kleinen Wirkungen bzw. eine große Ursache auch zu großen Wirkungen führen würde. Charakteristisch sind hingegen auch nichtlineare Ursache-Wirkungs-Beziehungen, bei denen eine kleine Ursache eine große Wirkung nach sich zieht und umgekehrt.
6. Weiterhin können sich Änderungen in den Anfangsbedingungen oder Störvariablen vollkommen unterschiedlich auf den Prozess auswirken:

Eine Möglichkeit ist, dass die möglichen Endzustände eines komplexen Systems in unerwarteter Weise gar nicht von einer Änderung der Anfangsbedingungen oder Störungen im Verlauf beeinflusst werden (sofern diese nicht zu groß sind), sondern die Endzustände „unbeirrt“ angestrebt werden [...]. In diesem Fall sind die Details auf der Mikroebene in einem bestimmten Rahmen völlig unerheblich, obwohl die Systembestandteile durch ihre Wechselwirkungen die unerwarteten makroskopischen Endzustände erst erzeugen. Dieser ersten Möglichkeit diametral entgegengesetzt besteht die weitere Möglichkeit darin, dass der Endzustand hochsensibel von Änderungen der Anfangsbedingungen reagiert. (Kuhlmann, 2007, S. 315)
7. Hinzu kommt, dass sich beim Überschreiten von kritischen Punkten (auch als Instabilitätspunkte bezeichnet), das auch durch kleinste Änderungen verursacht werden kann, unerwartete Systemzustände, also neue Formen und Eigenschaften auf der makroskopischen Systemebene (neue Ordnungen oder Chaos), einstellen können:

Genauer gesagt kann die Dynamik einen unerwarteten Verlauf zeigen [...]. Das lokale Zusammen- und Wechselspiel der vielen Systembestandteile kann auf komplexe Weise dazu führen, dass sich ein Gleichklang der Einzelteile einstellt, der so aussieht, als wäre durch einen externen Befehl erzeugt. (Kuhlmann, 2007, 314 f.)
8. **Neue Systemzustände sind dabei kein „additives Aggregat von Systemelementen“** (Mainzer, 2008, S. 42). Es entstehen neue Ordnungen und Zusammenhänge zwischen den Systemelementen (Mainzer, 2008). Diese Systemzustände werden als Emergenzen bezeichnet.
9. Bedingt dadurch weisen komplexe Phänomene eine hohe Variabilität auf (Bar-Yam, 1997).

Am Beispiel des Gehirns lassen sich die meisten dieser Charakteristika gut illustrieren: Das Gehirn besteht aus unzähligen Neuronen verschiedener Art, die über Synapsen miteinander verbunden sind. Unabhängig davon, dass noch nicht alle Arten von Neuronen hinsichtlich ihrer genauen Funktion im System des Gehirns beschrieben werden konnten, haben sie eine vereinende Eigenschaft: Neuronen können durch Aktionspotenziale ausgelöst werden und selbst wiederum solche auslösen, in dem sie in einer ‚Alles-oder-nichts‘ Reaktion Neurotransmitter in den Synaptischen Spalt entlassen. Das Entlassen des Neurotransmitters wird durch die Überschreitung eines kritischen Punktes, der an der

Kuhlmann (2007) das Doppelpendel an, dessen Bewegungen chaotisch sind und das somit eine hohe Komplexität der Dynamik aufweist.

Veränderung der Oberflächenspannung festgemacht wird, ausgelöst. Diese Prozesse gehören zur Mikrodynamik. Die feuernden Synapsen erzeugen auf der nächsthöheren Ebene spezifische Schaltmuster, die für unterschiedliche Wahrnehmungen stehen. Diese Wahrnehmungen sind selbst wieder Schaltmuster und können wiederum mit anderen Wahrnehmungen verknüpft werden, sodass auf einer höheren Ebene neue Schaltmuster von Schaltmustern entstehen: „Eine Wahrnehmung kann sich mit anderen Wahrnehmungen der Außenwelt, Selbstwahrnehmungen, Vorstellungen etc. verknüpfen und auf folgender Makroebene komplexe Schaltmuster von Schaltmustern von Schaltmustern etc. erzeugen – bis zu gespeicherten Erinnerungen, die z. B. das Langzeitgedächtnis mit unserer Lebensgeschichte und Ich-Identität einschalten“ (Mainzer, 2008, S. 68–69).

Die Verbindungen zwischen Neuronen sind dabei nicht unveränderbar, sondern können durch Lernen modifiziert werden und tragen damit zur Dynamik des Systems bei. Sie werden als neue Form der Selbstorganisation komplexer neuronaler Systeme beschrieben:

Nur die Voraussetzung und Fähigkeit des Lernens sind in hochentwickelten Organismen (wie z. B. dem Menschen) mit dem Aufbau eines Nervensystems genetisch vorgegeben. Was wir lernen, wie wir Probleme lösen, wie sich Gefühle, Gedanken und Einstellungen entwickeln, ist genetisch nicht im Einzelnen vorgegeben. In der Sprache der nichtlinearen Dynamik könnte also das Auftreten von Bewegungen, Wahrnehmungen, Gedanken, Gefühlen, Bewusstsein u.ä. als Emergenz von makroskopischen Gehirnzuständen aufgefasst werden, die nicht durch einzelne Neuronen, sondern nur durch ihre nichtlineare Wechselwirkung erklärbar werden. (Mainzer, 2008, S. 67)

Insgesamt werden die Charakteristika zur Herausforderung, wenn es um die Rekonstruktion bzw. das Verstehen der Systeme, das heißt um die Erfassung und Durchdringung der beteiligten Elemente und Wechselwirkungen, geht. Durch den hohen Grad an Vernetzung, nichtlineare Ursache-Wirkungs-Zusammenhänge sowie die Möglichkeit von Emergenzen stößt eine reduktionistische Herangehensweise – der Versuch, das System in kleinere Einheiten bis hin zur kleinstmöglichen Einheit zu zerlegen und diese isoliert zu betrachten – an ihre Grenzen (Kuhlmann, 2007; Mainzer, 2008; Mayr, 2002). Bar-Yam (1997) beschreibt deshalb, dass es notwendig wird, neben der Betrachtung der Elemente und ihres Verhalten bzw. ihrer Eigenschaften auch den Blick auf das ganze System zu weiten und damit in Beziehung zu setzen:

[...] to understand the behavior of a complex system we must understand not only the behavior of the parts but how they act together to form the behavior of the whole. It is because we cannot describe the whole without describing each part, and because each part must be described in relation to other parts, that complex systems are difficult to understand. (Bar-Yam, 1997, S. 1)

Mayr (2002) formuliert darauf bezogen ebenfalls, dass ein Wechsel der Betrachtungsebenen für ein Verstehen komplexer Systeme notwendig wird:

In einem biologischen System dagegen treten so viele Wechselwirkungen zwischen den Bestandteilen auf, zum Beispiel zwischen den Genen eines Genotyps, daß selbst ein vollständiges Wissen über die Eigenschaften der kleinsten Bestandteile notgedrungen nur eine Teilerklärung bieten kann. Diesem Sachverhalt wird man nur durch ein *holistisches Denken* gerecht. Nichts ist charakteristischer für biologische Prozesse als die auf allen Ebenen stattfindenden Wechselwirkungen: Zwischen den Genen und Geweben, zwischen Zellen und anderen Komponenten des Organismus, zwischen dem Organismus und seiner unbelebten Umwelt und zwischen verschiedenen Organismen. Gerade diese Wechselwirkungen zwischen ihren Teilen verleihen der Natur als Ganzes, dem Ökosystem, der sozialen Gruppe oder den Organen eines

einzelnen Organismus ihre charakteristischen Ausprägungen. Und ich wiederhole das zuvor Gesagte, daß nämlich die Ablehnung des Reduktionismus kein Angriff auf die Analyse ist. Kein komplexes System kann ohne sorgfältige Analyse verstanden werden. Doch die Wechselwirkungen zwischen den Bestandteilen müssen ebenso untersucht werden wie die Eigenschaften der einzelnen Komponenten. (Mayr, 2002, S. 28–29)

Die genannten Charakteristika komplexer Systeme bedingen auch, dass die Prognose von Entwicklungen eines Systems – zum Beispiel bei der Erstellung von Entwicklungsprognosen von Ökosystemen oder Vorhersagen zum Ansteigen des Meeresspiegels im Kontext des Klimawandels – zur Herausforderung wird: Für komplexe Systeme lässt sich kaum der Zustand im nächsten Moment prognostizieren (Mainzer, 2008). Angestrebt wird eher die Beschreibung der generellen Dynamik des Systems. Betrachtet man die Dynamik, also die Änderung der Systemzustände in der Zeit, laufen sowohl komplexe als auch einfache, lineare Systeme auf einen Endzustand – einen sogenannten Attraktor – zu. **Dabei sind Attraktoren Zustände, in die ein System „langfristig hineingezogen wird“** (Mainzer, 2008, S. 118). Die Art der Attraktoren unterscheidet sich von komplexen und linearen Systemen jedoch grundlegend: Während einfache, lineare Systeme immer zu einem Fixpunktattraktor (=Gleichgewichtszustand) laufen, der sich nicht mehr verändert, können komplexe Systeme beispielsweise auch sogenannte „Grenzzyklen, in denen sich Zustände periodisch wiederholen“ (Mainzer, 2008, S. 47), oder sogenannte Chaosattraktoren anstreben, in denen sich die „Entwicklungslinien völlig irregulär und nicht-periodisch in einem begrenzten Gebiet des Zustandsraums verdichten“ (Mainzer, 2008, S. 47).

Der Grad der Komplexität eines Systems ist dabei variabel und misst sich an der Anzahl der involvierten Elemente, der Anzahl der Interaktionen zwischen den Elementen sowie der Art der Interaktionen (Linearität bzw. Nichtlinearität) und dem Auftreten von unerwarteten Systemzuständen (Kuhlmann, 2007; Mainzer, 2008). Bar-Yam (1997), der sich als Physiker mit komplexen Systemen beschäftigt, abstrahiert den Grad der Komplexität zu **folgender Aussage**: [...] the complexity of a system is the amount of information needed in order to describe it. The complexity depends on the level of detail required in the description (Bar-Yam, 1997, S. 12).

3.2 Wissenschaftliche Erklärungen von und in komplexen Systemen

Aus den genannten Charakteristika von Komplexität bzw. komplexen Systemen resultieren Charakteristika für Erklärungen – für wissenschaftliche Erklärungen, jedoch folglich auch für instruktionale Erklärungen im Biologieunterricht, mit denen Lehrkräfte die wissenschaftlichen Fakten und Zusammenhänge an Schüler:innen vermitteln bzw. die wissenschaftlichen Erklärungen von Phänomenen nahebringen (Letzteres siehe Abschnitt 3.6).

Aufgrund der Anzahl involvierter Elemente sind Erklärungen von komplexen Systemen zunächst als multikausal zu charakterisieren, wobei es eben keine zentrale Ursache gibt, von der aus eine einzelne Wirkung oder eine Kette von Wirkungen in einer linearen Art und Weise gedacht werden kann. Auch der Art der Wechselwirkungen muss dabei Rechnung getragen werden, da es sich häufig um nichtlineare Ursache-Wirkungs-Zusammenhänge handelt (Bartelborth, 2007; Mainzer, 2008; Potochnik, 2013; siehe auch in Kapitel 2.3.1 im Abschnitt zu kausalen Erklärungen).

Weiterhin weisen komplexe Systeme typischerweise eine hohe Dynamik auf – sie zeigen eine Veränderung in Abhängigkeit von der Zeit, die unvorhersehbar sein kann (Mainzer, 2008). Dabei ist die Vorhersage des Zustands des Systems meist nicht möglich, sondern es werden Entwicklungsverläufe abgeschätzt. Wahrscheinlichkeitsabhängige Vorhersagen bzw. Erklärungen

sind daher gängige bzw. notwendige Varianten (Bar-Yam, 1997). Dabei sind zahlreiche Parameter von bekannten Elementen (z. B. Art und Stärke der Wechselwirkungen) abzuschätzen, sodass unterschiedliche Modelle respektive Erklärungen und Vorhersagen resultieren (Kampourakis & McCain, 2020; Rahmstorf & Schellnhuber, 2019).

An dieser Stelle wird die im Eingang zu Kapitel 3 erwähnte Verbindung von Komplexität und Nichtwissen, unsicherem Wissen sowie Risiko evident: Dem Versuch, (komplexe) Wirkungsgefüge aufzudecken, sind epistemologische Grenzen gesetzt. Die Erklärungen und Vorhersagen enthalten einen Anteil an unsicherem Wissen, der nicht beziffert werden kann, oder berechenbare Abschätzungen, die mit Risiko bezeichnet werden können (Unterscheidung siehe Kapitel 3.3). So ergeben sich beispielsweise bei der Prognose für die Höhe des Anstiegs des Meeresspiegels verschiedene Szenarien (Rahmstorf & Schellnhuber, 2019). Hinzu kommt, dass die Grenzen eines (komplexen) Systems sowie die darin enthaltene Elemente und deren Wechselwirkungen nicht direkt ersichtlich oder beobachtbar sind, sondern anhand empirischer Daten und auf Grundlage von (bisher bekanntem) Wissen rekonstruiert werden müssen. Folglich ist es mitunter nicht möglich, Elemente des Systems als solche oder die Relevanz der Elemente für das System zu erkennen (Kampourakis & McCain, 2020; Mitchell, 2008). Damit ergibt sich ein Anteil an Nichtwissen in Erklärungen und Vorhersagen. Als Beispiel kann der Einsatz von FCKW (Fluor-Chlor-Kohlenwasserstoff-Verbindungen) als Kühlmittel in Kühlschränken und Klimaanlage oder Treibgas in Rasierschaum- oder Haarspraydosen ab den 1930er Jahren genannt werden: Der Zusammenhang von FCKW und der Vergrößerung des Ozonlochs – also die Relevanz des Elements des FCKW für die Erdatmosphäre – waren nicht bekannt, sodass der Einsatz zunächst nur wesentlich positive Effekte erzielte (Wehling, 2009). Nichtwissen und unsicheres Wissen sind in Erklärungen komplexer Phänomene demnach charakteristisch.²⁷

Auch wenn das Gesamtsystem komplex ist, muss nicht jede Erklärung eines biologischen Phänomens diese Komplexität widerspiegeln, da es auf die Art der Frage²⁸ und damit einhergehend auf die betreffende Systemebene ankommt. So kann innerhalb des komplexen Systems ‚menschlicher Organismus‘ der Mechanismus zur Synthese des RNA-Primers durch die Primase an einem DNA-Strang im Zuge der DNA-Replikation beschrieben werden, ohne die Komplexität des Gesamtorganismus zu berühren. Um den Aspekt der Art der Frage und damit auch der Systemebene als Faktor zu verdeutlichen, kann ein Beispiel aus der Genetik angeführt werden. So ist die Frage nach der Ursache für die Symptomatik des Downsyndroms durch die Angabe, dass das Chromosom 21 dreimal statt zweimal vorhanden ist, beantwortbar. Die Ursache kann jedoch auch auf Ebene der Gene – also unter anderem mit der Expression von Genen – und den Zusammenhängen zwischen den entstandenen Proteinen und deren Funktionen im Organismus erklärt werden. Die Komplexität nimmt weiter zu, wenn nicht allgemeine Symptome des Downsyndroms, sondern Ursachen für die Ausprägung der Symptome bei einem Individuum erklärt werden sollen. Letzteres leitet über zu einem weiteren Aspekt, durch den die Komplexität biologischer Phänomene in Erklärungen nicht in Erscheinung tritt.

Auch durch Verallgemeinerung kann die Komplexität des realen Phänomens reduziert werden, wobei hierdurch Informationen reduziert werden und damit das beobachtbare Phänomen nur noch

²⁷ Nichtwissen und unsicheres Wissen sind nicht genuin mit Erklärungen komplexer Phänomene verknüpft. Der Anteil an Nichtwissen und unsicherem Wissen ist hier lediglich höher, da mehrere Elemente involviert sind und die Art der Wechselwirkungen höher ist.

²⁸ Für den Biologieunterricht ist die Auswahl der Frage ein entscheidender Aspekt, mit dem die Lehrkraft einen Beitrag dazu leistet, welche Inhalte die Schüler:innen lernen und welches Bild (implizit) von der Disziplin der Biologie vermittelt wird.

in Teilen repräsentiert wird.²⁹ Als Beispiel kann hier der Zusammenhang zwischen Populationsdichte von Räuber- und Beutetieren dienen, der die Populationsdynamik in einem Ökosystem als zyklische Schwankungen beschreibt: Aufgrund der Abhängigkeit der Räuber von der Nahrungsquelle ‚Beute‘ nimmt die Dichte der Räuberpopulation nach dem Anstieg der Dichte der Beutepopulation ebenfalls zu, da wieder vermehrt Nahrung verfügbar ist. Ist die Beutepopulation durch die Räuber stark bejagt worden, fällt die Dichte der Räuberpopulation zeitversetzt mit der Dichte der Beutepopulation auch wieder ab. Diese Schwankungen wiederholen sich zyklisch (Campbell & Reece, 2009). Das sogenannte Lotka-Volterra-Modell beschreibt modellhaft zyklische Populationsschwankungen, wobei Erklärungen und Prognosen nur annähernd zutreffen, wenn die Art der Räuber hauptsächlich von einer Art der Beute abhängig ist, was in den meisten als stabil geltenden Ökosystemen nicht der Fall ist. Gleichzeitig wirken abiotische Faktoren ebenso auf die Dichte der Räuber- und Beutepopulation, über die das Modell keine Aussagen trifft, sodass das Modell für Erklärungen und Prognosen auch hier meist nur einen ersten Anhaltspunkt bieten kann. Beschrieben wird folglich eher ein allgemeines Prinzip, das nur in Einzelfällen (z. B. beobachtete Zusammenhänge zwischen Luchs- und Schneeschuhhasenpopulation in Alaska) hinreichend zutrifft (Campbell & Reece, 2009).

3.3 Nichtwissen und unsicheres Wissen als Teil von Erkenntnis

In einer Wissensgesellschaft, in der Expert:innen ein Status aufgrund ihres Wissensstandes zugeschrieben wird, sowie im Kontext von wissensvermittelndem Unterricht und instruktionalem Erklären, von Nichtwissen und unsicherem Wissen zu sprechen, scheint zunächst etwas kontraintuitiv zu sein. Im vorhergegangenen Abschnitt (3.2) wurde jedoch bereits an einigen Stellen aufgezeigt, dass wissenschaftliche Erklärungen unter Bedingungen von unsicherem Wissen und Nichtwissen gegeben werden. Das vielschichtige Phänomen ‚Nichtwissen und unsicheres Wissen‘ soll deshalb in den folgenden Abschnitten (3.3.1 bis 3.3.3) differenziert betrachtet werden, um zur Klärung beizutragen.

Unsicheres Wissen und Nichtwissen spielen für die Wissenschaft keine marginale Rolle, sondern machen einen Großteil der Wissenschaften bzw. des wissenschaftlichen Arbeitens und des naturwissenschaftlichen Wissens aus: Zum einen sind Nichtwissen und unsicheres Wissen die Quelle neuer wissenschaftlicher Fragen und Untersuchungen, die wiederum zu mehr Wissen führen können. Zum anderen erzeugt die Schaffung neuen Wissens immer neue Fragen und damit neues Nichtwissen sowie unsicheres Wissen; Luhmann (1997) und Ravetz (1987) nehmen die Zunahme **des Nichtwissens bei der Wissensschaffung sogar als „überproportional“** an (Luhmann, 1997, S. 1106). Schließlich werden Nichtwissen und unsicheres Wissen respektive deren Auswirkungen und Bedeutungen dann besonders deutlich, wenn unterschiedliche Erklärungen für ein Phänomen nebeneinander existieren, Erklärungen durch neue Befunde angepasst oder sogar verworfen werden oder Prognosen von Wissenschaftler:innen unterschiedliche Szenarien hervorbringen (Kampourakis & McCain, 2020; Rahmstorf & Schellnhuber, 2019).

Insbesondere seit den sogenannten Risikokrisen der 1980er Jahre beschäftigen sich zunächst Wissenschaftssoziolog:innen³⁰ mit den Dimensionen, Auswirkungen auf Gesellschaft, Politik und Wissenschaft des Nichtwissens und unsicherem Wissen (Wehling, 2008). Dazu zählt zum Beispiel

²⁹ Eine Verallgemeinerung erfolgt ebenfalls graduell. Sie kann nicht per se als gut oder schlecht beurteilt werden, sondern die Angemessenheit muss an ihrem Zweck beurteilt werden.

³⁰ Ein Diskurs über Nichtwissen bestand zuvor vor allem in den Sozialwissenschaften. Hier ist Nichtwissen – konträr zur wissenschaftstheoretischen Diskussion – eher positiv besetzt, da beispielsweise Luhmann (1997) das Nichtwissen als (positiv) funktional für die Gesellschaft bzw. eine Gruppe beschreibt.

die Verursachung der Vergrößerung des Ozonlochs durch industriell hergestellten Fluor-Chlor-Kohlenwasserstoff (FCKW). Die Folgen des weitverbreiteten Einsatzes von FCKW waren lange Zeit nach der Einführung in den 1930er Jahren unbekannt (Wehling, 2009). Auch die Reaktorkatastrophe in Tschernobyl verstärkte den Blick auf Risikoabschätzungen und den Umgang bzw. die Kommunikation von Nichtwissen sowie unsicherem Wissen (Wehling, 2009).

Die zwei kurzen Abschnitte deuten bereits an, dass es sich bei Nichtwissen und unsicherem Wissen um einen vielschichtigen Themenkomplex handelt, bei dem es nicht nur um ein objektiv und neutral bestimmbares, graduelles Nichtvorhandensein von sicherem Wissen geht, sondern der Umgang mit diesem eine große Rolle für Individuen bzw. die Gesellschaft spielt. So werden Wissensstände bewertet, (wahrgenommene) Wissensstände auf verschiedene Arten und Weisen kommuniziert und Handlungen unter diesen Umständen geplant und durchgeführt.

Die folgende Grafik (Abb. 1) strukturiert die Perspektiven auf den Themenkomplex und schafft zunächst einen Überblick über zentrale Begrifflichkeiten und unterscheidbare Ebenen, die zur Beschreibung von Nichtwissen und unsicherem Wissen beziehungsweise damit verbundenen Phänomenen herangezogen werden können.

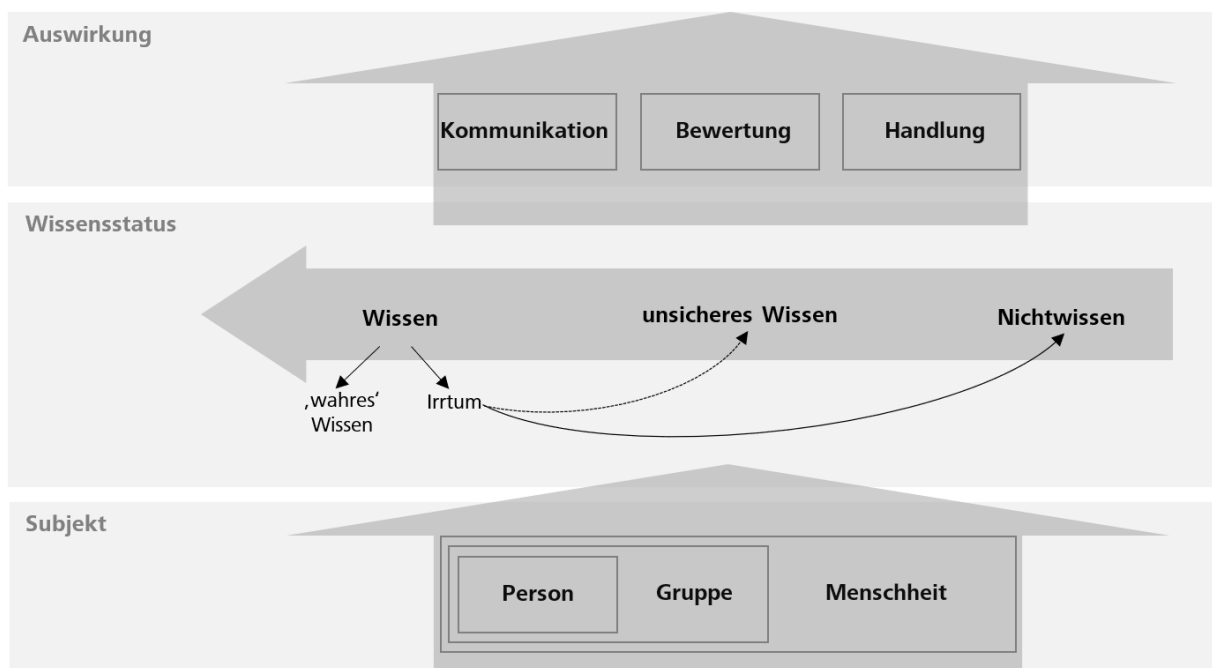


Abbildung 1: Zentrale Begriffe und unterscheidbare Ebenen im Kontext von verschiedenen Wissensstati

Auf der Ebene der Subjekte³¹ kann unterschieden werden, wer einen Wissensstatus³² bezüglich eines Phänomens bzw. eines Zusammenhangs oder Faktums besitzt – also etwas nicht weiß, sich unsicher ist oder etwas weiß. Das können entweder einzelne Personen oder Gruppen von Menschen sein oder es betrifft alle Personen gleichermaßen (siehe Kap. 3.3.1).

³¹ Der Begriff Subjekt wird hier in Anlehnung an eine Verwendung in der Soziologie verstanden. Endruweit (2014, S. 521) definiert, dass der Begriff des sozialen Subjekts eine „Sammelbezeichnung für alle Gegenstände der Soziologie [ist], die oder deren Mitglieder physisch existieren und die selbst oder durch Repräsentationen handeln können bzw. denen soziales Handeln zugeschrieben wird. Zu den sozialen Subjekten [...] gehören also z. B. einzelne Menschen, Familien, Klassen, Gruppen, Organisationen“.

³² Der Begriff Wissensstatus wird hier und im Folgenden als Oberkategorie für Nichtwissen, unsicheres Wissen und Wissen verwendet – nicht für Wissensarten wie zum Beispiel prozedurales und deklaratives Wissen.

Die darauffolgende Ebene differenziert die möglichen Wissensstadien – den Status des Wissens, den eine Person, eine Personengruppe oder die gesamte Menschheit besitzen kann – auf einem quantitativen Kontinuum von einem vollständigen ‚blinden Fleck‘ Nichtwissen über unsicheres Wissen bis zu Wissen über ein Phänomen (siehe Kap. 3.3.2; Wehling, 2008). Dabei ist das diskriminierende Kriterium zwischen den Wissensstadien nicht die individuelle Wahrnehmung, sondern die wissenschaftliche Evidenzlage.³³ Das Wissen kann sich im Laufe der Zeit als ‚wahres‘ Wissen oder Irrtum – also fälschlich angenommene Zusammenhänge und Fakten – herausstellen (Wehling, 2008). Die Unterscheidung ist mitunter gar nicht oder nur mit zeitlichem Abstand möglich,³⁴ führt jedoch im Falle eines Irrtums durch das Ablehnen oder Verwerfen aller oder eines Teils der Annahmen zunächst zu einem weniger sicheren Wissensstatus. Dass es sich tatsächlich um wahres Wissen handelt, ist nicht vollständig sicherzustellen: Mit einer zunehmenden Menge an Daten, die für die Erklärung oder die Theorie sprechen, wird lediglich die Wahrscheinlichkeit für das Vorhandensein von wahrem Wissen erhöht. Ein endgültiger Beweis ist hingegen nicht möglich (Kampourakis & McCain, 2020; siehe Kap. 3.3.2).

Die letzte Ebene der Grafik bezieht sich auf die Auswirkungen des Wissensstatus. Damit sind insbesondere soziale Phänomene angesprochen, die sich durch die Art der Kommunikation von und über Nichtwissen und unsicherem Wissen, die unterschiedliche Wahrnehmung und Bewertung des Wissensstatus sowie bei Entscheidungsprozessen und Handlungen unter Bedingungen unterschiedlicher Wissensstände ergeben. Die Ebene umfasst damit eine Vielzahl an Phänomenen, die von der Sprachwissenschaft, Kommunikationswissenschaft, Soziologie, Psychologie sowie den Didaktiken (insbesondere Naturwissenschaftsdidaktiken) untersucht werden (siehe Kapitel 3.3.3).

3.3.1 Betroffene Subjekte – Wer weiß etwas (nicht)?

Der Begriffsbestimmung von Nichtwissen und unsicherem Wissen muss zunächst eine grundlegende Unterscheidung in Bezug auf das Subjekt, das nicht weiß oder über unsicheres Wissen verfügt, vorausgehen: Träger:innen von Nichtwissen und unsicherem Wissen können einzelne Personen oder Gruppierungen sein (z. B. Schüler:innen einer Klasse, Forscher:innen, Forschungsgruppe) oder das Nichtwissen und die Unsicherheit betrifft alle Personen gleichermaßen, da entsprechendes Wissen bzw. gesichertes Wissen (noch) nicht geschaffen wurde.³⁵ Dieser Unterschied – Personen und Gruppen von Personen auf der einen Seite und alle Personen gleichermaßen auf der anderen Seite – ist für das Phänomen ‚Nichtwissen‘ von besonderer Bedeutung.

Das Nichtwissen auf individueller Ebene ist für alle Personen erfahrbar und für Ausgangssituationen in Lehr-Lern-Kontexten charakteristisch. Auch im wissenschaftlichen Bereich ist Nichtwissen in Bezug auf einzelne Personen allgegenwärtig, wobei das „Nichtwissen ein normaler, in Diskussionen unter Umständen für das eigene Image aber problematischer Zustand ist und ebenso in der Forschungspraxis eine wichtige Rolle spielt“ (Rhein, 2018, S. 72; siehe auch Kap. 3.3.3). Unter

³³ Die Präzisierung rekurriert auf Aussagen von beispielsweise wissenschaftsskeptischen Gruppierungen, dass wissenschaftliches Wissen bzw. Evidenz grundlegend nicht sicher sei.

³⁴ Als ein (historisches) Fallbeispiel kann hierfür der Schritt vom geozentrischen zum heliozentrischen Weltbild angeführt werden. Während zunächst angenommen wurde, dass die Erde im Zentrum der Bewegungen umliegender Himmelskörper steht, stellte Kopernikus die Vermutung auf, dass die Sonne im Zentrum steht. Dafür fehlende Daten und entsprechende Berechnungen lieferte Kepler (Mittelstrass, 2007).

³⁵ Wehling (2009) bezeichnet das individuelle Nichtwissen auch als „Unwissenheit“ (Wehling, 2009, S. 97).

unsicherem Wissen wird in Bezug auf eine einzelne Person verstanden, wenn diese etwas zu wissen glaubt, sich der Korrektheit dessen jedoch nicht sicher ist (Taddicken, Reif & Hoppe, 2018).^{36 37}

Neben einzelnen Personen oder Personengruppen kann die gesamte Menschheit dasjenige Subjekt sein, das Nichtwissen oder unsicheres Wissen besitzt; das bedeutet, dass (noch) kein Wissen geschaffen wurde. Davon ausgehend, dass insbesondere Wissenschaftler:innen zur Schaffung von Wissen beitragen, wird dieses Nichtwissen im Folgenden als gesamtwissenschaftliches Nichtwissen bezeichnet. Dabei sind Nichtwissen und Unsicherheit grundlegende, konstituierende Aspekte von Wissenschaft und Forschung: Sie sind die Quellen für neue wissenschaftliche Fragen und Untersuchungen, die ihrerseits wieder Nichtwissen und Unsicherheiten generieren und zu neuen Fragen und Untersuchungen führen. Damit sind sie zunächst als neutral oder „normal“ (Janich & Rhein, 2018, S. 7) zu betrachten bzw. tragen „Nichtwissen und Unsicherheiten [...] ganz grundlegend zur Entwicklung der Wissenschaft und zum Forschungsfortschritt bei“ (Rhein, 2018, S. 77).

Die Unterscheidung der Subjekte ist insbesondere aufgrund ihrer verschiedenen Folgerungen relevant: Während gesamtwissenschaftliches Nichtwissen – sofern es als solches erkannt wird – Forschungslücken markiert und damit neue Forschungsarbeiten initiieren kann (aber nicht muss – siehe dazu Kapitel 3.2.2), verweist das individuelle Nichtwissen auf einen Informationsrückstand, der eine Recherche im Sinne einer Wissensbeschaffung notwendig macht.

Im Folgenden wird der Fokus hauptsächlich auf das gesamtwissenschaftliche Wissen, Nichtwissen und unsicheres Wissen gelegt, da diese Facette der fachlichen Konzeptualisierung von Erklärungen vorrangig dienlich ist – das individuelle Wissen, Nichtwissen und unsicheres Wissen würde sich im Rahmen der Adressatenorientierung von Erklärungen wieder finden.

3.3.2 Wissensstati – Wie viel wissen wir über die Welt? Wie viel können wir über die Welt wissen?

Um die Wissensstati differenzieren zu können, sollen zunächst der Pol ‚Wissen‘ grob skizziert und anschließend die weiteren Formen (Nichtwissen, unsicheres Wissen), die Abstufungen von nicht vorhandenem Wissen darstellen, beschrieben werden.

Die wohl bekannteste Definition aus der Philosophie, von der die meisten Argumentationen und Diskussionen rund um Wissen ausgehen (Anacker, 2007; Gettier, 1963), stammt von Platon (abgehandelt in den Werken Theaitetos und Meno; siehe Anacker, 2007): Wissen ist wahre, begründete Überzeugung.³⁸ Die Definition besteht in ihrer Kürze aus drei Komponenten: „einer objektiven (Wahrheit), einer subjektiven (Glaube/Meinung/Überzeugung) und einer, die diese beiden irgendwie verbinden soll (Begründung/Rechtfertigung)“ (Anacker, 2007, S. 356).

³⁶ Besonders relevant ist diese Frage zur Sicherheit des eigenen Wissens bei Datenerhebungen, die den Wissensstand von Personen abfragen (Taddicken, Reif & Hoppe, 2018).

³⁷ Auf individueller Ebene kann weiterhin eine relevante Unterscheidung bezüglich möglicher Wissensstati getroffen werden: die Abgrenzung zwischen Wissen, Nichtwissen und unwahrem Wissen. Die Autor:innen Taddicken, Reif und Hoppe (2018) bezeichnen das falsche bzw. unwahre Wissen von Personen, das diese als eigenes, sicheres Wissen angeben, dass sich jedoch nicht mit dem in der Wissenschaft akzeptierten Wissen deckt, als „Unwissen“ (Taddicken, Reif & Hoppe, 2018, S. 119).

³⁸ Durch verschiedene Übersetzungen manchmal auch: „Wissen ist wahre, gerechtfertigte Meinung“ (Anacker, 2007, S. 356).

- i) Die Bedingung der Wahrheit ist für die Definition essenziell – nur was wahr ist, kann auch Wissen sein. Sollte sich zu einem späteren Zeitpunkt herausstellen, dass dieses Wissen nicht wahr bzw. falsch ist, handelte es sich um vermeintliches Wissen. Die Wahrheit ist eine notwendige, wenngleich schwierige Bedingung, da sie nicht überprüfbar ist. Sie berührt den Kern der Frage nach der Ontologie und der Epistemologie: Was ist wirklich? Kann man die Welt tatsächlich erkennen, wie sie ist?³⁹
- ii) Weiterhin ist die Bedingung einer Begründung grundlegend.
 - Vor dem Hintergrund des erkenntnistheoretischen Hiatus zwischen Subjekt und Objekt klingt das sehr plausibel, eine unmittelbare Konfrontation einer Überzeugung mit >der Wahrheit< ist ja per definitionem unmöglich, und so liefert die Begründung das einzig zugängliche Kriterium, anhand dessen die Wahrheit einer Überzeugung überprüft werden kann. (Anacker, 2007, S. 360)⁴⁰
- iii) Die Bedingung der Überzeugung setzt voraus, dass etwas nur als Wissen gelten kann, wenn es auch für wahr gehalten wird, das heißt, wenn eine Person oder Personengruppe davon überzeugt ist, dass der Fakt oder der Zusammenhang wahr ist.

Da die Wahrheit nicht bestätigt (nur widerlegt) werden kann, formuliert Wehling (2008), dass es für die Wissenssoziologie produktiver sei, Wissen als „*erfahrungsbasierte Überzeugung und Erwartungen*“ zu verstehen, deren Wahrheitsgehalt jeweils klärungsbedürftig ist und strittig sein kann“ (Wehling, 2008, S. 21–22). Andernfalls würde sich eine Situation ergeben, in der nicht mehr zwischen Wissen und Nichtwissen unterschieden werden kann. Dadurch ergibt sich weiterhin die Möglichkeit bzw. Notwendigkeit, bei Wissen zwischen Irrtum und wahren Wissen (siehe linker Pol auf Ebene der Wissensstati in der Grafik) zu unterscheiden.

Mit dem Blick auf den rechten Pol des Kontinuums der Wissensstati beschreibt Wehling (2008) das Nichtwissen als „*Abwesenheit kognitiver Überzeugungen*“ (Wehling, 2008, S. 22). Damit wird auch eine Trennlinie zwischen Nichtwissen und den weiteren Wissensstati gezogen – also zwischen dem rechten Pol des Kontinuums in der Grafik und allen weiteren möglichen Abstufungen, die weiter links davon liegen. Die Trennung erfolgt an der Stelle, an der „*eine Form von wie auch immer begrenztem, hypothetischem und unsicherem Wissen*“ (Wehling, 2006, S. 110) vorhanden ist.⁴¹ Die theoretische Trennung ist jedoch häufig schwer in realen Phänomenen zu bestimmen (Wehling, 2008).

Nichtwissen als komplementären Begriff zu Wissen zu beschreiben, reicht dabei nicht aus, um das Phänomen des Nichtwissens zu erfassen. Wehling (2006, 2008) erarbeitet deshalb drei Dimensionen von Nichtwissen, die dazu genutzt werden können, den vielschichten Charakter von Nichtwissen zu beschreiben und dessen Erscheinungsformen kenntlich zu machen:

- a) (Nicht-)Wissen des Nichtwissens
- b) Intentionalität des Nichtwissens
- c) Zeitliche Stabilität des Nichtwissens

³⁹ Entsprechend wird in der Philosophie auch diskutiert, ob es überhaupt Wissen geben kann (Anacker, 2007)

⁴⁰ An der Bedingung der Begründung wird bereits von Sokrates kritisiert, dass sie nichts zur Wahrheit hinzufügen kann. Wenn es sich um wahres Wissen handelt, was kann dann eine Begründung noch hinzufügen? Die Bedingung der Wahrheit nimmt also alles voraus und macht ein Begründen mehr oder weniger unnötig und führt mitunter zu einem Zirkelschluss, dass eine Begründung auch wieder aus Wissen, dessen Wahrheitsgehalt infrage steht, bestehen muss (Anacker, 2007). Für Gettier (1963) ist die Bedingung der Begründung hingegen nicht hinreichend und stellt damit die Definition grundlegend infrage. Mit zwei Fallbeispielen belegte er eine logische Lücke zwischen der Formulierung einer Begründung und deren Beziehung zum Wahrheitsgehalt.

⁴¹ Es gibt auch Positionen, die unsicheres Wissen und Nichtwissen zusammenfassen (z. B. Rhein, 2018).

Die genannten Dimensionen sind in den Beschreibungen jeweils bipolar angelegt, wobei die Pole über ein nicht näher bestimmtes Kontinuum verbunden sind (Wehling, 2008). Die drei Dimensionen sollen in den nächsten Abschnitten erläutert werden.

a) (Nicht-)Wissen des Nichtwissens

Dargestellt werden die zwei Pole dieser Dimension einerseits durch das explizit gewusste Nichtwissen (man weiß genau, was man nicht weiß) und andererseits das unerkannte Nichtwissen, von dem man sich keine Vorstellung machen kann, dass dieses überhaupt existiert (man weiß nicht, was man nicht weiß) (Wehling, 2008). Die Unterscheidung ist besonders brisant,

[...] denn auf der Ebene empirischer Beobachtungen sind positives Wissen und unerkanntes Nichtwissen nicht voneinander zu unterscheiden: Wissen wir, dass eine neue Technologie keine schädlichen Folgen hat, wenn hierfür keinerlei empirische Hinweise vorliegen? Oder bedeutet dies lediglich, dass wir ahnungslos sind, wo, in welchen Zeiträumen und in welcher Form sich negative Konsequenzen zeigen könnten – oder sogar schon eingetreten sind. (Wehling, 2008, S. 23)

Es kann folglich nicht mit Sicherheit bestimmt werden, ob tatsächlich kein unerkanntes Nichtwissen vorliegt, sondern lediglich mit zunehmender Forschung die Wahrscheinlichkeit verkleinert wird, dass Risiken oder negativen Konsequenzen übersehen werden (Kampourakis & McCain, 2020; Wehling, 2008). Diese Unterscheidung macht deutlich, dass die Positionierung bzw. Einschätzung des Nichtwissens von Personen oder Personengruppen einen großen Einfluss auf die subjektive Risikobewertung in gesellschaftspolitischen Diskussionen haben kann.

b) Intentionalität des Nichtwissens

Hier steht die Intention des Nichtwissens im Sinne einer Zurechnung zum „Handeln oder Unterlassen sozialer Akteure (Individuen, Gruppen, Organisationen)“ (Wehling, 2008, S. 23) als Kriterium im Vordergrund. Dabei werden die Extrempole auf der einen Seite von einem intendierten, absichtsvollen Zurückweisen von Wissen (Nicht-Wissen-Wollen) und auf den anderen Seiten von einem in Gänze unbeabsichtigten und daher unvermeidbarem Nichtwissen gebildet (Wehling, 2008).

Das intendierte, absichtsvolle Nichtwissen umfasst differenzierbare Facetten, die sich mithilfe von Beispielen veranschaulichen lassen:

- Individuen (oder Gruppen) können sich aktiv für ihr eigenes Nichtwissen entscheiden (Wehling, 2008). Beispielsweise können sich Paare gegen die Durchführung von pränataldiagnostischen Maßnahmen entscheiden und damit ihr Recht auf Nichtwissen einlösen. Wissenschaftler:innen können sich ebenso aktiv für Nichtwissen oder Wissen entscheiden – so kann oder muss diskutiert werden, ob oder unter welchen Bedingungen an der Entwicklung von Waffen geforscht werden darf. Im Kontext eines rasant anwachsenden Wissensbestandes, der zur selektiven Verarbeitung der produzierten Daten und Ergebnisse zwingt, wird auch zunehmend über Strategien und Handlungsempfehlungen zum Schutz vor einer Wissensüberflutung nachgedacht (Wehling, 2006).
- Individuen (oder Gruppen) können sich aktiv für das Nichtwissen anderer Individuen (oder Gruppen) entscheiden und Wissen geheim halten oder verschleiern. So wird vermutet, dass das Wissen über eine Verunreinigung von Babypuder der Marke Johnson & Johnson mit Asbest den Produktverantwortlichen wohl bekannt war (Kort, 2021).

Die Frage ‚Ab wann ist Nichtwissen fahrlässig in Kauf genommen worden?‘ markiert eine wichtige Form zwischen den beschriebenen Polen beabsichtigtes (Nicht-Wissen-Wollen) und unbeabsichtigtes Nichtwissen und zeigt erneut die gesellschaftspolitische Relevanz bzw. Brisanz. Wehling (2008) führt hier den Contergan-Skandal an, bei dem es auch juristisch zu beurteilen galt, ob die Herstellenden wissen hätten müssen, dass der Wirkstoff Thalidomid schwere Fehlbildungen bei Föten zur Folge haben kann. Ein weiteres Beispiel, das sich zwischen den zwei Polen positionieren lässt, ist der Einfluss geldgebender Instanzen auf die Forschungsthemen, die bearbeitet werden, und darauf, ob bzw. wie damit aufgrund fehlender finanzieller Anreize andere Forschungsthemen in den Hintergrund rücken.

c) *Zeitliche Stabilität*

Die letzte Dimension erfasst die Möglichkeit bzw. die Unmöglichkeit, den Nichtwissensbestand zu ändern, die durch die Pole ‚Noch-nicht-wissen‘, also reduzierbares Nichtwissen, und ‚Nicht-wissen-können‘, also irreduzibles Nichtwissen, beschrieben werden (Wehling, 2008). Die Zuordnung zu einer der Formen ist vorab meist nicht abzusehen, sodass zunächst meist offen ist, ob das Nichtwissen in Zukunft aufgelöst bzw. in Wissen überführt werden kann.

In der aktuellen Wissensgesellschaft besteht jedoch die Annahme, dass es sich meist um zeitlich begrenztes Nichtwissen handelt, das mit weiterer Forschung in Wissen überführt werden kann (Wehling, 2008). In diesem Kontext stellt Wehling (2008) auch die offene Frage, ob die Wissenschaft alle relevanten Fragen in einem bestimmten Gebiet in absehbarer Zeit beantworten können wird oder bei komplexen Wirkungszusammenhängen mit unauflösbarem ‚Nicht-wissen-können‘ gerechnet werden muss.

In der Grafik wird der Bereich zwischen den Polen Wissen und Nichtwissen mit unsicherem Wissen überschrieben. Damit ist nicht der Aspekt gemeint, dass der Erkenntnis stets eine Unsicherheit anhaftet, weil jede Erkenntnis nur durch die verfügbaren Erkenntnismethoden entstehen kann, also ein vermittelter Zugang zur Wirklichkeit besteht – also die Frage nach der Ontologie, dem tatsächlichen Sein. Hingegen sind darunter verschiedenartige Phänomene zu fassen, bei denen keine vollständigen Erklärungen oder sichere Vorhersagen getroffen werden können. Dabei können auf der einen Seite empirische Daten fehlen oder auf der anderen Seite widersprüchliche empirische Daten oder Ansichten vorhanden sein. „Unsicheres Wissen ist also auch Ergebnis einer Kontroverse, wenn verschiedene Evidenzen einander widersprechen, wenn agonale Positionen vertreten werden und dadurch Wissenskonflikte entstehen“ (Janich & Rhein, 2018, S. 8).

Mit einer Unterscheidung des Psychologen Gigerenzer wird eine Differenzierung in Hinblick auf die Beschaffenheit von unsicherem Wissen im Sinne eines Fehlens empirischer Daten möglich: auf der einen Seite ‚Risiko‘, wobei auf der Grundlage von Modellen eine Wahrscheinlichkeit für das Eintreten eines Resultats beziffert werden kann, und auf der anderen Seite Ungewissheit, bei der eine solche quantitative Abschätzung nicht möglich ist (Gigerenzer, 2019). Kampourakis und McCain (2020) greifen die Unterscheidung ebenfalls – jedoch unter den Begriffen „**known uncertainties**“ (Risiko) und „**unknown uncertainties**“ (Ungewissheit; Kampourakis & McCain, 2020, S. 43) – auf und beziehen sich damit stärker auf den Wissensbegriff.

Die Unterscheidung zwischen den Ebenen ‚Subjekte – Wissensstati – Auswirkungen‘ wie sie im Modell in Kapitel 3.3 getroffen wird, gewinnt hier noch einmal besondere Bedeutung, da die Unsicherheit von Wissensstati – zum Beispiel über den Klimawandel (Kampourakis & McCain, 2020; Rahmstorf & Schellnhuber, 2019) – in ihrer Wahrnehmung von Personen oder Personengruppen

stark variieren kann. Dabei können die Unsicherheiten politisch genutzt werden, indem sie in der Kommunikation überspitzt und damit instrumentalisiert werden, um Ziele zu erreichen.

3.3.3 Auswirkungen – Kommunizieren, Bewerten und Handeln

Nichtwissen und unsicheres Wissen haben Auswirkungen auf die Kommunikation über Phänomene und Problemlagen sowie auf deren Bewertung und beeinflussen das Handeln der Individuen und der Gesellschaft. Die ersten zwei Ebenen – Subjekte und Wissensstati – würden den Themenkomplex deshalb nicht repräsentieren. Wehling (2008) formuliert dazu aus Sicht eines forschenden Wissenssoziologen:

Zu untersuchen ist nicht mehr „nur“, wer was nicht weiß und warum nicht, worüber Wissen produziert wird und worüber nicht, wie das vorhandene Wissen verteilt und zugänglich gemacht wird oder welchen sozialen Akteuren von anderen Unkenntnis und Ahnungslosigkeit zugeschrieben wird. Von zentraler Bedeutung ist vielmehr auch: In welcher Form wird das Nichtwissen wahrgenommen (und anderen zugeschrieben), weshalb wird es so und nicht anders wahrgenommen, zum Beispiel vorwiegend als erkannte und temporäre Wissensdefizite und nicht als unerkanntes oder unüberwindliches Nichtwissen? (Wehling, 2008, S. 24)

Die Auswirkungen werden in verschiedenen Disziplinen in den Blick genommen und unter verschiedensten Perspektiven beleuchtet. Sie sollen an dieser Stelle in Auszügen erwähnt werden, um die Vielgestaltigkeit und Bedeutung der Ebene zu skizzieren.

Beispielsweise untersuchen Sprachwissenschaftler:innen, wie und in welchen Situationen Nichtwissen sowie unsicheres Wissen von Wissenschaftler:innen versprachlicht werden und welche positiven oder negativen Auswirkungen (z. B. Anerkennung oder Imageprobleme) sich dabei für die involvierten Wissenschaftler:innen ergeben: Wird von den Forschenden selbst auf Wissensgrenzen, Nichtwissen, Unsicherheiten, Limitationen und damit mitunter implizit auf ein ‚Noch-nicht-wissen‘ verwiesen, trägt dies positiv zur wahrgenommenen Kompetenz der Forschenden bei, da dies einer guten wissenschaftlichen Praxis entspricht (Rhein, 2018). Im wissenschaftlichen Diskurs – z. B. auf Tagungen und Konferenzen – kann die Offenbarung von Nichtwissen und Unsicherheit bezüglich **des Wissens jedoch auch imageschädigend oder gesichtsbedrohend wirken, wenn „Wissen erwartet wird und wenn (nur) Wissen als Ausdruck von Kompetenz und Autorität gilt“** (Wehling, 2012, S. 74–75). Rhein (2018) kann auf der Grundlage von Gesprächsdaten differenzieren, dass es sich dabei vor allem um Situationen handelt, in denen zwei oder mehr Akteur:innen mit konfligierenden Ansichten aufeinandertreffen:

An Diskussionen beteiligte Akteure haben unter Umständen konfligierende Ansichten, was dazu führt, dass Argumentationen, die auf vermeintlich sicherem Wissen beruhen, durch die Deklaration dieses vermeintlich sicheren Wissens als unsicheres Wissen oder Irrtum ausgehebelt werden. Dadurch schreiben sich die Kritisierenden selbst Wissen und Kompetenz, dem Kritisierten im gleichen Atemzug weniger Wissen und unter Umständen (partielle) Inkompetenz zu. (Rhein, 2018, S. 78)

Die Kommunikation von Wissen, Nichtwissen und unsicherem Wissen in die Öffentlichkeit hinein erfolgt häufig nicht durch die Wissenschaftler:innen selbst, sondern durch Journalist:innen. Dabei kommunizieren die Journalist:innen nicht unvermittelt die Informationen, sondern wählen zielgruppenorientiert und kontextspezifisch aus. Brechman, Lee und Cappella (2009) untersuchen im Bereich der Wissenschaftskommunikation beispielsweise die Berichterstattung über neue

Forschungsergebnisse aus der Genetik und können berichten, dass etwas mehr als zwei Drittel der journalistischen Aufbereitungen, die für die Studie herangezogen wurden, deterministisch und stark vereinfacht wurden, wodurch eine hohe Eindeutigkeit auch durch die Reduktion des unsicheren Wissens entsteht. Dabei ist in den Daten von Guenther und Ruhrmann (2016) die Vorstellung von Journalist:innen über die Adressat:innen der Berichterstattung ein wichtiger Prädiktor dafür, wie über Unsicherheiten gesprochen wird. Von den befragten Journalist:innen wird auch angenommen, dass die Adressat:innen mit der Komplexität der wissenschaftlichen Befunde leicht zu überfordern sind.

Gleichzeitig berichten Wissenschaftsjournalist:innen von der Befürchtung, durch die Kommunikation von unsicherem Wissen und Nichtwissen zu dem Eindruck beizutragen, dass die Wissenschaft nichts weiß und eine gewisse Beliebigkeit in den Aussagen liegt (Guenther, 2017). In öffentlichen Diskussionen rund um den Klimawandel sind diese Tendenzen wahrnehmbar. An einem Beispiel beschreiben Kampourakis und McCain (2020) die Gefahr:

[...] there is the danger of thinking that since science is uncertain, anything goes. One might be tempted to think that since scientists are not certain that any particular theory is absolutely correct, all theories are in the same boat. [...] As we have mentioned in earlier chapters, the prevailing view is that climate change is largely the result of human emissions of greenhouse gases such as CO₂. Some have attempted to undermine support for this explanation of rising average global temperature by proposing the increase of solar activity is instead responsible for the temperature increase. If we were to take the fact that scientific claims are never certain to mean that anything goes, we might be inclined to think that increased solar activity is just as good of an explanation of increased temperatures as human greenhouse gas emissions.” (Kampourakis & McCain, 2020, S. 210–211)

Weiterhin wird häufig von einem dramatisch aufgeladenen Diskurs über Nichtwissen und unsicheres Wissen in den Medien berichtet (Kampourakis & McCain, 2020; Varwig, 2020).

Damit einhergehend wird häufig nicht nur ein Anstieg von Unsicherheit bei den Bürger:innen (naturwissenschaftliche Laien) vermutet, sondern auch ein Vertrauensverlust in die Wissenschaft als mögliche Folge der Bewertung der Kommunikation von Nichtwissen und unsicherem Wissen beschrieben (Bösch, Kastenhofer, Rust, Soentgen & Wehling, 2008; Maier, Guenther, Ruhrmann, Berkela & Milde, 2018; Varwig, 2020).

Die Kommunikationspsychologinnen Retzbach und Maier (2015) untersuchten deshalb empirisch Effekte der Kommunikation von unsicherem Wissen am Beispiel von Medienberichten über Nanotechnologie auf das Interesse von Laien an Naturwissenschaft insgesamt sowie neuen Technologien, Überzeugungen über das Wesen der Naturwissenschaften und das Vertrauen in Wissenschaftler:innen. Im Studiendesign wurden die 945 Teilnehmer:innen drei Gruppen zugewiesen: Gruppe 1 erhielt erst drei Medienberichte, die vor allem die Chancen der Nanotechnologie adressierten, und anschließend drei Berichte, die die Risiken hervorhoben. Gruppe 2 bekam die Berichte in umgekehrter Reihenfolge. Gruppe 3 erhielt sechs andersartige Berichte, in denen jeweils Chancen und Risiken differenziert betrachtet und abgewogen wurden. Die Forscherinnen konnten weder Effekte auf die Überzeugungen über das Wesen der Naturwissenschaften noch einen Vertrauensverlust in Wissenschaftler:innen in den Daten nachweisen. Hinsichtlich des Interesses gab es sogar einen leicht positiven Effekt (Retzbach & Maier, 2015).

Soziolog:innen widmen sich jedoch auch der Frage, wie sich wissenschaftliche Disziplinen und Forschungsfelder selbst im Umgang mit und in der Bewertung von Nichtwissen und unsicherem Wissen unterscheiden. Anhand von Interviews und Schlüsseltexten zu den Themen ‚grüne Gentechnik‘ und ‚Mobilfunk‘ verglichen Böschen und Kolleg:innen (2008) beispielsweise die Molekularbiologie, die Ökologie sowie die praktische Umweltmedizin anhand von sieben theoretisch fundierten Dimensionen von Nichtwissenskulturen (in Anlehnung an das Konzept der epistemischen Kulturen; Böschen et al., 2008). Durch den Vergleich konnten drei Typen von Nichtwissenskulturen rekonstruiert werden, die sich unter anderem durch den Umgang mit Nichtwissen unterscheiden: Die Molekularbiologie entspricht eher einem kontrollorientierten Typ, bei dem das Noch-nicht-wissen im Fokus steht, die Ökologie eher einem komplexitätsorientierten Typ, der auch unerkanntes Nichtwissen anerkennt, sowie die praktische Umweltmedizin einem erfahrungsorientierten Typ, der das Nichtwissen personalisiert und als Infragestellung professioneller Autorität wahrnimmt (Böschen et al., 2008, S. 204). Eine zweite Dimension, hinsichtlich derer sich die drei Typen unterscheiden ist die des Umgangs mit Komplexität. Während der kontrollorientierende Typ eher dazu neigt, die Komplexität zu reduzieren und bestrebt ist, unkontrollierte Bedingungen zu vermeiden, erkennt der komplexitätsorientierte Typ die Komplexität an und tendiert – vermutlich daher – auch zu einer theoretischen und methodischen Offenheit. Der erfahrungsorientierte Typ denkt häufig in Einzelfällen und berücksichtigt dadurch die Komplexität im Sinne einer begrenzten Verallgemeinerbarkeit (Böschen et al., 2008).

Dennoch müssen Bürger:innen und politische Entscheidungsträger:innen unter der Bedingung von Nichtwissen und unsicherem Wissen Entscheidungen treffen. Sie müssen entschluss- und handlungsfähig bleiben und trotz fehlendem Wissen und Gewissheit für die eigene Person oder für die Gesellschaft entscheiden.

Den Aspekten des Bewertens und Handelns in komplexen Problemlagen und unter den Bedingungen von unsicherem Wissen und Nichtwissen widmen sich im schulischen Bildungskontext Fachdidaktiker:innen – insbesondere Geografiedidaktiker:innen beschäftigen sich im Kontext von Bewertungs- und Handlungskompetenz und im Themenbereich der Umwelt- und Nachhaltigkeitsbildung mit unsicherem Wissen und Nichtwissen (Mehren et al., 2015; Ohi, 2018). Schüler:innen müssen hier die Wissenslücken in ihren eigenen Recherchen oder in der wissenschaftlichen Evidenzlage erkennen, aushalten und dennoch produktiv damit umgehen, um Entscheidungen treffen zu können und handlungsfähig zu sein bzw. zu werden. Dabei werden unsicheres Wissen und Nichtwissen nicht gesondert betrachtet, sondern sind als eine Facette (über-)komplexer Problemlagen erfasst, die durch „unüberschaubar viele Einflussgrößen, Intransparenz und (wissenschaftliches) Nichtwissen beziehungsweise einen Mangel an wissenschaftlicher Evidenz sowie fachliche Kontroversen“ (Mehren et al., 2015, S. 4) gekennzeichnet sind. Neben diese faktische Komplexität tritt die ethische Komplexität, das heißt „konkurrierende Wertemaßstäbe und unterschiedliche Auffassungen darüber, welche Handlungsalternative „richtig“ ist“ (Mehren et al., 2015, S. 4). Mehren und Kolleg:innen (2015) beschreiben Themen der Umwelt- und Nachhaltigkeitsbildung in Anlehnung an Bögeholz und Barkmann (2005) mit dem Begriff der „doppelten Komplexität“ (Mehren et al., 2015, S. 4).

3.4 Herausforderungen bei der Vermittlung und dem Verständnis komplexer Phänomene

Während in den vorhergehenden Abschnitten zur Begriffsbestimmung von Komplexität und Nichtwissen bzw. unsicherem Wissen insbesondere die wissenschaftstheoretische, philosophische und wissenssoziologische Perspektive in den Fokus gestellt wurde, geht es im Folgenden um Herausforderungen bei der Vermittlung und dem Anbahnen eines Verständnisses komplexer Phänomene, sodass vor allem die didaktische und psychologische Perspektive adressiert wird.

Jacobson (2000) benennt das Unterrichten bzw. das Lernen über Komplexität und komplexe Systeme als große Herausforderung. Dass es sich lohnt bzw. notwendig ist, ein Verständnis darüber anzubahnen, um Schüler:innen eine adäquate Vorstellung der belebten Welt zu vermitteln, wird in Kapitel 3.5 gezeigt. Herausforderungen bei der Thematisierung von komplexen Phänomenen im Unterricht können in zwei Kategorien eingeteilt werden: Herausforderungen bei der Vermittlung auf der einen Seite und solche beim Verständnis auf der anderen Seite. Ein Verständnis komplexer Phänomene betrifft nicht im engeren Sinne Schüler:innen, sondern Noviz:innen (Jacobson & Wilensky, 2006), zu denen auch Studierende oder Lehrende zählen können. Im Folgenden sollen Herausforderungen entlang der zwei Kategorien skizziert werden.

Resnick (1995) beschreibt beispielsweise, dass Noviz:innen im Umgang mit komplexen Systemen Herausforderungen erfahren, da das Denken in komplexen Systemen ein anderes mentales Modell voraussetzt (das sogenannte ‚decentralized mindset‘) als das Denken in einfachen, linearen Systemen (das sogenannte ‚centralized mindset‘).

The sciences of complexity present not only scientific challenges but educational and epistemological ones, raising questions about how people view and make sense of the world. In particular, these new sciences tend to highlight *decentralized* models of the world – in contrast to traditional scientific models based on centralized control and centralized causes. Scientists are reexamining their models of biological systems, social systems, and technological systems. For example, traditional theories of bird flocking were based on the idea of a "leader" bird at the front of the flock. Newer theories view bird flocks as decentralized [...]. Each bird in the flock follows a set of simple rules, reacting to the movements of the birds nearby it. Orderly flock patterns arise from these simple, local interactions. None of the birds has a sense of the overall flock pattern. The bird in front is not a "leader" in any meaningful sense—it just happens to end up there. But even as the influence of decentralized ideas grows in the scientific community, there is a deep-seated resistance to such ideas. Most people seem to have strong attachments to centralized ways of thinking. When people see patterns in the world (like a flock of birds), they generally assume that there is some type of centralized control (a leader of the flock). According to this way of thinking, a pattern can exist only if someone (or something) creates and orchestrates the pattern. Everything must have a single cause, an ultimate controlling factor. (Resnick, 1995, S. 1–2)

Um diese und weitere Hürden beim Verstehen von komplexem Systemen bzw. beim Denken in komplexen Systemen zu eruieren, vergleicht Jacobson (2000) Noviz:innen und Expert:innen im Umgang mit Problemen in komplexen Systemen. Dabei zeigt sich, dass Noviz:innen dazu neigen, Faktoren eines Systems zu reduzieren, und versuchen, eine schrittweise Abfolge zu rekonstruieren. Sie nehmen eher eine einzelne, zentrale Ursache im System an und denken dabei vor allem in linearen Zusammenhängen (kleine Ursache erwirkt einen kleinen Effekt), wohingegen Expert:innen von Multikausalität, dezentralen Systeminteraktionen und auch nicht-linearen Ursache-Wirkungs-Beziehungen (kleine Ursachen können große Effekte erwirken) ausgehen. Ein weiterer Unterschied

zwischen Noviz:innen und Expert:innen im Umgang mit Problemen in komplexen Systemen bezieht sich auf die Vorhersagbarkeit („agent actions“; Jacobson, 2000, S. 17): Während Noviz:innen von vollständiger Vorhersagbarkeit der Effekte ausgehen („completely predictable“; Jacobson, 2000, S. 17), denken Expert:innen auch an Zufall und probabilistische Zusammenhänge, die keine vollständige Vorhersagbarkeit nach sich ziehen.

In Hinblick auf Schwierigkeiten beim Verstehen von komplexen Systemen nehmen Mambrey et al. (2020) die Art der Zusammenhänge, die im System bestehen, in den Blick. In einer empirischen Studie zum Verstehen komplexer Systeme durch Schüler:innen am Beispiel von Nahrungsbeziehungen in Ökosystemen zeigen ihre Daten, dass insbesondere die Art der Zusammenhänge im System und daraus resultierende Effekte (direkte oder indirekte Zusammenhänge)⁴² Einfluss auf das Verstehen bzw. auf das Erreichen von höheren Leveln des systemischen Denkens haben. Dabei sind indirekte Zusammenhänge und Effekte schwieriger zu verstehen als direkte, wobei indirekte Zusammenhänge wesentliche Charakteristika von komplexen Systemen sind.

Our results revealed that understanding direct relations seems to be easier than understanding indirect relations. In the case of indirect effects, however, we did not find significant differences based on the distinction between linear and complex relations. Since the findings are evident and very similar for all dimensions, we assume a significant tendency. We think this highlights that grasping indirect relations is more challenging than grasping direct effects, at least in the case of food webs. Since these indirect effects are decisive for a comprehensive insight into complex ecological systems, we consider the difficulty of understanding these effects to be an obstacle to the development of systems thinking. (Mambrey et al., 2020, S. 1647)

Weiterhin ist beim Umgang mit komplexen Phänomenen – also bei dem Versuch, diese zu verstehen, oder beim Aufstellen von Vermutungen über die Entwicklung – der ‚cognitive load‘⁴³ häufig als wesentlich höher als bei einfachen, linearen Erklärungen einzuschätzen (Grotzer, Solis, Tutwiler & Cuzzolino, 2017).

Mit Fokus auf die Vermittlung komplexer Phänomene stellt sich unter anderem die Frage, wie die zahlreichen gleichzeitig stattfindenden Prozesse und die hohe Dynamik für eine Darstellung in eine ‚erzählbare‘ Reihenfolge gebracht werden können. Die Sprache als Mittel der Informationsweitergabe funktioniert stets linear, sodass Informationen – zum Beispiel Erklärungen zu Auswirkungen einer Veränderung auf ein Systemelement, dessen Auswirkungen auf Veränderungen vieler weiterer Systemelemente und deren direktes oder indirektes Zurückwirken auf Systemelemente – nur nacheinander und Element für Element gegeben werden können, obwohl beim realen Phänomen die Prozesse zeitlich parallel stattfinden.

⁴² Direkte Effekte werden durch einfache Ursache-Wirkungs-Beziehungen beschrieben. Im Falle von Nahrungsbeziehungen kann das beispielsweise eine Räuber-Beute-Beziehung sein, bei der eine hohe Anzahl an Tieren der Räuberpopulation die Beutepopulation dezimiert, woraufhin das sinkende Nahrungsangebot für die Räuberpopulation auch zur Reduktion der Räuberpopulation führt. Indirekte Effekte ergeben sich für weitere Arten im Ökosystem (Mambrey et al., 2020).

⁴³ Der Begriff ‚cognitive load‘ – in der deutschen Übersetzung als kognitive Belastung – stammt aus der ‚cognitive-load-theory‘ (Sweller, 2011) Innerhalb der Theorie wird angenommen, dass die begrenzte Kapazität des Arbeitsgedächtnisses für effektives Lernen berücksichtigt werden muss. Die kognitive Belastung des Arbeitsgedächtnisses ergibt sich aus einer intrinsischen kognitiven Belastung (Verhältnis von Komplexität des Lerninhalts und Vorwissen) und der extrinsischen kognitiven Belastung (Belastung durch die günstige oder ungünstige Gestaltung des Lernmaterials). Ein Teil der gesamten kognitiven Belastung ist erwünscht (der sogenannte ‚germane cognitive load‘), da er für höherstufige kognitive Prozesse steht und Lernen ermöglicht.

Eine weitere Herausforderung bei der Vermittlung ist die der Visualisierung komplexer Phänomene für den Vermittlungszweck. Ablaufdiagramme und Mindmaps oder Concept-Maps können Systemelemente und ihre Zusammenhänge strukturell abbilden (Mambrey et al., 2020). Je nach Komplexitätsgrad des Phänomens – vor allem in Abhängigkeit von der Anzahl der Elemente und Wechselwirkungen – können Visualisierungen schnell unübersichtlich wirken. Um das Verhalten komplexer Systeme anschaulich und erfahrbar zu machen, argumentieren Jacobson und Wilensky (2006) für den Einsatz von Computerprogrammen. Mit entsprechenden Simulationen kann die Veränderbarkeit von komplexen Systemen zugänglich gemacht werden. Dabei liegt der Fokus jedoch auf dem Erfahren der Veränderbarkeit des Systems und nicht auf der Erklärung von Zusammenhängen.

Nichtwissen und unsicheres Wissen zu thematisieren, geht auch mit einer Reduktion von Eindeutigkeit und Erhöhung von Vagheit einher, da Leerstellen und damit Freiheitsgrade entstehen, die Lehrende und Lernende aushalten müssen.

Aufseiten von Vermittelnden – in diesem Fall Lehrkräften – konnten Krüger et al. (2013) das Streben nach Eindeutigkeit in empirischen Daten finden. Dabei wurden Gymnasiallehrkräften in Gruppeninterviews als Stimuli Videovignetten gezeigt, in denen Lehrkräfte mit unsicherer Evidenz bei Schülerexperimenten im Chemie- bzw. Physikunterricht oder bei der Verwendung von wissenschaftlichen Quellen im Chemie- bzw. Physikunterricht konfrontiert sind. „Insgesamt wurde bei fast allen Teilnehmer/innen ein hohes Bedürfnis nach Eindeutigkeit im Unterricht deutlich. Für sie ist es entscheidend, dass am Ende das „eine richtige“ Ergebnis festgehalten werden kann“ (Krüger et al., 2013, S. 729). Entsprechend nahm die überwiegende Anzahl der Lehrkräfte die gezeigten Unterrichtssituationen als störend wahr.

Sowohl bei Lehrenden als auch bei Lernenden kann das Bedürfnis nach Abgeschlossenheit von Wissen („cognitive closure“; Webster & Kruglanski, 1994, S. 1049) eine Rolle bei der Verarbeitung von komplexen Phänomenen und deren unsicherem Wissen sowie Nichtwissen spielen. Das Persönlichkeitsmerkmal „need for cognitive closure“ (Webster & Kruglanski, 1994, S. 1049) wird als latentes Konstrukt durch verschiedene manifeste Variablen repräsentiert: die Präferenz von Ordnung und Struktur, den Wunsch nach Vorhersagbarkeit, Entschlussfreudigkeit, Unbehagen mit Ambiguität und Engstirnigkeit.⁴⁴ Bei komplexen Phänomenen überwiegen jedoch Antworten, die diesen Bedürfnissen weniger nachkommen.

In Kapitel 3.3.3, das die Auswirkungen von Nichtwissen und unsicherem Wissen skizziert hat, wurde bereits auf einen befürchteten Vertrauensverlust in die Wissenschaft durch die Thematisierung von Nichtwissen und unsicherem Wissen hingewiesen (Böschen et al., 2008; Maier et al., 2018; Varwig, 2020).⁴⁵ Die wahrgenommene Gefahr kann mitunter auch einen Effekt auf die Thematisierung im Unterricht haben.

In Bezug auf die Thematisierung von Nichtwissen und unsicherem Wissen könnten für Lehrkräfte auch die Rahmenbedingungen von Schule als Herausforderung gesehen werden: Neben instruktionalen Erklärungen ist auch schulischer Unterricht allgemein auf die Vermittlung von Wissen ausgerichtet, das anschließend in Form von Leistungsnachweisen abgefragt wird. Dabei hat das Schulfach Biologie sogar das Image eines Auswendiglernfachs (Dittmer, 2010). Die Orientierung an

⁴⁴ Im Original von Webster und Kruglanski (1994, S. 1053): „preference for order“, „preference for predictability“, „decisiveness“, „discomfort with ambiguity“, „closed mindedness“.

⁴⁵ Dieser Aspekt wird in der didaktischen und psychologischen Literatur nicht aufgegriffen.

Fakten und deren Abprüfen in Wissenstest gibt dabei Anlass zur Frage, wie Nichtwissen und unsicheres Wissen damit vereinbar sind.

Hinzu kommt, dass der Wissenskanon der Schule von Lehrkräften eher als gesichertes und weniger als veränderbares Wissen eingeschätzt wird (Krüger et al., 2013), sodass die Thematisierung von unsicherem Wissen und Nichtwissen nicht naheliegt.⁴⁶

3.5 Hinweise zum Stellenwert sowie zur Bedeutung eines Verständnisses von Komplexität und Nichtwissen im Biologieunterricht

In den Kapiteln 3.1 bis 3.3 wurde gezeigt, dass Komplexität und damit einhergehend auch Nichtwissen bzw. unsicheres Wissen Eigenschaften biologischer Gegenstände und Phänomene sind. Die Komplexität biologischer Phänomene erfährt von Wissenschaftler:innen verschiedener Disziplinen vermehrte Aufmerksamkeit und kann als überspannendes Prinzip beschrieben werden (siehe Kapitel 3). Sie bedingt veränderte Erkenntnismethoden und Modelle, die die Phänomene der Welt beschreiben können, bzw. verlangt nach diesen (Mitchell, 2008). Nichtwissen und unsicheres Wissen erlangten insbesondere durch direkte Erfahrungen mit der Unterschätzung des Nichtwissens bzw. invers durch die Überschätzung des bestehenden Wissens, das weitreichende negative Folgen für die Gesellschaft oder Umwelt hatte (z. B. FCKW), eine empirische Evidenz und damit einhergehend auch eine politische Bedeutsamkeit (Böschen et al., 2008). Das Nichtwissen wird von Forschenden häufig mitgedacht, wenn auch eher im Sinne eines ‚noch nicht wissen‘.

Von dieser Grundlage aus ist die Vermittlung eines grundlegenden Verständnisses für Komplexität und Nichtwissen bzw. unsicherem Wissen in der schulischen Bildung ein fachlicher und wissenschaftspropädeutischer Inhalt. Auf den Stellenwert sowie die Bedeutung eines Verständnisses soll im Folgenden eingegangen werden.

Hinweise auf den Stellenwert von Komplexität und Nichtwissen bzw. unsicherem Wissen können in den bundesweiten Bildungsstandards als Richtlinien für die im Unterricht zu vermittelnden Kompetenzen bzw. die zu erreichenden allgemeinen Bildungsziele sein, die Schüler:innen eine „verantwortungsvolle gesellschaftliche Teilhabe“ (KMK, 2020, S. 9) ermöglichen sollen. Die Bildungsstandards für die allgemeine Hochschulreife (KMK, 2020) beziehen sich in den Ausführungen der vier Kompetenzbereiche (Sachkompetenz, Erkenntnisgewinnungskompetenz, Kommunikationskompetenz und Bewertungskompetenz) an einzelnen Stellen auf Komplexität und unsicheres Wissen. Auf Letzteres wird dabei unter Bezug auf die Limitationen der naturwissenschaftlichen Methoden und die Vorläufigkeit naturwissenschaftlichen Wissens verwiesen (also im Sinne von ‚noch nicht wissen‘). Das Nichtwissen und der Umgang damit werden nicht explizit aufgegriffen. Verwiesen wird jedoch an einigen Stellen auf eine sachgerechte Darstellung, wodurch sowohl die Komplexität der biologischen Phänomene als auch das Spektrum zwischen Wissen und Nichtwissen beinhaltet sind, da sie Charakteristika der biologischen Gegenstände sind. Die folgenden Auszüge aus den Bildungsstandards sollen dies exemplarisch verdeutlichen.

⁴⁶ Ravetz (1987) schreibt dazu: „Almost all the facts learned as students were uncontested and incontestable; only during research did we discover that scientific results can vary in quality; later we may have come across scientific problems that could not be solved; and only through participation in the grooving of science do we learn of choices and criteria“ (Ravetz, 1987, S. 91).

In den Bildungsstandards heißt es zur Definition der Sachkompetenz allgemein:

Lernende erhalten die Möglichkeit, im Bereich der Sachkompetenz fundiertes Wissen über biologische Sachverhalte wie beispielsweise Phänomene, Konzepte, Theorien und Verfahren zu erwerben und Kompetenzen im Sinne einer vertieften Allgemeinbildung aufzubauen. Diese Kompetenzen ermöglichen es ihnen, u. a. theoriegeleitet Fragen zu stellen sowie anspruchsvolle Problemstellungen im Zusammenhang mit biologischen Sachverhalten zu bewältigen bzw. Alltagsfragen zu naturwissenschaftlichen Sachverhalten zu beantworten. (KMK, 2020, S. 13)

Das hier enthaltene Verständnis von Sachkompetenz bezieht sich auf das Wissen (also das bereits wissenschaftlich geschaffene Wissen), das aufgebaut werden soll. Das Stellen theoriegeleiteter Fragen kann als ein Verweis auf Nichtwissen und unsicheres Wissen, das Fragen offenlässt, interpretiert werden, wobei dies im Schulkontext vorrangig auf das individuelle Nichtwissen bezogen ist (Osborne, Ratcliffe, Collins, Millar & Duschl, 2001). Das Stichwort der Komplexität als Eigenschaft biologischer Phänomene wird nicht explizit genannt, ist aber in einer Aussage zur Vernetztheit von Systemebenen implizit erhalten:

Dabei werden Zusammenhänge strukturiert sowie qualitativ und quantitativ erläutert sowie Vernetzungen zwischen Systemebenen von der molekularen Ebene bis zur Ebene der Biosphäre aufgezeigt. Jede der Systemebenen beinhaltet häufig Eigenschaften, die in der vorherigen Ebene nicht erkennbar sind. (KMK, 2020, S. 14)

In den Formulierungen der Standards heißt es schließlich, dass Lernende „biologische Sachverhalte sowie Anwendungen der Biologie sachgerecht [beschreiben]“ (KMK, 2020, S. 14), worunter auch Komplexität, Nichtwissen und unsicheres Wissen gefasst werden könnten, da es sich dabei um Eigenschaften der Sachverhalte handelt.

Die Kommunikationskompetenz wird von der KMK wie folgt definiert:

Die Kommunikationskompetenz der Lernenden zeigt sich in der Kenntnis von Fachsprache, fachtypischen Darstellungen und Argumentationsstrukturen und in der Fähigkeit, diese zu nutzen, um fachbezogene Informationen zu erschließen, adressaten- und situationsgerecht darzustellen und auszutauschen. Biologisch kompetent Kommunizieren bedingt ein Durchdringen der Teilkompetenzbereiche Erschließen, Aufbereiten und Austauschen. (KMK, 2020, S. 16)

Auch hier wird die Kommunikation über Nichtwissen oder auch unsicheres Wissen nicht explizit angesprochen, sondern der Fokus liegt auf fachbezogenen Informationen, womit wiederum eher positives Wissen fokussiert wird.

Die Bewertungskompetenz wird von der KMK allgemein wie folgt definiert:

Die Bewertungskompetenz der Lernenden zeigt sich in der Kenntnis von fachlichen und überfachlichen Perspektiven und Bewertungsverfahren und in der Fähigkeit, diese zu nutzen, um Aussagen bzw. Daten anhand verschiedener Kriterien zu beurteilen, sich dazu begründet Meinungen zu bilden, Entscheidungen auch auf ethischer Grundlage zu treffen und Entscheidungsprozesse und deren Folgen zu reflektieren. (KMK, 2020, S. 17)

Um Aussagen einordnen zu können und daraufhin auch unter unsicheren Bedingungen eine Bewertung abgeben zu können, ist Wissen über verschiedene Wissensstadien hilfreich (siehe unten). Wissensstadien wie Nichtwissen oder unsicheres Wissen werden jedoch im Kompetenzbereich der Bewertung nicht explizit angesprochen.

Zur Erkenntnisgewinnungskompetenz als Ziel des Biologieunterrichts heißt es mit einem Fokus auf die Kenntnis und Beherrschung der Methoden allgemein:

Die Erkenntnisgewinnungskompetenz der Lernenden zeigt sich in der Kenntnis von naturwissenschaftlichen Denk- und Arbeitsweisen und in der Fähigkeit, diese zu beschreiben, zu erklären und zu verknüpfen, um Erkenntnisprozesse nachvollziehen oder gestalten zu können und deren Möglichkeiten und Grenzen zu reflektieren. Sie zeichnen sich insbesondere dadurch aus, dass sie der hohen Komplexität biologischer Systeme Rechnung tragen sowie dem Umstand, dass es sich um lebende Systeme handelt. Dies wirft neben wissenschaftspropädeutischen auch ethischen Fragen auf. (KMK, 2020; S.14)

Dabei wird die doppelte Komplexität (Mehren et al., 2015) biologischer Phänomene bzw. Systeme adressiert.

Die zu reflektierenden Möglichkeiten und Grenzen der naturwissenschaftlichen Erkenntnisgewinnung werden in den Standards präzisiert: „Die Lernenden beurteilen die Gültigkeit von Daten und ermitteln mögliche Fehlerquellen“ (KMK, 2020, S. 15) und „diskutieren Möglichkeiten und Grenzen von Modellen“ (KMK, 2020, S. 16). Zu den Merkmalen wissenschaftlicher Aussagen heißt es weiter: Lernende „reflektieren Möglichkeiten und Grenzen des konkreten Erkenntnisgewinnungsprozesses sowie der gewonnenen Erkenntnisse (z. B. Reproduzierbarkeit, Falsifizierbarkeit, Intersubjektivität, logische Konsistenz, Vorläufigkeit)“ (KMK, 2020, S. 16). Die angesprochenen Punkte können unter der kritischen Betrachtung der wissenschaftlichen Erkenntnismethodik zusammengefasst werden. Dabei wird der Fokus auch hier auf bereits bestehendes Wissen, das mitunter unsicher ist, und die Möglichkeit, dass durch neues Wissen – durch neue empirische Belege – das bisherige, vermeintliche Wissen verändert werden muss, gelegt.

Mit der Vorläufigkeit von Wissen nehmen die Bildungsstandards der KMK den Diskurs der ‚Nature of Science‘ auf. In dem Diskurs, der sich der Frage nach den Charakteristika der Naturwissenschaft als wissenschaftliche Disziplin und danach, welche dieser Charakteristika in der Schule vermittelt werden sollen, widmet, wird auch Bezug auf die Beschaffenheit des naturwissenschaftlichen Wissens genommen. McComas und Olson (1998) sowie Lederman (2007) betonen – wie in den Bildungsstandards der KMK explizit erwähnt – auch die Vorläufigkeit und Veränderbarkeit des wissenschaftlichen Wissens. Osborne et al. (2001) richten bei der Facette der Vorläufigkeit den Fokus auf die Sicherheit als Status des Wissens und problematisieren dabei indirekt die Auswahl an Phänomenen im Biologieunterricht bzw. das damit vermittelte Verständnis von wissenschaftlichem Wissen und dessen Status:

Pupils should appreciate why much scientific knowledge, particularly that taught in school science, is well-established and beyond reasonable doubt, and why other scientific knowledge is more open to legitimate doubt. It should be explained that current scientific knowledge is the best we have but may be subject to change in the future, given new evidence or new interpretations of old evidence. (Osborne et al., 2001, S. 59–60)

Zusammenfassend werden in den Bildungsstandards der KMK die Komplexität biologischer Gegenstände und unsicheres Wissen, das durch die Limitationen der verfügbaren Erkenntnismethoden entsteht, adressiert. Erkanntes oder unerkanntes Nichtwissen kann an einigen Stellen mitgedacht werden, wird jedoch nicht direkt erwähnt. Dabei würden Nichtwissen und unsicheres Wissen als Teil komplexer biologischer Phänomene eine Schnittstelle zwischen Sachkompetenz und Erkenntnisgewinnungskompetenz bilden, denn das Nichtwissen oder

unsicheres Wissen kommen zum Teil durch die Methoden der naturwissenschaftlichen Erkenntnisgewinnung zustande, sind aber in jedem Fall Teil des fachlichen Gegenstandes bzw. Teil des Wissens um den fachlichen Gegenstand.

Wie Komplexität biologischer Gegenstände und insbesondere die Facette des Nichtwissens in unterrichtlichen Erklärungen aufgegriffen bzw. allgemein im Unterricht durch Lehrkräfte realisiert werden, ist nicht durch empirische Studien belegt.⁴⁷ Kattmann (1995) bescheinigt dem Biologieunterricht zwar, eine stark reduzierte Sichtweise auf die Phänomene zu vermitteln, bezieht dies jedoch auf die Ausklammerung der Evolution als historische Perspektive, die allen Gegenständen zugrunde liegt. Potochnik (2013) konstatiert für den naturwissenschaftlichen Unterricht in den USA, dass lineare Erklärungen, das heißt jene, die nicht auf Komplexität eingehen und eine Abgeschlossenheit suggerieren sowie folglich nicht auf Nichtwissen eingehen, überwiegen.⁴⁸ Jacobson und Wilensky (2006) stellen ebenfalls fest, dass trotz steigender Beachtung von Komplexität bzw. der Komplexitätstheorie in den Wissenschaften und der sich ändernden Sicht auf die Welt der Aspekt ‚Komplexität‘ kaum Eingang in die Curricula findet.⁴⁹ Auch Braaten und Windschitl (2011) formulieren die Gefahr, dass kausale Modelle im Unterricht hauptsächlich zu eher linearen Erklärungen führen und nicht zu Wirkungsgefügen ausgebaut werden. Bei der Durchschau von Schulbüchern für den Biologieunterricht entsteht der Eindruck, dass dieses Bild auch für Deutschland gelten kann – eine systematische Überprüfung steht aus.

Dabei können der Grad an Komplexität und die Darstellung von Nichtwissen bzw. unsicherem Wissen im Unterricht allgemein und in Erklärungen im Speziellen zwar im Rahmen der didaktischen Reduktion zugunsten einer Prinzipienorientierung reduziert werden. Die Entscheidung, nicht nur den Grad der Komplexität zu reduzieren, sondern auch – bekannte – Wissenslücken (sektoral) zureduzieren, sollte jedoch bewusst erfolgen, da sie ein Merkmal des wissenschaftlichen Gegenstandes sind, denn die Darstellung von Komplexität und Nichtwissen bzw. unsicherem Wissen könnte neben der fachlichen Korrektheit weitere erstrebenswerte Bildungsziele potenziell fördern. So scheint ein grundlegendes Verständnis der Konzepte Komplexität, Nichtwissen und unsicheres Wissen für ein Verständnis und eine adäquate Deutung aktueller Problemlagen der Gesellschaft und Diskussionen darüber sinnvoll (Grotzer, 2012; Jacobson & Wilensky, 2006).⁵⁰

Geht man davon aus, dass Nichtwissen und unsicheres Wissen im Unterricht nur selten thematisiert werden und es eine Tendenz hin zu Eindeutigkeit von Erklärungen gibt (Grotzer, 2012; Jacobson & Wilensky, 2006; Potochnik, 2013), ist zu fragen, welche Vorstellungen Lernende durch die Ausklammerung von Nichtwissen und die Fokussierung auf Eindeutigkeit und Abgeschlossenheit von Wissenschaft bzw. wissenschaftlichem Wissen erhalten. An dieser Stelle wird die wissenschaftspropädeutische Bedeutung besonders deutlich. Daran anknüpfend ist fraglich, welche Auswirkungen diese Vorstellungen bei der Konfrontation mit unsicherem Wissen in akuten

⁴⁷ Explizit wird die Komplexität im Biologieunterricht hauptsächlich im Bereich der Ökologie gemacht. Empirische Untersuchungen nehmen vermutlich auch deshalb insbesondere im Bereich des Ökologieunterrichts den Umgang der Schüler:innen mit komplexen Nahrungsnetzen in den Blick (Mambrey, Timm, Landskron & Schmiemann, 2020).

⁴⁸ Das Spannungsverhältnis zwischen dem Aufzeigen von Komplexität und Nichtwissen in Erklärungen und der Reduzierung auf lineare Ursache-Wirkungs-Beziehungen, die auch durch fehlende Verweise auf Nichtwissen und unsicheres Wissen Eindeutigkeit vermitteln, wird in den Videovignetten aufgegriffen (siehe Kap. 5.2.2.2 bzw. 5.2.2.5)

⁴⁹ Eine mögliche Ursache für eine solche Ausklammerung der Aspekte ‚Komplexität‘ und ‚Nichtwissen‘ aus dem Biologieunterricht kann die schulische Prüfungskultur bzw. die Outputorientierung darstellen (Dittmer, 2010; Grotzer, 2017).

⁵⁰ Jacobson und Wilensky (2006) greifen ein Verständnis von Komplexität als Teil der Scientific Literacy auf.

Problemlagen – wie zum Beispiel der Covid-19-Pandemie, die davon in hohem Maße geprägt war – haben.⁵¹

Eine explizite Thematisierung von Nichtwissen kann auch als Teil der Bewertungskompetenz der Schüler:innen bedeutsam sein: Wenn Schüler:innen für das Phänomen Nichtwissen sensibilisiert werden und lernen, dass es unterschiedliche Wahrnehmungen und Bedeutungszuschreibungen gibt, die Akteur:innen in Diskussionen einbringen, kann das Orientierung in der Analyse gesellschafts-politischer Diskussionen beispielsweise über den Klimawandel oder die Covid-19-Pandemie geben. Kontroversen oder Konflikte über Nichtwissen, dessen Risikopotenziale und Schuldzuschreibungen (Intention des Nichtwissens) können so auf einer fachlichen Ebene betrachtet und besser verstanden werden. Böschen und Kolleg:innen (2008) machen unter Rückgriff auf die von Wehling (2009) differenzierten Dimensionen von Nichtwissen die Komplexität einer Diskussion, in der auch auf Nichtwissen als Faktor zur Entscheidungsfindung hingewiesen wird, deutlich:

Wo die eine Seite von begrenzten und eindeutig identifizierbaren Wissenslücken spricht, verweist die andere auf die Gefahren des unerkannten Nichtwissens, wenn man also noch nicht einmal weiß, was man nicht weiß; und während die einen von bloß temporären Informationsdefiziten ausgehen, heben andere die Möglichkeit eines anhaltenden, wenn nicht unüberwindlichen <<Nicht-Wissen-Könnens>> hervor. (Böschen et al., 2008, S. 198)⁵²

Erstrebenswert wäre eine Thematisierung von Komplexität im Biologieunterricht, die den Schüler:innen ein grundlegendes Verständnis für die Vernetzung von Systemelementen vermittelt. Dabei muss es nicht in erster Linie um ein Durchdringen der komplexen Phänomene gehen, sodass Schüler:innen – wie in der Beschreibung der Sachkompetenz der KMK angeführt – fähig sind, „anspruchsvolle Problemstellungen im Zusammenhang mit biologischen Sachverhalten zu bewältigen bzw. Alltagsfragen zu naturwissenschaftlichen Sachverhalten zu beantworten“ (KMK, 2020, S.13). Jacobson und Wilensky (2006) formulieren dazu, dass für ein Verständnis der Aspekte von Komplexität (z. B. keine zentrale Ursache, sondern dezentrale Organisation) ein *conceptual change* notwendig wird und einhergehend damit auch Überzeugungen und Intuitionen über die Zusammenhänge in der Welt notwendig werden:

Students learning to think in terms of complex systems ideas may need to go through a process of “strong” or “radical” conceptual change. Consequently, pedagogies and curricula for learning complex systems ideas will need to focus not only on the conceptual aspects of these ideas but also on enriching the cognitive network of beliefs and intuitions students have about the world

⁵¹ Wie in Kapitel 3.3.3 bereits erwähnt, führen zum Beispiel Böschen und Kolleg:innen (2008) die Möglichkeit eines Vertrauensverlustes in wissenschaftliches Wissen und auf Grundlage dieses Wissens gefällten politischen Entscheidungen durch eine nicht vorhandene Kommunikation von Nichtwissen in der Vergangenheit an. Während der Covid-19-Pandemie ließ sich ebenfalls eine Irritation gegenüber divergierenden Aussagen von Wissenschaftler:innen und sich mit wachsendem Wissensstand ändernden Einschätzungen zu notwendigen Schutzmaßnahmen beobachten. Habermas äußerte sich in einem Interview mit dem Journalisten Markus Schwering zu Beginn der Pandemie zu den Herausforderungen für Individuen und die Gesellschaften in Bezug auf den Wissensstand über das Virus wie folgt: „Eines kann man sagen: So viel Wissen über unser Nichtwissen und über den Zwang, unter Unsicherheit handeln und leben zu müssen, gab es noch nie“ (Habermas, zitiert nach Schwering, 2020). Die unterschiedlichen Einschätzungen der Virolog:innen, die unter anderem auch als Grundlage für die politischen Entscheidungen für Schulschließungen bzw. Schulöffnungen, Hygiene- und Abstandsregeln sowie Reisewarnungen oder -verbote dienten, stießen dabei teilweise in der Bevölkerung auf Unverständnis. Ein Vertrauensverlust in Wissenschaft und Wissenschaftler:innen wurde insbesondere in Demonstrationen der sogenannten ‚Querdenker:innen‘ deutlich zum Ausdruck gebracht.

⁵² Recht aktuelle Beispiele lassen sich auch in den Diskussionen um den SARS-CoV-2 Virus finden. So wurden insbesondere die Langzeitfolgen einer Infektion von verschiedenen Gruppen unterschiedlich beurteilt: Während sie zunächst in der breiten Bevölkerung als unerkanntes Nichtwissen nicht diskutiert wurden – viel entscheidender war zunächst die Frage nach Risikofaktoren für eine schwere Erkrankung, die zu langen Krankenhausaufenthalten und mitunter dem Tod endet –, wurde das Thema mit zunehmender Dauer der Pandemie beforscht und die Ergebnisse wurden durch Medien der Bevölkerung zugänglich gemacht.

and about knowledge so as to bridge to a complex systems perspective. (Jacobson & Wilensky, 2006, S. 15)

Weiterhin ist eine Benennung von Wissenslücken – also die Thematisierung von Nichtwissen und unsicherem Wissen als Teil der biologischen Phänomene – bedeutungsvoll, um Schüler:innen eine undramatische Kommunikation darüber zu vermitteln und damit einhergehend Nichtwissen und unsicheres Wissen als normale Zustände der Wissenschaft darzustellen und so, um einen Vertrauensverlust in Wissenschaft zu vermeiden sowie eine Befähigung zur politischen Teilhabe und Mitgestaltung der Gesellschaft zu ermöglichen.

3.6 Zusammenfassung und Fazit: instruktionale Erklärungen komplexer biologischer Phänomene

In Kapitel 3.1 wurde unter Bezugnahme auf wissenschaftstheoretische Überlegungen gezeigt, dass komplexe Phänomene bzw. komplexe Systeme charakteristische Eigenschaften aufweisen: Neben einer meist hohen Anzahl an Elementen, die über zahlreiche lineare und nicht-lineare Wechselwirkungen vernetzt sind und eine hohe Dynamik aufweisen, ist auch die Entstehung neuer Systemzustände auf einer höheren Systemebene, die nicht durch die einzelnen Elemente und deren Wechselwirkungen auf der darunterliegenden Systemebene erklärt werden können – sogenannte Emergenzen –, charakteristisch.

Die wissenschaftlichen Erklärungen solcher komplexen Phänomene sind deshalb multikausal und ohne zentrale Ursache, beziehen sich häufig auf verschiedene Systemebenen, sind probabilistisch sowie mit Nichtwissen und unsicherem Wissen in besonderer Weise assoziiert (siehe Kap. 3.2).

Die wissenssoziologischen und wissenschaftstheoretischen Überlegungen und Untersuchungen zu Nichtwissen und unsicherem Wissen können in betroffene Subjekte (einzelne Personen oder die Gesamtwissenschaft), die verschiedene Wissensstadien bezüglich eines Phänomens oder eines Sachverhaltes aufweisen (wahres Wissen, unsicheres Wissen, Nichtwissen), strukturiert werden. Daraus resultieren vielgestaltige Auswirkungen auf die Kommunikation und die Bewertung von Nichtwissen und unsicherem Wissen sowie folglich die Handlungen unter diesen Bedingungen (siehe Kap. 3.3).

Für die Biologiedidaktik bzw. in Hinblick auf instruktionale Erklärungen im Biologieunterricht im Speziellen lassen sich aus den Überlegungen zur Komplexität sowie zu Nichtwissen und unsicherem Wissen folgende Punkte resümieren und ableiten:

Es gibt eine hohe Anzahl komplexer Phänomene im Gegenstandsbereich der Biologie, was nicht bedeutet, dass bei allen Unterrichtsgegenständen die Komplexität in vollem Umfang aufgezeigt werden muss. Insbesondere instruktionale Erklärungen – also relativ kurz gehaltene, monologische Erkärfunktionen der Lehrkraft – können kein volles Verständnis von Komplexität bzw. kein umfängliches Darlegen der Komplexität des Gegenstandes möglich machen. Vielmehr können dabei einzelne Aspekte der Konzepte Komplexität sowie Nichtwissen und unsicheres Wissen gezielt herausgegriffen werden und dabei kann ein grundlegendes Verständnis dafür angebahnt werden. Die wissenschaftstheoretischen Ausführungen zur Komplexität können dabei eine neue Perspektive für Überlegungen im Zuge der didaktischen Reduktion in der Unterrichtsvorbereitung sein, inwieweit Charakteristika der Komplexität reduziert oder aufgenommen werden sollten und als

überspannendes Prinzip bei der Unterrichtsplanung respektive der Auswahl von Unterrichtsgegenständen Berücksichtigung finden.

Auch der Aspekt des Nichtwissens und des unsicheren Wissens in Form des gesamtwissenschaftlichen Nichtwissens liefert für die didaktische Reduktion einen neuen Reflexionsanlass: An welchen Stellen wird der fachliche Gegenstand nicht adäquat abgebildet, wenn Nichtwissen und unsicheres Wissen keine Erwähnung finden, weil die dabei entstehende Eindeutigkeit keine angemessene Vorstellung über den Gegenstand repräsentiert? An welchen Stellen ergeben sich Lernpotenziale, wenn Nichtwissen und unsicheres Wissen des fachlichen Lerngegenstands als Unterrichtsgegenstand Eingang finden?

Der Erklärungs-begriff (siehe Kap. 2.4) kommt dabei strenggenommen an seine Grenzen, denn ein Zurückführen auf Bekanntes, ein Vermitteln von Wissen oder Verstehen von Zusammenhängen ist an den Stellen, die von Nichtwissen und auch unsicherem Wissen geprägt sind, nicht mehr möglich. Diejenigen Phänomene bzw. Erklärgegenstände, die gesamtwissenschaftliches Nichtwissen und unsicheres Wissen – auch in hohem Maße – enthalten, sind jedoch nicht weniger erklärungsbedürftig.

Die Aspekte Komplexität und Nichtwissen sowie unsicheres Wissen konnten in den Kapiteln 3.1 bis 3.3 als manifester Teil der Wissenschaft Biologie belegt werden, sodass sie Teil des Biologieunterrichts sein sollten, um ein angemessenes Verständnis des Fachs bei den Schüler:innen zu erreichen und sie bei der Orientierung in Diskussionen sowie politische Teilhabe rund um aktuelle, komplexe Problemlagen wie beispielsweise den Klimawandel zu unterstützen. Dabei kann auch eine undramatische Kommunikation über Wissenslücken modelliert werden. In den Kapiteln 3.4 und 3.5 wird jedoch deutlich, dass die Vermittlung mit Herausforderungen verbunden ist und wahrscheinlich noch keinen Eingang in den Biologieunterricht gefunden hat.

Die Herausforderungen bei der Vermittlung von komplexen Phänomenen betreffen Noviz:innen und somit Lehrkräfte wie Schüler:innen: Die Tendenz zum linearen Denken („centralized mindset“; Resnick, 1995), zur Reduktion von Faktoren im System und zur Rekonstruktion einer schrittweisen Abfolge behindert mitunter ein Verständnis der komplexen Systeme. Die Tendenz zur Eindeutigkeit und Abgeschlossenheit („need for cognitive closure“; Webster & Kruglanski, 1994) steht der Vermittlung und dem Verständnis der Phänomene entgegen. Die Erwartungen an Wissenschaft – die im wörtlichen Sinne Wissen und nicht Nichtwissen schaffen – sowie die Schule als wissensvermittelnde Institution können ebenfalls als Faktoren für Schwierigkeiten bei der Vermittlung und dem Verständnis in Betracht gezogen werden (siehe Kap. 3.4).

In den Bildungsstandards der KMK (2020) finden sich jedoch nur zwei Verweise auf die Komplexität (Sachkompetenz und Erkenntnisgewinnungskompetenz), die nur im Falle der Sachkompetenz etwas weiter ausgeführt wird, indem auf die Systemebenen und deren Vernetzung sowie Emergenzen (ohne diesen Begriff zu nennen) verwiesen wird. Auf Nichtwissen und unsicheres Wissen wird nicht explizit eingegangen, jedoch wird bei der Erkenntnisgewinnungskompetenz die Vorläufigkeit von wissenschaftlichem Wissen angesprochen. Zwischen der Bedeutung eines Verständnisses und des anhand der Bildungsstandards rekonstruierten Stellenwerts besteht somit eine Diskrepanz (siehe Kap. 3.5).

Für die letzten beiden Kapitel lässt sich daher als Fazit festhalten, dass es für instruktionale Erklärungen (oder allgemein Unterrichtseinheiten), die die Komplexität der Gegenstände aufgreifen und Nichtwissen sowie unsicheres Wissen offen thematisieren, zunächst Lerngelegenheiten und

Anlässe braucht – sowohl für Lehrkräfte unter anderem in Form von empirisch fundierten Unterrichtsmaterialien, Fortbildungen und Bezügen in Dokumenten der Bildungsadministration als auch für Schüler:innen im Biologieunterricht, die die speziellen Verstehensschwierigkeiten und die bekannten, fachlich problematischen Schülervorstellungen berücksichtigen.

4 Kriterien guten Erklärens

Um Kriterien guten Erklärens zu identifizieren, gibt es verschiedene Ansätze. Zunächst können auf Grundlage bisheriger Erkenntnisse zur Gestaltung von Lernprozessen theoretische Grundüberlegungen angestellt werden. Dabei werden allgemeine Erkenntnisse zum Wissenserwerb und Verstehensprozessen auf die Sprechhandlung des Erklärens übertragen und Empfehlungen für die Gestaltung von instruktionalen Erklärungen ausgesprochen (z. B. Hargie, 2013; Kiel, 1999; Pauli, 2015; Wagner & Wörn, 2011; Wellenreuther, 2013; Wittwer & Renkl, 2008). Des Weiteren können Expert:innen auf dem Gebiet des Lehrens bzw. des Erklärens zu entsprechenden Kriterien befragt werden. Damit werden auf der Grundlage des professionellen Wissens der Teilnehmer:innen die Erfahrungen in der Konstruktion von lernförderlichen Erklärungen empirisch abgebildet (z. B. Schopf & Zwischenbrugger, 2015). Zudem können Erklärungen, die von Lehrkräften oder Studierenden gegeben wurden, analysiert und auf Grundlage der empirischen Daten Kriterien identifiziert werden (z. B. Geelan, 2013; Leinhardt, 1987; Sevian & Gonsalves, 2008; Treagust & Harrison, 2000; Wragg, 1993).

Mit spezifischem Fokus auf den Biologieunterricht sind bislang kaum Studien zu Kriterien für instruktionale Erklärungen entstanden.⁵³ Wragg (1993) vergleicht beispielsweise die Erklärung von zwei vorgegebenen Erkläregegenständen aus dem Englisch- und Biologieunterricht von insgesamt 32 Lehrkräften.⁵⁴ Aufgrund eines angeschlossenen Verstehenstests für Schüler:innen können anschließend die Erklärungen untersucht werden, die bei ihnen hohe Lernerfolge erzielten. Die abgeleiteten Kriterien können jedoch als wenig spezifisch für die jeweiligen Disziplinen angesehen werden, da beispielsweise auf der Ebene der Erkläregegenstände keine Analysen hinsichtlich deren Eigenschaften vorgenommen werden. Mehrere Studien beschäftigen sich hingegen mit dem Erklären im Mathematikunterricht (Leinhardt, 1987; Wagner & Wörn, 2011) oder Physikunterricht⁵⁵ (Geelan, 2013; Treagust & Harrison, 2000). Für den Physikunterricht liegen weiterhin Arbeiten von Kulgemeyer und Schecker (2013) sowie Kulgemeyer und Tomczyszyn (2015) vor, die einen Vorschlag für die Modellierung der Fähigkeit des Erklärens für Physiklehrkräfte sowie darauf aufbauend für die Berechnung der Erklärensfähigkeit der Physiklehrkräfte vorlegen. Im Zuge der Datenerhebung analysieren Kulgemeyer und Schecker (2013) auch Erklärungen, die sich durch das entwickelte Maß als gelungen einstufen lassen, hinsichtlich der enthaltenen Elemente (z. B. Einleitung, Beispiele). Auch Odora (2014) verfolgt die Messung der Fähigkeit zum Erklären von Technologielehrkräften (**im Original „technology teachers“**; Odora, 2014, S. 71). Hierzu befragt er

⁵³ Die Arbeit von Trommler, Gresch und Hammann (2017) muss hier dennoch erwähnt werden – die Kolleg:innen untersuchen zwar nicht explizit Kriterien von Erklärungen, sie nehmen jedoch Aspekte des Erklärens im Biologieunterricht in den Blick, die auch Rückschlüsse auf Kriterien zulassen. Sie untersuchen die Präferenz teleologischer Erklärungen gegenüber kausalen Erklärungen von Schüler:innen und kategorisieren dabei Begründungsmuster für die Bevorzugung.

⁵⁴ Da die Untersuchung von Wragg (1993) im englischsprachigen Raum verortet ist, bezieht sich der Begriff ‚Englischunterricht‘ in diesem Fall nicht auf den Fremdsprachenunterricht.

⁵⁵ Auch hier wird die Fachspezifität lediglich durch den Einsatz entsprechender Erkläregegenstände repräsentiert, jedoch keine vergleichende Perspektive eingenommen, sodass die Kriterien allgemein für Erklären im Unterricht gelten können.

Schüler:innen direkt zur Einschätzung unterschiedlicher kommunikativer Fähigkeiten ihrer Lehrkräfte. Diese Ansätze und Ergebnisse zur Messung der Fähigkeit zu erklären können ebenfalls in die Erarbeitung von Kriterien für instruktionale Erklärungen einbezogen werden.

Darüber hinaus können Untersuchungen hinsichtlich zweier Grundpositionen unterschieden werden: Einerseits sind dies diejenigen, die die Qualität einer Erklärung an den zu erzielenden Effekten – also dem Lernerfolg, dem Wissenszuwachs oder dem Verstehen – festmachen (z. B. Geelan, 2013; Leinhardt, 1987; Wragg, 1993). Dadurch ist es möglich, zwischen effektiven und weniger effektiven Erklärungen zu unterscheiden. Andererseits gibt es diejenigen Untersuchungen, die die Qualität hinsichtlich aktueller Standards der Disziplin und normativen Ansichten, die von Angehörigen des Feldes geteilt werden, untersuchen (z. B. Schopf & Zwischenbrugger, 2015; Treagust & Harrison, 2000). Der Unterteilung Berliners (2005) folgend ist die erste Position bzw. der erste Untersuchungsansatz mit dem Adjektiv ‚effektiv‘ (im Original „**effective teaching**“) und der zweite als ‚gut‘ (im Original „**good teaching**“) zu bezeichnen (Berliner, 2005, S. 207).⁵⁶ In Anlehnung an diese Unterscheidung werden die Begriffe im Folgenden verwendet.

Eine weitere systematisierende Unterscheidung treffen Schopf und Zwischenbrugger (2015) in ihrer Arbeit. Sie differenzieren dabei zwischen Elementen und Merkmalen des Erklärens.⁵⁷ „Während Elemente Bausteine darstellen, die eine verständliche Erklärung beinhalten sollte, sind Merkmale **Attribute, die die einzelnen Elemente bzw. die gesamte Erklärung erfüllen sollte/n**“ (Schopf & Zwischenbrugger, 2015, S.12). So kann beispielsweise eine Erklärung das Element ‚Beispiel‘ enthalten, das wiederum die Merkmale ‚passend‘ oder ‚interessant‘ erfüllt. Auch diese begriffliche Differenzierungsmöglichkeit wird in der vorliegenden Arbeit aufgegriffen.

Die Kriterien guten bzw. effektiven Erklärens, die in der genannten Literatur Beachtung finden, sind vielfältig. Sie reichen von globalen Forderungen, Erklärungen anschaulich zu gestalten, bis hin zu detaillierten Charakterisierungen von geeigneten Beispielen. Für die Aufarbeitung und Darstellung der Befunde zu gutem bzw. effektivem Erklären wurden die Kriterien zu inhaltlich nahen Kategorien zusammengefasst. Diese sind: Adressatenorientierung (Kap. 4.1), Strukturiertheit (Kap. 4.2), Repräsentationen, Veranschaulichungen und Beispiele (4.3), verständliche Sprache (4.4), Sprech- und Körperausdruck (4.5) sowie fachliche Aspekte (4.6). Im Folgenden werden die genannten Kriterien zunächst beschrieben und die zentralen Befunde aus den oben aufgeführten Studien und theoretischen Überlegungen dargestellt. Zur Vertiefung dort weniger ausführlich behandelter Kriterien – insbesondere zu Repräsentationen sowie zum Einsatz des Sprech- und Körperausdrucks – wird nicht nur Literatur einbezogen, die sich mit Erklären befasst, sondern der Literaturkanon erweitert. Für die spezifische Ausrichtung auf das instruktionale Erklären im Biologieunterricht wird ebenfalls auf weitere biologiedidaktische Literatur (z. B. zum Umgang mit Vorstellungen von Schüler:innen, zu Besonderheiten der Fachsprache, zum Einsatz von Metaphern oder zu fachspezifischen Visualisierungen) zurückgegriffen. Nach der kriterienspezifischen Darstellung werden in Kapitel 4.7 die Ergebnisse der Literaturrecherche in Hinblick auf die sich daraus ergebenden Anforderungen an Lehrkräfte zusammengefasst. Da im empirischen Zugang dieser Arbeit das *gute* Erklären und nicht die Frage nach der Effizienz der Erklärungen im Fokus steht – in

⁵⁶ Qualitativer Unterricht bzw. qualitative Erklärungen bestehen aus den beiden konzeptuell unterscheidbaren Teilen ‚gut‘ und ‚effektiv‘ (Berliner, 2005).

⁵⁷ Eine ähnliche Einteilung findet sich bei Findeisen (2018), die zwischen niedrig-inferenten (beobachtbares Verhalten, z. B. Häufigkeit oder Dauer) und hochinferenten (Interpretation abstrakter Sachverhalte z. B. Qualität) Merkmalen unterscheidet. Die Begriffe finden in dieser Arbeit keine Verwendung, da sie mit der Wortwahl in der Messmethodik verortet ist und mit dem Vorschlag von Schopf & Zwischenbrugger (2015) eine inhaltsnahe Alternative vorliegt.

Anlehnung an die begrifflicher Unterscheidung von Berliner (2005) –, werden in Kapitel 4.8 abschließend die an der Konstruktion dieser normativen Setzung beteiligten Perspektiven skizziert.

4.1 Adressatenorientierung

Die Adressatenorientierung als die Anpassung des zu vermittelnden Inhalts an die Voraussetzungen der Rezipient:innen wird häufig als das zentrale Kriterium instruktionaler Erklärungen beschrieben. Hargie (2013) schreibt aus **kommunikationstheoretischer Sicht**: „Die vielleicht elementarste Regel lautet, dass Erklärungen, wie auch Kommunikation ganz allgemein, auf die Bedürfnisse, Fähigkeiten und Hintergrundkenntnisse der Zuhörer zugeschnitten sein müssen“ (Hargie, 2013, S.265). Auch Wittwer und Renkl (2008) halten aus der Sicht der Lehr-Lern-Psychologie fest, dass Instruktionen ohne die Orientierung an den Adressat:innen nicht effektiv – im Sinne eines Wissensaufbaus – sein können:

In general, the importance of adaptive teaching for learning has long been acknowledged in educational psychology. There is widespread agreement that for any instruction to be effective and efficient, it should be adaptive and built on the learners' individual characteristics. (Wittwer & Renkl, 2008, S. 51).

Aus einer theoretischen Perspektive bezeichnen Schilcher et al. (2017) die Adressatenorientierung und das damit verbundene Ziel des Verstehens durch die Rezipient:innen als ein Merkmal schulischer, instruktionaler Erklärungen, das diese von wissenschaftlichen Erklärungen abgrenzt.

Mitunter wird durch diese fundamentale Bedeutung der Adressatenorientierung, die – wie in den folgenden Abschnitten gezeigt wird – die Ausgestaltung der Erklärung unter anderem auf inhaltlicher, sprachlicher, sprecherischer und visueller Ebene beeinflusst, in einigen Texten die Adressatenorientierung als solche nicht direkt als gesondertes Kriterium angesprochen, wenn Qualitätskriterien von instruktionalen Erklärungen aufgelistet werden (z. B. Kiel, 1999; Odora, 2014). Daraus ist jedoch nicht abzuleiten, dass die Orientierung an den Voraussetzungen der Adressat:innen für einige Autor:innen keine Rolle spielt, sie wird vielmehr in den genannten Aspekten verwoben.⁵⁸ Die Adressatenorientierung ist den Bemühungen des Erklärens – dem Verstehen des Gesagten durch die Adressat:innen – immanent und wird deshalb mitunter weniger explizit dargestellt.

Die Eingangsvoraussetzungen und Eigenschaften der Adressat:innen, die berücksichtigt werden sollen, sind umfassend. Wittwer und Renkl (2008) resümieren, dass Instruktionen allgemein an das Vorwissen, die kognitiven Fähigkeiten sowie die Fertigkeiten der jeweiligen Rezipient:innen anknüpfen müssen, um effektiv sein zu können bzw. zu einem Wissenszuwachs führen zu können. Treagust und Harrison (2000) erweitern dieses Set an Voraussetzungen mit der Berücksichtigung der Motivation und der Interessen der Rezipient:innen. Im Modell von Kulgemeyer und Tomczyszyn (2015) zum Prozess des Erklärens im Physikunterricht mit seinen relevanten Merkmalen wird ebenfalls das Interesse neben dem Vorwissen und Vorstellungen als relevante Variable der Adressatenorientierung erfasst. Die von Schopf und Zwischenbrugger (2015) in Interviews befragten Lehrenden des Lehrstuhls für Wirtschaftspädagogik an der Universität Wien verstehen unter den zu

⁵⁸ Schopf und Zwischenbrugger (2015) bemerken diesen Aspekt in vergleichbarer Weise auch in ihren Befragungen von Wirtschaftslehrenden der Universität Wien zu Merkmalen und Elementen guter Erklärungen. Dabei wurde das Merkmal **guter Erklärungen** „Beachten der Eingangsvoraussetzungen“ (Schopf & Zwischenbrugger, 2015, S.16) von neun von elf Expert:innen genannt, wobei die Interpretation weiterer Interviewabschnitte für die Autorinnen den Schluss nahelegen, dass dieses Merkmal implizit vorausgesetzt und deshalb seltener explizit erwähnt wird.

berücksichtigenden Eingangsvoraussetzungen neben „Vorwissen und Erfahrungen [...] auch Aspekte wie Intelligenz, Interesse, Erwartungen oder auch Alter, soziales Umfeld etc“ (Schopf & Zwischenbrugger, 2015, S.18). Während der hier genannte Aspekt ‚Intelligenz‘ in der Unterteilung von Wittwer und Renkl (2008) zu den kognitiven Fähigkeiten subsummiert werden kann, und der Aspekt ‚Interesse‘ von Treagust und Harrison (2000) sowie Kulgemeyer und Tomczyszyn (2015) aufgenommen wird, sind die genannten Aspekte ‚Alter‘ und ‚soziales Umfeld‘ nicht direkt zuzuordnen. Sie sind vielmehr Hintergrundvariablen, die entsprechend die Aspekte Vorwissen, Fertigkeiten, kognitive Fähigkeiten, Interesse und Motivation beeinflussen können. Für die erklärenden Personen können sie jedoch ein erster Anhaltspunkt sein, die Eingangsvoraussetzungen abzuschätzen. Die genannten Eingangsvoraussetzungen können näher bestimmt werden, sodass die Aspekte, hinsichtlich derer eine Erklärung an die Adressat:innen angepasst werden soll, ebenfalls genauer beschrieben werden können.

Der Aspekt des Vorwissens umfasst zum einen eine fachinhaltsbezogene Dimension. Neben dem schulisch erworbenen Wissen und Vorstellungen in Zusammenhang mit dem Erklärgegenstand sind auch Erfahrungen aus der Lebenswelt und dabei erworbenes Wissen sowie Vorstellungen der Schüler:innen darunter zu verstehen (Schopf & Zwischenbrugger, 2015; Pauli, 2015; Wellenreuther, 2013). Zum anderen wird der Aspekt des Vorwissens auch auf Wissen über Sprache (Kulgemeyer & Schecker, 2013; Kulgemeyer & Tomczyszyn, 2015; Schopf & Zwischenbrugger, 2015; Sevan & Gonsalves, 2008) sowie über (fachspezifische) Darstellungsformen bzw. Repräsentationen (Kulgemeyer & Schecker, 2013; Kulgemeyer & Tomczyszyn, 2015; Leinhardt, 1987; Sevan & Gonsalves, 2008; Wagner & Wörn, 2011) bezogen.

Hinsichtlich der fachinhaltsbezogenen Dimension kann die biologiedidaktische Forschung auf einen umfassenden Korpus zu Schülervorstellungen zurückgreifen. In zahlreichen Studien wurden diese zu unterschiedlichen Themen erhoben und in Übersichtswerken praxisnah zusammenfassend dargestellt (z. B. Hammann & Asshof, 2014; Kattmann, 2015). Die Ergebnisse können dazu genutzt werden, fachlich nicht angemessene Vorstellungen vorab zu identifizieren und die inhaltliche Konzeption der Erklärungen daraufhin anzupassen. Darüber hinaus kann das Vorwissen durch die Auswahl des Inhalts der Erklärung sowie die Erklärungstiefe berücksichtigt werden (Schopf & Zwischenbrugger, 2015; Wagner & Wörn, 2011; Wellenreuther, 2013). Bezüglich der Anpassung an das Vorwissen der Lernenden konnten Nückles, Wittwer und Renkl (2005) in einem experimentellen Setting zeigen, dass erklärende Personen, die Informationen über das Vorwissen ihrer Lernenden erhielten, die Erklärung hinsichtlich der Informationsmenge und Art der Informationen variierten. Die an die Lerner adaptierten Erklärungen erzielten bessere Lernergebnisse in nachfolgenden Tests als die Kontrollgruppen mit falschen Informationen über die Lernenden. Kotthoff (2009) verweist dazu auch auf eine Studie von Erickson und Shultz (1982), die in Beratungsgesprächen das Phänomen des ‚Übererklärens‘ identifizierten: Studierende, die von Personen beraten wurden, die deren Vorwissen sowie die Verstehenssignale nicht angemessen berücksichtigten und das Vorwissen dadurch unterschätzten, fühlten sich „für dumm verkauft“ (Kotthoff, 2009, S. 122) und herablassend behandelt.

Das Wissen über Sprache (z. B. geläufiger Wortschatz, verständlicher Satzbau, Formeln als Sprache der Mathematik) und Darstellungsformen bzw. Repräsentationen (z. B. Tabellen, Diagrammen, Analogien und Metaphern) wird in der Literatur in einer Vielzahl an Veröffentlichungen aufgegriffen (z. B. Kulgemeyer & Schecker, 2013; Kulgemeyer & Tomczyszyn, 2015; Schopf & Zwischenbrugger, 2015; Sevan & Gonsalves, 2008). Die genauere Betrachtung der beiden Aspekte findet in

gesonderten Kapiteln (Kap. 4.2 ‚Repräsentationen, Veranschaulichungen und Beispiele‘ und Kap. 4.4 zum Aspekt ‚Sprache‘) statt.

Herausgegriffen werden soll an dieser Stelle lediglich das Modell von Kulgemeyer und Tomczyszyn (2015) zum sach- und adressatengerechten Erklären im Physikunterricht. In ihrem Modell kann die „Adressatengemäßheit“ (Kulgemeyer & Tomczyszyn, 2015, S.5) durch vier Variablen hergestellt werden: den allgemeinen sprachlichen Code, die Analogien und Beispiele, die Mathematisierungen sowie die grafischen Darstellungsformen. Mittels der vier Variablen kann die erklärende Person die Erklärung an die Voraussetzungen der Rezipient:innen (Vorwissen, Interessen, Vorstellungen) anpassen, gegebenenfalls auch nach deren Rückmeldung. Dass die vier Variablen Anpassungsprozesse beschreiben können, zeigen sich Autor:innen anhand von Daten aus Expert:innen-Noviz:innen-Dialogen. Das beschriebene Modell ist an dieser Stelle erwähnenswert, da die Adressatengemäßheit ausschließlich über sprachliche Merkmale und Merkmale der Repräsentationen, die von den Autor:innen zur Fachsprache gezählt werden, modelliert wird.⁵⁹

Bei der Beschreibung der Eingangsvoraussetzungen wurden auch die kognitiven Fähigkeiten genannt. Darunter können beispielsweise das Abstraktionsvermögen (Treagust & Harrison, 2000; Wagner & Wörn, 2011) sowie die Kapazität des Arbeitsgedächtnisses (Wellenreuther, 2013) verstanden werden. Die Berücksichtigung des Abstraktionsvermögens der Adressat:innen impliziert eine Anpassung des Abstraktionsgrades des Gegenstandes, wofür dieser von der Lehrkraft erfasst und anschließend adaptiert werden muss. Zu den Anpassungsmöglichkeiten kann der von Wellenreuther (2013) im Zuge seiner Beschreibungen zur Entlastung des Arbeitsgedächtnisses geforderte Einsatz von konkreten Beispielen angeführt werden. Treagust und Harrison (2000) kommen bei der Analyse von Richard Feynmans Lernmaterial „Atoms in Motion“ (als ein Teil der Reihe „Six Easy Pieces“ (1994)) zu dem Schluss, dass teleologische Formulierungen zur Reduktion des Abstraktionsgrades eingesetzt werden können und die Erklärung dadurch an Effektivität gewinnen kann.⁶⁰ Dabei sollten sie im Verlauf jedoch nicht als Ersatz für die wissenschaftlichen Erklärungen stehen bleiben (Treagust & Harrison, 2000). Eine Diskussion über die Verwendung und Angemessenheit teleologischer Formulierungen wird auch in der Biologiedidaktik geführt. Dabei kann festgehalten werden, dass Schüler:innen diese häufig aus verschiedenen Gründen nutzen (Trommler et al., 2017). Insofern kann die Aufnahme dieser Formulierungen in die Unterrichtssprache der Lehrkraft als eine Hinwendung zu den Lernenden verstanden werden. Unter anderem von Kattmann (2005) werden sie deshalb auch als förderlich für fachliches Lernen beschrieben. Er bezeichnet sie als „mental nahe liegendes Instrument für das Lernen“ (Kattmann, 2005, S.166). Gleichzeitig versprachlichen teleologische Formulierungen eine Ziel- bzw. Zweckorientierung, die als fachlich inkorrekt beurteilt werden kann, da sich biologische Phänomene oder Systeme nicht aktiv zielorientiert entwickeln oder verhalten (Mayr, 2002). Teleologische Formulierungen können daher auch lernhemmend wirken, da sie fehlerhafte Vorstellungen evozieren können. Es kann zusammengefasst werden, dass teleologische Formulierungen ein adressatenorientiertes sprachliches Mittel sind, die möglichen teleologischen Vorstellungen vor dem Hintergrund fachlich angemessener Konzepte jedoch hinterfragt werden müssen. Insofern kann die

⁵⁹ Kulgemeyer und Tomczyszyn (2015) unterscheiden in ihrem Modell, das aus dem von Kulgemeyer und Schecker (2013) entwickelten, allgemeinen Modell zur Kommunikation von wissenschaftlichen Inhalten hervorgeht, jedoch nicht zwischen Adressatengemäßheit und Sachgemäßheit. Beide Aspekte können von der Lehrkraft durch die Gestaltung bzw. die Adaption der vier Variablen erreicht werden. In ihrer Studie zeigte sich, dass sich die vier Variablen gut zur Beschreibung der Adaption von Erklärungen auf Grund von Nachfragen Seitens der Rezipient:innen eignen.

⁶⁰ Die teleologischen Formulierungen Feynmans (1994) dienen laut Treagust und Harrison (2000) auch der Motivation der Schüler:innen und verhelfen der Erklärung zur Kohärenz, da weniger lange Teilerklärungen zu Abstraktem oder schwer Verständlichem nötig werden.

Forderung von Treagust und Harrison (2000), dass teleologisch formulierte (Teil-)Erklärungen am Ende des Erklärprozesses oder der Unterrichtssequenz nicht als Endprodukt bestehen bleiben dürfen, sondern reflektiert werden sollten, auch für den Biologieunterricht gelten.

Die von Schopf und Zwischenbrugger (2015) befragten Expert:innen gaben weiterhin an, dass der Komplexitätsgrad des Erklärgegenstandes an die Adressat:innen angepasst werden muss.

Als weitere Charakteristika der Lernenden werden deren Interesse und Motivation genannt. Beide Aspekte tragen dazu bei, Aufmerksamkeit als ersten Schritt für einen erfolgreichen Lernprozess zu generieren und Bereitschaft dafür zu entwickeln, sich aktiv mit dem Erklärgegenstand auseinanderzusetzen (Kulgemeyer & Schecker, 2013; Pauli, 2015). Die kognitive Aktivität der Rezipient:innen ist dabei Grundlage für die Sinnkonstruktion bzw. die Verstehensleistung und damit Voraussetzung für den Wissenserwerb. Da die Schüler:innen frei sind, diese Aktivität zu leisten, wird eine interessante und motivierende Gestaltung von Erklärungen als relevantes Merkmal beschrieben (Kulgemeyer & Schecker, 2013). Um Interesse⁶¹ und Motivation herzustellen, empfiehlt Pauli (2015), dass die Fragestellung, die der Erklärung vorausgeht, relevant für die Lernenden sein sollte. Um die Aktivität der Adressat:innen weiterhin zu erhöhen, wird von Kulgemeyer und Schecker (2013) die direkte Ansprache der Lernenden empfohlen. Auch der Einbezug der Lebenswelt der Schüler:innen kann Interesse und Motivation erhöhen.

Geelan (2013) konnte bei der Beschreibung von Erklärungen erfolgreicher Lehrkräfte auch den Aspekt ‚Humor‘ herausarbeiten, der zum Aufbau einer persönlichen Beziehung zu den Rezipient:innen genutzt werden und dadurch die Offenheit für erklärte Inhalte steigern kann:

Different teachers had different personal styles, and physics teacher humor tends to be quite dry, but many of the teachers used humor with their students to help build relationships and improve classroom climate. Some used quite sophisticated “physics humor” to test and reinforce physics concepts. (Geelan, 2013, S. 1759-1760).

Die Adaptionen sind während des Planungsprozesses, also vor dem Erklären des Sachverhaltens, durch die methodischen Überlegungen und eine entsprechende didaktische Reduktion seitens der Lehrkraft zu leisten. Die didaktische Reduktion ist zentral, um den Umfang des Erklärinhalts zu reduzieren (sektorale Reduktion) und die Komplexität (auch im Sinne des Detailreichtums; Aussparung von Teilaspekten, die wiederum einer Erklärung bedürftigen) des Erklärinhalts zu reduzieren (strukturelle Reduktion; Kattmann, 2008a). Angepasst werden dadurch zum einen der Schwierigkeitsgrad und zum anderen der Informationsgehalt der Erklärung.

Während und direkt nach der Erklärung kann aufgrund von Rezipient:innensignalen die Erklärung ebenso angepasst bzw. eine erneute Erklärsequenz für unverstandene Teile abgegeben werden (Treagust & Harrison, 2000). Rezipient:innensignale, die die Lehrkraft für die Notwendigkeit einer Anpassung nutzen kann, können in nonverbale und verbale Signale unterschieden werden (Schopf & Zwischenbrugger, 2015; Sevan & Gonsalves, 2008; Pauli, 2015). Zu den nonverbalen Rückmeldungen zählen beispielsweise die Mimik und Gestik der Rezipient:innen (Schopf & Zwischenbrugger, 2015). Sie müssen von Lehrkräften dementsprechend interpretiert werden, um Aufschluss über das Verstehen oder Nichtverstehen der Erklärung zu erlangen. Zu den verbalen

⁶¹ Es ist davon auszugehen, dass hier nicht von Interesse als einer Persönlichkeitseigenschaft, also einer langanhaltenden, relativ stabilen individuellen Präferenz gegenüber einem Lerngegenstand oder Themengebiet gesprochen wird, sondern eher von einem situationalen, kurzzeitigen motivationalen Zustand, der vor allem durch die attraktive Gestaltung des Lerngegenstandes bzw. allgemein der Umwelt hervorgerufen wird (Krapp & Weidenmann, 2006).

Rückmeldungen gehören beispielsweise Fragen von Schüler:innen, die das Verstehensdefizit unmittelbar verdeutlichen.

Um ein Nichtverstehen im Nachgang der Erklärung oder nach Abschluss von Teilschritten zu diagnostizieren, können Lehrkräfte auch explizite Fragen stellen und damit das Verstehen – zumindest einzelner Rezipient:innen – überprüfen (Kennedy, 1996; Kulgemeyer & Tomczyszyn, 2015; Schopf & Zwischenbrugger, 2015; Sevan & Gonsalves, 2008). Laut den von Schopf und Zwischenbrugger (2015) befragten Wirtschaftslehrenden an der Universität Wien können dies auch dichotom angelegte, direkte Nachfragen zum Verstehen sein (z. B. ‚Habt ihr das verstanden?‘). Dagegen verweisen Kulgemeyer und Tomczyszyn (2015) darauf, dass Nachfragen dieser Art nicht zu diagnostischen Zwecken geeignet sind, sondern dabei auf Wissensfragen zurückgegriffen werden sollte, bei denen das Erklärte angewendet werden muss. Dementsprechend werden auch ausschließlich Wissensfragen im Codiermanual von Kulgemeyer und Tomczyszyn (2015) sowie zur Erklärens-fähigkeit zuträglich codiert. In den empirischen Daten von Kulgemeyer und Schecker (2013) zeigte sich das Einholen einer Verstehensbestätigung auch als einer von sechs signifikanten Einflussfaktoren auf die Erklärens-fähigkeit angehender Lehrkräfte. Weiterhin werden Handlungsaufforderungen und die Diagnose ihrer Ergebnisse als Möglichkeit beschrieben (Kulgemeyer & Tomczyszyn, 2015).

Dadurch, dass Teilschritte der Erklärung oder die Erklärung als Ganzes nicht verstanden werden können, müssen von der erklärenden Person weitere Erklärvarianten mit gegebenenfalls anderen Beispielen, variierender sprachlicher Darstellung oder anderen Herleitungen zur Verfügung gestellt werden können. Von Treagust und Harrison (2000) sowie Wagner und Wörn (2011) wird deshalb die Verfügbarkeit mehrerer Erklärvarianten als Merkmal der Fähigkeit, gut erklären zu können, **verstanden** (bei Treagust und Harrison als „**explanatory flexibility**“ bezeichnet (Treagust & Harrison, 2000, S. 1161)). In den Studien von COACTIV⁶² (Kunter et al., 2011) und FALKO⁶³ (Krauss et al., 2017) wurde dies als Teil der Facette ‚Wissen über Erklären und Repräsentieren von Inhalten‘ konzeptualisiert und operationalisiert.

Insgesamt stellt sich auf Basis der genannten Literatur die Adressatenorientierung als umfassender Aspekt dar. Die Einschätzung der kognitiven und motivationalen Charakteristika der Zielgruppe (z. B. Vorwissen, Schülervorstellungen, kognitive Fähigkeiten, Interesse, Motivation) ist die Grundlage für didaktische und methodische Überlegungen, um den Inhalt in Bezug auf dessen Abstraktions- und Komplexitätsgrad, die sprachliche Darstellung sowie die Repräsentationen passend auszuwählen respektive zu gestalten. Durch die Verbindung mit diesen Aspekten sind eine disjunkte Konzeptualisierung und Operationalisierung der Adressatenorientierung herausfordernd und werden von verschiedenen Forschenden unterschiedlich gelöst (Kulgemeyer & Tomczyszyn, 2015; Sevan & Gonsalves, 2008; Wagner & Wörn, 2011; zur Umsetzung in dieser Arbeit siehe Kap. 5.2.4.1).

⁶² Akronym für ‚Cognitive Activation in the Classroom‘; Forschungsprojekt mit dem Ziel, einen Professionswissenstests für Mathematiklehrkräfte zu entwickeln und mit der Gestaltung von Mathematikunterricht sowie dem Lernen der Schüler:innen in Verbindung zu bringen.

⁶³ Akronym für ‚Fachspezifische Lehrerkompetenzen‘; ein Forschungsprojekt an der Universität Regensburg mit dem Ziel Professionswissenstests für die Fächer Deutsch, Englisch, Latein, Physik, Musik, Evangelische Religion und Pädagogik zu entwickeln.

4.2 Strukturiertheit

Anhand von Unterrichtsbeobachtungen wurde die Prozessstruktur von instruktionalen Erklärungen untersucht und dabei deren Einbettung in den Unterrichtsverlauf sowie abgrenzbare Teilschritte dieses Prozesses beschrieben (z.B. Hargie, 2013; Vogt, 2009; Wagner & Wörn, 2011). Wagner und Wörn (2011) gliedern den Ablauf einer Erklärung allgemein vierphasig in den Erklär Anlass (zum Beispiel Wissens- oder Verständnislücken aufseiten der Schüler:innen), die Erklärinitiation (signalisiert das Bedürfnis einer Erklärung für alle Beteiligten), den Erklärprozess (die Bearbeitung des Explanandum) und die Erklärcoda (der Abschluss der Erklärung zum Beispiel durch Verstehensäußerungen der Schüler:innen). Innerhalb dieses Modells präzisieren Wagner und Wörn (2011) auch, dass die Phase des Erklärprozesses nicht zwingend linear verlaufen muss:

Bei komplexeren Explanandi kommt es innerhalb des eigentlichen Erklärprozesses dagegen oft zu neuen, eingeschobenen (Teil-)Erklärungen [...]. Auch innerhalb der Teilerklärungen können sich wieder neue Erklärgegenstände ergeben, sodass auch hier neue (Teil-)Erklärungen nötig werden. In diesen Fällen handelt es sich nicht um einen linearen Verlauf, sondern vielmehr um vielschichtige, verzweigte Verläufe, die von ineinander verschachtelten Erklärschleifen geprägt sind. (Wagner & Wörn, 2011; S. 26).⁶⁴

Zur Güte der inhaltlichen Strukturierung – also zur Phase des Erklärprozesses – finden sich sowohl theoretische Überlegungen als auch empirische Daten. Dabei heben einige Arbeiten die klare Herausarbeitung von Schlüsselementen oder Verstehenselementen hervor, wobei darunter jene zentralen Elemente, Fakten oder Konzepte verstanden werden, die das Verstehen des Erklärgegenstandes ermöglichen (Brown, 2006; Hargie, 2013; Wagner & Wörn, 2011; Wragg, 1993). Diese müssen von der Lehrkraft bereits beim Planungsprozess herausgearbeitet und bei Bedarf reduziert werden (Hargie, 2013). Wittwer und Renkl (2008) formulieren mit besonderer Aufmerksamkeit auf Erklärungen im Mathematikunterricht die Notwendigkeit der Fokussierung auf Konzepte und Prinzipien: „[...] **instructional explanations that specifically focus on concepts and principles**“ (Wittwer & Renkl, 2008, S. 54).⁶⁵ Da sich Erklärungen im Mathematikunterricht auch ausschließlich auf Regeln und damit auf eine algorithmische Vorgehensweise beim Lösen von Aufgaben, die ohne ein Verständnis für entsprechende Konzepte auskommt, beziehen können (Weinhuber, Lachner, Leuders & Nückles, 2019), ist dieses Kriterium für mathematische Erklärungen mitunter besonders relevant. Auch bei der Expert:innenbefragung von Schopf und Zwischenbrugger (2015) aus dem Bereich der Wirtschaftspädagogik wird die Herausstellung des zugrunde liegenden Prinzips als Voraussetzung für den Aufbau transferfähigen Wissens genannt. Für den Biologieunterricht liegen beispielsweise durch die Formulierung von Basiskonzepten Vorschläge für grundlegende Konzepte vor (siehe Kap. 4.6).

Die verschiedenen Teile der Erklärung müssen des weiteren Schritt für Schritt eingeführt und schließlich in eine logische Anordnung gebracht werden (Schopf & Zwischenbrugger, 2015; Sevan & Gonsalves, 2008; Pauli, 2015; Wragg, 1993).

⁶⁴ Gerade mit Blick auf die unter Kapitel 3 beschriebene Komplexität allgemein und die zeitliche Parallelität von Prozessen im Speziellen, die biologischen Phänomenen und damit Erklärgegenständen im Biologieunterricht inhärent sind, liefert dieses Modell auch theoretische Ansatzpunkte für Schwierigkeiten bei der Erklärung komplexer Erklärgegenstände.

⁶⁵ Sie tun das mit Verweis auf Studien zur Struktur von Expert:innenwissen: Expert:innen besitzen im Gegensatz zu Noviz:innen in ihrer Domäne mehr elaborierte Konzepte und Schemata, die eine rasche und tiefe Durchdringung von Problemstellungen und damit auch eine Identifikation von aussichtsreichen Problemlösestrategien ermöglichen.

In dem von Sevian und Gonsalves (2008) auf empirischer Basis entwickelten Kategoriensystem ist weiterhin die Organisation vom Einfachen zum Komplexen aufgeführt. Aus ihren empirischen Daten leiten sie auch einen weiteren Aspekt zur Strukturiertheit ab: Neben der logischen Organisation der Teile (Einführung, Problemstellung, Beispiele etc.) einer Erklärung sollten die Teile auch in einem ausgewogenen zeitlichen Verhältnis innerhalb des Erklärprozesses aufgegriffen werden.

Für eine kohärente Anordnung der Teile der Erklärung sind dabei vor allem die Beziehungen und Zusammenhänge zwischen den Verstehenselementen entscheidend. Dabei muss sich bei der Anordnung sowohl an der Sachstruktur des Erklärgegenstandes als auch an den vorhandenen Wissensstrukturen der Adressat:innen orientiert werden (Ehlich, 2009). Die Anordnung der Teile ist damit für jede Erklärsituation individuell.

Die Zusammenhänge zwischen den Verstehenselementen müssen für die Rezipient:innen klar sein oder durch die Lehrkraft explizit deutlich gemacht und Sprünge vermieden werden. So entsteht ein roter Faden, dem die Adressat:innen folgen können. Die Metapher des roten Fadens bleibt als solche eher abstrakt und gibt damit wenig Aufschluss für die Erhöhung der Qualität einer instruktionalen Erklärung. Anhaltspunkte zur Realisierung eines roten Fadens können sich im Bereich der Textanalyse – also im Bereich der Sprachwissenschaft – finden lassen: Der Begriff der Kohärenz umfasst die Verknüpfung von Aussagen und deren Widerspruchsfreiheit (Friedrich, 2017; Kintsch & van Dijk, 1978; Schnotz, 1994). Unterschieden werden kann dabei zwischen lokaler Kohärenz, wobei Sinnzusammenhänge zwischen aufeinanderfolgenden Aussagen hergestellt werden, und globaler Kohärenz, wobei Sinnzusammenhänge über den gesamten Text hinweg deutlich gemacht werden. An diesem Beispiel wird deutlich, dass die Strukturiertheit einer Erklärung auch durch sprachliche Mittel erzeugt werden kann (siehe Kap. 4.4).

Um die Klarheit der Struktur zu erhöhen, wird häufig gefordert, den Aufbau der Erklärung für die Adressat:innen transparent zu gestalten (Brown, 2006; Kennedy, 1996; Leinhardt, 1987). Zur Klarheit kann zum einen die explizite Offenlegung der Struktur beim Einstieg in den Erklärprozess beitragen (Brown, 2006; Sevian & Gonsalves, 2008). Zum anderen kann durch strukturierende Äußerungen der Lehrkraft während der Erklärung die Klarheit erhöht werden. Die Äußerungen können beispielsweise den jeweiligen Anfang und das Ende von Themen und Einschüben in der Erklärung markieren, sogenannte ‚framing statements‘ – (Brown, 2006; Hargie, 2013; Kennedy, 1996), oder den Fortschritt der Erklärung in Hinblick auf das zu erreichende Ziel verdeutlichen (Leinhardt, 1987). Weiterhin können sich die Äußerungen auf inhaltliche Aspekte der Erklärung beziehen und wichtige Punkte der Erklärung betonen (Brown, 2006; Schopf & Zwischenbrugger, 2015).⁶⁶ Ohne dabei auf das sprachwissenschaftliche Konzept der Kohärenz einzugehen, wird dieses in der hier genannten didaktischen Literatur ebenfalls implizit genutzt. So können die ‚framing statements‘ und die Hinweise auf den Fortschritt der Erklärung als Kohärenzbildungshilfen bezeichnet werden (Schnotz, 1994).

Auch implizit kann Struktur an die Hörer:innen vermittelt werden, indem beispielsweise entsprechende Gestiken dazu genutzt werden, den Beginn eines neuen Punktes zu markieren, oder

⁶⁶ Die strukturierenden Äußerungen während der Erklärung werden von einigen Autor:innen eher dem Aspekt der Strukturiertheit einer Erklärung zugeordnet (z. B. Brown, 2006; Kennedy, 1996; Leinhardt, 1987), von anderen Autor:innen (z. B. Hargie, 2013) zu den sprachlichen Aspekten der Erklärung. Die Zuordnung zur Struktur der Erklärung ist eher offensichtlich, da derartige sprachliche Äußerungen inhaltlich die Struktur betreffen und sie der Kommunikation derer von der erklärenden Person an die Rezipient:innen dienen. Die Modellierung bzw. Konzeptualisierung als sprachliches Merkmal ist jedoch ebenso gut begründbar: Einschübe und Verweise, wie der aufbereitete Inhalt gegliedert ist, können auf Ebene der Textanalyse betrachtet werden und strukturieren eine zusammenhängende sprachliche Äußerung (siehe Kap. 4.4).

am Ende eines Punktes die Stimme gesenkt und mit entsprechender Pausensetzung ein Abschluss kenntlich gemacht wird (Hargie, 2013).

Einen weiteren Aspekt arbeiten Wagner und Wörn (2011) theoretisch heraus: „**Eine klare Struktur** wird auch dadurch erreicht, dass die Erklärung sich auf die wesentlichen Aspekte des jeweiligen Erklärgegenstandes konzentriert und den eigentlichen Kern herausarbeitet. Insbesondere bei komplexen Erklärungen [...] besteht die Gefahr, das Wesentliche aus den Augen zu verlieren und abzuschweifen“ (Wagner & Wörn, 2011, S. 28; ähnlich auch bei Pauli (2015) und Hargie (2013)). Begründet werden können die Begrenzung der Informationsmenge innerhalb der Erklärung und die Fokussierung auf wesentliche Kerninhalte dadurch, dass die Rezipient:innen lediglich eine definierte und begrenzte Menge an Informationen in einer vorgegebenen Zeit verarbeiten können (Hargie, 2013; Wellenreuther, 2013). Insbesondere der Umstand, dass bei mündlichen nicht wie schriftlichen Erklärungen nach hinten gesprungen und in unverständlichen Textpassagen verweilt werden kann, betont die Bedeutung der Reduktion unwesentlicher Details in mündlichen instruktionalen Erklärungen. Auch in einer Expert:innenbefragung wird dieser Aspekt für gute Erklärungen angeführt (Schopf & Zwischenbrugger, 2015).

Neben diesen Kriterien, die auf die jeweilige Ausgestaltung referieren, werden auch Elemente als förderlich für die Qualität bzw. Effektivität einer Erklärung beschrieben. Die Herausarbeitung des Erklärziels (auch als Aushandlung des Erklärgegenstandes) und die klare Zielformulierung zu Beginn der Erklärung werden ebenfalls als relevante Elemente beschrieben (Wagner & Wörn, 2011) und sind Teil effektiver Erklärungen (Kulgemeyer & Schecker, 2013; Leinhardt, 1987). Kulgemeyer und Schecker (2013) heben als Funktion des Einstiegs in die Erklärung auch die Herausarbeitung der Relevanz des Themas hervor. Eng damit verbunden ist die Einbettung in einen größeren Zusammenhang. Sowohl bei der Expert:innenbefragung von Schopf und Zwischenbrugger (2015) zu guten Erklärungen als auch bei der Beschreibung effektiver Erklärungen von Sevian und Gonsalves (2008) wird dieses Kriterium empirisch fundiert herausgearbeitet. Dabei geht es um die Vernetzung des Erklärgegenstandes mit weiteren Inhalten des eigenen Faches oder anderer Fächer, um diesen in einen größeren Kontext einbetten zu können.

Das Ende des Erklärprozesses kann mit einer Zusammenfassung des Inhalts schließen (Kulgemeyer & Tomczyszyn, 2015; Odora, 2014), die bei einer Untersuchung von Wragg (1993) auch ein Element effektiver Erklärungen war.

Insgesamt kann die Struktur einer Erklärung demnach durch Elemente wie klare Zielformulierung, Offenlegung der Struktur und eine Zusammenfassung verbessert werden. Die Strukturiertheit als Merkmal einer guten Erklärung ergibt sich aus inhaltlichen, sprech-sprachlichen und visuellen Faktoren, die Stringenz erzeugen und den Adressat:innen ermöglichen, dem Gesagten zu folgen. Dabei sollte der Gegenstand klar im Fokus stehen, gegebenenfalls auf besonders relevante Punkte verwiesen werden und die notwendigen Verstehenselemente sollten Schritt für Schritt eingeführt werden.

4.3 Einsatz von Repräsentationen

Der Begriff der Repräsentationen wird vor allem im Bereich des Multimedia Learning genutzt und stammt aus dem Forschungsbereich der Kognitionspsychologie. Dabei werden unter externen Repräsentationen unter anderem Fotos, Schemazeichnungen, Tabellen, Reaktionsgleichungen sowie alle Formen von Texten und verbalen Ausdrücken verstanden (Nitz, Nerdel & PrechtI, 2012;

Schnotz, 2003). Sie repräsentieren reale Objekte, Zusammenhänge oder Sachverhalte stellvertretend. Wenn sie von den Adressat:innen verstanden werden, können aus den externen Repräsentationen interne, mentale Repräsentationen über die repräsentierten Objekte, Zusammenhänge oder Sachverhalte aufgebaut werden (Schnotz, 2003).

Im Diskurs um gutes Erklären wird meist von Visualisierungen oder Veranschaulichungen statt von externen Repräsentationen gesprochen. Der Begriff der externen Repräsentationen bietet jedoch den Vorteil, dass er Teil eines größeren Konzepts ist, bei dem die Verarbeitung der externen Repräsentationen bis hin zum Aufbau interner, mentaler Repräsentationen untersucht wurde (Schnotz, 2003). Die Untersuchung und Beschreibung der Verarbeitungsprozeduren basieren auf einer vielrezipierten Einteilung externer Repräsentationen, die von Schnotz (2003) vorgenommen wurde. Er unterscheidet zwischen realistischen Bildern, logischen Bildern und Texten. Da diese Unterteilung auch in den folgenden Abschnitten aufgegriffen wird, ist in der nachstehenden Tabelle (Tab. 3) für eine bessere Orientierung im Text zunächst eine Übersicht über die Formen externer Repräsentationen sowie konkrete Beispiele aus dem Biologieunterricht dargestellt.

Tabelle 3: Formen externer Repräsentationen sowie Beispiele für die jeweiligen Formen

Formen externer Repräsentationen	Verwendungen im Biologieunterricht
Realistische Bilder	Fotografien, mikroskopische Aufnahmen, Zeichnungen in verschiedenen Abstraktionsstufen (z. B. Skizzen, Schemazeichnungen)
Logische Bilder	Tabellen, Flussdiagramme, Strukturdiagramme, Säulendiagramme oder Liniendiagramme
Texte	Mündliche und schriftliche Texte, Metaphern und Analogien, Symbole und Formeln

Die drei Formen externer Repräsentation werden im Folgenden hinsichtlich ihrer Charakteristika skizziert. Dabei wird auch auf die Verarbeitungsprozesse bei den Adressat:innen hingewiesen. Anschließend werden Gestaltungsrichtlinien und -empfehlungen für den Einsatz der externen Repräsentationen in instruktionalen Erklärungen beschrieben. Da im Kontext des Biologieunterrichts auch dem Einsatz lebender Objekte eine große Relevanz zukommt, werden diese als Form der Veranschaulichung am Ende des Kapitels ebenfalls kurz aufgegriffen.

An dieser Stelle kann vorweggenommen werden, dass der Einsatz von externen Repräsentationen (bzw. Visualisierungen oder Veranschaulichungen) grundsätzlich als Merkmal guter bzw. effektiver Erklärungen angegeben wird, da sie dem Aufbau interner Repräsentationen respektive der Veranschaulichung des Erklärgegenstandes dienen und damit potenziell dem Verstehen zuträglich sind. Grundlegend gilt, dass die externen Repräsentationen die Voraussetzungen der Adressat:innen – z. B. das Vorwissen zur Informationsentnahme aus der gezeigten Tabelle – sowie die fachlichen Merkmale des Erklärgegenstandes berücksichtigen müssen.

Realistische Bilder als externe Repräsentationen

Zur Kategorie der realistischen Bilder zählen beispielsweise Fotos, mikroskopische Aufnahmen, Skizzen und Schemazeichnungen. An dieser Aufzählung lässt sich bereits erkennen, dass realistische Bilder bezüglich ihres Abstraktionsgrades variieren können. Sie haben deshalb eine mehr oder weniger große Ähnlichkeit mit dem repräsentierten Gegenstand, repräsentieren aber stets

gleichartig konkrete Strukturmerkmale: So wird ein räumliches Merkmal des Gegenstands auch durch räumliche Merkmale im Bild repräsentiert (Schnotz, 2001). Bei logischen Bildern (z. B. Diagrammen – siehe folgender Abschnitt) ist dies nicht mehr der Fall. Die kognitive Verarbeitung realistischer Bilder wird von Weidenmann (1994) in zwei Phasen unterteilt. Sie besteht aus der perzeptiven Encodierung und basiert auf unbewussten Prozessen (präattentive Prozesse) der visuellen Wahrnehmung (Weidenmann, 1994), verläuft automatisiert und ist überwiegend unabhängig vom Vorwissen und den Zielen der Betrachtung (Schnotz, 2001; Weidenmann, 1994). Für den Verstehensprozess relevant ist die anschließende semantische Encodierung (attentive Prozesse), bei der die eigentliche Interpretation des Gesehenen stattfindet (Schnotz, 2001; Weidenmann, 1994). Die Verarbeitung der Informationen ist dabei vom Vorwissen und von den Zielen der Betrachtung abhängig. Die Autor:innen Brandstetter, Florian und Sandmann (2016) konnten in ihrer Studie zum Bildverstehen im Kontext des Biologieunterrichts zeigen, dass das Bildverstehen (also die semantische Encodierung) in hohem Maße vom Vorwissen der Schüler:innen abhängig ist. Zwischen dem Abstraktionsgrad der Darstellungen unter Kontrolle des Vorwissens der getesteten Schüler:innen und dem Bildverstehen bestehen jedoch keine gleichgerichteten Zusammenhänge. So verdeutlichen die Studienergebnisse der Autor:innen, dass das Bildverstehen für verschiedene Darstellungen zum Kniesehenreflex höher ausfiel, wenn die Bilder einen höheren Abstraktionsgrad aufwiesen. Für Abbildungen zum Blutkreislauf erwies sich jedoch eine weniger abstrakte Darstellung als verstehensförderlich (Brandstetter et al., 2016).

Logische Bilder als externe Repräsentationen

Zu den logischen Bildern zählen hingegen grafische Darstellungen, die sehr viel abstrakter sind und über das direkt Wahrnehmbare hinausgehen. Zum Beispiel sind dies Tabellen, Flussdiagramme, Strukturdiagramme, Säulendiagramme oder Liniendiagramme (Schnotz, 2001). Hierbei ist keine direkte Übereinstimmung von Strukturmerkmalen gegeben, sondern es besteht eine Analogierelation zwischen der Eigenschaft des repräsentierten Gegenstands (z. B. Anzahl einer Art in einem Ökosystem zu verschiedenen Zeitpunkten) und der externen Repräsentation (z. B. Höhe des Balkens im Diagramm zu einem Zeitpunkt; ein nicht-räumliches Merkmal wird in der Repräsentation räumlich dargestellt). Die kognitive Verarbeitung wird auch hier in prä-attentive und attentive Prozesse bzw. perzeptiv und semantische Encodierung unterteilt. Die prä-attentiven Prozesse beziehen sich auf die Unterscheidung und das Erkennen grafischer Elemente der Darstellung, also von Punkten, Linien und Flächen sowie die grafische Konfiguration derer (Schnotz, 2001). Die verstehensrelevanten, attentiven Prozesse beziehen sich anschließend auf eine konzeptgeleitete Analyse der grafischen Konfigurationen, also eine Interpretation der visuell-räumlichen Relationen (z. B. die Bedeutung der Fläche des dargestellten Balkens in einem Balkendiagramm; Schnotz, 2001). Durch die Repräsentation von analogen Relationen statt der Abbildungen von gleichartigen Strukturmerkmalen, wie dies bei realistischen Bildern der Fall ist, müssen Lernende hier über alltägliche Wahrnehmungsschemata hinaus Vorwissen zum Lesen bzw. zur Informationsentnahme aus der jeweiligen Darstellungsform abrufen.

Texte als verbale, externe Repräsentationen

Verbale Repräsentationen umfassen allgemein sprachlich basierte Repräsentationen in Form von Sprache, die visuell oder auditiv dargeboten sein können. Hinsichtlich der höheren kognitiven Prozesse zum Verstehen des Gesagten oder des Geschriebenen und zur Konstruktion mentaler

Repräsentationen werden keine wesentlichen Unterschiede bei der Verarbeitung visueller und auditiver Darbietung angenommen (Schnotz, 2001). In beiden Fällen müssen semantische Zusammenhänge (lokale und globale Kohärenz; siehe Abschnitt 4.4) rekonstruiert werden. Durch Forschungsvorhaben im Bereich der Lernpsychologie wurde der Vorgang bei der Konstruktion mentaler Repräsentationen weiter beschrieben:

Demnach bildet der Leser eines schriftlich-visuell dargebotenen Texts oder der Hörer eines mündlich-auditiv dargebotenen Texts eine mentale Repräsentation der sprachlichen Oberflächenstruktur, generiert auf dieser Grundlage eine propositionale Repräsentation⁶⁷ des semantischen Gehalts des Texts (die sog. Textbasis) und konstruiert auf dieser Grundlage schließlich ein mentales Modell des dargestellten Sachverhalts. (Schnotz, 2001, S. 301–302)

Wie auch bei realistischen und logischen Bildern scheint das Vorwissen, das im Langzeitgedächtnis verortet wird, bei der Konstruktion von mentalen Repräsentationen aus Texten eine Rolle zu spielen: Schnotz (2005) geht davon aus, dass das Vorwissen zunächst für den Aufbau der mentalen Repräsentationen notwendig ist, aber auch einen Informationsmangel, fehlende Kapazität im Arbeitsgedächtnis sowie Defizite in den propositionalen Repräsentationen ausgleichen kann (Schnotz, 2005).

Ein zentraler Unterschied zwischen visuell und auditiv dargebotenen, verbalen externen Repräsentationen findet sich bei der Informationsaufnahme: Während visuell, also schriftlich dargebotene Texte meist in einem individuellen Tempo rezipiert werden können, wird die Informationsaufnahme bei auditiven, also mündlich dargebotenen Texten meist durch den bzw. die Sprecher:in gesteuert. Werden Informationen bei auditiver Darbietung nicht sofort von den Zuhörer:innen aufgenommen und verarbeitet, ist ein Nachholen aufgrund der Flüchtigkeit nicht mehr möglich, sodass eine kontinuierliche Aufmerksamkeit von den Zuhörer:innen gefordert ist (Schnotz, 2001).

Teil von biologischen Texten – also verbalen externen Repräsentationen – sind auch Formeln (z. B. Reaktionsgleichungen, Strukturformeln für Elemente), Zeichen und Symbole (z. B. Darstellung von Männlichkeit und Weiblichkeit in Stammbäumen; Elemente von Blütendiagrammen) und Abkürzungen (CO₂, MRT, ADP). Nitz et al. (2012) bezeichnen diese Form von Repräsentationen als symbolische Repräsentationen. Formeln, Zeichen, Symbole und Abkürzungen sind aufgrund ihrer Kontext- bzw. Fachspezifität Schüler:innen insgesamt wenig geläufig und können damit Schwierigkeiten bei der Interpretation und Aufbau mentaler Modelle darstellen.

Der sprachlichen Verständlichkeit eines Textes bzw. dem Sprachverstehen auf Wort-, Satz- und Textebene kommt große Bedeutung zu und dies wird daher gesondert im hierauf folgenden Kapitel (Kap. 4.4) aufgegriffen.⁶⁸

Eine besondere Form der externen, verbalen Repräsentationen bilden Analogien und Metaphern. Als Form sprachlicher Bilder werden sie in diesem Kapitel aufgegriffen und auch im Folgenden unter externe Repräsentationen gefasst. Dabei ist eine Unterscheidung von Analogien und Metaphern viel

⁶⁷ Im integrativen Modell des Text- und Bildverständnisses nach Schnotz (2005) werden die eingegangenen Informationen in Form von propositionalen Repräsentationen im Arbeitsgedächtnis weiterverarbeitet. Die propositionalen Repräsentationen bestehen aus einer geringen Anzahl an Propositionen, wobei Propositionen – analog zum Begriff in der Satzsemantik – die kleinsten Informationseinheiten sind, die als wahr oder falsch eingeordnet werden können. Es handelt sich dabei um abstrakte Wissenseinheiten.

⁶⁸ Die genannten Formen externer Repräsentationen (realistische Bilder, logische Bilder, verbale Repräsentationen) sind Bestandteile der naturwissenschaftlichen bzw. biologischen Fachsprache (Kulgemeyer & Tomczyszyn, 2015; Nitz, Nerdel & Prechtel, 2012).

diskutiert (Aubusson, Harrison & Ritchie, 2006; Lakoff & Johnson, 2011). Beiden sprachlichen Phänomenen ist gemein, dass bei deren Verwendung die unmittelbaren, konkreten Erfahrungen aus der Lebenswelt der Adressat:innen (= Herkunftsbereich) und deren aufgebaute Wissensbestände zunutze gemacht werden, um unbekannte, abstrakte Konzepte (= Übertragungsbereich/Zielbereich) fassbar und vorstellbar zu machen. Zum Beispiel analysiert Ohlhoff (2002) die Metaphorik zur Darstellungen der Immunreaktion: Die sprachlichen Beschreibungen beruhen dabei auf den Konzepten des Kampfes und der Verteidigung gegen fremde Eindringlinge (= Herkunftsbereich). Deutlich wird dies in Ausdrücken wie dem „Eindringen von Fremdkörpern“ in ein scheinbar abgrenzbares Gebiet (Ohlhoff, 2002, S. 90). Glynn (1994) fokussiert sich in seiner Arbeit hingegen auf die Verwendung von Analogien im naturwissenschaftlichen Unterricht und nennt als Beispiel die nutzbare Analogie zwischen einer Kamera (= Herkunftsbereich) und dem menschlichen Auge (= Zielbereich) in Hinblick auf den Aufbau. Einer der zahlreichen Unterschiede, die für den Versuch einer trennscharfen Abgrenzung zwischen Analogien und Metaphern beschrieben wurden (siehe Aubusson et al., 2006), bezieht sich auf die Bewusstheit der Beziehung zwischen Herkunftsbereich und Übertragungsbereich: Die Ähnlichkeiten und Unterschiede zwischen Herkunftsbereich und Zielbereich sind bei Metaphern unbewusst, wohingegen sie bei Analogien explizit gemacht werden können. Insgesamt wird Metaphern daraus folgernd eine eher unbewusste Verwendung, Analogien hingegen eher eine bewusste Verwendung zugeschrieben (Aubusson et al., 2006).⁶⁹

Im Diskurs um gutes Erklären wird hauptsächlich auf Analogien und ihr Potenzial für erfolgreiches Lernen eingegangen (Geelan, 2013; Kiel, 1999; Kulgemeyer & Tomczyszyn, 2015; Sevia & Gonsalves, 2008; Wragg, 1993). Sie werden als Kennzeichen effektiver Erklärungen beschrieben (Geelan, 2013; Treagust & Harrison, 2000; Wragg, 1993) und in Kriterienkatalogen zur Erklärbarkeit aufgenommen (Kulgemeyer & Tomczyszyn, 2015; Sevia & Gonsalves, 2008). Glynn (1994) beschreibt, dass die Qualität der Analogie an sich umso höher ist, je mehr Übereinstimmungen und Übertragungsmöglichkeiten es gibt – die Analogie zwischen Auge und Kamera wird von ihm deshalb als sehr gewinnbringend angesehen. Dabei sollte bei der Nutzung von Analogien sowohl auf die übertragbaren Bereiche von Herkunftsbereich und Übertragungsbereich als auch explizit auf die Grenzen dieser Übertragung eingegangen werden (Glynn, 1994). Dadurch kann unangemessenen Interpretationen und Übertragungen begegnet werden, die aufseiten der Schüler:innen möglich sind (Aubusson et al., 2006; Glynn, 1994). Geelan (2013) stellt bei seinen Unterrichtsbeobachtungen jedoch fest, dass dieser Schritt nur von wenigen Lehrkräften vollzogen wurde, wohingegen Treagust und Harrison (2000) bei der Analyse von Richard Feynmans Lernmaterial ‚Atoms in motion‘ (Feynman, 1994) häufig den Hinweis auf Limitationen feststellen. Insgesamt muss der Herkunftsbereich den Lernenden vertraut und bekannt sein, um lernwirksam sein zu können (Treagust & Harrison, 2000).

Metaphern finden – vielleicht aufgrund ihrer unbewussten Nutzung – kaum Berücksichtigung (Ausnahme bei Treagust und Harrison (2000)). Für einen reflektierten Umgang mit Metaphern im Biologieunterricht und eine gezielte Bewusstmachung der transportierten Bilder empfiehlt Langlet (2004), Metaphern auf deren Ursprung zurückzuführen, sie weiter bis ans Ende zu denken oder sie zu zeichnen.

Insgesamt wird die gemeinsame Verwendung von Texten und Bildern als sinnvoll beschrieben, da sie sich beim Aufbau mentaler Repräsentationen unterstützen können (Schnotz & Bannert, 2003).

⁶⁹ Da die Verwendung von Analogien und Metaphern in instruktionalen Erklärungen des Biologieunterrichts in dieser Studie nicht weiter analysiert werden, soll diese Unterscheidung an dieser Stelle genügen.

Begründet werden kann dies beispielsweise im Gedächtnismodell von Baddeley und Graham Hitch (1974), innerhalb dessen die Verarbeitung von visuellen und verbalen Informationen in unterschiedlichen Bereichen des Arbeitsgedächtnisses angenommen wird (visuell-räumlicher Notizblock und phonologische Schleife⁷⁰). In ähnlicher Weise nimmt Paivio (1986) eine getrennte Verarbeitung visueller und auditiv-verbaler Reize an (Dual-Coding-Theory). Da in beiden Bereichen eigene Ressourcen bei der Verarbeitung bereitstehen, können von den Rezipient:innen mehr Informationen gleichzeitig im Arbeitsgedächtnis präsent gehalten und verarbeitet werden. Höhere Lerneffekte werden jedoch nur erzielt, wenn sich visuelle und auditive Repräsentationen ergänzen und gleichzeitig präsentiert werden. Dabei sollte die Länge des Textes kurzgehalten werden und die Komplexität gering ausfallen, um die Integrationsleistung erbringen zu können. Die realistischen oder grafischen Bilder sollten ebenso so einfach wie möglich gestaltet sein (Sweller, 2011).

Werden externe Repräsentationen in einem visuellen Modus dargestellt, werden sie in pädagogisch-didaktischen Texten häufig als Visualisierungen bezeichnet (Wagner & Wörn, 2011). Dabei wird ihnen neben einer lernunterstützenden Funktion weiterhin eine aufmerksamkeitslenkende und den Vortrag strukturierende Funktion zugeschrieben (Hargie, 2013). Visualisierungen können aus einzelnen externen Repräsentationen bestehen oder sich aus mehreren Formen externer Repräsentationen zusammensetzen. So kann ein Prozessablauf mit Skizzen, Pfeilen und Stichworten bzw. Wortkarten dargestellt werden. Sie kommen häufig im Biologieunterricht vor und sind insbesondere bei biologischen Prozessabläufen hochkomplex, da sie verschiedene Organisations- und Systemebenen unter Nutzung vieler Symbole darstellen können (Kattmann, 2008b).

Insgesamt werden Visualisierungen in den rezipierten Texten als Merkmale effektiver Erklärungen auf empirischer Basis identifiziert (Sevian & Gonsalves, 2008; Wragg, 1993) bzw. bei Befragungen von Expert:innen als Merkmal guter Erklärungen angeführt (Schopf & Zwischenbrugger, 2015). Bei Kulgemeyer und Schecker (2013) sowie Kulgemeyer und Tomczyszyn (2015) beinhaltet das Modell zur Erklärens-fähigkeit die Nutzung von grafischen Darstellungsformen (neben Beispielen und Analogien), um adressatengemäß zu erklären.

Die eingesetzten Visualisierungen müssen dabei für die Rezipient:innen sichtbar und verständlich sein. Während die Sichtbarkeit durch eine ausreichende Größe oder Vergrößerung, die Platzierung im Raum im Verhältnis zu den Schüler:innen sowie die Position der Lehrkraft im Raum gewährleistet werden kann, betrifft die Verständlichkeit von Visualisierungen konkrete Gestaltungsmerkmale. Bezüglich konkreter Gestaltungsempfehlungen wird angeführt, dass die Visualisierungen sowohl den Adressat:innen als auch dem Gegenstand gerecht werden müssen (Kiel, 1999; Wagner & Wörn, 2011), wobei in Bezug auf die Adressatenorientierung insbesondere der Abstraktionsgrad der gewählten Veranschaulichung als relevant hervorgehoben wird (Wagner & Wörn, 2011). Weiterhin sollten die visuell dargebotenen Informationen gut in die verbale Erklärung integriert werden (Kulgemeyer & Tomczyszyn, 2015; Schopf & Zwischenbrugger, 2015). Außerdem wird die Nutzung verschiedener Visualisierungen als vorteilhaft beschrieben, wobei die Verbindungen zwischen ihnen sowie Vor- und Nachteile dieser klar herausgestellt werden sollen (Kulgemeyer & Schecker, 2013; Leinhardt, 1987). Ein interessanter Punkt – gerade in Hinblick auf die Angemessenheit der Visualisierungen gegenüber dem fachlichen Gegenstand – findet sich bei Kattmann (2008b):

⁷⁰ Der visuell-räumliche Notizblock sowie die phonologische Schleife sind Bezeichnungen für Substrukturen in einigen Modellen des Arbeitsgedächtnisses. Der visuell-räumliche Notizblock (auch Notiztafel) speichert visuell und räumlich aufgenommene Informationen, die phonologische Schleife speichert sprachliche Informationen in Form von Lauten (also noch nicht den propositionalen Repräsentationen).

Bei allen Vorteilen, die Medien für den Erkenntnis- und Lernprozess haben, sollte beachtet werden, dass ihr Einsatz eine Methodenentscheidung ist, bei der bestimmte Vorentscheidungen bereits getroffen wurden: Mit der didaktischen Aufbereitung dieser Medien werden gegenüber der Realität Veränderungen vorgenommen und dadurch Vorleistungen für die Erkenntnisgewinnung erbracht. Anscheinend Unwesentliches wird abgetrennt, anscheinend Wichtiges in der bestmöglichen Perspektive dargestellt. Was in der Natur oft nicht gleichzeitig vorkommt (z. B. Blüten und Früchte, Sommer- und Winterpelz), wird in einem Medium im nächsten Augenblick gezeigt. Dabei kann ein verfälschtes Bild entstehen: Alles erscheint den Lernenden als wunderbar geordnet und ist doch in der Natur so nicht zu beobachten. (Kattmann, 2008b, 295)

Vor allem die Komplexität biologischer Phänomene stellt Lehrende vor die Aufgabe, die zeitgleich ablaufenden, ineinander verwobenen Prozesse in eine geordnete Darstellung zu bringen, die es in der Natur so nicht gibt (siehe auch Kapitel 3.4). Die zunächst einfach klingende Forderung nach einer Angemessenheit der Visualisierungen gegenüber dem fachlichen Gegenstand stellt sich damit bei näherer Betrachtung als facettenreicher und mitunter schwieriger dar.

Gestaltungsprämissen für verständliche Visualisierungen können aus der ‚Cognitive Load Theory‘ bzw. den darin enthaltenen Prinzipien der Verarbeitung von Visualisierungen abgeleitet werden: Text und Bild sollten gleichzeitig präsentiert werden, Bilder und grafische Elemente sollten so einfach wie möglich gehalten werden, die Visualisierungen sollten in kleinere Segmente geteilt und nacheinander präsentiert werden, die Blickrichtung der Betrachtenden sollte durch die Visualisierungen geführt sein (dabei sollte unter anderem die Leserichtung von links nach rechts beachtet werden), Beschriftungen von grafischen Elementen sollten in räumlicher Nähe platziert werden, wesentliche Informationen sollten hervorgehoben werden, redundante Informationen in Text und Bild sollten vermieden werden (Bay, Thiede & Wirtz, 2016; Mayer & Moreno, 2003).

Veranschaulichende Wirkung und Bedeutung für das Lernen im Biologieunterricht werden auch realen Objekten zugeschrieben. Gehlhaar (2008) fasst zusammen, dass sie als sogenannte Primärerfahrungen – also unvermittelte, direkte Begegnungen – gerade beim unterrichtlichen Einsatz von lebenden Tieren und Pflanzen Interesse und Motivation der Schüler:innen steigern und alle Sinne ansprechen können. Damit bieten sie die Merkmale der Objekte in einer anderen Qualität dar, als dies externe Repräsentationen tun (beispielsweise sind Größe und Gewicht, Geruch und Farbe direkt repräsentiert). Die realen, mitunter lebenden Objekte weisen jedoch eine hohe Komplexität auf, die von der Lehrkraft nicht wie beim Einsatz von Bildern reduziert werden kann.

Häufig wird auch die Nutzung von Beispielen zusammen mit Repräsentationen oder Visualisierungen genannt (Schopf & Zwischenbrugger, 2015; Sevia & Gonsalves, 2008; Pauli, 2015; Wellenreuther, 2013). Sie werden in Modellen zur Einteilung von Repräsentationen oder deren Verarbeitung durch Rezipient:innen aus dem Bereich des Multimedia Learning nicht explizit berücksichtigt (Schnotz, 2003). Im Kontext der Überlegungen zur Qualität von Erklärungen werden Beispiele als Möglichkeit der Veranschaulichung – ähnlich wie bei Bildern oder Metaphern, Analogien und Vergleichen – betrachtet und als Element guter bzw. effektiver Erklärungen geführt (z. B. Brown, 2006; Kiel, 1999; Kulgemeyer & Schecker, 2013; Kulgemeyer & Tomczyszyn, 2015; Leinhardt, 1987; Schopf & Zwischenbrugger, 2015). Sevia & Gonsalves (2008). Beispiele können dabei das eher abstrakte allgemeine Grundprinzip konkretisieren und so für die Adressat:innen greifbar machen (Schopf & Zwischenbrugger, 2015). Sevia und Gonsalves (2008) nutzen in diesem Zusammenhang auch den **Begriff der „mental images“** (mentale Bilder), die durch Beispiele bei den Lernenden erzeugt werden können (Sevia & Gonsalves, 2008, S. 1453).

Bei Wragg (1993) werden Beispiele als ein Merkmal von Erklärungen effektiver Lehrkräfte in den Fächern Biologie und Deutsch genannt. Kulgemeyer und Schecker (2013) stellten für den Physikunterricht fest, dass das Geben von Beispielen signifikant positiv mit der Erklärens-fähigkeit zusammenhängt. Auch aus der Sicht von Lehrenden scheint der Einbezug von Beispielen ein relevantes Merkmal zu sein, da alle der von Schopf und Zwischenbrugger (2015) befragten Lehrenden des Wirtschaftspädagogiklehrstuhls auf die Nutzung von Beispielen in Erklärungen als Element verwiesen. Dabei unterschieden diese zwischen Beispielen aus der Lebenswelt der Schüler:innen und jenen aus der Praxis der eigenen Disziplin (Schopf & Zwischenbrugger, 2015).

Für die Gestaltung der Beispiele formulierten die in der Studie befragten Lehrenden Kriterien. Sie sollten:

- Passend sein, um das Grundprinzip verdeutlichen zu können
- Typische Fälle aus der Disziplin und zu Beginn weniger Randphänomene aufgreifen
- Deckend sein, um das Grundprinzip vollständig abzubilden
- Konkrete Fälle heranziehen (z. B. „**Unternehmen, Produkte, Zahlen etc.**“ (Schopf & Zwischenbrugger, 2015, S. 15)
- Realitätsnah und den Schüler:innen nah (im Sinne von vorstellbar, bei Bedarf vereinfacht und an Vorwissen anknüpfend) sein

Relevant (Odora, 2014) und interessant (Odora, 2014; Treagust & Harrison, 2000) sind weitere Charakteristika, mit denen geeignete Beispiele beschrieben werden. Neben der Funktion der Veranschaulichung wird ihnen dabei auch bei der Motivation und Erzeugung von Interesse eine Rolle zugeschrieben (Treagust & Harrison, 2000).

Zusammenfassend gilt die Nutzung von externen Repräsentationen als Facette guter instruktionaler Erklärungen, da diese eine veranschaulichende Funktion erfüllen und damit das Verstehen des Gesagten erhöhen können. Die Auswahl der geeigneten Repräsentationsform – also verbale Repräsentationen wie Analogien, Metaphern oder Beispiele; realistische oder logische Bilder wie Fotografien, Skizzen oder Diagramme – sowie deren Ausgestaltung für die Adressat:innen oder die Auswahl originaler Objekte müssen dabei in Passung zum jeweiligen Erkläregegenstand und zu den Voraussetzungen der Lernenden erfolgen.

4.4 Verständliche Sprache

Sprache in ihrer mündlichen und schriftlichen Realisierung wird als das zentrale Medium beschrieben, Informationen zu vermitteln (Morek & Heller, 2012; Wellenreuther, 2013). In verschiedenen Kommunikationssituationen werden jedoch unterschiedliche Varianten von Sprache (sogenannte Sprachregister) genutzt, die sich hinsichtlich sprachlicher Merkmale differenzieren lassen. So unterscheiden sich beispielsweise Wortschatz, Satzbau und Struktur eines Gesprächs zwischen Schüler:innen am Pausenhof, zwischen Lehrkräften und Schüler:innen im Unterricht sowie zwischen Wissenschaftler:innen auf Tagungen und Konferenzen. Häufig werden diese oder ähnliche Beispiele für Kommunikationssituationen mit den Sprachregistern ‚Alltagssprache‘, ‚Bildungssprache‘ und ‚Fachsprache‘ überschrieben.⁷¹ Schüler:innen rezipieren und produzieren dabei am häufigsten das Sprachregister der Alltagssprache, das im täglichen Umgang genutzt wird,

⁷¹ Für die Bezeichnung von Sprachregistern findet sich eine Vielzahl von Begriffen: Alltagssprache, Schulsprache, Unterrichtssprache, Bildungssprache, Fachsprache (Gogolin & Michel, 2010; Leisen, 2010; Rincke, 2010; Schmörlzer-Eibinger, 2013). Auf eine Diskussion der Begriffe soll an dieser Stelle verzichtet werden.

sodass dessen Wortschatz, Satzbau und Struktur am geläufigsten sind. Demgegenüber wird bei der Kommunikation von Fachinhalten im Unterricht auf ein eher formalisiertes, fachlich orientiertes Sprachregister (Bildungssprache bzw. mit Fokus auf den Fachunterricht das Sprachregister der Fachsprache) mit charakteristischen Begriffen und Sprachmustern zurückgegriffen (Morek & Heller, 2012). Dieses ist den Schüler:innen weniger geläufig, sodass die Sprache auch eine Verstehenshürde sein kann.

Um Merkmale einer verständlichen Sprache in instruktionalen Erklärungen des Biologieunterrichts zu skizzieren, sind daher zwei Diskurse zu berücksichtigen: Zum einen sind dies allgemeine, sprachwissenschaftliche Untersuchungsergebnisse und Überlegungen zu einer verständlichen Sprache. Beispielsweise fasst Schnotz (1994) Ergebnisse von Studien zur Textverständlichkeit zusammen und führt die Geläufigkeit der Wörter, die Komplexität der Syntax, die lokale und globale Kohärenz sowie Kohärenzbildungshilfen an.⁷² Zum anderen handelt es sich um biologiedidaktische bzw. naturwissenschaftsdidaktische Ergebnisse und Überlegungen zu Merkmalen der unterrichtlichen Fachsprache und deren Verstehenshürden.⁷³ Der Begriff der Fachsprache bezieht sich im naturwissenschaftsdidaktischen Diskurs dabei meist nicht nur auf die verbale Sprache – also gesprochenen oder geschriebenen Text –, sondern auch auf Mathematisierungen, Symbole und grafische Darstellungsformen (z. B. Skizzen, Versuchsaufbauten, Flussdiagramme, Balkendiagramme, Tabellen) sowie deren Anordnung im Text oder Vortrag (Kulgemeyer & Tomczyszyn, 2015; Ziepprecht, Jäger & Schwanewedel, 2017). Da die genannten Elemente abgesehen von der verbalen Sprache bereits unter dem Punkt der Repräsentationen beleuchtet wurden, sollen hier die verbale Sprache und deren Verständlichkeit im Fokus stehen.

Vergleicht man Fachsprache und Alltagssprache hinsichtlich ihrer Charakteristika, können unter anderem auf Wort- und Satzebene Merkmale beschrieben werden, die der Fachsprache eigen sind, sich aber in der Alltagssprache nur selten finden lassen. Systematische Analysen zu Charakteristika der Bildungs- bzw. Fachsprache im naturwissenschaftlichen Unterricht beziehen sich meist auf schriftliche Realisierungen von Sprache (siehe Leisen, 2010; Rincke, 2010; Ziepprecht et al., 2017). Auch wenn der Fokus hier auf mündlichem Erklären liegt, werden die Analyseergebnisse zunächst zur Beschreibung der Charakteristika von Fachsprache und potenziellen Verstehenshürden herangezogen, da sich die mündliche Fachsprache durch eine starke Orientierung an der schriftlichen Konzeption⁷⁴ – also deren Merkmalen – auszeichnet.

Auf Ebene einzelner Worte⁷⁵ weist die Fachsprache unter anderem eine Häufung von substantivierten Infinitiven (z. B. das Beobachten, das Skizzieren, das Messen), von Adjektiven auf -bar, -los, -reich, -arm, -fest (z. B. nährstoffarm, energiereich, sequenzierbar) von Adjektiven mit Präfix (vererbbar, multifaktoriell), von mehrgliedrigen Komposita (z. B. Lichtreaktion,

⁷² Neben diesen textuellen Merkmalen, die Prädiktoren für das Textverstehen sein können, gibt es auch personenbezogene Prädiktoren (z. B. das Interesse am Thema, die Größe des Arbeitsgedächtnisses, das Vorwissen zum Inhalt sowie zum verwendeten Wortschatz sowie Verarbeitungsstrategien; Friedrich, 2017). Diese sollen an dieser Stelle nicht ausgeführt werden, da sie durch das Kriterium der Adressatenorientierung bereits repräsentiert sind.

⁷³ Da die Fachsprache eben solche Verstehenshürden erzeugen kann, die es im Unterricht zu überwinden gilt, wird die Sprache im naturwissenschaftlichen Diskurs nicht nur als Medium betrachtet, sondern auch als eigener Lerngegenstand, der in den Unterricht integriert und explizit thematisiert werden muss (Gogolin & Michel, 2010; Rincke, 2010).

⁷⁴ Koch und Oesterreicher (1985) unterscheiden in ihrem Modell zur Beschreibung von mündlicher und schriftlicher Sprache hinsichtlich zweier Dimensionen: dem Medium (grafischer Kode (schriftlich)/phonischer Kode (mündlich)) sowie der Konzeption (auftretende Charakteristika des Mündlichen oder Schriftlichen in der sprachlichen Äußerung bzw. dem Text). So kann eine Äußerung zwar medial mündlich gegeben werden, jedoch konzeptionelle Schriftlichkeit aufweisen, wie dies beispielsweise in öffentlichen Reden oder Vorträgen der Fall ist.

⁷⁵ In der Biologie- bzw. Naturwissenschaftsdidaktik wird häufiger von ‚Fachbegriffen‘ als von ‚Fachwörtern‘ gesprochen (Schmiemann, 2011). Damit soll darauf verwiesen werden, „dass hinter den mehr oder weniger austauschbaren Wörtern Bedeutungen, also Begriffe, stehen“ (Schmiemann, 2011, S. 117).

Umweltbedingungen, Nahrungsmittelverfügbarkeit), von Mehrwortkomplexen (z. B. der in den Epithelzellen abgesonderte zähe Schleim; die steigende Bindungsaffinität des Hämoglobin-Moleküls), von Wortbildungen mit Eigennamen (z. B. Bergmann-Regel, Erlenmeyerkolben) oder von fachspezifischen Akronymen (z. B. CG-Dimere, ATPase; Rincke, 2010) auf.

Das hier bereits mehrfach verwendete Konzept der ‚Geläufigkeit‘ wird im sprachwissenschaftlichen Diskurs von Schnotz (1994) als textbezogener Prädiktor für das Textverstehen auf Wortebene verwendet. Dabei ist die Geläufigkeit der Wörter ein Indikator dafür, wie häufig die Wörter durchschnittlich in der Sprache vorkommen. Durch eine hohe Geläufigkeit kann durch die Rezipient:innen der sprachlichen Äußerung eine leichtere Zuordnung einer Bedeutung erfolgen – das Textverstehen ist dementsprechend hoch. Die dahinterstehende Leitfrage (‚Wie geläufig ist das jeweilige Wort?‘) ist dabei aus der Sicht der Rezipient:innen zu beantworten.

Verbindet man die Überlegungen aus der Naturwissenschaftsdidaktik und der Sprachwissenschaft, so können die im Text von Rincke (2010) beschriebenen Merkmale auf Wortebene als insgesamt wenig geläufig für Schüler:innen bezeichnet werden, da sie selten in der Alltagssprache zu finden sind. Das individuelle Vorwissen von Schüler:innen ist hier zu berücksichtigen.

In der Literatur zu Kriterien guten bzw. effektiven Erklärens wird in Bezug auf die Wortebene gefordert, während der Erklärung unnötige Fachbegriffe zu meiden (Schopf & Zwischenbrugger, 2015) sowie auf für die Schüler:innen klar definierte Begriffe zurückzugreifen oder entsprechend zu umschreiben (Brown, 2006; Hargie, 2013). Bei den Untersuchungen von Seviau und Gonsalves (2008) sowie Kulgemeyer und Tomczyszyn (2015) wird die Verwendung klar definierter Begrifflichkeiten auch in das Messinstrument zur Beschreibung von Erklärens-fähigkeit aufgenommen. Des Weiteren geben die von Schopf und Zwischenbrugger (2015) befragten Lehrenden an, dass Abkürzungen Erklärungen unnötig abstrakt machen und daher nicht verwendet werden sollten. Insgesamt sollte der verwendete Wortschatz den Schüler:innen bekannt sein bzw. bekannt gemacht werden, um ein Verstehen initiieren zu können (Kulgemeyer & Schecker, 2013; Kulgemeyer & Tomczyszyn, 2015; Schopf & Zwischenbrugger, 2015; Seviau & Gonsalves, 2008; Wagner & Wörn, 2011).

Auf Satzebene werden in der Fachsprache deutlich häufiger als in der Alltagssprache Funktionsverbgefüge (z. B. in Zusammenhang stehen, eine Beobachtung anstellen), Nominalisierungsgruppen (z. B. das Ablesen der Populationsgröße, die Aufnahme der Vitamine), Satzglieder anstelle von Gliedsätzen (z. B. nach dem Kauen der Nahrung) und komplexe Attribute statt Attributsätze (z. B. die im Rahmen der primären Immunreaktion aktivierten T-Helferzellen) verwendet (Rincke, 2010). Auch diese sprachlichen Merkmale treten vor allem in der mündlichen Alltagssprache kaum auf und können so Verständnisschwierigkeiten für Schüler:innen darstellen. Hinzu kommt die unpersönliche Formulierung als stilistisches Merkmal (Rincke, 2010).

Aus sprachwissenschaftlicher Perspektive kommt das Kriterium der Komplexität der Syntax hinzu, die das Textverstehen hemmen kann (Schnotz, 1994).

Die Komplexität der Syntax [...] gibt an, wie viele verschiedene syntaktische Konstruktionen die Sätze enthalten, wie stark sie ineinander verschachtelt sind und wie weit die Referenten und die auf sie referierenden Teile im Satz auseinanderstehen. Eine hohe Komplexität entsteht z. B. durch Parenthesen (Einschübe), unnötige Passiv-Konstruktionen, verschachtelte Satzkonstruktionen und sogenannte „Satzgirlanden“, bei denen die Hauptsätze durch mehrere Nebensätze unterbrochen werden. (Friedrich, 2017, S. 67)

Im Diskurs um gutes Erklären werden für die Ebene der Syntax kaum konkrete Empfehlungen auf Basis der sprachwissenschaftlichen oder didaktischen Überlegungen ausgesprochen oder Forderungen für sprachlich verständliche Erklärungen auf Satzebene gestellt. Es findet sich lediglich die Nutzung eines einfachen Satzbaus (Kiel, 1999; Schopf & Zwischenbrugger, 2015). Ein solcher würde in Anlehnung an die dargestellten Merkmale der Fachsprache und die mögliche Komplexität der Syntax insbesondere Folgendes bedeuten: die Vermeidung hypotaktischer Satzstrukturen zugunsten von kürzeren Sätzen mit weniger Einschüben und Nebensätzen sowie die Bevorzugung von Attributsätzen im Gegensatz zu komplexen Attributen.

Darüber hinaus und mit Blick auf den Inhalt der Sätze wird von einigen Autor:innen darauf verwiesen, dass Aussagen sprachlich nicht vage, sondern eindeutig und präzise (Brown, 2006; Kiel, 1999; Odora, 2014; Sevian & Gonsalves, 2008; Wagner & Wörn, 2011) sowie allgemein klar sein sollen (Leinhardt, 1987; Schopf & Zwischenbrugger, 2015). Wagner und Wörn (2011) führen dazu aus, dass sprachliche Vagheit (im Original als „sprachliche Unschärfen“ bezeichnet (Wagner & Wörn, 2011, S. 29)) die sachliche Richtigkeit beeinträchtigen kann, da sie Raum für falsche Vorstellungen lässt. Hargie (2013, S. 278) zählt Faktoren auf und nennt entsprechende Beispiele, aus denen Unbestimmtheit und damit Vagheit resultieren können, die bei Erklärungen vermieden werden sollen. Zum Beispiel:

- Unklare Bezeichnungen (*Dinge dieser Art, dies alles*)
- Nichtdefinierte Vergleiche (*Unsere Zahlen sprechen für ein deutliches Wachstum*)
- Negative Verstärkungspartikel (*war kaum mehr als, nicht selten*)
- Näherungen (*ungefähr so viele wie, fast jeder, beinahe*)
- Unbestimmte Zahlenangaben (*ein paar, eine ziemlich große Anzahl*)
- Begriffsgruppen (*Aspekte, Faktoren, Dinge*)
- Möglichkeit und Wahrscheinlichkeit (*sind nicht unbedingt, es könnte sein, dass ..., wahrscheinlich*)

Werden im Unterricht komplexe Phänomene aufgegriffen, die sich als wenig eindeutig oder unzureichend empirisch untersucht darstellen (siehe Kapitel 3), sind auch im sprachlichen Ausdruck Unbestimmtheit und Vagheit enthalten, die nicht reduzierbar sind. Dies betrifft insbesondere die Formulierungen, die von Hargie (2013, S. 278) mit „Möglichkeit und Wahrscheinlichkeit“ überschriebenen sind.

Auf Textebene kann das Textverstehen durch eine hohe lokale und globale Kohärenz verbessert werden, insbesondere dann, wenn die Rezipient:innen über wenig Vorwissen verfügen oder nur wenig Motivation mitbringen (Schnotz, 1994). Kohärenz bezeichnet dabei allgemein die Verknüpfung von Aussagen sowie deren Widerspruchsfreiheit (Friedrich, 2017; Kintsch & van Dijk, 1978). Lokale Kohärenz kann von Sprecher:innen oder Schreiber:innen zwischen aufeinanderfolgenden Aussagen hergestellt werden. Globale Kohärenz beschreibt hingegen die Kohärenz über den gesamten Text hinweg. Je höher lokale und globale Kohärenz ausfallen, desto besser kann es den Rezipient:innen gelingen, die Verknüpfungen der Aussagen bzw. der Textteile zu entnehmen und Informationen zu entnehmen bzw. den Sinn zu rekonstruieren (Friedrich, 2017).

Um die Kohärenz zu erhöhen, können Kohärenzbildungshilfen eingesetzt werden. Dies sind Hilfestellungen, die die Rezipient:innen zum einen dabei unterstützen, die Relevanz der verschiedenen Inhalte innerhalb der Erklärung zu entnehmen, und zum anderen Hinweise geben, „in welchem Zusammenhang die größeren Sinneinheiten zueinander stehen“ (Friedrich, 2017, S. 68)

und damit strukturgebend sind. Während bei schriftlichen Texten neben Hervorhebungen oder Einrahmungen Wichtigkeit signalisieren können und die Kohärenzbildungshilfen damit keinen wörtlichen Ausdruck finden müssen, sind die Sprecher:innen bei der mündlichen Kommunikation im Wesentlichen auf eine explizite Verbalisierung angewiesen (Schnotz, 1994). So können Sätze wie ‚XY ist besonders wichtig, wenn wir uns nun YX ansehen‘ oder ‚nebenbei bemerkt‘ die Relevanz des Inhalts andeuten und damit gezielt den Aufmerksamkeitsfokus der Zuhörer:innen lenken. Eine verbalisierte Kohärenzbildungshilfe, die die globale Kohärenz erhöht und eine Hilfe bei der Wissensstrukturierung darstellt, ist beispielsweise ‚So viel zu XY, nun kommen wir zu YX‘. Diese Signale können in der mündlichen Kommunikation so weit reduziert werden, dass der Abschluss einer Aussage bzw. eines Abschnitts zum Beispiel nur noch mit einzelnen Worten wie ‚genau‘ oder ‚so‘ markiert wird. Wie bereits in Kapitel 4.2 angesprochen, kann durch diese Einschübe zur Erhöhung der Kohärenz die Strukturiertheit beeinflusst und im Idealfall erhöht werden.

Häufig wird die sprachliche Gestaltung in der Literatur zu gutem bzw. effektivem Erklären recht allgemein beschrieben. So soll ein angemessener Schwierigkeitsgrad der Sprache beachtet werden (Kiel, 1999; Sevan & Gonsalves, 2008; Wagner & Wörn, 2011), wobei keine Zuordnung auf eine sprachliche Ebene (Wort, Satz oder Text) erfolgt. Es zeigt sich, dass eine direkte Verbindung zur Adressatenorientierung besteht. Im Modell von Kulgemeyer und Schecker (2013) sowie Kulgemeyer und Tomczyszyn (2015) wird in Verbindung mit dem Sprachniveau die Variation der Sprachregisters⁷⁶ – insbesondere von einer stark wissenschaftsorientierten Sprache zu einer eher am alltäglichen Wortschatz der Schüler:innen orientierten Sprache – als Qualitätskriterium beschrieben. In Hinblick auf den Schwierigkeitsgrad der Sprache findet sich sowohl in empirischen Texten (Schopf & Zwischenbrugger, 2015) als auch in theoretischen Grundüberlegungen (Wellenreuther, 2013) die allgemeine Forderung nach einer möglichst einfachen Sprache. Dabei ist das Adjektiv ‚einfach‘ für die Beschreibung einer geeigneten Unterrichtssprache mitunter nicht angemessen: Gerade in Hinblick auf die Tatsache, dass Sprache im Fachunterricht auch als Lerngegenstand betrachtet wird bzw. betrachtet werden muss (Gogolin & Michel, 2010; Harren, 2009; Rincke, 2010; entsprechende Vorgaben finden sich hierzu auch in den Bildungsstandards), sind Aussagen, die eine sprachliche Einfachheit fordern, zu reflektieren. Als Teil einer naturwissenschaftlichen Grundbildung leistet die Beherrschung der Fachsprache bzw. als Teil der allgemeinen Grundbildung der Bildungssprache auch einen wichtigen Beitrag zur gesellschaftlichen Partizipationsfähigkeit der Schüler:innen (Gogolin & Michel, 2010; Rincke, 2010). Das bisher vorangestellt Adjektiv ‚verständlich‘ schließt diese Aufgabe der Lehrkräfte nicht aus und trifft darüber hinaus im Kern das Anliegen sprachlicher Mitteilung – nicht ausschließlich von instruktionalen Erklärungen.

An dieser Stelle sollen noch einige Merkmale der medialen Mündlichkeit genannt werden, die Einfluss auf die Verständlichkeit ausüben können. Zunächst ist das Medium des Mündlichen durch seine Flüchtigkeit gekennzeichnet (Koch & Oesterreicher, 1985). Damit einhergehend müssen von den Zuhörer:innen die Informationen so lange im Arbeitsgedächtnis präsent gehalten werden, bis sie mit weiteren Informationen zu einem sinnvollen Ganzen zusammengefügt werden können. Die Abfolge dieser Informationen ist für die Zuhörer:innen dabei jedoch nicht vorab abschätzbar, sodass gewissermaßen ‚gewartet‘ werden muss, bis diese durch den oder die Sprecher:in gegeben werden. Weiterhin muss durch die Flüchtigkeit des Gesagten der Inhalt bereits beim ersten Mal dekodiert

⁷⁶ Im Original verwenden die Autor:innen den Begriff „code“ (Kulgemeyer & Schecker, 2013, S. 2239) bzw. „sprachlicher Code“ (Kulgemeyer & Tomczyszyn, 2015, S. 116). Den Beschreibungen in den Artikeln ist jedoch ein synonyme Gebrauch zum Begriff ‚Sprachregister‘ zu entnehmen.

werden, da ein Zurückspringen und wiederholendes Lesen oder Hören meist nicht gegeben ist (eine Ausnahme bilden hier individuell rezipierte Videos).

Typisch für mediale Mündlichkeit ist – je nach Sprecher:in – zwar mitunter die Reduktion der Komplexität und Variation der Satzstrukturen, jedoch können lange Satzreihungen unter anderem durch die Prozesshaftigkeit der mündlichen Kommunikation und die meist geringe Planungszeit der Sprachproduktion entstehen (Koch & Oesterreicher, 1985). Dies kann dazu führen, dass die für ein Verständnis notwendigen Inhalte nicht alle gleichzeitig im Arbeitsgedächtnis präsent gehalten werden können.

Wie bereits in Kapitel 2.1 sowie 2.3 dargestellt wurde, ist die verbale Kommunikation bei Verfügbarkeit eines direkten Gegenübers meist mit nonverbaler Kommunikation sowie familienähnlichen Aktivitäten wie beispielsweise dem Zeigen und Demonstrieren verwoben (Koch & Oesterreicher, 1985; Stukenbrock, 2009). Diese Aktivitäten können die Verständlichkeit der Inhalte erhöhen.

Insgesamt zeigt sich für das Kriterium der verständlichen Sprache eine Verbindung zu den Kriterien der Adressatenbezogenheit, ohne die die Verständlichkeit der gewählten Sprache nicht eingeschätzt werden kann, sowie zur Strukturiertheit einer instruktionalen Erklärung, die auch Eigenschaft eines Textes ist und durch sprachliche Äußerungen erzeugt werden kann. Dabei sind Merkmale guter instruktionaler Erklärungen auf Ebene der einzelnen Worte (Verwendung bekannter Begriffe oder deren Definition oder Erklärung), der Sätze (Anpassung der syntaktischen Komplexität sowie Vermeidung von langen Satzreihungen) sowie des Textes (hohe lokale und globale Kohärenz z. B. durch die Nutzung von Kohärenzbildungshilfen) zu finden. Diese müssen im Kontext der angestrebten Fachsprache und des Vorwissens der Rezipient:innen reflektiert werden. Das Gesagte sollte dabei so klar und präzise wie möglich formuliert sein.

4.5 Sprech- und Körperausdruck

Zusätzlich zum versprochenen Inhalt, also dem, *was* gesagt wird, vermitteln Sprecher:innen bei mündlicher Kommunikation auch über ihren Sprech- und Körperausdruck, also das, *wie* es gesagt wird, Informationen, die von den Zuhörer:innen aufgenommen werden (Allhoff & Allhoff, 2010). Diese sogenannte nonverbale Kommunikation findet meist unbewusst statt, ist jedoch ein elementarer Teil der Kommunikation auch in schulischen Kontexten (Allhoff & Allhoff, 2010; Babad, 2009).

Der nonverbale Anteil der Kommunikation besteht aus paraverbalen und extraverbalen Mitteln bzw. kann durch diese beschrieben werden (Allhoff & Allhoff, 2010; Neuber, 2016). Mit der Beschreibung und der wissenschaftlichen Untersuchung dessen beschäftigt sich die Sprechwissenschaft.⁷⁷

Unter paraverbalen Mitteln werden artikulatorische und prosodische Merkmale verstanden, die über den auditiven Kanal aufgenommen werden. Sie sind Teil der Stimme und werden deshalb auch als stimmlich-sprecherische Mittel bezeichnet (Neuber, 2016). Der Begriff Prosodie bzw. der prosodischen Merkmale umfasst dabei „Intensität (Lautstärke), Temporalität (Rhythmisierung,

⁷⁷ Die Sprechwissenschaft beschäftigt sich insgesamt mit der Beschreibung und Analyse von Sprechkommunikation sowie der Beeinflussung dieser. Die Sprechsituation wird dabei als ein komplexes Gefüge von „kommunizierenden Menschen („Wer mit Wem“), de[m] Kommunikationsgegenstand („Worüber?“), de[m] Kommunikationsinhalt (Was?) sowie [den] Modalitäten (Wie?)“ (Neuber, 2016, S. 1) angenommen.

durchschnittliche Sprechgeschwindigkeit, Tempowechsel, Pausierung), Melodisierung (Sprehtonhöhe, Sprechtbereich, Melodieintervalle) und Stimmklang (individuelle stimmliche Merkmale sowie unbewusste und bewusste Modifikationen der Stimme)" (Neuber, 2016, S. 135).

Die folgenden Beispiele können die Funktion und Wirkung der genannten Merkmale auf die Kommunikation von Inhalten bzw. das Erklären verdeutlichen. Eine deutliche Artikulation (die sogenannte Artikulationsschärfe) ist zunächst die Basis, um Schüler:innen eine Sinnkonstruktion zu ermöglichen (Allhoff & Allhoff, 2010).

Empirische Daten zum Einfluss des Stimmklangs auf das Lernen von Schüler:innen wurden beispielsweise von Morton und Watson (2001) in einem experimentellen Setting erhoben. Dabei wurden Kindern Audiodateien mit kurzen Passagen vorgespielt. Eine der Gruppen erhielt Aufnahmen, die von einer weiblichen Person mit unauffälliger, gesunder Stimme eingesprochen wurden. Die andere Gruppe bekam hingegen die Aufnahme von einer weiblichen Person, die mit einer als auffällig, heiser oder als rau zu bezeichnenden Stimme spricht. Anschließende Aufgaben zum Erinnerungsabruf einzelner Wörter und zum Ziehen von Schlussfolgerungen aus dem Gesagten wurden von denjenigen, die den Inhalt durch die auffällige, heisere Stimme gehört hatten, deutlich weniger gut bearbeitet. In ähnlicher Weise berichtet Voigt-Zimmermann (2011) von schlechteren Behaltensleistungen mitunter in Folge von Verständnisschwierigkeiten, die von teilnehmenden Kindern geäußert wurden. Besonders hervorzuheben ist, dass die unauffällige, gesunde Stimme häufiger als angenehm beschrieben und ihr lieber zugehört wurde.

Die Sprechgeschwindigkeit wird meist durch Wörter (Silben oder Laute) pro Minute einschließlich der realisierten Sprechpausen angegeben. Die Befunde zum Zusammenhang zwischen Sprechgeschwindigkeit und Verstehensleistung sind eher uneinheitlich (Apel, 2009). Dabei kann zu langsames Sprechen die Sinnerfassung erschweren und zu schnelles Sprechen zusätzlich die akustische Verständlichkeit herabsetzen (Apel, 2009). Werden bei schnellerem Sprechen aussagekräftige Betonungen vorgenommen und ausreichend Pausen nach Sinneinheiten gesetzt, kann die Aufnahme und Behaltensleistung erhalten bleiben (Apel, 2009) oder sogar etwas erhöht werden (Allhoff & Allhoff, 2010). Bei erhöhter Sprechgeschwindigkeit ist außerdem auf eine deutliche Artikulation zu achten (Raithel & Wrede, 2003).

Weiterhin sollten die Sprechpausen bedeutungsunterstützend sein, das heißt, keine zusammenhängenden Sinneinheiten trennen, sondern an linguistischen (sprachlichen) oder prosodischen (sprecherischen/melodischen) Äußerungsgrenzen gesetzt werden (Allhoff & Allhoff, 2010; Sick, 2014). Durch adäquate Sprechpausen wird das Gesagte zusätzlich strukturiert und Zeit geschaffen, in denen das Gesagte von den Zuhörer:innen kognitiv verarbeitet werden kann, wodurch die Verständlichkeit insgesamt gesteigert werden kann (Sick, 2014).

Neuber (2002) legt des Weiteren eine Studie vor, bei der Zusammenhänge zwischen dem Komplex der prosodischen Merkmale und der Behaltensleistung von Zuhörer:innen überprüft werden. Auch hier wurden zwei Gruppen zwei unterschiedliche Audioaufnahmen desselben Textes zur Rezeption vorgelegt. Es handelte sich um ein eingesprochenes Märchen, das einmal mit monotoner Stimme und einmal mit variationsreicher, „eindringlicher“ Stimme (Neuber, 2002, S. 175) vorgelesen wurde. In der eindringlichen Version wurden unter anderem Sprechmelodie, Lautstärke und Tempo sinnhaft variiert. Die Variation zeigte deutliche Unterschiede in den Behaltensleistungen der Zuhörer:innen.

Zusammenfassend können durch stimmlich-specherische Merkmale sowohl die Aufmerksamkeit als auch die Verständlichkeit des Gesagten beeinflusst sowie die Behaltensleistung verbessert oder

verschlechtert werden. Sie haben daher auch eine große Bedeutung für unterrichtliche Kommunikationssituationen einschließlich des instruktionalen Erklärens.

Die extraverbalen Mittel werden über den visuellen Kanal aufgenommen und umfassen allgemein den Körperausdruck.⁷⁸ Der Körperausdruck einer Person in einer bestimmten Sprechsituation kann durch die Kinesik (Körperhaltung und -bewegung, Gestik, Mimik, Blickrichtung und Blickkontakt) und die Proxemik (Verhalten und Bewegung im Raum, Distanzverhalten) beschrieben werden (Allhoff & Allhoff, 2010; Neuber, 2016).

Insgesamt war der Körperausdruck bisher weniger als der Sprechausdruck Teil der sprechwissenschaftlichen Forschung (Neuber, 2016). Berücksichtigung findet er vor allem in der Forschungstradition der ‚teacher immediacy‘ Forschung (Babad, 2009). Unter ‚teacher immediacy‘ werden dabei in den Konzeptualisierungen diejenigen nonverbalen Verhaltensweisen der Lehrkraft berücksichtigt, die zum Abbau physischer oder psycho-sozialer Distanz beitragen und Nähe zwischen Lehrkraft und Schüler:innen schaffen.⁷⁹ Teil der nonverbalen ‚teacher immediacy‘ sind zum Beispiel der Einsatz von Gesten beim Sprechen, der Blickkontakt mit den Schüler:innen, das Lächeln oder die einladende Mimik, eine entspannte Körperhaltung und Bewegungen im Raum (Babad, 2009). Babad (2009) resümiert aus zwei Metastudien zum nonverbalen Verhalten von Lehrkräften (Harris & Rosenthal, 2005; Witt, Wheelless & Allen, 2004):

Both meta-analyses showed that teachers’ NVI [=nonverbal immediacy] was strongly related to many positive student outcomes: Liking for the course and the instructor, willingness to take more classes with the instructor and more classes on that subject, and students’ perceptions that they have learned a lot in the class. (Babad, 2009, S.821)

Unklar bleibt ein empirischer Zusammenhang mit den Lernleistungen der Schüler:innen (Babad, 2009). Neuber (2016) beschreibt, dass der Körperausdruck den Kommunikationsinhalt unterstützen kann, unpassende Gestik und Mimik hingegen dem Gesagten widersprechen und damit das Verstehen hemmen können.

In den Untersuchungen zum Erklären – ob in Hinblick auf gutes oder effektives Erklären – werden das stimmlich-sprecherische und insbesondere das körperausdrucksbezogene Verhalten der Lehrkraft selten und wenig ausführlich besprochen.⁸⁰ Die von Schopf und Zwischenbrugger (2015) befragten Expert:innen erwähnen den Sprech- und Körperausdruck der erklärenden Person nicht als wichtiges Element oder Merkmal guter Erklärungen. In der Untersuchung von Seviau und Gonsalves (2008) wird der Sprech- und Körperausdruck (im Original „**presentation style**“ (Seviau & Gonsalves, 2008, S. 1448); enthält auch die Nutzung von Medien) noch im Kategoriensystem zur allgemeinen Beschreibung von videografierten Erklärungen repräsentiert, tritt dann jedoch im Codiermanual für die Messung von Erklärkompetenz nur in der Dimension zur Nutzung und Umgang mit Medien auf.

⁷⁸ Neuber (2016) verweist darauf, dass häufig auch der Begriff der ‚Körpersprache‘ verwendet wird, die Autor:innen des Sammelbandes jedoch auf diesen verzichten möchten, da der Begriff ‚Sprache‘ im Kompositum auch ein der Sprache ähnliches stabiles und kodifizierbares System von Regeln impliziert. Dieses ist jedoch nicht gegeben.

⁷⁹ Ein konzeptionell sehr ähnliches Konstrukt stellt der ‚teacher enthusiasm‘ dar (Babad, 2009; siehe Kap. 5.2.5).

⁸⁰ Gordon, Druckman, Rozelle und James (2006) verweisen darauf, dass die Bedeutungsinterpretation und damit auch die Angemessenheit der para- und extraverbalen Merkmalen nicht absolut festgelegt werden kann, sondern sich aus der Situation und ihrem Kontext ergibt. So kann beispielsweise die Lautstärke der Umgebung, in der gesprochen wird, die Lautstärke der Stimme beeinflussen oder geteilte Erfahrungen der Interaktionspartner:innen formen die genutzte Gestik. In besonderer Weise spielen bei der Interpretation dieser Merkmale die kulturellen Hintergründe der Personen sowie die Beziehung der Personen in Hinblick auf den zugeschriebenen Status eine Rolle.

Auch Kulgemeyer und Tomczyszyn (2015) modellieren die Erklärens-fähigkeit ohne Berücksichtigung der sprecherischen und körperausdrucksbezogenen Merkmale.⁸¹

In den folgenden Arbeiten im Diskurs um gutes respektive effektives Erklären finden sich Hinweise auf den Aspekt des Sprech- und Körperausdrucks: Auf die Bedeutung der Modulierung der Stimme weisen Hargie (2013), Kennedy (1996) sowie Kiel (1999) hin. Die von Wragg (1993) analysierten instruktionalen Erklärungen, die durch einen Verstehenstest von Schüler:innen als effektiv beurteilt werden können, teilten unter anderem das Merkmal einer variationsreichen Stimme. Er führt die Wirkung dieser Variation auch auf ein gesteigertes Interesse zurück (Wragg, 1993). Weiterhin sollte die Sprechgeschwindigkeit flüssig, jedoch nicht zu schnell sein (Hargie, 2013). Der Verständlichkeit können auch angemessene Pausensetzungen zuträglich sein (Brown, 2006; Hargie, 2013). Hargie (2013) nimmt an, dass die Sprechpausen den Zuhörer:innen ermöglichen, das Gesagte zu strukturieren und zu verarbeiten, sodass sich die Verständlichkeit erhöht. Auch Odora (2014) nimmt das Kriterium der Pausensetzung **unter dem Aspekt „communication skills“** (Odora, 2014, S. 77) in seinen Kriterienkatalog zur Messung der Erklärkompetenz von Techniklehrkräften auf.

Auf die eingesetzte Gestik zur Unterstützung des Gesagten nimmt zum einen Brown (2006) Bezug. Er hält zusammenfassend aus eigenen empirischen Untersuchungen fest, dass Gestik ebenso wie die stimmliche Gestaltung variieren sollte. Zum anderen greift Odora (2014) das Kriterium in seinem Kriterienkatalog zur Feststellung der Erklärkompetenz auf. Ihm folgend trägt das Nutzen von Gestik dazu bei, Punkte zu betonen und zu illustrieren. Hargie (2013) differenziert weiter und beschreibt, dass erfahrene Sprecher:innen auch ihre Gestik und Mimik aktiv für die Gestaltung der Erklärung nutzen: „Insbesondere werden verschiedenste Bewegungen von Augen, Kopf, Gesicht, Fingern, Händen und des ganzen Körpers zweckgerichtet, fokussiert und passend zu der Information **eingesetzt, die betont werden soll**“ (Hargie, 2013, S. 278).

Zusammenfassend kann der Sprech- und Körperausdruck als relevantes Merkmal guter bzw. effektiver Erklärungen (und allgemein für Lehr-Lern-Settings) bezeichnet werden, dessen Ausprägungen Einfluss auf die Aufmerksamkeit, Verständlichkeit, Verstehens- und Behaltensleistung von Schüler:innen ausüben können. Babad (2009) formuliert dazu:

But the NV [(=nonverbal)] is the critical factor in the delivery of teaching, mediating teachers' success in attaining the primary goals of education. Bad teachers most often fail in their NV delivery, whereas excellence in teaching is always characterized by teachers' positive expressive style. (Babad, 2009, S. 819)

4.6 Fachliche Kriterien

Fachliche Kriterien einer guten instruktionalen Erklärung losgelöst vom Fach zu bestimmen, ist nur in wenigen Punkten möglich. In den Texten, die zwar Erklärungen im Fachunterricht, jedoch ohne explizierten Schwerpunkt auf die Besonderheiten des Faches untersuchen, werden deshalb zur Konkretisierung des fachlichen Aspekts lediglich die fachliche Korrektheit sowie die Vollständigkeit genannt. Beide Kriterien sind für fachliches Lernen basal und nur auf den ersten Blick (siehe unten) trivial.

⁸¹ Ebenso unberücksichtigt bleibt der Sprech- und Körperausdruck bei Untersuchungen von Leinhardt (1987) und Geelan (2013) sowie bei theoretisch fundierten Gestaltungsempfehlungen von Wagner und Wörn (2011), Wellenreuther (2013) und Pauli (2015).

In der Studie von Schopf und Zwischenbrugger (2015), in der Hochschuldozierende nach Kriterien guten Erklärens befragt wurden, nannten 16 von 22 Teilnehmenden die fachliche Korrektheit. Sie wurde als „**unabdingbare Voraussetzung für eine gute Erklärung**“ (Schopf & Zwischenbrugger, 2015, S. 18) beschrieben. Die Vermittlung fachlich falscher oder zumindest fachlich nicht angemessener Inhalte steht dem Ziel instruktionaler Erklärungen – der Initiierung eines Wissens-, Verstehens- oder Fähigkeitserwerbs – entgegen.

Wahrscheinlich dadurch bedingt wird das Kriterium der fachlichen Korrektheit in Arbeiten zum guten Erklären eher implizit berücksichtigt. Kulgemeyer und Schecker (2013) unterscheiden in ihrem Modell zur ‚Scientific Communication Competence‘ zwar zwischen einer Orientierung am Inhalt und einer Orientierung an den Adressat:innen, untersuchen jedoch ausschließlich die Ausprägung der Kommunikation mit den Adressat:innen. Sevian und Gonsalves (2008) nehmen das Kriterium der fachlichen Korrektheit bei der Beschreibung der Facetten von Erklärkompetenz mit auf: Als höchstes Level des Fachwissens **wird deshalb die Ausprägung „presents scientifically correct content“** (Sevian & Gonsalves, 2008, S. 1451) aufgeführt. Wittwer und Renkl (2008) beschreiben instruktionale Erklärungen als Möglichkeit, die fachlich nicht angemessenen Vorstellungen von Schüler:innen in korrekte zu überführen, sodass das Kriterium der fachlichen Korrektheit instruktionaler Erklärungen angenommen werden kann.

Zur fachlichen Korrektheit zählt dabei nicht nur die Richtigkeit der Fakten, Zusammenhänge und Prinzipien, sondern auch die Berücksichtigung der adäquaten Erklärungsform (siehe Kap. 2.3.1 zu wissenschaftlichen Erklärungsformen).

Von Wagner und Wörn (2011) wird darüber hinaus die fachliche Vollständigkeit als Kriterium aufgenommen. Sie beschreiben das Kriterium anhand des Vorhandenseins aller notwendigen Verstehenselemente (siehe Kap. 4.2). Die Notwendigkeit der Verstehenselemente bemisst sich den Autor:innen zufolge jedoch am Vorwissen der Adressat:innen, sodass eine Verbindung zum Kriterium der Adressatenorientierung besteht. Unter Vollständigkeit einer instruktionalen Erklärung wird von Findeisen (2017) weiterhin verstanden, dass ergänzende Ausführungen zum Zweck eines Erklärgegenstandes (*Wozu braucht man das Erklärte?*) oder den Hintergründen (*Warum ist das so? Warum wird das so gemacht?*) gegeben werden. Dabei stehen jedoch weniger der Gegenstand an sich bzw. die Frage, ob ein Mechanismus, eine Funktion oder ein Zusammenhang fachlich vollständig mit der Erklärung gefasst wird (im Sinne von ‚alle Merkmale des Gegenstands erfassend‘), sondern das Lernen bzw. das Wissen der Schüler:innen, das mit einer Erklärung aufgebaut werden soll, und der Blick auf den Einsatz dieses Wissens beim selbstständigen Problemlösen in anderen Anwendungssituationen im Fokus.

Eingangs wurde bereits angemerkt, dass beide Kriterien – die fachliche Korrektheit und die fachliche Vollständigkeit – nur auf den ersten Blick trivial erscheinen. In Bezug auf die fachliche Korrektheit muss die Frage gestellt werden, ob sich dieses Kriterium auf das konkret Gesagte und damit die ausgewählten Inhalte der Erklärung bezieht oder das Kriterium der fachlichen Korrektheit auf das natürliche, reale Phänomen an sich bezogen wird. Ersteres sollte in jedem Fall gewährleistet sein. Letzteres – die korrekte Darstellung des natürlichen Phänomens, die diesem auch gerecht wird – reflektiert die Lehrkraft im Rahmen der didaktischen Reduktion. Bezieht man die Überlegungen zur Komplexität biologischer Phänomene und zum gesamtwissenschaftlichen Nichtwissen sowie zum unsicheren Wissen über diese Phänomene mit ein (siehe Kap. 3), so stellt sich die Frage, ob ein Phänomen auch noch korrekt dargestellt wird, wenn so viele Faktoren reduziert werden, dass die Komplexität nicht mehr deutlich in Erscheinung tritt und Nichtwissen sowie unsicheres Wissen keine Erwähnung finden. An dieser Stelle zeigt sich auch die Herausforderung zur Beurteilung der

fachlichen Vollständigkeit: Berücksichtigt man die Ausführungen zur Komplexität, zu gesamtwissenschaftlichem Nichtwissen sowie unsicherem Wissen über diese Phänomene, dann kann (muss) angenommen werden, dass instruktionale Erklärungen die natürlichen Phänomene – insbesondere diejenigen mit hohem Komplexitätsgrad – nicht vollständig abbilden können.

Das Messen fachlicher Vollständigkeit am Vorhandensein aller notwendigen Verstehenselemente für Schüler:innen, wie es Wagner und Wörn (2011) konzeptualisieren, ist deshalb aus fachlicher Sicht nicht gerechtfertigt. Das Fehlen von Hinweisen auf weitere unbekannte Faktoren, unsichere Wirkungszusammenhänge etc. kann nicht von Schüler:innen erkannt werden – es trägt nicht zum unmittelbaren Verstehen eines Ursache-Wirkungs-Zusammenhangs, sondern zu einem korrekten Verständnis über das natürliche Phänomen, das Lehrkräfte verantworten, bei.

Aus fachdidaktischer Sicht geht es hier um ein Abwägen zwischen der Orientierung an allgemeinen Prinzipien und Mechanismen auf der einen Seite und der Orientierung am natürlichen Phänomen auf der anderen Seite. An einem Beispiel illustriert: Es ist die didaktische Entscheidung zu treffen, ob entweder allgemein eine Räuber-Beute-Beziehung in einem Ökosystem thematisiert wird, bei dem das Ökosystem und die gewählten Tierarten eher als Kontext fungieren, um Anschaulichkeit zu erzeugen. Dabei wird der Ausschnitt möglichst klein und spezifisch gewählt, sodass das zu zeigende Prinzip, das modellhaft für weitere Phänomene gelten und schließlich übertragbar sein soll, unter der Wahrung fachlicher Korrektheit deutlich zum Tragen kommt. Die Erklärung bzw. das Phänomen besitzt im Kern eine Modellfunktion. Oder es wird ein konkretes Beispiel aus einem realen Ökosystem betrachtet, bei dem kein Kontext als solcher existiert, sondern ein individuelles, natürliches Phänomen mit möglichst umfassender Betrachtung aller Faktoren sowie dem darüber existierenden Nichtwissen und unsicheren Wissen im Fokus steht.

Als Facette des fachlichen Aspekts guter instruktionaler Erklärungen kann auch die Berücksichtigung der Auswirkung des Gesagten gefasst werden – hier zeigt sich ein weiteres Mal innerhalb des fachlichen Aspekts guter Erklärungen die Verbindung zur Adressatenorientierung, die für instruktionale Erklärungen von hoher Relevanz ist. Von Dittmer und Ehras (im Druck) wird unter Bezug auf die transformatorische Bildungstheorie (Koller, 2007)⁸² die Bedeutung biologischer Erkenntnisse für das Selbst-Weltverhältnis erläutert, wobei der Zusammenhang als wissenschaftspropädeutischer Inhalt formuliert und zugleich als Ausdruck der ethischen Dimension der Biologie gedeutet wird:

Hierauf bezogen greift der *Nature of Bioscience*-Ansatz die bildungsphilosophische Perspektive der transformatorischen Bildungstheorie auf (Koller 2007) und betrachtet den Einfluss biologischer Forschung auf das Selbst- und Weltverhältnis der Akteur*innen, die sich biologisches Wissen aneignen oder als Expert*innen weitergeben. Biologische Erklärungen und Theorien verändern die Art und Weise, wie der Mensch sich selbst, andere Lebewesen sowie die Natur versteht und bewertet. Ein prominentes Beispiel hierfür sind bioethische Debatten über den moralischen Status leidensfähiger Organismen, bei denen auch auf biologische Daten zurückgegriffen wird. In unserem Kulturkreis profitieren Säugetiere davon, dass wir Ihnen evidenzbasiert und aufgrund von Analogieschlüssen Leidensfähigkeit zuschreiben können. Invertebraten scheinen es dagegen schwerer zu haben, da es uns aus biologischer Perspektive

⁸² Gerade in Bezug auf komplexe biologische Phänomene scheinen Irritationen und mitunter Krisen, die in der transformatorischen Bildungstheorie als Ausgangspunkt für Bildung gelten, häufig möglich. Insbesondere Nichtwissen und unsicheres Wissen, die bei vielen komplexen Phänomenen (z. B. dem Klimawandel, der Wirkung von Medikamenten im menschlichen Organismus, den Auswirkungen von Eingriffen in Ökosysteme) besonders hervortreten, scheinen als Kristallisationspunkt für Irritationen und Krisen relevant zu sein.

noch deutlich schwerer fällt, sich vorzustellen, wie es beispielsweise ist, als Insekt oder Muschel in der Welt zu sein. (Dittmer & Ehras, 2024, S. 13)

Damit wird eine Auswirkung des Gesagten auf die Person deutlich, die über eine Veränderung von Wissensbeständen hinausgeht.

Im Kontext der Kommunikation von Nichtwissen und unsicherem Wissen werden ebenfalls Auswirkungen des Gesagten diskutiert (siehe Kap. 3.3.3), die im Unterricht von der Lehrkraft antizipiert bzw. aufgefangen und berücksichtigt werden müssen: Neben empfundener Unsicherheit als Folge eines Mangels an Orientierung wird auch der Vertrauensverlust in die Wissenschaft an sich thematisiert (Bösch et al., 2008; Kampourakis & McCain, 2020; Maier et al., 2018; Varwig, 2020). Auf der anderen Seite können sogar positive Wirkungen hervorgerufen werden: So konnten Retzbach und Maier (2015) in ihrer Studie zeigen, dass Personen, die Zeitungsartikel und Fernsehbeiträge zur Nanotechnologie rezipierten, in denen Unsicherheit explizit thematisiert wurde, keinen nachweisbaren Vertrauensverlust in Wissenschaftler:innen, sondern ein erhöhtes Interesse in Wissenschaft und neue Technologien berichteten.

4.7 Fazit – Anforderungen an eine erklärende Lehrkraft

Die vorhergehenden Kapitel 4.1 bis 4.6 zu den sechs Kategorien bzw. Aspekten – Adressatenorientierung, Strukturiertheit, Einsatz von Repräsentationen, verständliche Sprache, Sprech- und Körperausdruck sowie fachliche Aspekte – geben einen Einblick in die Vielzahl an Kriterien und damit auch der Anforderungen und Aufgaben für Lehrkräfte im Erklärprozess. Tabelle 4 zeigt die genannten Kriterien noch einmal im Überblick.

Die erarbeiteten Kriterien sind Voraussetzungen für die Konstruktion guter instruktionaler Erklärungen. Sie sind jedoch nicht als Garantie für eine effektive bzw. qualitätsvolle Erklärung (Berliner, 2005), die ein Verstehen oder einen Wissenserwerb nach sich zieht, zu verstehen: Die Lehrkraft stellt mit der instruktionalen Erklärung ein entsprechendes Unterrichtsangebot bereit, wobei die eigentliche Wissens- oder Sinnkonstruktion von den Schüler:innen aktiv selbst zu leisten ist (vgl. Angebots-Nutzungsmodell nach Helmke, 2012). Weiterhin konnten Webb et al. (1995) in diesem Kontext Daten erheben, die zeigen, dass die konstruktiven Aktivitäten, die einer Erklärung nachfolgen – in diesem Fall Erklärungen von Schüler:innen und nicht von Lehrkräften –, meist bessere Prädiktoren für das Abschneiden in Lernerfolgstest der Schüler:innen darstellen als Kriterien, die an die Qualität der Erklärung an sich gestellt werden. Auch Wittwer und Renkl (2008) verweisen auf die Wichtigkeit der Verarbeitung bzw. Anwendung der Erklärinhalte als wichtiges Element effektiver Erklärungen, um beispielsweise eine Verstehensillusion zu verhindern.⁸³

⁸³ Eine Verstehensillusion bezeichnet das Gefühl einer instruktionalen Erklärung, den Inhalt verstanden zu haben. Erst bei der Anwendung des vermeintlich erworbenen Wissens wird dem Rezipient:innen deutlich, dass Wissenslücken vorhanden sind bzw. der Inhalt nicht ausreichend durchdrungen und verstanden wurde.

Tabelle 4: Übersicht zu den Kriterien guter instruktionaler Erklärungen

Aspekt	Facette	Kriterium
Adressatenorientierung	Anpassung durch didaktische Reduktion	Anpassung des Informationsgehalts durch sektorale didaktische Reduktion
		Anpassung des Schwierigkeitsgrads durch strukturelle didaktische Reduktion
	Vorwissen und Vorstellungen	Berücksichtigung des Vorwissens über den Erklärgegenstand
		Berücksichtigung von Vorstellungen über den Erklärgegenstand
		Berücksichtigung des Vorwissens über (fachspezifische) Repräsentationen
Kognitive Fähigkeiten	Berücksichtigung des Abstraktionsvermögens	
	Berücksichtigung der Kapazität des Arbeitsgedächtnisses	
Affektiv-motivationale Voraussetzungen	Berücksichtigung des Interesses z. B. durch Auswahl des Beispiels oder Lebensweltbezug	
	Berücksichtigung der Motivation z. B. durch Auswahl des Beispiels oder Lebensweltbezug	
Strukturiertheit	Inhaltliche Fokussierung	Fokussierung auf Wesentliches
		Klare Herausarbeitung von Verstehenselementen
	Gliederung und Aufbau	Fokussierung auf Konzepte und Prinzipien
Logische Anordnung der Verstehens-elemente und Teilerklärungen; Herstellung eines roten Fadens		
Orientierung und Transparenz	Schrittweise Einführung der Verstehenselemente	Zusammenfassung des Inhalts bei längeren Erklärungen
		Klare Zielformulierung
	Offenlegung der Struktur	
Einsatz von externen Repräsentationen	Auswahl passender externer Repräsentationen	Einbettung in einen größeren Zusammenhang
		Berücksichtigung der Passung zum Gegenstand
		Berücksichtigung des Vorwissens der Adressat:innen
	Veranschaulichung durch reale Objekte	Explikation des Ziels der Betrachtung
		Berücksichtigung der Komplexität der realen Objekte
Nutzung von realistischen und logischen Bildern	Berücksichtigung von Empfindungen gegenüber den realen Objekten (z. B. Ekel, Angst)	Einsatz von Bildern in Erklärungen
		Anleitung der Adressat:innen bei der Identifikation relevanter Elemente und der Interpretation
	Zeitgleiche Präsentation von Text und Bild	Integration von Text und Bild
		Einhaltung von Gestaltungsprämissen zur Verständlichkeit von Visualisierungen/ logischen Bildern
		Sichtbarkeit der Visualisierung
Nutzung von Analogien, Metaphern und Vergleichen	Einsatz von Analogien, Metaphern und Vergleichen in Erklärungen	

		Herausarbeitung des Herkunftsbereichs und des Übertragungsbereichs
		Aufzeigen der Grenzen der Übertragung
	Nutzung von Beispielen	Einsatz von Beispielen in Erklärungen Auswahl oder Gestaltung der Beispiele nach den Eigenschaften: passend, typisch, deckend, konkret, lebensweltnah, interessant und relevant
	Herstellung von Verbindungen zwischen verschiedenen Repräsentationen	
Sprech- und Körperausdruck	Artikulation	Deutliche Artikulation
	Prosodie	Gesunder Stimmklang
		Angemessene Sprechgeschwindigkeit
		Bedeutungsunterstützende Sprechpausen
		Variationsreiche Stimme (Sprechmelodie, Lautstärke, Tempo)
Kinesik	Blickkontakt mit den Adressat:innen	
	Einladende Mimik	
	Offene und zugewandte Körperhaltung	
	Gestik und Mimik, die das Gesagte unterstützt	
	Entspannte Körperhaltung	
	Proxemik	Angemessene Bewegung im Raum
Verständliche Sprache	Wortebene	Nutzung geläufiger Wörter
		Aussparung, Paraphrasierung oder Erklärung nicht geläufiger Wörter
	Satzebene	Anpassung der Komplexität der Syntax Eindeutige und präzise Formulierung der Aussagen
	Textebene	Herstellung lokaler und globaler Kohärenz (z. B. durch Kohärenzbildungshilfen)
Fachliche Aspekte	Fachliche Korrektheit	Richtigkeit der Fakten, Zusammenhänge und Prinzipien
		Wahl der adäquaten fachlichen Erklärungsform
		Angemessene Darstellung des aktuellen Wissensstandes
	Berücksichtigung der Auswirkung des Gesagten	Gefühl der (Un-)Sicherheit Faszination und Interessiertheit
	Auswahl repräsentativer und geeigneter Phänomene (Phänomen- vs. Prinzipienorientierung)	Orientierung am Phänomen und Darstellung dessen Komplexität
		Orientierung an allgemeinen, übertragbaren Prinzipien der Biologie
	Fachliche Vollständigkeit	Berücksichtigung aller notwendiger Verstandeselemente

An mehreren Stellen ist bereits in den vorhergegangenen Kapiteln aufgegriffen worden, dass die Kategorien teilweise miteinander verwoben sind. Dabei haben einzelne Kriterien bzw. Handlungsmöglichkeiten der Lehrkräfte auch Effekte auf die Qualität weiterer Kategorien. So kann

die Strukturiertheit einer instruktionalen Erklärung auch durch den Einsatz von Kohärenzhilfen für die leichtere Erfassung lokaler und globaler Zusammenhänge erhöht werden. Damit können bestimmte sprachliche Merkmale zur Erhöhung der wahrgenommenen Strukturiertheit (z. B. Erkennen des roten Fadens, Offenlegung der Struktur) der Erklärung beitragen. Auch der Einsatz des Sprech- und Körperausdrucks kann die Strukturiertheit erhöhen, indem Sinnabschnitte beispielsweise durch Handbewegungen und gezielte Sprechpausen abgesetzt und dadurch für die Zuhörer:innen besser kenntlich gemacht werden. Die Adressatenorientierung – die wie eingangs beschrieben häufig als *das* zentrale Merkmal bezeichnet wird – nimmt implizit stets Einfluss auf Gestaltungsentscheidungen in den fünf weiteren Kategorien bzw. Aspekten. Beispielsweise besteht für die Verständlichkeit der Sprache eine Abhängigkeit zwischen dem von der Lehrkraft gewählten Wortschatz von dem sprachlichen Vorwissen der Lernenden. Auch bei der Wahl von Repräsentationen muss die Lehrkraft das Vorwissen der Lernenden in Hinblick auf die Bekanntheit der gewählten Darstellungsform einbeziehen. Ein Bezug besteht auch bei der Ausgestaltung der Struktur: Zum Beispiel ist der rote Faden der Erklärung für Schüler:innen mit höherem Vorwissen mitunter leichter erkennbar als für Schüler:innen mit geringerem Vorwissen, da Zusammenhänge selbstständig gefunden werden können und damit der Aufbau der Erklärung durch die Lehrkraft nachvollzogen werden kann. Insgesamt stellt sich die Qualität einer instruktionalen Erklärung damit als komplexes Konstrukt dar.

Wendet man den Blick von den zahlreichen Kriterien, also den Merkmalen guter instruktionaler Erklärungen, auf die Lehrkraft als Konstrukteur:in dieser Erklärungen, muss festgestellt werden, dass ihr nicht nur eine zentrale Rolle, sondern auch eine umfassende Aufgabe zukommt: Sie muss in Hinblick auf das zu erreichende Ziel der Erklärung die Charakteristika des zu erklärenden Phänomens (Komplexität; Abstraktionsgrad; zugrunde liegendes, allgemeines Prinzip) und fachlich angemessene Erklärungen des Phänomens auf der einen Seite und die Merkmale der Adressat:innen (Vorwissen und Vorstellungen; kognitive Fähigkeiten; Interesse und Motivation) auf der anderen Seite kennen und berücksichtigen, indem sie angemessen didaktisch reduziert, die Inhalte geeignet strukturiert, passende Repräsentationen auswählt und einbettet, auf eine verständliche Sprache achtet und den Sprech- und Körperausdruck bedeutungs- und aufmerksamkeitsunterstützend einsetzt. Die folgende Grafik (Abb. 2) fasst die genannten Einflussfaktoren auf die Ausgestaltung des instruktionalen Erklärens aus Sicht der Lehrkraft zusammen und veranschaulicht diese. Dabei verdeutlicht der Pfeil, der von der instruktionalen Erklärung zu den Adressat:innen zeigt, dass die instruktionale Erklärung gegebenenfalls in Hinblick auf das erreichte bzw. nicht erreichte Ziel nochmals an die Rezipient:innen angepasst werden muss. Diese Anpassung an die Rezipient:innen – z. B. durch Veränderung der Struktur der Erklärung oder Nutzung anderer Repräsentationen – kann während der instruktionalen Erklärung oder nach dem Abschluss der Erklärung erfolgen. Insgesamt entsteht dabei das Produkt der instruktionalen Erklärung in einer Kokonstruktion zwischen Adressat:innen und erklärenden Personen (z. B. Harren, 2015; Spreckels, 2009; Stukenbrock, 2009). Initiationspunkte für die erneute Anpassung können von der Lehrkraft aktiv gesucht – zum Beispiel durch Verständnissnachfragen, die sich direkt an die Schüler:innen richten – oder von den Schüler:innen ohne Zutun der Lehrkraft gegeben werden. Dabei können Schüler:innen verbale oder nonverbale Signale geben. Unter verbale Rezipientensignale fallen beispielsweise kurze Aussprüche wie ‚Hä?‘ oder ‚Wie?‘ genauso wie vollständige Sätze und Fragen wie ‚Das habe ich nicht verstanden‘ oder ‚Können Sie das noch einmal erklären?‘. Nonverbale Rezipientensignale sind hingegen der Mimik und Gestik der Schüler:innen zu entnehmen. Treagust und Harrison (2000, S. 1161) bezeichnen die allgemeine Fähigkeit von Lehrkräften, eine Erklärung an die Rezipient:innen anzupassen, als „*explanatory flexibility*“.

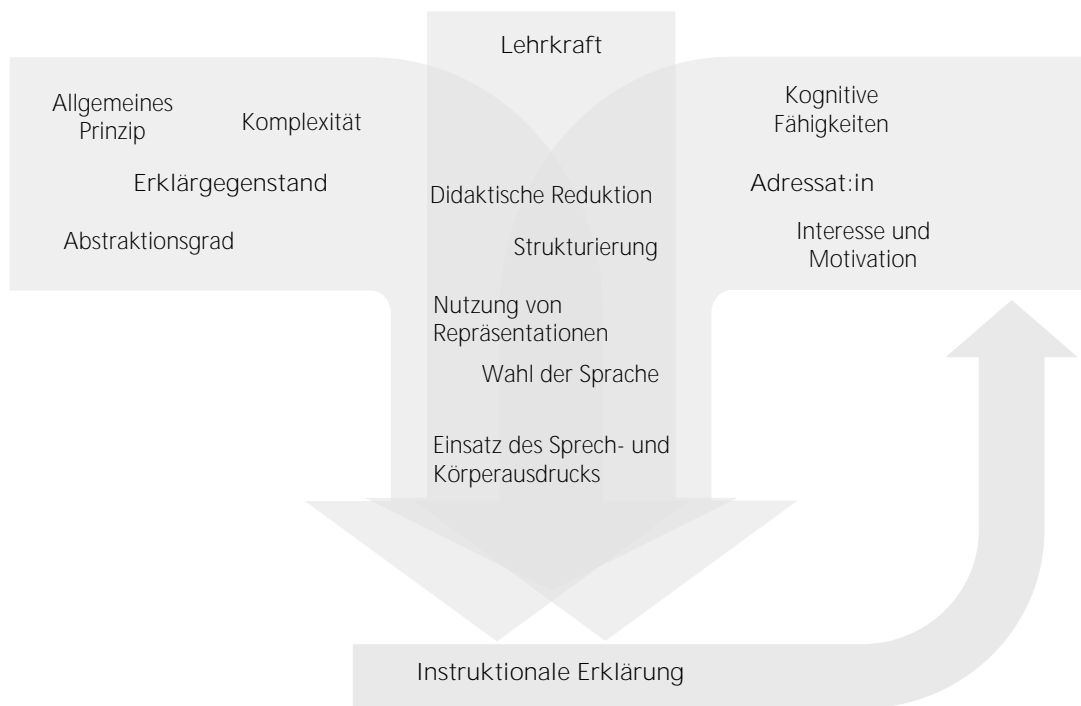


Abbildung 2: Faktoren zur Gestaltung einer (guten) instruktionalen Erklärung aus Sicht der Lehrkraft

Die Vielzahl an Faktoren bildet im Ansatz ab, wie komplex die Kommunikationssituation bzw. das Geben instruktionalen Erklärungen ist, bei der Lehrkräfte zahlreiche Anforderungen bewältigen müssen und ein umfassendes Professionswissen verlangt wird.⁸⁴ Die Faktoren können zum Teil von der Lehrkraft in der Unterrichtsvorbereitung abgeschätzt und einbezogen werden, müssen zum Teil aber auch ad hoc im Unterricht bewältigt werden.

Die Kommunikationssituation prägen dabei auch Merkmale der Lehrkraft selbst, die nicht in Abbildung 2 enthalten sind. Hierzu überlegen Treagust und Harrison (2000), dass beispielsweise neben den Vorstellungen der erklärenden Lehrkraft von der Natur der Naturwissenschaften, deren pädagogischer Expertise und dem Fachwissen über den Erklärgegenstand (die bisher genannten Aspekte sind Teile des Professionswissens von Lehrkräften) auch individuelle Erklärstile und ästhetische Präferenzen einen Einfluss auf die Gestaltung einer instruktionalen Erklärungen haben können. Ogborn (2011) fügt diesen Merkmalen noch allgemein die bisher gemachten Erfahrungen der Lehrkraft mit Erklärungen hinzu.

Insgesamt handelt es sich bei instruktionalen Erklärungen aufgrund der zahlreichen Freiheitsgrade, die den Erklärprozess formen können, stets um einen individuellen Konstruktionsprozess bzw. einen individuellen Kokonstruktionsprozess zwischen Lehrkräften und den Adressat:innen, der in seinen Details vermutlich einmalig ist. So resümiert auch Wragg (1993) am Ende seiner Analyse von

⁸⁴ Das notwendige Wissen für die erfolgreiche Bewältigung einer Erklärsituation wird dabei von verschiedenen Forscher:innen unterschiedlich in Beziehung zur den drei differenzierbaren Arten des Professionswissens von Lehrkräften (CK: content knowledge; PCK: pedagogical content knowledge; PK: pedagogical knowledge) gesetzt. Das Codiersystem von Sevan und Gonsalves (2008), das die Bewertung von Erklärungen ermöglichen soll, nimmt notwendiges Wissen in allen drei Arten des Professionswissens an. Findeisen (2017) nimmt aufgrund der hohen Inhaltsspezifität von Erklärprozessen an, dass es sich beim Erklärungswissen von Wirtschaftslehrkräften „um ein domänen- bzw. sogar ein inhaltspezifisches Wissen“ (Findeisen, 2017, S.112) handeln könnte. Bei Wagner und Wörn (2011) findet sich dieser Gedanke ebenso im Kontext von Mathematikunterricht. Treagust und Harrison (2000) ordnen das notwendige Wissen dem PCK zu.

Erklärungen aus dem Deutsch- und Naturwissenschaftsunterricht, dass selbst zwischen den gut bewerteten Erklärungen zahlreiche Unterschiede zu finden sind:

As is often the case, however, despite some communalities, high-scoring explanations were also different from each other, with their own unique qualities. The general conclusions above are not universal prescriptions, but rather interesting illuminations from one intensive study. (Wragg, 1993, S. 136)

4.8 Gutes instruktionales Erklären – eine Frage der Perspektive

Es kann davon ausgegangen werden, dass aufgrund unterschiedlicher Erfahrungen, Wissensbestände, Ziele und Bedürfnisse Personen⁸⁵ Unterricht bzw. instruktionale Erklärungen differierend wahrnehmen und schließlich beurteilen (Helmke, 2007). Für unterschiedliche Personengruppen bedeutet ‚gutes Erklären‘ folglich nicht das Gleiche, sodass sich die Frage, welche Kriterien guter instruktionaler Erklärungen kennzeichnen, vor diesem Hintergrund differenziert.

Zwei weitere theoretische Bezüge aus der Bildungsforschung sowie aus der qualitativen Methodik führen die Gedanken weiter. In Hinblick auf die Ausführungen Berliner (2005) und Flick (2011) verlangt die Beschreibung von Kriterien guten instruktionalen Erklärens den Einbezug verschiedener Perspektiven:

- i) Die Differenzierung zwischen gutem und effektivem Unterrichten nach Berliner (2005) wurde eingangs (siehe Kap. 1) bereits aufgegriffen und auf das instruktionale Erklären übertragen: Während effektives instruktionales Erklären am Effekt gemessen wird, also am Lernerfolg bzw. an den erreichten Lernzielen wird gutes instruktionales Erklären an normativen Standards gemessen, die die Beteiligten des Feldes anlegen. Die Beteiligten des Feldes definieren damit in ihrer Übereinstimmung, was als ‚gut‘ gelten kann.
- ii) Aus der Sicht von Denzin (1970) und Flick (2011) ist es aus methodischer Sicht notwendig, unterschiedliche Perspektiven bei der Untersuchung eines Forschungsgegenstandes gleichberechtigt einzubeziehen, um ein tieferes Verständnis des Gegenstands zu erzielen.

Die zwei Ausführungen schaffen die Grundlage dafür, dass für die Bestimmung von Kriterien guter instruktionaler Erklärungen nicht ausschließlich aus einer Perspektive heraus erfolgen kann. Daraus ergibt sich die Frage, welche Perspektiven für die Bestimmung guten instruktionalen Erklärens relevant und wie diese charakterisiert sind. Im Folgenden werden mögliche Gruppen, die als Statusgruppen bezeichnet werden, aufgeführt und hinsichtlich ihrer Perspektive auf gutes instruktionales Erklären skizziert.

Abgeleitet aus dem Merkmal der asymmetrischen Wissensverteilung bei instruktionalen Erklärungen bzw. aus den Rollen der involvierten Personen (siehe Kap. 2.1; Ehlich, 2009; Hohenstein, 2006) sind die Perspektiven von Lehrkräften (als vermittelnde Person) und Schüler:innen (als Rezipient:innen des Inhalts) als relevant anzusehen.

Die Einschätzungen der Schüler:innen sind dabei von besonderer Bedeutung, da sich das Lernangebot, das ihnen Lehrkräfte im Zuge der instruktionalen Erklärung bereitstellen, an sie richtet. Nur wenn das Angebot angenommen wird, können nachfolgende Lernaktivitäten erfolgreich sein

⁸⁵ Nicht berücksichtigt werden hier die individuellen Perspektiven auf den Gegenstand – Personen werden zu Personengruppen und für diese allgemein charakterisierende Merkmale zusammengefasst.

und gewünschte Lernergebnisse resultieren (Helmke, 2012; Seidel, 2014). Hattie (2015) formuliert zur Bedeutung der Perspektive der Lernenden Folgendes:

Es sind die Lernenden, die in den Klassen sitzen, und merken, ob ihre Lehrperson das Lernen mit ihren Augen sieht [...]. Lernen muss von den Lehrpersonen aus der Perspektive der Lernenden betrachtet werden, damit sie besser verstehen, wie das Lernen aus der Sicht der Lernenden aussieht und wie es sich für sie anfühlt. (Hattie, 2015, S. 139)

Im Kontext von instruktionalen Erklärungen nehmen Schüler:innen die einzige Perspektive ein, für die die Erklärung eine echte Erklärung sein kann, da sie über individuelles Nichtwissen verfügen. Schüler:innen können die Passung des Angebots für sich selbst beurteilen und sind daher insbesondere für die Bewertung des Adressatenbezug geeignet (Kunter & Baumert, 2007). Baumert (2002) bezeichnet – aus theoretischer Sicht – Lernende aufgrund ihrer Unterrichtserfahrungen sogar als Expert:innen für die Beurteilung von Unterricht: „Die jungen Erwachsenen der gymnasialen Oberstufe – und das gilt auch für Schülerinnen und Schüler der Mittelstufe – sind Experten für Unterrichtsbeurteilung. 15 000 Unterrichtsstunden eines Schülerlebens gehen nicht spurlos vorbei“ (Baumert, 2002, S. 136). Gautschi (2011) nimmt in seiner empirischen Untersuchung von Aspekten eines guten Geschichtsunterrichts ebenfalls die Beurteilungen von Schüler:innen auf. Er kommt auf Basis seiner Daten zu dem Schluss, dass die Beurteilungen durch Lernende „**reichhaltig**“ (Gautschi, 2011, S. 267) sind, und liefert damit ein Indiz dafür, dass der Einbezug von Schüler:innen einen Beitrag zum Verständnis guter instruktionaler Erklärungen leisten kann.

Beurteilungen durch Schüler:innen werden im Kontext der Messungen von Unterrichtsqualität hingegen häufig kritisch gesehen, weil diese für viele Aspekte als weniger valide gelten (z. B. Clausen, 2002; Fauth, Göllner, Lenske, Praetorius & Wagner, 2020; Kleickmann, Praetorius & Riecke-Baulecke, 2019; Kunter & Baumert, 2007). Clausen (2002) zieht aus seinen Daten zur Beurteilung von Unterrichtsqualität mit Fokus auf die Validität der Beurteilungen von Schüler:innen, Lehrenden und externen Beobachter:innen das Fazit, dass „[f]ür eine neutrale wissenschaftlich-differenzierende Betrachtung des Unterrichtsgeschehens [...] Schülerurteile aufgrund der Tendenz zur affektiven geprägten generalisierenden Beurteilung weniger geeignet [sind]“ (Clausen, 2002, S. 188). Gleichzeitig räumt er der Beurteilung durch Schüler:innen eine große Bedeutung ein, da gerade diese „**wahrgenommene affektive Qualität des Unterrichtens**“ (Clausen, 2002, S. 188) mit der Entwicklung der Leistung und des Interesses der Lernenden einen engen Zusammenhang aufweist.

Lehrkräfte besitzen im Gegensatz zu Schüler:innen entsprechendes Wissen – pädagogisches Wissen, fachdidaktisches Wissen und Fachwissen (Shulman, 1986) –, um Unterricht differenziert wahrnehmen und beurteilen zu können (Blömeke, Gustafsson & Shavelson, 2015; Sherin, 2001). Sie verfügen weiterhin über eine große Menge an Erfahrungen über gelungene und weniger gelungene Erklärungen aus ihrem eigenen Unterricht. Herausfordernd ist für Lehrkräfte jedoch, dass sie eher selten Unterricht oder Erklärungen beurteilen müssen, wenn sie nicht in die Ausbildung von (angehenden) Lehrkräften eingebunden sind, sodass die meisten Lehrkräfte wenig Erfahrung darin besitzen (Clausen, 2002).⁸⁶ Der Perspektive von Schüler:innen und Lehrkräften ist gemein, dass sie Teil der unterrichtlichen Praxis sind und so bei ihren Einschätzungen als Referenzrahmen diesen Unterricht heranziehen.

⁸⁶ Eine Ausnahme bilden mitunter Seminarlehrkräfte (Lehrkräfte in Bayern, die Referendar:innen ausbilden), Betreuungslehrkräfte (Lehrkräfte in Bayern, die Lehrkräfte im Vorbereitungsdienst begleitend betreuen) sowie Fachbetreuer:innen (Lehrkräfte in Bayern, die an ihren jeweiligen Schulen die Betreuung der Fachschaft innehaben, dabei gegebenenfalls Unterrichtsbesuche durchführen und Lehr-Lern-Materialien in Hinblick auf eine Beschaffung beurteilen), die regelmäßig im Kontext ihrer Funktion in Beurteilungsprozesse involviert sind.

Bezieht man das Kompetenzmodell von Blömeke und Kolleg:innen (2015) mit ein, in dem sich die Kompetenz als ein Kontinuum darstellt, sind Lehramtsstudierende Noviz:innen in Bezug auf die Praxis des Unterrichtens sowie in Bezug auf professionelle Wissensbestände. Die Perspektive von Studierenden auf die Frage nach Kriterien guter instruktionaler Erklärungen kann jedoch ebenso interessant und für einen Erkenntnisgewinn darüber relevant sein: Studierende verfügen über eine bessere Ausdrucksfähigkeit als Schüler:innen und können damit Beurteilungen aus sprachlicher Sicht differenzierter darstellen (Asen-Molz, Ehras & Krauss, im Druck), sie haben als angehende Lehrkräfte bereits professionsbezogenes Wissen erworben und können sich auf dieses bei der Beurteilung beziehen, sodass die Beurteilungen nicht nur sprachlich, sondern auch inhaltlich differenzierter sein können als diejenigen von Schüler:innen (Blömeke et al., 2015; Seidel, Blomberg & Stürmer, 2010; Sherin, 2001). Darüber hinaus ist das Erfahrungswissen aus dem Unterricht als Schüler:in noch eher zugänglich als bei bereits länger praktizierenden Lehrkräften, sodass sie mitunter neben Anteilen professioneller Wissensbestände auch Erfahrungswissen als Adressat:innen in die Beurteilungen einfließen lassen.

Hochschuldidaktiker:innen prägen als Lehrkräftebildner:innen, was unter gutem Unterricht bzw. gutem Erklären verstanden wird. Gautschi (2011) bezieht in seine Datengrundlage für die Beantwortung der Frage nach gutem Geschichtsunterricht ebenfalls Hochschuldidaktiker:innen mit ein (er bezeichnet die Gruppe jedoch als Expert:innen bzw. Kontrollexpert:innen; sie besteht aus lehrerfahrenen Geschichtsdidaktiker:innen). Die Beurteilungen aus dieser Statusgruppe zeigen in Gautschis Studie ein heterogenes Bild: Neben Bezügen auf die Persönlichkeitsmerkmale der unterrichtenden Person finden sich auch Beurteilungen in Hinblick auf die äußeren Rahmenbedingungen des Unterrichts (Größe der Klasse, Ausstattung des Klassenzimmers) – also Kategorien, die mit dem eigentlichen Inhalt des Unterrichts nicht in Zusammenhang stehen. Gleichzeitig rekurren einige Hochschuldidaktiker:innen in ihren Beurteilungen auf Theorien und Modelle aus der Pädagogik, Psychologie oder Fachdidaktik. Aufgrund ihrer Tätigkeiten an der Universität ist eine solche eher theorie- und forschungsbasierte Perspektive erwartbar (Gautschi, 2011).

Die aufgeführten Gruppen (im Folgenden Statusgruppen genannt) können bei der Frage nach Kriterien guten instruktionalen Erklärens eine spezifische Perspektive einbringen und zu einem tieferen und umfassenden Verständnis beitragen:

Die vier Statusgruppen repräsentieren den Blick von Adressat:innen (Schüler:innen), die kein professionelles Wissen besitzen, sowie die Metaperspektive derjenigen (Studierende, Lehrkräfte, Didaktiker:innen), die über professionelles Wissen verfügen. Während Studierende eher als Noviz:innen eingestuft werden können, sind Lehrkräfte und Didaktiker:innen eher Expert:innen in Hinblick auf professionelles Wissen über Unterrichten bzw. Erklären, sodass diese verschiedene Professionalisierungsniveaus repräsentieren. Didaktiker:innen repräsentieren gegenüber Lehrkräften einen anderen Professionalisierungsschwerpunkt – während Hochschuldidaktiker:innen aufgrund ihres primären Tätigkeitsfeldes eher mit Theorien und Modellen des Unterrichtens vertraut sind, verfügen Lehrkräfte aufgrund ihres Tätigkeitsbereiches vermehrt über Erfahrungswissen zum Unterrichten.

5 Zielsetzung und Konzeption der empirischen Untersuchung zum instruktionalen Erklären im Biologieunterricht

5.1 Zielsetzung und Fragestellungen

Die Studie zielt insgesamt darauf ab, eine Konzeptualisierung für instruktionales Erklären im Biologieunterricht zu erarbeiten sowie empirisch Kriterien guten instruktionalen Erklärens im Biologieunterricht unter Berücksichtigung verschiedener Perspektiven zu untersuchen. In Hinblick auf den Biologieunterricht und dessen Bedeutung bei der Vermittlung eines angemessenen Verständnisses biologischer Phänomene besteht insbesondere ein Interesse, Unterschiede bzw. Schwierigkeiten beim Erklären komplexer Phänomene im Vergleich zu Erklärungen von linearen Ursache-Wirkungs-Beziehungen herauszuarbeiten.

Für eine grundlegende Begriffsbestimmung, was instruktionales Erklären im Biologieunterricht ist, wurden zunächst in Kapitel 2 Überlegungen der Sprachwissenschaft, der Didaktik und Lehr-Lern-Theorie sowie der Wissenschaftstheorie der Biologie zusammengeführt. Ausgehend von einer Betrachtung der in den jeweiligen Diskursen formulierten Ziele und Funktionen sowie differenzierbaren Typen von Erklärungen wurde eine Definition instruktionaler Erklärungen im Biologieunterricht erarbeitet (siehe Kap. 3.6):

Instruktionale Erklärungen im Biologieunterricht sind vielgestaltige Sprachhandlungen von Lehrkräften, die an Lernende gerichtet sind und die Initiierung einer Wissensvermittlung, eines Verstehensprozesses oder eines Fähigkeitserwerbs zum Ziel haben, wobei sich die Erklärgegenstände stets der direkten Wahrnehmung entziehen. Bei der Erklärung biologischer Phänomene orientieren sich Lehrkräfte an wissenschaftlichen Erklärungsformen der Biologie und modellieren damit implizit, welche Erklärungen aus Sicht des Faches angemessen sind.

In Kapitel 4 wurde dargestellt, welche fachübergreifenden Kriterien in Diskussionen zu guten bzw. effektiven instruktionalen Erklärungen bisher aufgegriffen werden, sowie biologiespezifische Kriterien aus verschiedenen Diskursen abgeleitet. Ergänzend wurde dargelegt, dass für die grundlegende Frage nach Kriterien *guter* instruktionaler Erklärungen die normativen Ansichten von Angehörigen des Feldes (Schüler:innen als Adressat:innen, Studierende als Noviz:innen des Lehrens, Lehrkräfte und Hochschuldidaktiker:innen) zu berücksichtigen sind und davon ausgegangen werden kann, dass sich diese aufgrund differierender professionellen Wissens unterscheiden. Bei der empirischen Bestimmung bzw. Überprüfung von Kriterien guter instruktionaler Erklärungen von Biologielehrkräften müssen diese Perspektiven gleichberechtigt aufeinander bezogen werden.

Folgende Forschungsfragen können abgeleitet werden:

1. Welche Kriterien werden zur Bewertung instruktionaler Erklärung herangezogen?

1.1. Welche Kriterien werden statusgruppenübergreifend zur Bewertung instruktionaler Erklärungen herangezogen?

Vorannahmen: Zur Bewertung instruktionaler Erklärungen werden die in der Literatur gefundenen Aspekte Adressatenorientierung, Strukturiertheit, Repräsentationen, Sprech- und

Körperausdruck, verständliche Sprache sowie fachliche Aspekte herangezogen. Insbesondere die Aspekte des Sprech- und Körperausdrucks sowie der verständlichen Sprache werden jedoch seltener und eher global genannt, da diese Aspekte in der Lehrkräftebildung und darüber hinaus wenig thematisiert werden und so für eine differenzierte Bewertung der Aspekte entsprechendes Wissen zur bewussten Wahrnehmung und schließlich zur Bewertung fehlt.

1.2. Welche Kriterien werden von verschiedenen Statusgruppen (Schüler:innen, Studierende, Lehrkräfte, Hochschuldidaktiker:innen) zur Bewertung instruktionaler Erklärungen herangezogen?

Vorannahmen: Schüler:innen geben weniger differenzierte Bewertungen ab, nennen also weniger Kriterien und führen häufiger globale Aspekte an. Auch die Lehrperson und ihre Merkmale werden in eine Bewertung der Erklärung einbezogen. Weiterhin wird angenommen, dass sich die Bewertungen von Studierenden, Lehrkräften und Hochschuldidaktiker:innen aufgrund ihres professionellen Wissens von denen der Schüler:innen unterscheiden. Sie geben häufiger differenzierte Bewertungen ab, geben also häufiger spezifische Kriterien und weniger häufig globale Aspekte an und führen selten Merkmale der Lehrperson als Kriterien für die Bewertung der instruktionalen Erklärung an. In den Bewertungen von Hochschuldidaktiker:innen lassen sich häufiger Bezüge zu fachdidaktischen und lernpsychologischen Theorien sowie zur Wissenschaftstheorie finden als in den übrigen Gruppen.

2. Wie unterscheiden sich die Bewertungen instruktionaler Erklärungen durch verschiedene Statusgruppen (Schüler:innen, Studierende, Lehrkräfte, Hochschuldidaktiker:innen) in Hinblick auf vorgegebene Kriterien?

Vorannahmen: Auf Grundlage der Theorie können keine konkreten Vorannahmen hinsichtlich auftretender Differenzen in der Bewertung entlang fester Kriterien – also die bessere oder schlechtere Bewertung einzelner Kriterien oder Facetten – von instruktionalen Erklärungen getroffen werden. Aufgrund der Expertiseunterschiede kann vermutet werden, dass die meisten bzw. die größten Unterschiede in den Einschätzungen zwischen Schüler:innen und Lehrkräften sowie zwischen Schüler:innen und Didaktiker:innen liegen. Wenn hingegen zwischen Lehrkräften und Schüler:innen wenig unterschiedliche Einschätzungen zu finden sind, spricht das eher für ein geteiltes Verständnis zu gutem instruktionalen Erklären sowie für eine gute Einschätzung der Fähigkeiten sowie des Wissens von Schüler:innen durch Lehrkräfte.

3. Welche Aspekte sind für die Bewertung von instruktionalen Erklärungen große Einflussfaktoren?

3.1. Welche Aspekte sind statusgruppenübergreifend für die Bewertung von instruktionalen Erklärungen große Einflussfaktoren?

Vorannahmen: In theoretischen Arbeiten wird häufig die Adressatenorientierung als grundlegender Aspekt für instruktionaler Erklärungen beschrieben, sodass angenommen werden kann, dass die Adressatenorientierung ein großer Einflussfaktor für die Bewertung instruktionaler Erklärungen ist. Weitere theoretisch fundierte Annahmen über die Gewichtung einzelner Aspekte guten instruktionalen Erklärens können nicht getroffen werden. Insgesamt wird angenommen, dass alle implementierten Aspekte Einflussfaktoren auf die Qualität instruktionaler Erklärungen darstellen.

3.2. Welche Aspekte sind aus der Perspektive der Schüler:innen im Unterschied zur Perspektive der Lehrenden für die Bewertung von instruktionalen Erklärungen große Einflussfaktoren?

Vorannahmen: Aufgrund der Unterschiede bezüglich des professionellen Wissens im Bereich des fachlichen Wissens und des fachdidaktischen Wissens wird davon ausgegangen, dass zusätzlich zur Adressatenorientierung für Lehrende der fachspezifische Aspekt einen größeren Einfluss auf die Bewertung hat als für Schüler:innen. Da Schüler:innen eher eine oberflächliche Beurteilung von Unterricht zugeschrieben wird, kann angenommen werden, dass zusätzlich zur Adressatenorientierung der Sprech- und Körperausdruck und/oder die Wirkung der erklärenden Person Einflussfaktoren darstellen. Weitere theoretisch fundierte Annahmen über die statusgruppenspezifischen Gewichtungen maßgeblicher Einflussfaktoren können auch hier nicht getroffen werden.

In Kapitel 3 wurde schließlich auf Grundlage wissenschaftstheoretischer Literatur ein Einblick in die Komplexitätstheorie gegeben und auf die Biologie bezogen: Wissenschaftliche Erklärungen komplexer Phänomene sind multikausal, die Formulierung und Anwendung von Gesetzen ist begrenzt und sie sind von Nichtwissen und unsicherem Wissen geprägt. Einzelne Befunde können zeigen, dass Noviz:innen Verständnisschwierigkeiten in Bezug auf Merkmale komplexer Phänomene haben und Lehrende damit vor Herausforderungen stehen, komplexe Phänomene verständlich zu unterrichten. In Hinblick auf den Biologieunterricht und dessen Bedeutung bei der Vermittlung eines angemessenen Verständnisses natürlicher Phänomene besteht damit insbesondere ein Interesse, Unterschiede bzw. Schwierigkeiten beim Erklären komplexer Phänomene im Vergleich zu Erklärungen von linearen Ursache-Wirkungs-Beziehungen herauszuarbeiten.

Deshalb sind folgende Forschungsfragen von Interesse:

4. Wie unterscheiden sich die globalen Bewertungen und deren Begründungen zwischen linearen und komplexen instruktionalen Erklärungen in Abhängigkeit von der Statusgruppe?

4.1. Wie unterscheiden sich die globalen Bewertungen zwischen linearen und komplexen instruktionalen Erklärungen in Abhängigkeit von der Statusgruppe?

Vorannahmen: Da komplexe instruktionaler Erklärungen sowohl Lehrende als auch Lernende vor Herausforderungen stellen, wird angenommen, dass komplexe Erklärungen schlechtere Bewertungen erhalten.

4.2. Wie unterscheiden sich die Begründungen der globalen Bewertungen zwischen linearen und komplexen instruktionalen Erklärungen in Abhängigkeit von der Statusgruppe?

Vorannahmen: Spezifische Annahmen über das Heranziehen einzelner Kriterien bei linearen instruktionalen Erklärungen im Unterschied zu komplexen instruktionalen Erklärungen können nicht formuliert werden.

5.2 Forschungszugang und Untersuchungsdesign

Die durchgeführte Studie wurde – wie in Kapitel 1.1 beschrieben – im Rahmen des inter- und transdisziplinären Projektes FALKE (FACHspezifische LEhrkräftekompetenzen im Erklären)⁸⁷ als ein Teilprojekt von KOLEG (KOoperative LEhrkräftebildung Gestalten) an der Universität Regensburg mit einer Förderung des Bundesministeriums für Bildung und Forschung durchgeführt. In Kooperation mit den zehn beteiligten Fachdidaktiken sowie der deutschen Sprachwissenschaft, der Sprecherziehung und der Methodik der empirischen Bildungsforschung wurde das grundlegende Design der Untersuchung konzipiert. Dadurch sind neben fachspezifischen Fragestellungen und Analysen, die in besonderer Weise relevant für den Biologieunterricht sind, auch Vergleiche zwischen den Fächern möglich.⁸⁸ Im Folgenden wird das Untersuchungsdesign zunächst skizziert und in den darauffolgenden Abschnitten weiter ausgeführt.

Die hauptsächlich quantitativ ausgerichtete Untersuchung ist als quasi-experimentell zu beschreiben und besitzt explorativen Charakter, weshalb keine konkreten Hypothesen vorab formuliert wurden (Bortz & Döring, 2006). Vor allem in Bezug auf die Perspektiven der Statusgruppen liegen wenige und in Bezug auf den fachspezifischen Schwerpunkt zur Komplexität biologischer Erklärungen kaum Ergebnisse vor, sodass eine fundierte Formulierung von Hypothesen nicht möglich ist. Damit folgt der Untersuchungsansatz nicht der für quantitative Forschung häufig prototypisch formulierten theorieprüfenden und deduktiven Vorgehensweise (Lamnek, 2010).⁸⁹

Kern des Untersuchungsdesigns ist ein computerbasierter *Fragebogen*, in den *Videovignetten* unterrichtsnaher, instruktionaler Erklärungen als Stimuli implementiert wurden. In der Teilstudie FALKE-Biologie greifen die sechs Videovignetten sechs verschiedene, biologische Phänomene auf, wobei diese im Grad der dargestellten Komplexität systematisch variiert wurden. Drei der Erklärungen führen das Phänomen auf eine eindeutige, lineare Ursache-Wirkungs-Beziehung zurück, wohingegen in drei weiteren Erklärungen ein komplexes Ursache-Wirkungs-Gefüge darstellen, also mehrere Ursachen angeführt werden und gleichzeitig von der Lehrkraft hervorgehoben wird, dass nicht alle Ursachen sowie Wechselwirkungen und Zusammenhänge gänzlich bekannt sind, sodass Nichtwissen und unsicheres Wissen beim Erklären mit angesprochen werden.

Die Erklärungen wurden von vier verschiedenen Statusgruppen (Schüler:innen, Studierende, Lehrkräfte, Didaktiker:innen) hinsichtlich ihrer Güte eingeschätzt. Damit sind verschiedene Perspektiven – Lehrende aus der Metaperspektive und Lernende aus der Rezipient:innenperspektive – repräsentiert, die im Sinne einer *Perspektivtriangulation* für ein umfassendes Verständnis des guten instruktionalen Erklärens zusammengeführt werden.

Im Laufe des Fragebogens geben die Teilnehmer:innen zunächst ein holistisches Globalurteil bezüglich in Form einer Schulnote zu jeder der sechs gesehenen Erklärung ab und können ihre

⁸⁷ Dank gilt an dieser Stelle allen Projektbeteiligten von FALKE (Arne Dittmer, Oliver Tepner, Michael Elmer, Anita Schilcher, Lisa Gaier, Petra Kirchhoff, Maria Gastl-Pischetsrieder, Michael Fricke, Renate Murmann, Harriet Rudolph, Josef Memminger, Anna-Maria Ruck, Birgit Eiglsperger, Matthias Weich, Magnus Gaul, Mario Frei, Astrid Rank, Katharina Asen-Molz, Stefan Krauss, Simone Röhr, Karsten Rincke, Jana Heinze, Wieland Kranich, Christian Gegner, Eileen Läger-Gunga, Christiane Thim-Mabrey, Eva-Maria Meier, Alfred Lindl, Sven Hilbert) für die gemeinsamen Diskussionen und geteilten Anregungen über die gesamte Projektlaufzeit hinweg.

⁸⁸ Eine Übersicht zu Ergebnissen der FALKE-Gesamtstudie zu Unterschieden und Gemeinsamkeiten zwischen den beteiligten Fächern ist in Schilcher, Krauss, Lindl und Hilbert (im Druck) nachzulesen.

⁸⁹ Lamnek (2010) stellt vor der polarisierenden Abgrenzung von quantitativen und qualitativen Forschungsparadigmen jedoch auch heraus, dass diese Dichotomie in der Forschungspraxis weit weniger starr ist und den vielfältigen Methoden, die innerhalb beider Paradigmen zur Anwendung kommen, nicht gerecht wird.

Einschätzung danach in einem offenen Textfeld begründen. Die Einschätzung erfolgt an dieser Stelle ohne die Vorgabe bestimmter Kriterien (= selbstgewählte Kriterien). Die Äußerungen der Teilnehmer:innen aus der offenen, unstandardisierten Frage können dazu herangezogen werden, bisher Kriterien zu validieren, mit Beispielen zu konkretisieren sowie weitere Kriterien zu finden. Dazu werden die Daten in Anlehnung an die inhaltlich strukturierende qualitative Inhaltsanalyse (Kuckartz, 2018) ausgewertet und anschließend quantifiziert.

Der Fragebogen enthält außerdem geschlossene Items, sodass offene und geschlossene Antworten im Sinne einer *Methodentriangulation* miteinander verschränkt werden können. Die geschlossenen Items werden von den teilnehmenden Personen nach erneuter Betrachtung der Erklärungen eingeschätzt und leiten so zu einer kriteriengeleiteten Bewertung an (= vorgegebene Kriterien). Hierfür wurden innerhalb der Projektgruppe FALKE die Aspekte Adressatenorientierung und Strukturiertheit gewählt. Die *Konzeptualisierung und Operationalisierung* wurden in der Projektgruppe gemeinsam vorgenommen und die entstandenen Items in allen Fächern analog in den Fragebogen implementiert. Des Weiteren wurden als fachübergreifende Aspekte die Verständlichkeit der Sprache sowie der Sprech- und Körperausdruck aufgenommen. Die Aspekte wurden von den jeweiligen Fachvertreter:innen im Projekt leitend erarbeitet.⁹⁰ Auch diese Items wurden im Fragebogen der biologiedidaktischen Studie aufgenommen und werden bei der Auswertung der Ergebnisse berücksichtigt. Neben diesen fachübergreifend relevanten Aspekten guten Erklärens steht die Wahrnehmung der dargestellten Komplexität im Fokus der biologiedidaktischen Studie. Für die Aspekte lagen keine Skalen vor, sodass die Items auf Grundlage der in Kapitel 4 dargestellten Befunde zur Qualitätskriterien und des in Kapitel 3 vorgestellten fachspezifischen Schwerpunktes selbst erstellt wurden. Die Items sowie die informierenden und instruktionalen Anteile des Fragebogens durchliefen eine *Pilotierung* anhand von kognitiven Interviews. An die Pilotierung schloss sich eine Überarbeitungsphase unter Berücksichtigung der gewonnenen Ergebnisse an.

Insgesamt ist das dargestellte Studiendesign quantitativ ausgerichtet, enthält jedoch durch die Begründung zur Notenvergabe auch einen qualitativen Anteil, der eine Öffnung für ein breiteres bzw. vertieftes Verständnis generiert und Einblicke in selbstgewählte Kriterien der Teilnehmer:innen gibt.

In den Kapiteln 5.2.1 bis 5.2.6 werden die Elemente des Studiendesigns (Fragebogen, Videovignetten, Perspektiventriangulation, Methodentriangulation, Konzeptualisierung und Operationalisierung der Aspekte guten Erklärens, Pilotierung) ausführlicher dargelegt. Dabei sollen zunächst je allgemeine Verortungen vorgenommen und theoretische Bezüge hergestellt werden. Anschließend wird die konkrete Realisierung in der vorliegenden Studie beschrieben und begründet. Das Kapitel schließt mit der Beschreibung der erhobenen Stichprobe und der Durchführung (Kap. 5.2.7).

5.2.1 Aufbau und Ablauf des Fragebogens

Der eingesetzte Fragebogen besteht aus zwei Teilen, die Messzeitpunkte 1 und 2 genannt werden. Durch die Teilung sollte die Zumutbarkeit für die Teilnehmer:innen als ein Nebengütekriterium erhöht werden, da die gesamte Bearbeitungsdauer insgesamt im Mittel über alle Statusgruppen

⁹⁰ Ausführlich nachzulesen im erschienenen Projektband bei Thim-Mabrey und Lindl (im Druck) sowie Lägel-Gunga, Schilcher, Kranich und Gegner (im Druck).

78:21 Minuten betrug. Die zwei Teile konnten unmittelbar hintereinander oder an verschiedenen Tagen bearbeitet werden.

Abbildung 3 zeigt zunächst einen grafischen Überblick über den Ablauf des Fragebogens hinsichtlich der Videobetrachtung und der Reihenfolge der einzuschätzenden Aspekte.

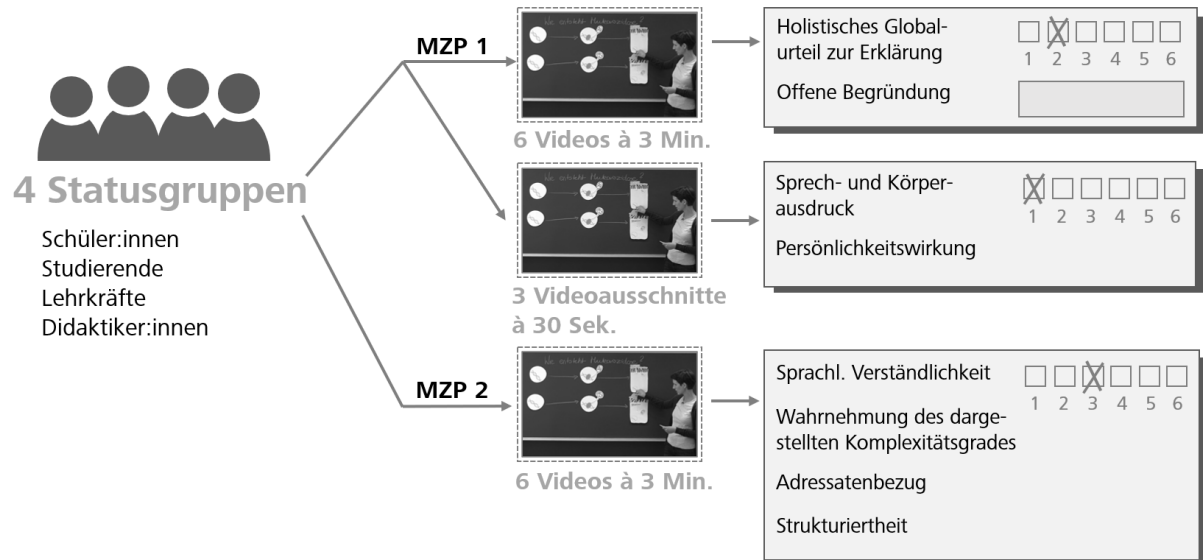


Abbildung 3: Aufbau und Ablauf des eingesetzten Fragebogens

In den folgenden Punkten wird der Ablauf des Fragebogens ausführlicher beschrieben, wobei die Punkte 1 bis 9 im ersten Messzeitpunkt sowie die Punkte 10 bis 12 im zweiten Messzeitpunkt implementiert waren:

1. *Informationen zu Zielen und Ablauf:* Alle Statusgruppen erhalten zunächst die Informationen, dass es thematisch um gute Erklärungen im Biologieunterricht gehen soll und im Folgenden Videos hinsichtlich ihrer Qualität in einem ersten Teil global und in einem zweiten Teil der Umfrage kriteriengeleitet eingeschätzt werden sollen. Ebenfalls wurde über die freiwillige Teilnahme aufgeklärt sowie klargestellt, dass keine Rückschlüsse auf die eigene Person möglich sind. Ein Zeitrahmen von ca. 45 Minuten für die Bearbeitung eines Teils wurde als Richtwert angegeben.
2. *Teilnehmer:innencode:* Anschließend wird von den Teilnehmer:innen ein individueller Code erstellt – bestehend aus jeweils den zwei ersten Buchstaben des mütterlichen Vornamens, den zwei ersten Buchstaben des väterlichen Vornamens und den ersten zwei Ziffern der Hausnummer –, um eine Zusammenführung der beiden Messzeitpunkte zu ermöglichen.
3. *Demografische Angaben:* Die Studierenden, Lehrkräfte und Hochschuldidaktiker:innen beantworten jeweils statusgruppenspezifische Fragen zur Demografie bzw. zu interessierenden Größen (z. B. Berufserfahrung in Jahren, Seminarbesuch zum Thema ‚Erklären‘ an der Universität Regensburg). Schüler:innen sind aus datenschutzrechtlichen Gründen davon ausgenommen, sodass hier keine weiteren Angaben (z. B. zum Geschlecht, Alter, Biologienote) vorliegen.
4. *Überprüfung der Audioeinstellungen:* Auf dieser Seite hören die Teilnehmer:innen eine Audioaufnahme, die zur Einstellung der Lautstärke am Computer auffordert, sodass das erste Video von Beginn an mit einer adäquaten Lautstärke gestartet werden kann.
5. *Informationen zu den Rahmenbedingungen der Erklärungen:* Den Teilnehmer:innen aus den Statusgruppen der Lehrenden wurde vorab erläutert, dass bei der Aufnahme der Erklärungen

Rahmenbedingungen (ein ungefährender Zeitrahmen von drei Minuten; ohne direkte Ansprache von Schüler:innen während der Erklärung) vorgegeben waren⁹¹.

6. *Globale Bewertung der Erklärungen:* Vor der Videovignette sehen alle Teilnehmer:innen eine identische Kontextualisierung (siehe allgemein in Kap. 5.2.2.3 und im Speziellen je Videovignette in Kap. 5.2.2.5). Anschließend wird die Erklärung angesehen, wobei kein Pausieren oder Zurückspulen möglich ist. Auf der folgenden Seite können dann eine Schulnote als globale Bewertung (Notenskalierung von 1 = sehr gut bis 6 = ungenügend) und optional eine Tendenz (+/-) zur Präzisierung angegeben werden. Direkt darunter kann in einem Textfeld die globale Bewertung begründet werden („Warum hast du diese Note vergeben?“/„Warum haben Sie diese Note vergeben?“). Diese Vorgehensweise wiederholt sich für alle weiteren fünf Videovignetten identisch.
7. *Selbsteinschätzung des Vorwissens bei Schüler:innen:* Die teilnehmenden Schüler:innen werden nach jedem Video gebeten, zu beurteilen, inwieweit der Inhalt der Erklärung ihnen vorher schon vollständig, teilweise oder gar nicht bekannt war. Dieses Item dient als Kontrollvariable (siehe Kap. 5.2.5 und 6.1.1).
8. *Bewertung des Sprech- und Körperausdrucks:* Um den Sprech- und Körperausdruck einzuschätzen, wurden aus Gründen der Zumutbarkeit hinsichtlich der Bearbeitungsdauer drei repräsentative Ausschnitte aus den Videovignetten von etwa 30 Sekunden Länge ausgewählt. Babad, Avni-Babad und Rosenthal (2003) konnten zeigen, dass die Einschätzungen der nonverbalen Kommunikation anhand von zehn Sekunden langen Ausschnitten gute Prädiktoren für das Gesamtevaluationsergebnis darstellen. Die Repräsentativität hinsichtlich des sprecherischen Stils und des Körperausdrucks wurde von einer Sprechwissenschaftlerin vorab eingeschätzt. Da sowohl die stimmliche als auch die körpersprachliche Gestaltung von den situativen Umständen abhängen können, sodass auch bei einer erklärenden Person bezüglich dieses Merkmals Varianz auftreten kann (Geißner, 1988; Neuber, 2016), werden trotz gleicher Sprecherin drei Ausschnitte gezeigt. Damit zu allen Vignetten Daten bezüglich des Sprech- und Körperausdrucks vorliegen, sind die Videos zwischen den Versionen des Fragebogens (A und B) aufgeteilt, wobei je ein Set von drei Vignetten eingeschätzt wird.
9. *Sympathie, Enthusiasmus und Natürlichkeit:* Anhand der drei Videoausschnitte werden auf derselben Fragebogenseite die empfundene Sympathie sowie der wahrgenommene Enthusiasmus und die Natürlichkeit der erklärenden Person auf einer sechsstufigen Ratingskala eingeschätzt (siehe Kap. 5.2.5). Die drei Items werden zum Aspekt ‚Persönlichkeitswirkung‘ zusammengefasst (siehe Kap. 6.1.2).
10. *Informationen zum Ablauf:* Alle Statusgruppen erhalten nun die Information, dass dieselben Videos noch einmal nacheinander präsentiert und nach jeder Erklärung Fragen zu deren Qualität gestellt werden. Auch hier wird ein Richtwert von 45 Minuten als Zeitrahmen für die Bearbeitung angegeben.
11. *Teilnehmer:innencode:* An dieser Stelle sollen alle Teilnehmer:innen den zum ersten Messzeitpunkt erstellten Code erneut eingeben.
12. *Einschätzung hinsichtlich weiterer Kriterien:* Nach der Kontextualisierung und Betrachtung der Videovignette (identisch zu Messzeitpunkt 1) werden die Teilnehmer:innen aufgefordert, die Items zur sprachlichen Verständlichkeit, zur Wahrnehmung des dargestellten Komplexitätsgrades, zur Strukturiertheit und zur Adressatenorientierung zu bearbeiten. Alle

⁹¹ In den Pilotierungsinterviews zeigte sich, dass vor allem der monologische Erklärungsstil erläutert werden muss, indem er als Vorgabe vom Forschungsprojekt an die erklärende Lehrkraft eingeführt wird. Der überwiegende Konsens über die Notwendigkeit eines Einbezugs von Schüler:innen zum Beispiel in Form von kurzen Redebeiträgen dominierte sonst das Antwortverhalten.

Items können auf einer sechsstufigen Ratingskala eingeschätzt und nicht ausgelassen werden. Diese Vorgehensweise wiederholt sich ebenso für die folgenden fünf Videovignetten.

Um etwaige Reihenfolgeeffekte aufgrund der Abfolge der Videovignetten kontrollieren zu können, wurden für alle Statusgruppen durch systematische Randomisierung der Erklärungen zwei Versionen des Fragebogens (A und B) erstellt (siehe Kap. 5.2.5 und 6.1.1).

5.2.2 Videovignetten als Stimulus

Allgemein werden als Videovignetten meist kurze Ausschnitte aus Videografien realer Situationen oder fiktionaler, realitätsnahe Szenarien bezeichnet (Seifried & Wuttke, 2017). Sie finden aktuell nicht nur Verwendung als Lerngelegenheiten – beispielsweise in Form von Unterrichtsausschnitten zum Aufbau handlungsnaher Kompetenzen in der Aus- und Weiterbildung von Lehrkräften –, sondern auch in unterschiedlichen Formaten in der empirischen Bildungswissenschaft. Seidel und Thiel (2017) beschreiben für letzteres Einsatzgebiet die Dokumentation von Lehr-Lern-Prozessen, die anschließend umfassend ausgewertet werden können, und die Verwendung der Videos als Itemstimuli in standardisierten Tests zur Messung von Professionalisierungsprozessen und Kompetenzen von Lehrpersonen. Als Itemstimuli werden sie vor allem in Forschungsansätzen zur Modellierung professioneller Wahrnehmung eingesetzt – z. B. als ‚professional vision‘⁹² (Sherin, 2001; van Es & Sherin, 2002), als ‚professionelle Wahrnehmung‘⁹³ (Seidel et al., 2010) oder als ‚perception, interpretation, decision-making‘⁹⁴ (Blömeke et al., 2015) –, um authentische Kontexte und Anforderungssituationen abbilden zu können und damit eine valide Kompetenzmessung anzustreben.

In der vorliegenden Studie wurden die Videovignetten hingegen nicht als Stimuli zur Messung der Kompetenzen der Teilnehmenden, sondern als exemplarische Beispiele für Erklärungen aus dem Biologieunterricht verwendet, die die Teilnehmer:innen zu einer Reaktion evozieren: zur Bewertung der gesehenen Erklärungen hinsichtlich verschiedener Aspekte anhand geschlossener Items und Äußerungen selbstgewählter Kriterien guten Erklärens, die ohne Stimuli schwer ad hoc abrufbar und explizierbar wären. Um diese Einschätzungen gezielt hinsichtlich der Unterschiede zwischen linearen und komplexen Erklärungen zu erhalten, wurden sogenannte ‚staged-videos‘ verwendet. Damit sind solche Videovignetten gemeint, die nicht direkt aus einer Unterrichtsvideografie entstehen, sondern anhand eines vorab erstellten Drehbuchs nachgestellt werden (Piwowar, Barth, Ophardt & Thiel, 2018). So konnten der spezifische inhaltliche Schwerpunkt unter gezielter Kontrolle von variablen und konstanten Merkmalen der Erklärung umgesetzt und sechs vergleichbare Videovignetten für die Erhebung produziert werden. In realem Unterricht treten die gewünschten Bedingungen hingegen so selten auf, dass die Möglichkeit des Filmens nicht gegeben ist (L. Meyer, Seidel &

⁹² Das Modell geht auf die Untersuchungen von Goodwin (1994) zum professionellen Handeln zurück und unterscheidet für das Handlungsfeld von Lehrkräften zwei Prozesse: das selektive Wahrnehmung unterrichtsrelevanter Merkmale (noticing) und das wissensgestützt Schlussfolgern (knowledge-based reasoning). Letzteres unterteilt sich weiter in Beschreiben beobachteter Unterrichtsaspekte, Bewerten derer hinsichtlich der Qualität und Interpretieren als Ziehen von Rückschlüssen auf Wirkmechanismen.

⁹³ Die Autorinnen modellieren das ‚knowledge-based reasoning‘ als differenziertes Beschreiben (description), theoriebezogenes Erklären (explanation) und Vorhersagen zum weiteren Lehr-Lern-Prozess (prediction).

⁹⁴ Das Modell der drei Autor:innen unterscheidet neben kognitiven und affektiv-motivationalen Dispositionen (disposition) auch situationsspezifische Fähigkeiten (situation-specific skills), zu denen die Wahrnehmung (perception), Interpretation (interpretation) und Schlussfolgerung (decision making) gehört. Beide Aspekte tragen zum beobachtbaren Verhalten, also zur Performanz, bei.

Prenzel, 2006; Seifried & Wuttke, 2017). Gleichzeitig sind durch nachgedrehte Videos die datenschutzrechtlichen Vorgaben besser zu erfüllen (Piwowar et al., 2018).

„Staged-videos“ weisen im Gegensatz zu ebenfalls häufig verwendeten Textvignetten zahlreiche Vorteile auf: Die enthaltenen Informationen werden simultan und in ihrem entsprechenden Symbolsystem – visuell oder akustisch – dargeboten, wodurch die Anzahl an enthaltenen Informationen und deren Dichte erhalten bleibt (Kleinknecht & Schneider, 2013). Zudem werden keine Struktur und Relevanz von Aspekten vorgegeben, die in einem Text durch den Autor oder die Autorin redigiert werden müssen (Krammer & Reusser, 2005). Im Kontext der Authentizität und damit der Validität der Untersuchung war besonders relevant, dass nicht nur die Adressatenorientierung und Struktur der Erklärung, sondern eben auch sprachliche und vor allem sprecherische Aspekte beurteilt werden können.

5.2.2.1 Auswahl der Erklärgegenstände

Die sechs erstellten Videovignetten beinhalten unterschiedliche Erklärgegenstände. Gewählt wurden je zwei Erklärgegenstände aus den Themenbereichen Ökologie (Ö), Physiologie (P) und Genetik (G), die kausal erklärt werden. Je eine Erklärung eines Themenbereichs zeigt eine lineare Ursache-Wirkungs-Beziehung (l) und die andere ein komplexes Ursache-Wirkungs-Gefüge (k) (zur Konzeptualisierung und Operationalisierung siehe Kap. 5.2.2.2). Die zwei Erklärungen eines Themenbereichs werden im Folgenden als Videopaar und die drei Erklärungen mit derselben Ausprägung der UV als Erklärungsvarianten bezeichnet. Die sechs Erklärgegenstände sind in Tabelle 5 einzusehen.

Tabelle 5: Übersicht zu den Themen der Videovignetten

Themenbereich	Genetik	Ökologie	Physiologie
Erklärungsvariante			
Lineare Darstellung eines Phänomens (lineare Erklärungsvariante)	Entstehung von Mukoviszidose (G _l)	Entstehung der Populationsschwankungen bei Schneeschuhhasen (Ö _l)	Entstehung eines Vitamin-A-Mangels (P _l)
Komplexe Darstellung eines Phänomens (komplexe Erklärungsvariante)	Entstehung von Multipler Sklerose (G _k)	Entstehung der Populationsschwankungen bei Erdkröten (Ö _k)	Entstehung eines Folsäuremangels (P _k)

Für die Auswahl der Erklärgegenstände war leitend, dass diese für eine Jahrgangsstufe des bayerischen Gymnasiums geeignet sind, da die teilnehmenden Schüler:innen diese besuchen. Dabei sollten die Themen noch nicht behandelt worden, das Vorwissen soll als Basis für das Verstehen der Erklärung jedoch vorhanden sein. Daher sind die Erklärgegenstände so gewählt, dass sie an das in der 9. Jahrgangsstufe des bayerischen Gymnasiums erworbene Wissen anknüpfen. Für diese Einschätzung wurde der aktuell gültige Lehrplan der entsprechenden Jahrgangsstufe und Schulform herangezogen. Die Erklärgegenstände der Videovignetten sind in Themen der 10. und 11. Jahrgangsstufe zu verorten.

Neben den strukturellen Vorgaben sind Überlegungen bezüglich der Umsetzbarkeit der unabhängigen Variable (UV), also zur linearen und komplexen Darstellung, eingeflossen. Innerhalb eines Themengebiets mussten zwei ähnliche Erklärgegenstände gefunden werden, um eine entsprechende Vergleichbarkeit für das empirische Vorhaben gewährleisten zu können. Dabei

dienten für die linearen Erklärungen klassische Schulbuchbeispiele als Vorlage. Da Erklärungen, die die Komplexität und Nichtwissen sowie unsicheres Wissen gezielt aufzeigen, kaum Teil der schulischen Erklärkultur sind, sodass Schulbücher oder ähnliche Lehr-Lern-Materialien nicht als Quelle herangezogen werden können, wurden daraufhin die entsprechenden Erklärgegenstände in Abstimmung auf die linearen gewählt.

5.2.2.2 Konzeptualisierung und Operationalisierung der unabhängigen Variable ‚Grad der dargestellten Komplexität‘ in den Videovignetten

Für die Konzeptualisierung und Operationalisierung der dargestellten Komplexität (UV) in den Videovignetten wurden die Aspekte der Anzahl der einbezogenen Faktoren sowie die Thematisierung von Nichtwissen und unsicherem Wissen gewählt (theoretische Fundierung in Kapitel 3 und insbesondere 3.2). Im Folgenden wird zunächst die Auswahl der zwei Aspekte ‚Multikausalität‘ sowie ‚Nichtwissen und unsicheres Wissen‘ aus dem umfangreichen Konstrukt ‚Komplexität‘ begründet.

Multikausalität sowie ein Anteil an Nichtwissen und unsicherem Wissen sind zum einen zentral für das Konzept der Komplexität (Bar-Yam, 1997; Kuhlmann, 2007; Mainzer, 2008; Mitchell, 2008) sodass sie als „core concepts“ betrachtet werden können (Jacobson & Wilensky, 2006, S. 21). Zum anderen sind sie im Rahmen einer kurzen, instruktionalen Erklärung vermittelbar.

Obwohl die Art der Ursache-Wirkungs-Beziehungen bzw. die Nichtlinearität von Ursache-Wirkungs-Zusammenhängen ebenfalls ein bedeutendes Charakteristikum komplexer Phänomene ist (Bartelborth, 2007; Mainzer, 2008; Potochnik, 2013), wurde diese in den videografierten Erklärungen nicht näher charakterisiert, da Schüler:innen tendenziell in linearen Zusammenhängen denken (Jacobson, 2000), sodass eine grundlegende Veränderung im Denken (conceptual change) vollzogen werden müsste (Jacobson & Wilensky, 2006). Dabei stellt bereits das Zusammenwirken mehrerer Faktoren für Schüler:innen eine potenzielle Herausforderung dar (Jacobson, 2000), die die Erklärung für Schüler:innen anspruchsvoll machen kann.

Weiterhin wird auf den Aspekt der Emergenz nicht eingegangen, da dieser einen sehr hohen Abstraktionsgrad aufweist, für ein Verständnis Vorwissen über den Aufbau von Systemen notwendig ist und darüber hinaus die unterrichtliche Realisierung nur innerhalb einer größeren Zeitspanne an einem exemplarischen Beispiel möglich wäre.

Nachstehend wird auf die Operationalisierung der zwei Aspekte (Multikausalität; Nichtwissen und unsicheres Wissen) in den Erklärungen eingegangen.

In den drei Erklärungen, in denen ein linearer Ursache-Wirkungs-Zusammenhang aufgezeigt wird, wird stets nur eine Ursache aufgegriffen, die zur Phänomenentstehung führt. Die Erklärungen der komplexeren Variante führen indes mehrere Faktoren an, die zur Phänomenentstehung beitragen. In den komplexer gestalteten Vignetten wird damit kein einzelner Wirkmechanismus in den Vordergrund gestellt. Die Ausführungen zu den Zusammenhängen beschränken sich auf qualitative Aussagen, also ohne die Nennung von konkreten Verhältnissen, sodass die Faktoren nicht nur genannt werden, sondern – wo bekannt – auch der Zusammenhang bzw. Mechanismus skizziert wird (z. B. ist bei Personen, die Mukoviszidose haben, ein Gen verändert, aus dem ein Protein für den Wassertransport synthetisiert wird; durch die Veränderung des Proteins kann dieses nicht mehr in die Membran von Schleimhautzellen eingebaut werden; der Schleim wird nicht mehr durch Wasser verdünnt und ist deshalb zäher und folglich schwerer abzutransportieren).

Als weiterer Aspekt wurde die Thematisierung des Nichtwissens zwischen den Videovarianten systematisch verändert. In den linearen Erklärungen macht die erklärende Person Nichtwissen und unsicheres Wissen nicht explizit deutlich und verweist somit nicht auf unbekannte Faktoren oder Wirkmechanismen. Dadurch bekommt die Erklärung eher den Charakter, abgeschlossen und eindeutig zu sein. Diese Eindeutigkeit und Sicherheit werden in denjenigen Videovignetten aufgebrochen, die die Komplexität des Erklärgegenstandes aufzeigen, da Nichtwissen und unsicheres Wissen an einigen Stellen direkt angesprochen werden. Dies betrifft zum Beispiel unbekannte Einflussfaktoren, deren Existenz vermutet wird, weil das Phänomen mit den bisher bekannten Faktoren in seiner vielfältigen Erscheinungsform nicht gänzlich verstanden bzw. erklärt oder vorhergesagt werden kann. Weiterhin sind auch hier die Wirkmechanismen eine weitere Quelle für Nichtwissen und unsicheres Wissen. Dies kann zum Beispiel der Fall sein, wenn die Faktoren bzw. die konkrete Ausprägung der Faktoren aufgrund von Hypothesen in Untersuchungen mit aufgenommen und daraufhin (statistisch) häufig beobachtet werden, sodass ein Zusammenhang angenommen werden kann, ein konkreter Wirkmechanismus jedoch noch unbekannt ist.

Da Visualisierungen für das Verstehen eines Sachverhaltes eine zentrale Bedeutung zukommen (siehe Kap. 4.3), wurden in beiden Erklärvarianten stets Visualisierungen während der Erarbeitung der Inhalte entwickelt. Diese wurden in Hinblick auf die Vergleichbarkeit der Ergebnisse ähnlich gestaltet: Es wurden Abläufe mit Bild- und/oder Wortkarten sowie Mindmaps mit Wortkarten eingesetzt. Unsichere Zusammenhänge wurden dabei mit gestrichelten Verbindungslinien, sichere hingegen mit durchgezogenen Linien gekennzeichnet. Unbekannte Faktoren wurden durch Karten mit einem Fragezeichen symbolisiert.

Die Videovignetten werden in Kapitel 5.2.2.5 nacheinander beschrieben. Dort finden sich auch Abbildungen der Visualisierungen. Die Transkripte der verwendeten Videos sind im Anhang A einzusehen.

5.2.2.3 Gestaltungsprämissen – technische und methodische Umsetzung

Durch die steigende Anzahl an Studien mit implementierten Videovignetten liegen bereits einige Konstruktionskriterien hinsichtlich der technischen sowie methodischen Umsetzung vor (Dieker et al., 2009; Piwowar et al., 2018). Diese verweisen hinsichtlich der technischen Aufnahmebedingungen vor allem auf die qualitativ hochwertige Bild- und Audioaufzeichnung (Piwowar et al., 2018). Neben einer Kamera mit angemessener Auflösung (MP4-Format mit 1920 x 1080 / 50p)⁹⁵ wurde deshalb auch ein zusätzliches Ansteckmikrofon verwendet, um das Gesprochene störungsfrei aufzunehmen.⁹⁶

Für die Länge einer Videovignette existieren aktuell keine einheitlichen, empirisch gestützten Vorschläge. Die Dauer variiert in verschiedenen Studien zwischen eineinhalb Minuten (beispielsweise bei Schmelzing (2010) zur Testung des fachdidaktischen Wissens von Biologielehrkräften) und sechs Minuten (beispielsweise bei Friesen, Kuntze und Vogel (2018) zur Testung der Förderung von Analysefähigkeit bei Studierenden in Bezug auf den Umgang mit Repräsentationen im Mathematikunterricht). Für die Studie FALKE-B wurde eine Länge von drei Minuten als Richtwert

⁹⁵ Um ein störungsfreies Abspielen der Videovignetten während der Erhebung vor allem in Schulen zu ermöglichen, mussten die großen Dateiformate anschließend komprimiert werden (1920 x 1080/30p).

⁹⁶ Für die sprechwissenschaftlichen Analysen wurde eine zusätzliche Audiospur abgenommen und in unkomprimierten WAV-Dateien (96kHz/24bit) gespeichert, die durch die erhöhte Datenmenge eine feinere Analyse ermöglichen (Lägel-Gunga, Schilcher, Kranich & Gegner, im Druck).

festgelegt. Dadurch kann eine Erklärung gezeigt werden, die hinsichtlich der Ausführlichkeit und Tiefe angemessen ist.

Bezüglich der Kameraführung wird weiterhin empfohlen, die Perspektive des idealen Schülers oder der idealen Schülerin einzunehmen. Diese etablierte sich beispielsweise bei der TIMS-Studie (Stigler, Gonzales, Kawanaka, Knoll & Serrano, 1999). Dabei antizipiert die Kameraführung die Blickbewegungen eines Schülers oder einer Schülerin, der/die in idealer Weise dem Unterricht aufmerksam folgt (Piwowar et al., 2018). Da in der vorliegenden Studie die Wahrnehmung der Erklärung jedoch möglichst ohne Vorgaben beispielsweise durch Annahmen der Forscher:innen hinsichtlich der Aufmerksamkeitsfokussierung erfasst werden soll sowie nicht nur Schüler:innen die Videovignetten, sondern eben auch Lehrende, deren Blickführung von der des idealen Schülers abweichen kann, einschätzen, wurde die feste Einstellung gewählt.

Die Einstellungsgröße war stets die amerikanische Einstellung, bei der die Person etwa bis zum Knie zu sehen ist. Damit können sowohl größere Bewegungen der Person wie Gestik und Proxemik (die Bewegung im Raum) als auch deren Körperspannung und kleinere Bewegungen wie beispielsweise die Mimik gut erfasst werden. Unverändert bleibt über alle Erklärungen hinweg zudem das Setting: Alle Videovignetten wurden im gleichen schulähnlichen Raum der Universität Regensburg gedreht. Der untenstehende Ausschnitt (Abb. 4) zeigt die Aufnahmeperspektive in einem Standbild aus einer der Videovignetten und gewährt einen Eindruck zu Einstellung und Setting.

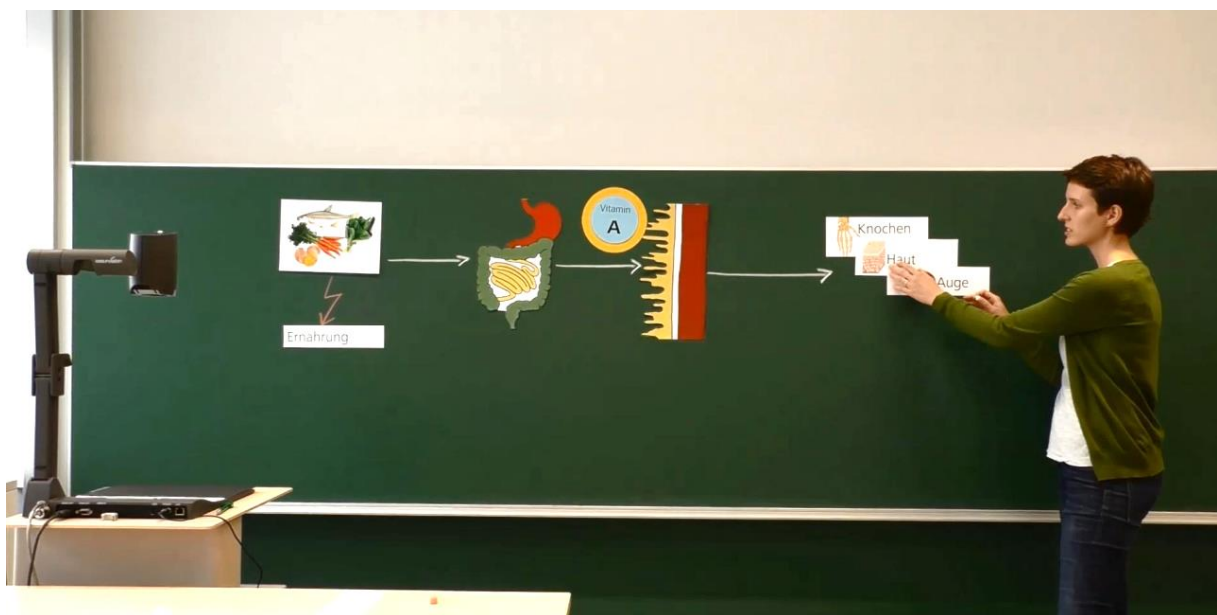


Abbildung 4: Standbild aus dem Video zur Entstehung von Vitamin-A-Mangel als Beispiel für das Setting aller Aufnahmen

Bei Pilotierungen anderer am Projekt beteiligter Fächer, deren Videovignetten zunächst direkte Interaktionen mit Schüler:innen aufwiesen, hat sich dahingehend gezeigt, dass Äußerungen der Schüler:innen als relevantes Merkmal für die Einschätzungen herangezogen werden. Da damit die Ergebnisse verzerrt werden können, zeigen die Aufzeichnungen keine Schüler:innen, sodass die Handlungen der Lehrkraft im Fokus der Einschätzungen stehen.

Rehm und Bölsterli (2014) empfehlen für die Entwicklung von nachgestellten Videovignetten, authentisches Material heranzuziehen. Auch van Vliet und Kolleg:innen (2013) betonen, dass die ökologische Validität bei ‚staged-videos‘ gegebenenfalls geringer ausfällt, wenn nicht ausreichend Informationen zur ‚best practice‘ eingeholt werden. Für die Fundierung der eingesetzten

Videovignetten wurden daher Schulbücher, Lehrerhandreichungen und vorgefertigte Arbeitsmaterialien verwendet sowie Gespräche mit Fachkolleg:innen geführt.

Weitere methodisch-konzeptionelle Überlegungen bezogen sich auf die Vergleichbarkeit der Videovignetten. Dies betrifft zum einen die gezielten Variationen in der Umsetzung der Erklärungsvarianten (UV), die sich hinsichtlich der genannten Faktoren und den Verweisen auf bestehende Wissenslücken systematisch unterscheiden. Zum anderen wurden weitere Elemente über alle Videovignetten hinweg so konstant wie möglich gehalten. Daher werden alle Erklärungen von einer Person gesprochen, um die Einflüsse personenspezifischer Merkmale wie beispielsweise Sympathie und Sprechausdruck weitestgehend zu minimieren. Zwischen den inhaltsähnlichen Videopaaren wurden weiterhin die Einstiege identisch gestaltet. So beginnt beispielsweise die Aufblende beider Videos aus dem Themengebiet Physiologie mit dem Aufgreifen der vorherigen Schüler:innenmeldung durch die Lehrkraft, indem für die Wortmeldung gedankt wird und zwei Punkte daraus wiederholt werden, an die die Problemfrage bzw. die Erklärung anknüpft. Zwischen den Themengebieten wurden die Einstiege hingegen variiert. Durch die identischen Einstiege kann die Vergleichbarkeit zwischen den Videovignetten mit variierender UV erhöht und mittels der Variation der Einstiege zwischen den Themengebieten der Ermüdung der Teilnehmer:innen durch zu hohe Gleichförmigkeit etwas entgegengewirkt werden.

Vorgeschlagen wird weiterhin der Einsatz professioneller Schauspieler:innen, da diese aufgrund ihrer Ausbildung in besonderer Weise dafür qualifiziert sind, längere Texte detailliert wiederzugeben und nach Anweisungen unterschiedliche Varianten zu realisieren (Seifried & Wuttke, 2017; Piwowar et al., 2018). In den für die Studie erstellten Videos wurde entgegen diesem Vorschlag nicht auf Schauspieler:innen zurückgegriffen, da die Erklärungen nicht wortgetreu vorformuliert werden sollten, um die sprachliche Authentizität zu erhöhen. Daher wurden lediglich grobe Skripte für den inhaltlichen Aufbau der Erklärung ausgearbeitet. Um gleichzeitig die Variation zwischen linearer und komplexer Darstellung korrekt umzusetzen, wurden die Erklärungen von der Autorin der Studie gesprochen.⁹⁷

Festgelegt wurden hingegen einige Begrifflichkeiten, auf deren Einsatz geachtet werden muss. **Beispielsweise musste der (vermutlich) neue Begriff „Autoimmunkrankheit“ als Charakterisierung der Multiplen Sklerose paraphrasiert werden.** Aus den Erklärungen zum Themenbereich Ökologie wurde der Populationsbegriff gänzlich reduziert, da angenommen werden muss, dass das Konzept der Population für Schüler:innen noch unbekannt ist. Gleichzeitig ist es eher schwer zugänglich und muss in der Regel anhand von mehreren Beispielen anschaulich gemacht werden. Für die Erklärungen wurde daher der Begriff mit „Anzahl der Tiere in einem Gebiet“ paraphrasiert.

Für die Verwendung der Videovignetten im Fragebogen wird schließlich die Anreicherung mit Informationen zum Kontext der Videosequenz empfohlen (Seidel & Prenzel, 2007). Dies gilt insbesondere für kurze Videoausschnitte, wie sie auch in diesem Projekt genutzt werden. Als Grund dafür wird genannt, dass die Teilnehmer:innen über das Geschehen in der Videovignette eine Erwartung aufbauen können. Umsetzbar ist dies zum Beispiel über einen kurzen, einführenden Text (Seidel & Prenzel, 2007). Die dafür konstruierten Texte geben den Teilnehmer:innen Informationen über das Rahmenthema, innerhalb dessen die Erklärung ausgehend von einer Problemfrage

⁹⁷ Auf persönlichen Kontakt zwischen Autorin und Teilnehmer:innen, der zu Verzerrungen der Ergebnisse beitragen könnte, wurde daher weitestgehend verzichtet. Dies betrifft vor allem die Erhebung an Schulen (siehe Kapitel 5.2.7). Zu Kolleg:innen bestand vorab bereits persönlicher Kontakt durch Tagungen oder Fortbildungen. Durch eine explizite Hervorhebung, dass die Erhebung anonymisiert durchgeführt wird, sollten auch mögliche Mildeeffekte bei der Einschätzung entgegengewirkt werden.

behandelt wird, die Aktivitäten der Klasse vor dieser Erklärung sowie einen kurzen Input zu relevantem Vorwissen. Die konkrete Kontextualisierung für das Video zur Entstehung von Multipler Sklerose sah beispielsweise folgendermaßen aus:

Die Klasse hat im Biologieunterricht gerade einen Text zur Krankheit „Mukoviszidose“ gelesen, in dem beschrieben wird, dass dabei von den Schleimhäuten – z. B. in der Lunge – ein viel zäherer Schleim gebildet wird als bei gesunden Menschen.

Insgesamt wurden so sechs Videovignetten erstellt. Die Anzahl der implementierten Vignetten ergab sich zum einen hinsichtlich der gesamten Bearbeitungsdauer respektive der Zumutbarkeit für die Teilnehmer:innen der Studie als ein Nebengütekriterium von Forschung (Bühner, 2011) und zum anderen aus der notwendigen Mindestanzahl, um ausreichende Varianz zwischen den Erklärungen abbilden zu können.

5.2.2.4 Entstehungsprozess der Videovignetten und Qualitätssicherung

Die Entwicklung der Videovignetten – von der Ideenfindung bis zum fertigen Produkt – lehnte sich an die systematische Zusammenschau von Verfahrensvorschlägen von Piwowar und Kolleg:innen (2018) sowie Dieker und Kolleg:innen (2009) an. Die schrittweise Genese ist hinsichtlich der wichtigsten Punkte in Abbildung 5 überblicksartig dargestellt und wird darunter kurz erläutert, wobei die grafisch abgegrenzten Phasen in der praktischen Arbeit teilweise überlappen.

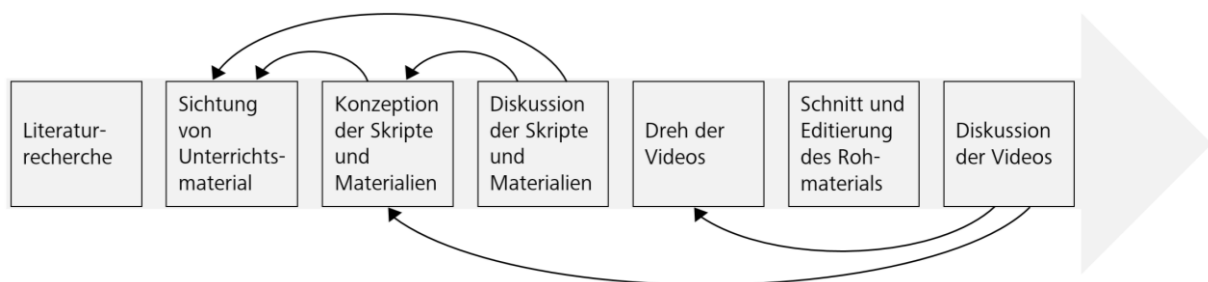


Abbildung 5: Durchlaufene Schritte der Videovignettenentwicklung in Anlehnung an Piwowar et al. (2018) und Dieker et al. (2009)

Die Literaturrecherche zu Beginn der Arbeit diente vor allen Dingen dazu, das Konstrukt der Komplexität biologischer Phänomene in seinen Facetten zu fassen, wofür insbesondere wissenschaftstheoretische Literatur herangezogen wurde (siehe Kapitel 3). Zudem wurden Texte rezipiert, die sich zum Beispiel mit dem Verstehen von komplexen Systemen und der Schwierigkeit der mentalen Repräsentation von komplexen Zusammenhängen beschäftigen (siehe Kapitel 3.4), um die Erklärungen evidenzbasiert aufbereiten zu können.

Für die Authentizität der Videovignetten wurde anschließend Unterrichtsmaterial (z. B. Schulbücher, Lehrerhandreichungen und vorgefertigte Arbeitsmaterialien) gesichtet. In dieser Phase wurden bereits geeignete und aus dem Schulalltag bekannte Beispiele für die Variation gesammelt sowie die entsprechenden Repräsentationen, die dazu verwendet werden, in Augenschein genommen.

Darauffolgend wurden geeignet erscheinende Beispiele als Ausgangspunkt genommen, um die überstehenden Themen fachlich vertiefend zu recherchieren und konkrete didaktische Umsetzungen auszuarbeiten. Dies mündete in der Erstellung großschrittiger Skripte. Bereits bei der Erstellung sowie auch in der anschließenden kritischen Diskussion der erstellten Skripte mit (Fach-)Kolleg:innen

wurden einige der Beispiele wieder vernachlässigt. Resultierend daraus wiederholten sich diese Phasen zyklisch, bis sechs den Voraussetzungen genügende Themen bzw. Beispiele vorlagen.

Nach einigen Probeaufnahmen, bei denen die technischen Einstellungen überprüft und modifiziert⁹⁸ wurden, wurden die Vignetten nach und nach aufgenommen. Da die Skripte zum Ablauf lediglich den Inhalt grob strukturierten, entstanden in mehreren Durchläufen verschiedene Varianten, die anschließend gesichtet wurden. Nach kritischer Begutachtung wurden die jeweils besten Varianten ausgewählt und bearbeitet. Die Ergebnisse wurden Kolleg:innen der Projektgruppe und der Biologiedidaktik zur Diskussion vorgelegt, um weitere Ansatzpunkte für Überarbeitungen zu erhalten. Feedback wurde sowohl von fachinternen als auch fachexternen Personen eingeholt, um ein möglichst breites Spektrum zu erhalten. Dabei ist vor allem der fachfremde Blick lohnenswert, da durch die Diskussionen auch im eigenen Fach routinierte Vorgehensweisen oder sprachliche Wendungen nochmals reflektiert werden. Um die inhaltliche Validität – die Übereinstimmung der Konzeptualisierung der unabhängigen Variable ‚Komplexität‘ mit den Realisierungen in den Erklärungen – zu prüfen, wurden in einer Pilotierungsphase die Videos noch einmal Biologiedidaktiker:innen und Lehramtsstudierenden mit Hauptfach Biologie vorgelegt (siehe Kap. 5.2.6). Insgesamt dienen die dargestellten Rückkopplungsschleifen im gesamten Planungsprozess dazu, die Qualität der Vignetten zu sichern bzw. herzustellen.

5.2.2.5 Beschreibung der Videovignetten

In diesem Kapitel werden die Videovignetten in ihrem Ablauf beschrieben sowie die wichtigsten Bezugspunkte im Sinne einer fachlichen Klärung angeschlossen. Anschließend werden die Vignetten bezüglich der Umsetzung der UV (Grad der dargestellten Komplexität) kommentiert. Die vollständigen Transkripte finden sich in Anhang A.

a) Entstehung von Mukoviszidose (G_i)

Der kontextualisierende Text, der allen Teilnehmenden vor dem Betrachten des Videos gezeigt wurde, lautet:

Die Klasse hat im Biologieunterricht gerade einen Text zur Krankheit „Mukoviszidose“ gelesen, in dem beschrieben wird, dass dabei von den Schleimhäuten – z. B. in der Lunge – ein viel zäherer Schleim gebildet wird als bei gesunden Menschen.

Zum Ablauf der Erklärung:

Der Einstieg in das Video zu Entstehung von Mukoviszidose erfolgt mittels Verweises auf diesen vorangegangenen Text, nach dessen Lesen die Frage offengeblieben ist, was die Ursache von Mukoviszidose ist:

Was der Text jetzt allerdings noch nicht erklärt hat, ist, was die Ursache der Mukoviszidose ist oder warum der Schleim auf einmal zäher ist. Und das schauen wir uns jetzt noch gemeinsam an.
(Zitat aus dem Video G_i)

Anschließend wird von der erklärenden Person die Frage „Wie entsteht Mukoviszidose?“ an der Tafel notiert (die Zeit für den Tafelanschrieb wurde durch einen Videoschnitt verkürzt). Die Frage

⁹⁸ Vor allem hinsichtlich der Aufnahmeperspektive musste die Wirkung und die Sichtbarkeit aller relevanten Elemente in einem dem Endprodukt ähnlichen Video überprüft werden. Dazu zählen beispielsweise die gleichzeitige Aufnahme von Gestik, Mimik und Proxemik.

markiert den Einstieg in die Erklärung. Abbildung 6 zeigt das während der Erklärung entwickelte Tafelbild, auf dessen Elemente in der folgenden Beschreibung Bezug genommen wird.

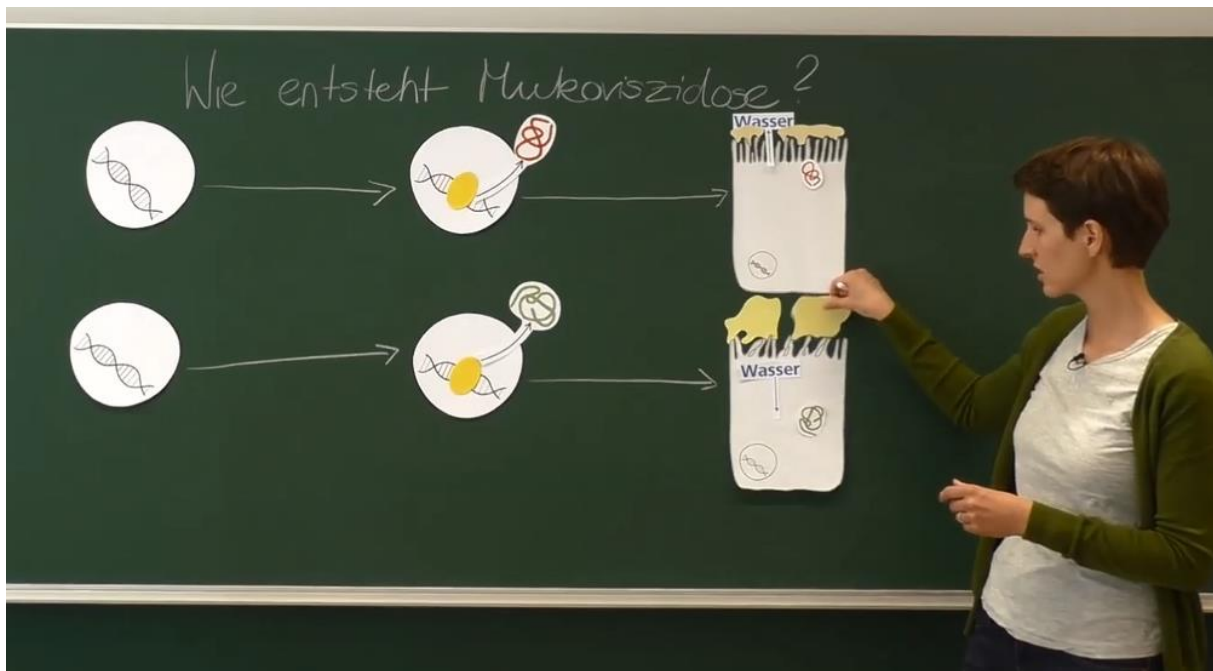


Abbildung 6: Screenshot aus dem Video zur Entstehung von Mukoviszidose (Minute 2:20)

Die erklärende Person ordnet Mukoviszidose zunächst als genetisch bedingte Krankheit ein. Darauf folgt eine Erklärung des grundlegenden Ablaufs der Genexpression: Gene tragen Informationen, die zur Herstellung von Proteinen dienen. **Der Begriff Gen wird dabei durch „Abschnitte auf der DNA“ paraphrasiert (verkürzte Worterklärung).** Der Begriff der Herstellung wird durch eine Paraphrase mit **dem Fachbegriff „Synthese“ bzw. „synthetisieren“ ergänzt (verkürzte Worterklärung).** Die Sprecherin nutzt die Erklärung zur Genexpression gleichzeitig dazu, Teile der Visualisierung einzuführen, indem sie diese währenddessen an die Tafel heftet und benennt: Es wird zunächst die Repräsentation einer Zelle mit einem vergrößerten DNA-Abschnitt angeheftet. Anschließend wird die Transkription und Translation verkürzt und ohne explizite Nennung der Begriffe durch eine zweite Repräsentation einer Zelle mit vergrößertem DNA-Abschnitt symbolisiert, bei der ein nicht näher bestimmtes Protein entsteht. Erst nach der Skizzierung der Proteinbiosynthese wird auf die Krankheit der Mukoviszidose eingegangen, wobei erst die Funktionsweise bei nicht erkrankten Personen beschrieben wird. Die erklärende Person beschreibt dazu zunächst die Funktion des Proteins: Es wird in die Schleimhautzelle der Lunge **„eingebaut“** und ist für einen **indirekten Wassertransport aus der Zelle** verantwortlich. Das Wasser verdünnt den Schleim außerhalb der Schleimhautzelle und macht ihn so durch die Flimmerhärchen gut transportierbar, bis er abgehustet werden kann. Der Abtransport durch die Flimmerhärchen wird mit der Metapher des Fließbandes verdeutlicht. Nach Abschluss der Erklärung zur Funktion des Proteins erklärt die Sprecherin die Dysfunktion bei an Mukoviszidose erkrankten Personen vergleichend und baut dabei eine analoge Repräsentation schrittweise auf: Es wird genannt, dass das Gen bei erkrankten Personen anders aussieht, aber dennoch abgelesen und auch ein Protein synthetisiert wird. Dieses Protein wird als verändert beschrieben und als Ursache dafür die veränderten Informationen des Gens benannt. Daraufhin wird das veränderte Aussehen des Proteins als Ursache für die Dysfunktion angeführt – das Protein kann nicht mehr in die Schleimhautzelle eingebaut werden. Als Folge werden das Ausbleiben des indirekten Wassertransports aus der Zelle und daraufhin das Resultat eines zäheren Schleims genannt, der nicht mehr so gut abtransportiert und schließlich nicht mehr vollständig abgehustet werden kann.

Die erklärende Person überträgt schließlich das Gesagte, indem sie anfügt, dass die Klärung von Ursachen – also das Finden von Genen und ihren Mutationen – für die Entwicklung von Medikamenten und Behandlungsmethoden von hoher Bedeutung ist. Sie schließt die Erklärsituation mit dem Verweis auf eine Folgeaktivität der Klasse, bei der Behandlungsmethoden von Mukoviszidose in den Blick genommen werden sollen.

Zum fachlichen Hintergrund der Erklärung:

Grundlegend entsteht Mukoviszidose durch eine Genmutation, wobei das betroffene Gen für ein Membranprotein (CFTR = cystic fibrosis transmembrane regulator) codiert. Am häufigsten ist die **Mutation $\Delta F508$** , durch die das Membranprotein derart strukturell verändert wird, dass dieses nicht mehr in den Membranen eingebaut werden kann – zum Beispiel in der apikalen Membran der Lungenepithelzellen. Da dieses Membranprotein unter anderem indirekt für den Wassertransport aus der Zelle verantwortlich ist, ist die Wasserabsorption unterbunden und es tritt sogar eine erhöhte Wasserrückresorption auf (Bosch & Boeck, 2016).

Die **Mutation $\Delta F508$** ist jedoch nicht bei allen Patient:innen nachweisbar, sondern es konnten bisher (Stand 2016) insgesamt 161 weitere Mutationen innerhalb des DNS-Abschnitts des CFTR Membranproteins gefunden werden (Bosch & Boeck, 2016; Savant & McColley, 2017). Die Art der Mutation ist zum einen für die Art der Dysfunktion des Membranproteins (Verhinderung des Einbaus in die Membran, Bildung von dysfunktionalen Rezeptoren) und zum anderen bedingt dadurch auch für den Verlauf und die Schwere der Symptomatik verantwortlich. Bisher fehlen Befunde, um einige der Mutationen einzuordnen und insbesondere deren Wirkmechanismen zu beschreiben (Bosch & Boeck, 2016).

Aufgrund variabler Symptome bei erkrankten Personen werden noch weitere exogene Faktoren vermutet, die die Effekte der Genmutation moderieren und damit das Krankheitsbild beeinflussen (Savant & McColley, 2017).

Kommentar zur Umsetzung der UV:

Die Erklärung ist insgesamt als linear zu beschreiben, da ein einzelnes Gen als Ursache für die Entstehung von Mukoviszidose angeführt wird. Im Zentrum steht ein einzelner kausaler Zusammenhang – der des Gens und des Wassertransports aus der Schleimhautzelle (in diesem Fall eine Epithelzelle der Lungenschleimhaut). Die Komplexität des natürlichen Phänomens wird insofern (sektoral) reduziert, als das weitere genetische Mutationen sowie exogene Faktoren, die die Schwere der Symptomatik moderieren, nicht aufgenommen werden. Verweise auf unsicheres Wissen oder Nichtwissen wie beispielsweise die Wirkmechanismen von Genmutationen und Veränderungen im CFT-Protein oder weitere Faktoren werden von der erklärenden Person nicht angegeben, sodass die Erklärung eher sicher und eindeutiger erscheint, als sich das natürliche Phänomen für die Wissenschaft aktuell darstellt.

b) Entstehung von Multipler Sklerose (G_k)

Der folgende kontextualisierende Text, konnte vor der Betrachtung von allen Teilnehmenden gelesen werden:

Die Klasse hat im Biologieunterricht gerade einen Text zur Krankheit „Multiple Sklerose“ gelesen, in dem beschrieben wird, dass dabei die Nervenzellen geschädigt werden.

Zum Ablauf der Erklärung:

Der Einstieg in das Video zur komplexen Erklärung im Themenbereich der Genetik ist so ähnlich wie möglich zum Einstieg in das Video zur linearen Erklärung aufgebaut und bezieht sich deshalb ebenso auf einen Text, den die Klasse vorab gelesen hat:

Was der Text noch nicht erklärt hat, ist, wie Multiple Sklerose eigentlich entsteht. Und das wollen wir jetzt gemeinsam noch machen. (Zitat aus dem Video G₄)

Die erklärende Person notiert anschließend auch hier die Frage „Wie entsteht Multiple Sklerose?“ an der Tafel (die Zeit für den Tafelanschrieb wurde ebenfalls durch einen Videoschnitt verkürzt). Sie markiert den Einstieg in die Erklärung. In Abbildung 7 ist das entwickelte Tafelbild ersichtlich, auf dessen Elemente in der darauffolgenden Beschreibung verwiesen wird.

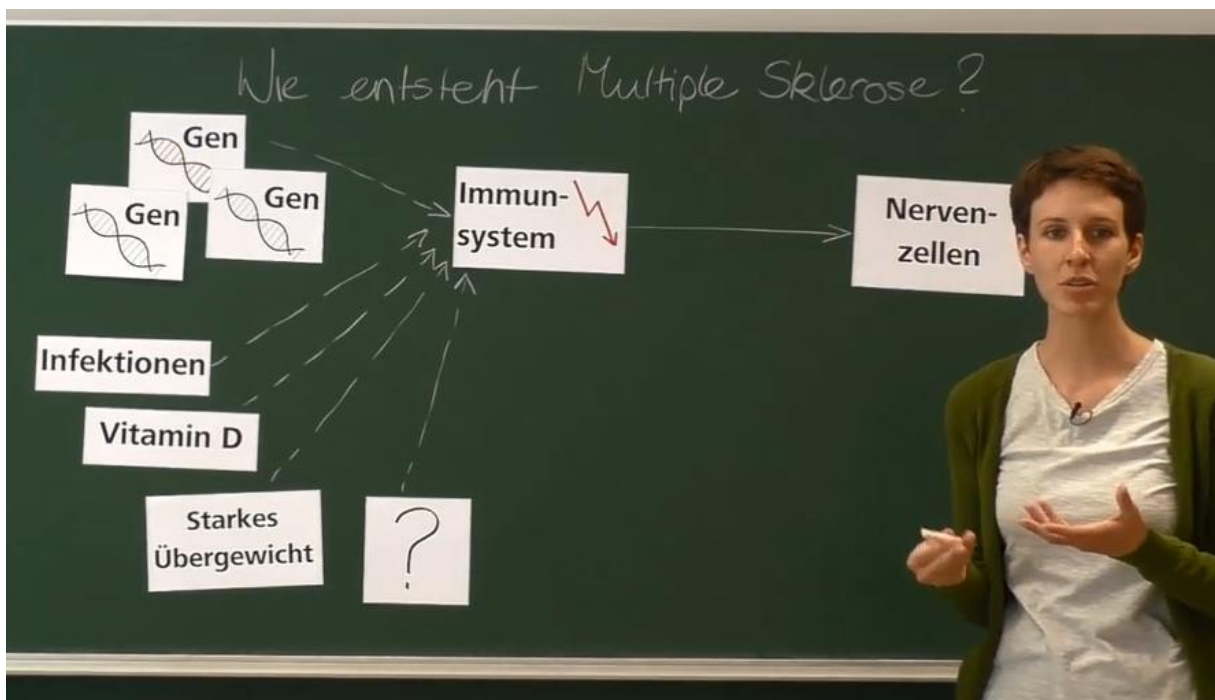


Abbildung 7: Screenshot aus dem Video zur Entstehung von Multipler Sklerose (Minute 2:36)

Multiple Sklerose wird zunächst als Autoimmunkrankheit eingeordnet. Daraufhin folgt eine allgemeine Erklärung, was eine Autoimmunkrankheit ist, bei der die Funktion des Immunsystems grob skizziert (zum Beispiel Viren und Bakterien unschädlich zu machen, die in den Körper eindringen) und anschließend die Dysfunktion des Immunsystems bei einer Autoimmunkrankheit allgemein (das Angreifen körpereigener Zellen) sowie bei Multipler Sklerose im Speziellen (das Angreifen von Zellen des Nervensystems) genannt wird. Die Sprecherin führt anschließend an, dass diese unmittelbare Ursache für das Krankheitsbild der Multiplen Sklerose, die Dysfunktion des Immunsystems, auch wiederum eine Ursache hat. Mit einem Bezug darauf, dass Forschende und Ärzt:innen daran interessiert sind, die Ursachen aufzuklären, leitet sie zu den Ursachen der Dysfunktion des Immunsystems über. Als erster Faktor wird angeführt, dass bereits Gene identifiziert wurden, die dazu beitragen, dass Multiple Sklerose entsteht. Es wird explizit daraufhin gewiesen, dass es sich hierbei nicht um ein einzelnes Gen, sondern um mehrere Gene handelt. Analog zur linearen Erklärung im Themenbereich der Genetik (Entstehung von Mukoviszidose) wird auch hier der Zusammenhang zwischen der Veränderung von Genen und der Synthese von veränderten Proteinen erklärt, die dadurch nicht mehr ihre eigentliche Funktion im Organismus erfüllen können. Daraufhin führt die erklärende Person als weiteren Faktor an, dass sich an Multipler Sklerose erkrankte Personen von gesunden Personen unterscheiden, indem sie früher mit anderen

Krankheiten infiziert waren. Sie nennt außerdem den Vitamin-D-Spiegel und ein starkes Übergewicht als untersuchte Faktoren. Da die Wirkmechanismen der Faktoren früherer Infektionen, des Vitamin D Spiegels und eines starken Übergewichts auf die Funktion des Immunsystems nicht beschrieben sind bzw. mehrere Hypothesen darüber existieren, die im Rahmen der kurzen instruktionalen Erklärung nicht dargestellt werden können, bleibt es bei einer Nennung der Faktoren. Die erklärende Person fügt nach der Nennung der Faktoren explizit an, dass die Wirkmechanismen dieser Faktoren nicht gänzlich bekannt sind und weitere Faktoren vermutet werden, die noch nicht bekannt sind.

Wie in der linearen Vignette im Themenbereich der Genetik (Entstehung von Mukoviszidose) verweist die erklärende Person auf die Relevanz, Faktoren zu erforschen, um neue und effektive Medikamente sowie Behandlungsmethoden zu entwickeln. Im Gegensatz zur linearen Erklärung wird an dieser Stelle jedoch darauf hingewiesen, dass die Behandlung durch das Nichtwissen deutlich erschwert ist.

[...] für Mediziner und Ärzte ist es natürlich unglaublich spannend, diese Faktoren aufzudecken, denn nur so kann man neue Behandlungsmethoden und Medikamente auf den Weg bringen. Aber im Fall der Multiplen Sklerose ist das gar nicht so einfach, weil es ebenso viele Unbekannte noch gibt. Also welche Faktoren noch dazu kommen oder wie das jetzt genau zusammenhängt. (Zitat aus dem Video G_k)

Im Anschluss wird auch in dieser Vignette auf die weiterführende Tätigkeit der Klasse verwiesen, bei der die Schüler:innen Behandlungsmethoden kennenlernen werden.

Zum fachlichen Hintergrund der Erklärung:

Zur ursächlichen Entstehung von Multipler Sklerose ist bekannt, dass verschiedene genetische Dispositionen (vor allem Gene, die bei der Immunregulation beteiligt sind) für den Ausbruch der Krankheit verantwortlich sind (Dendrou, Fugger & Friese, 2015; Melief et al., 2015). Dabei ist laut Dendrou und Kolleg:innen (2015) das Erkrankungsrisiko für Multiple Sklerose lediglich zu 30 % auf genetische Dispositionen zurückzuführen. Darüber hinaus werden weitere Faktoren untersucht, die zur Entstehung der Krankheit beitragen. Dazu zählen ein zu geringer Vitamin-D-Spiegel, der wahrscheinlich über die immunmodulierenden Eigenschaften des Vitamins Einfluss nehmen kann (da Costa et al., 2016) sowie eine genetisch beeinflusste Adipositas (Harroud et al., 2021). Als Risikofaktoren werden aktuell auch einige Vorerkrankungen durch virale Infektionen (Dendrou et al., 2015; Farez & Correale, 2011), die Sensitivität eines Glucocorticoid Rezeptors, der die Aggressivität der Multiplen Sklerose beeinflusst (Melief et al., 2015), und das Geschlecht (Dendrou et al., 2015) als Faktoren untersucht. Die konkreten Wirkmechanismen sind zum größten Teil nicht beschrieben.

Bei der Entstehung der Multiplen Sklerose ist nicht mit Sicherheit bekannt, welche Art von Faktor (endogen oder exogen) den maßgeblichen Anstoß für den Krankheitsausbruch gibt (Dendrou et al., 2015).

Kommentar zur Umsetzung der UV:

Die Erklärung nimmt zwei Charakteristika komplexer Phänomene auf: die Beteiligung mehrerer Faktoren an der Entstehung eines Phänomens sowie das Vorhandensein von unsicherem Wissen und Nichtwissen. So wird nicht nur eine Ursache – zum Beispiel ein Gen – für die Entstehung genannt, sondern es werden neben mehreren nicht konkreter bestimmten Genen drei weitere Faktoren aufgegriffen. Dennoch ist hierbei die Anzahl der Einflussfaktoren reduziert (z. B. die

Sensitivität des Glucocorticoid Rezeptors und des Geschlechts). Insgesamt wird keine einzelne zentrale Ursache angeführt und damit das Denken im ‚decentralized mindset‘ notwendig.

Der Aspekt des Nichtwissens wird aufgenommen, indem die unbekanntes Wirkmechanismen explizit als Wissenslücken benannt werden. Auch auf die Vermutung weiterer Faktoren wird verwiesen und damit das gesamtwissenschaftliche Nichtwissen angesprochen. Aufgrund des gesamtwissenschaftlichen Nichtwissens können die Erklärungen in Bezug auf die Wirkmechanismen nicht konkreter werden, sodass die Erklärung im Gegensatz zur Erklärung der Entstehung von Mukoviszidose weniger abschlossen und eindeutig wirkt.

c) Entstehung der Populationsschwankungen bei Schneeschuhhasen (Ö)

Vor dem Video erhielten alle Teilnehmenden den folgenden kontextualisierenden Text:

Die Klasse hat im Biologieunterricht gerade ein Diagramm besprochen, das die Anzahl der Schneeschuhhasen in der Hudson Bay (Kanada) von 1860 bis 1900 zeigt.

Zum Ablauf der Erklärung:

Beim Einstieg in das Video zur Entstehung der Populationsschwankungen bei Schneeschuhhasen ist das Liniendiagramm zu sehen, das die Anzahl der Schneeschuhhasen in diesem Zeitraum zeigt. Dieses wird aufgegriffen und es folgt ein Verweis auf die nicht konstante Anzahl an Tieren bzw. die überraschenden Schwankungen der Anzahl der Tiere, die so oder so ähnlich aus einem vorhergehenden Gespräch mit den Schüler:innen resultieren könnte:

Und was vielleicht jetzt ein bisschen verblüfft oder erstaunlich ist, dass gar nicht immer gleich viele Hasen dort vorkommen, sondern wir tatsächlich sehr starke Schwankungen haben. Wenn ihr guckt, hier, 1864, da haben wir einen ziemlichen Peak. Da sind sehr viele Hasen in diesem Gebiet gesichtet worden. [...] Und nur wenige Jahre später, hier 1870, gibt es nur noch sehr, sehr wenige Hasen. (Zitat aus dem Video Ö)

Darauffolgend leitet die erklärende Person in die Erklärung ein und formuliert mündlich die Frage:

Und was man sich jetzt fragen kann, ist: Wie kommt es eigentlich, dass es zunächst sehr viele Hasen gibt und es nur wenige Jahre später nur noch sehr, sehr wenige sind. Und genauso umgekehrt, warum steigt eigentlich die Anzahl dann auch wieder? (Zitat aus dem Video Ö)

Der Einstieg wird auch dazu genutzt, bereits einen Teil der Visualisierung zu benennen und aufzubauen. Die vollständige Visualisierung des modellhaften Populationszyklus ist in Abbildung 8 ersichtlich.

So heftet die erklärende Person bei der Nennung, dass es zu einem Zeitpunkt sehr viele Hasen in dem Gebiet gibt, eine Bildkarte als Repräsentation dafür an die linke Seite der Tafel, auf der symbolisch sechs Hasen abgebildet sind. Bei der Nennung, dass es ein paar Jahre später nur noch sehr wenige Hasen in diesem Gebiet gibt, wird auf die rechte Seite der Tafel als Repräsentation eine Bildkarte mit nur zwei Hasen angeheftet. Dabei wird außerdem ein durchgehender Pfeil von der Bildkarte mit vielen Hasen zu der Bildkarte mit wenigen Hasen über die obere Tafelhälfte und umgekehrt auf der unteren Tafelhälfte gezeichnet.



Abbildung 8: Screenshot aus dem Video zur Entstehung von Populationsschwankungen bei Schneeschuhhasen (Minute 3:06)

Die erklärende Person führt anschließend ein, dass man zur Ursachenklärung zunächst den Lebensraum der Schneeschuhhasen betrachten muss. Daraufhin informiert sie die Zuhörer:innen, dass es in der Hudson Bay auch Luchse gibt, woraufhin die Luchse und ihre Lebensweise grob skizziert werden: Die Art wird in die Familie der Katzen eingeordnet und als räuberisch lebend charakterisiert. Letzteres wird mit dem Begriff „fleischfressend“ in einer Paraphrase erläutert.

Genannt wird, dass die Schneeschuhhasen die Hauptnahrungsquelle der Luchse in diesem Gebiet sind. Die erklärende Person fordert daraufhin die Schüler:innen auf, sich vorzustellen, was es bedeutet, wenn Räuber ein großes Nahrungsangebot vorfinden. Nach einer kurzen Pause löst die erklärende Person auf und fügt an, dass bei hohem Nahrungsangebot – also vielen Schneeschuhhasen bzw. Beutetieren – viele Räubertiere überleben, da weniger Räubertiere an Nahrungsmangel verenden sowie Jungtiere gut ernährt werden können. Daraufhin wird eine Bildkarte mit vielen Luchsen in die Mitte des Pfeils geheftet, der von der Repräsentation vieler Schneeschuhhasen zu der weniger Schneeschuhhasen führt. Zusammenfassend und verallgemeinernd wird anschließend formuliert, dass zunächst umso mehr Räuber in diesem Gebiet zu finden sind, je mehr Beutetiere in einem Gebiet zur Verfügung stehen. Eine entsprechende Wortkarte („Je mehr Beute, desto mehr Räuber“) wird angeheftet.

Anschließend wird der nächste Abschnitt des Populationszyklus in den Blick genommen: Die Sprecherin stellt den Bezug zwischen vielen Räubertieren und sinkender Anzahl an Beutetieren durch die unmittelbare Nahrungsbeziehung her. Auch an dieser Stelle wird mit einer Wortkarte („Je mehr Räuber, desto weniger Beute“) verallgemeinert und der Bezug zur sichtbaren, abnehmenden Kurve im Liniendiagramm hergestellt.

Analog dazu wird der nächste Abschnitt im Populationszyklus zunächst mit der Frage eingeleitet, was anschließend mit der Anzahl der Luchse passiert. Die erklärende Person löst nach einer kurzen

Redepause auf, dass die Luchse weniger Nahrung finden und folglich die Anzahl der Tiere aufgrund von Nahrungsknappheit **in diesem Gebiet sinkt** („Je weniger Beute, desto weniger Räuber“). Der Populationszyklus wird vervollständigt, indem auch der letzte Abschnitt erst in Bezug auf die Anzahl von Schneeschuhhasen und Luchsen angeführt und dann allgemein formuliert wird: „Je weniger Räuber, **desto mehr Beute**“.

Die erklärende Person schließt die Erklärung zur Ursache der beobachteten Schwankungen in der Anzahl der Schneeschuhhasen, in dem sie zusammenfasst, dass diese Schwankungen durch die Abhängigkeit bzw. Nahrungsbeziehung entstehen und Vorhersagen zur weiteren Entwicklung von Interesse sein können.

Zum fachlichen Hintergrund der Erklärung:

Ein Modell, mit dem eine grundlegende Populationsdynamik beschrieben werden kann, ist das Lotka-Volterra-Modell (auch Lotka-Volterra-Gleichungen). Das mathematische Modell stellt die Abhängigkeit der Populationsgröße der Beutetiere von der Populationsgröße der Räuber idealisiert dar bzw. nimmt die Abhängigkeit derer idealisiert an. Voraussetzung für die Anwendbarkeit des Modells ist, dass sich Räuber ausschließlich von einer Art Beutetier ernähren und diese Beutetiere keiner weiteren Art Räuber als Nahrung dienen. Weiterhin wird angenommen, dass sich die Beutepopulation ohne die Anwesenheit der Räuber exponentiell entwickelt (Begon, Howarth & Townsend, 2017).

Das Modell wird allgemein durch drei Regeln beschrieben (die erste Regel wird im Video aufgegriffen):

- i) Periodizität: Die Populationsgrößen von Räuber und Beute schwanken periodisch.
- ii) Erhaltung der Mittelwerte: Die Populationsgrößen von Räuber und Beute schwanken konstant um einen jeweils festen Mittelwert.
- iii) Störung der Mittelwerte: Wenn die Anzahl der Räuber- und Beutetiere gleichermaßen und proportional zu ihren Ausgangswerten dezimiert werden, erholt sich die Population der Beutetiere schneller.

Das Lotka-Volterra-Modell liefert zwar eine wichtige Grundlage für die theoretische Ökologie, die Anwendbarkeit muss jedoch anhand der Annahmen des Modells im Einzelfall geprüft werden. Abiotische und weitere biotische Faktoren, die einen Einfluss auf die Populationsgrößen ausüben, werden nicht berücksichtigt (Begon et al., 2017). Eine Beschreibung der Schwankungen der Populationsgrößen sowie eine Vorhersage sind dann folglich nicht mehr möglich, da die Grundannahmen des Modells verletzt sind.

In nur wenigen Fällen sind die Voraussetzungen des Modells erfüllt, sodass die Lotka-Volterra-Gleichungen zur Erklärung und Vorhersage der Populationsgrößen verwendet werden können. Eines dieser Beispiele ist die Nahrungsbeziehung von Schneeschuhhasen (*Lepus americanus*) und Luchsen (*Lynx canadensis*) im Norden Kanadas (Hudson Bay), wobei sich die Luchse auf die Schneeschuhhasen als Beute spezialisiert haben (Campbell & Reece, 2009; Begon et al., 2017).⁹⁹ Die Populationsschwankungen durchlaufen innerhalb von zehn Jahren einen Zyklus. Zur Klärung, warum der Zyklus in einem Zeitintervall von zehn Jahren schwankt, wurden Experimente durchgeführt. Die Resultate der untersuchten Hypothesen zeigen, dass ein Nahrungsmangel für die herbivoren Schneeschuhhasen als Todesursache ausgeschlossen werden kann, jedoch 90 % der

⁹⁹ Selbst wenn zwischen Räuber- und Beutepopulation eine Ähnlichkeit in ihren Zyklen beobachtet werden kann, muss das nicht bedeuten, dass die Ursache eine bestehende Räuber-Beute-Beziehung ist (Begon, Howarth & Townsend, 2017).

Schneeschuhhasen starben, weil sie als Nahrung für Luchse endeten. Ein weiterer, wenn auch kleiner Effekt konnte für die Sonnenfleckenaktivität¹⁰⁰ nachgewiesen werden: Die Sonnenfleckenaktivität schwankt in einem ähnlichen Zyklus, wobei bei niedriger Sonnenfleckenaktivität weniger Ozon in der Atmosphäre entsteht und folglich mehr UV-Strahlung an die Erdoberfläche gelangt. Pflanzen regt die höhere UV-Strahlung zur Produktion von sie davor schützenden Substanzen an, wobei gleichzeitig weniger Substanzen zum Fraßschutz vor Herbivoren produziert werden (Campbell & Reece, 2009). Die Nahrungsqualität für die Schneeschuhhasen steigt dabei an.

Kommentar zu Umsetzung der UV:

Die Erklärung greift das spezifische Beispiel in der Hudson Bay auf, für das diese Art von Zusammenhang zwischen Räuber- und Beutepopulation, die in Ökosystemen nur in wenigen Fällen isoliert zu beobachten ist und deshalb eher Modellcharakter besitzt, einen erklärenden Wert hat, also die Populationsdynamik erklären kann. Die Erklärung spiegelt dabei einen niedrigen Grad an Komplexität wider, da vier analog geartete lineare Ursache-Wirkungs-Zusammenhänge aufgezeigt werden, die jeweils die Abhängigkeit der Populationsgröße von Räuber und Beutetieren über die bestehende Nahrungsbeziehung verdeutlichen. Dabei steht nicht die Vorgehensweise zur Ermittlung solcher Abhängigkeiten bzw. die Aufklärung der Populationsverläufe durch wissenschaftliche Erkenntnismethoden (z. B. Datenerfassung, Hypothesenformulierung) – sie werden in der Erklärung nur angerissen –, sondern die Illustration des allgemeinen Zusammenhangs bzw. des allgemeinen Prinzips der Abhängigkeit von Räuber- und Beutetierpopulation im Fokus.

Weitere Faktoren, die in diesem konkreten Fall als moderierende Effekte untersucht und teilweise nachgewiesen wurden (Verfügbarkeit von Nahrung im Winter; Einfluss der Sonnenfleckenaktivität auf die Nahrungsqualität für Schneeschuhhasen), werden nicht erwähnt, da die erste Regel des Lotka-Volterra-Modells im Vordergrund steht. Auf unsicheres Wissen und Nichtwissen oder auf Kritik bzw. die eingeschränkte Gültigkeit des Modells wird im Rahmen der Darstellung des allgemeinen Prinzips der Abhängigkeit von Räuber- und Beutepopulation ebenfalls nicht eingegangen. Die Erklärung ist so insgesamt linear gestaltet und wirkt inhaltlich abgeschlossen.

d) Entstehung der Populationsschwankungen bei Erdkröten (Ö_k)

Analog zum kontextualisierenden Text des linearen Videos aus dem Themenbereich der Ökologie erhielten die Teilnehmenden auch hier einen Text mit den entsprechenden Informationen:

Die Klasse hat im Biologieunterricht gerade ein Diagramm besprochen, das die Anzahl der Erdkröten in der Stadt Merklingen von 1900 bis 2015 zeigt.

Zum Ablauf der Erklärung:

Beim Einstieg in das Video zur Entstehung von Populationsschwankungen bei Erdkröten ist ebenfalls ein Liniendiagramm zu sehen, das die Anzahl von Individuen einer Art in einem bestimmten Gebiet zeigt (Erdkröten im Zeitraum von 1990 bis 2015 in der Stadt Merklingen). Auch hier folgt ein Verweis auf die nicht konstante Anzahl der Tiere, um die Fragestellung abzuleiten:

Und was vielleicht ein bisschen überrascht, ist, dass es nicht immer gleich viele Erdkröten in diesem Gebiet gibt. Sondern eben mal sehr viele, wenn ihr hier guckt, 1994 oder 1995 gibt es einen ganz

¹⁰⁰ Sonnenflecken sind auf der Oberfläche der Sonne als dunkle Flecken sichtbar und strahlen weniger Sonnenlicht ab als die übrige Fläche der Sonne, wodurch die Flecken kühler sind. Sie entstehen durch Verzerrungen des Magnetfelds der Sonne, das die Plasmaströme an der Oberfläche verändert.

großen Ausschlag, da hat man sehr viele Erdkröten gefunden. Und nur wenige Jahre später, so 2000, sind es deutlich weniger. (Zitat aus dem Video Ök)

Die erklärende Person formuliert daraufhin als Einstieg in die Erklärung die Frage:

Und genau damit wollen wir uns jetzt beschäftigen. Also warum es diese Schwankungen gibt, welche Ursachen oder welche Gründe dann da dahinterstecken, dass es mal mehr und mal weniger gibt. (Zitat aus dem Video Ök)

Die folgende Abbildung 9 zeigt das während der Erklärung entwickelte Tafelbild, auf das in der unten folgenden Beschreibung Bezug genommen wird.

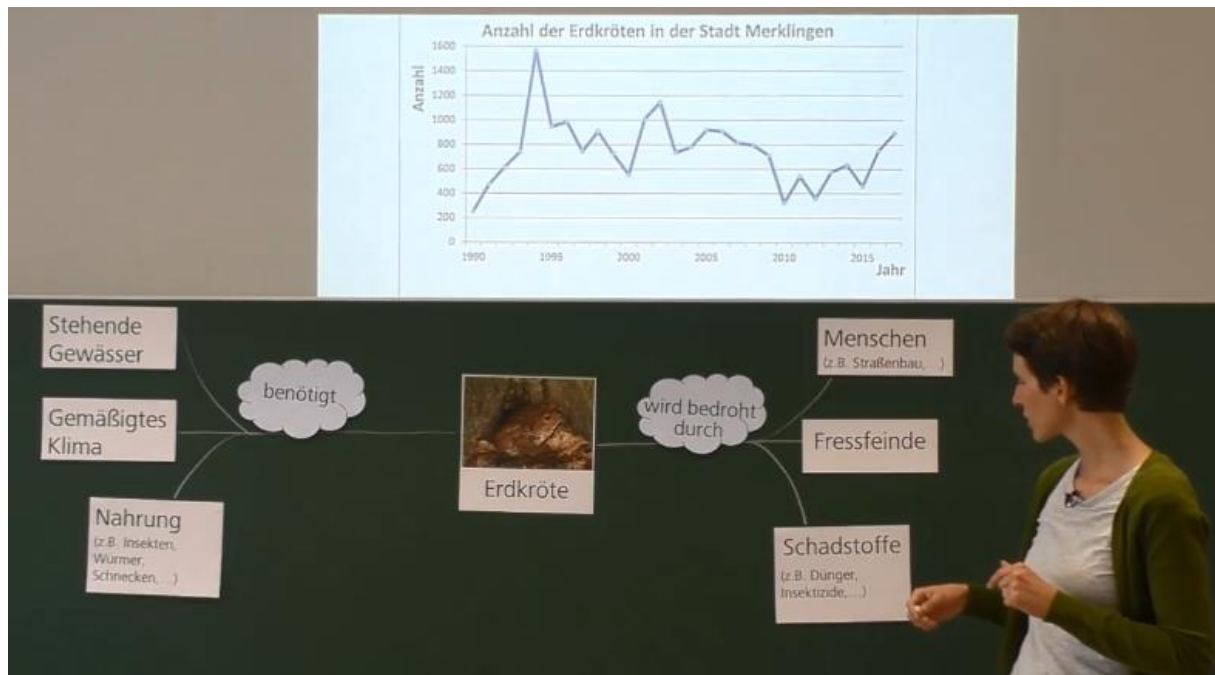


Abbildung 9: Screenshot aus dem Video zur Entstehung von Populationsschwankungen bei Erdkröten (Minute 2:43)

Auch in der komplexen Erklärung zur Entstehung von Populationsschwankungen führt die erklärende Person ein, dass man sich zur Ursachenklärung zunächst den Lebensraum der Erdkröte ansehen muss. Für die Erklärung der zyklischen Schwankungen wird in den folgenden Ausführungen der erklärenden Person vorrangig die Lebensphase des adulten Tieres betrachtet und anhand derer eine Auswahl an Mortalitätsfaktoren erläutert.¹⁰¹

Die Erdkröte wird dafür grundlegend in die Klasse der Amphibien eingeordnet und mit dem **Anheften der Wortkarte „benötigt“** auf der linken Seite der Tafel werden einige Charakteristika bzw. Lebensbedingungen genannt, entsprechende Wortkarten angeheftet und der Zusammenhang zwischen der Anzahl an Erdkröten und der Ausprägung des Faktors skizziert. So wird **„gemäßigtes Klima“** als möglicher Faktor bei der Erklärung der Schwankungen genannt, das mit **„nicht zu viele heiße Tage und ausreichend Feuchtigkeit oder Niederschlag“** paraphrasiert wird. Den Bezug zu den Schwankungen im Liniendiagramm stellt die Sprecherin her, indem sie exemplarisch zu viele heiße Tage in einem Sommer mit weniger überlebenden Erdkröten in Zusammenhang bringt.

¹⁰¹ Die Unterteilung in abiotische und biotische Faktoren wird nicht explizit vorgenommen. Zum einen, um die Einführung zweier weiterer Fachbegriffe zu vermeiden, und zum anderen, da die Analyse der Mortalitätsfaktoren grundlegend auch ohne die Unterteilung auskommt.

Anschließend greift sie auf, dass Erdkröten als Amphibien bzw. zur Familie der Kröten gehörend für die Eiablage auf stehende Gewässer in ihrem Lebensraum angewiesen sind, in denen sich aus den befruchteten Eiern die Kaulquappen und schließlich adulte Erdkröten entwickeln können. Verdeutlicht wird der Einfluss des Faktors auch hier beispielhaft, indem die Anzahl an verfügbaren stehenden Gewässern mit der insgesamt sinkenden Anzahl an Erdkröten in Verbindung gebracht wird.

Im Anschluss wird auf die Nahrung der Erdkröte eingegangen (Insekten, Schnecken und Würmer). Auch der Zusammenhang mit diesem Faktor wird skizziert, indem ein hohes Nahrungsangebot mit steigender Anzahl von Erdkröten und umgekehrt ein niedriges Nahrungsangebot mit sinkender Anzahl von Erdkröten durch Verenden vieler Tiere in Verbindung gebracht wird.

Daraufhin leitet die erklärende Person zu Bedrohungen der Erdkröten über, die zu einer Dezimierung der Anzahl führen können. Die Gruppe der Faktoren wird an der rechten Seite der Tafel platziert **werden und ist mit einer Wortkarte „wird bedroht durch“ überschrieben. Als erster der insgesamt drei Faktoren wird das Vorhandensein von Fressfeinden genannt und Ratten sowie größere Vögel werden als Beispiele angeführt. Der Zusammenhang zwischen steigender Anzahl von Fressfeinden und sinkender Anzahl von Erdkröten wird hergestellt.**

Als zweiten Faktor erwähnt die Sprecherin den Menschen und sein Verhalten. Beispielhaft nennt sie dafür den Straßenbau, durch den der Lebensraum von Erdkröten zerstört wird. Den Zusammenhang verdeutlicht sie, indem sie erläutert, dass der Bau von Straßen Erdkröten vom Zugang zu Laichgewässern abschneidet, wobei sie **das Wort „ablaichen“ durch eine kurze eingeschobene Worterklärung erklärt.**

Als letzte Ursache für eine abnehmende Anzahl von Erdkröten wird die Einbringung von Schadstoffen genannt und als Beispiele werden Dünger und Insektizide angeführt. Letztere werden wiederum durch eine eingeschobene Worterklärung erklärt. Der Zusammenhang zwischen den eingebrachten Schadstoffen und der sinkenden Anzahl von Erdkröten wird aufgegriffen, indem die Schadstoffe als giftig bezeichnet werden, sodass Tiere an diesen Stoffen verenden können.

Zusammenfassend nimmt die erklärende Person darauf Bezug, dass es viele Faktoren sind, die die Anzahl der Erdkröten beeinflussen bzw. die Bedingungen für die dort lebenden Tiere stellen, wobei sie die Faktoren noch einmal aufzählt. Es wird der Schluss gezogen, dass es aufgrund der Vielzahl von Faktoren schwierig ist, die Ursachen exakt zu benennen, warum die Anzahl von Erdkröten bis zweitausend kontinuierlich gesunden ist, sowie die Vielzahl von Faktoren eine wesentliche Herausforderung darstellt, wenn es um die Vorhersage der Populationsentwicklung geht.

Zum fachlichen Hintergrund der Erklärung:

Die Erklärung basiert auf der Vorgehensweise, dass für die Erklärung einer Abundanz (Anzahl der Individuen einer Population) deren Determinanten geklärt werden. Dafür werden zunächst alle Phasen des Lebenszyklus der interessierenden Art beleuchtet und Faktoren, die das Sterben von Individuen bestimmen (Mortalitätsfaktoren), abgeleitet (Begon et al., 2017). Durch die empirische Erfassung der Abundanz und der Mortalitätsfaktoren (z. B. Nahrungsangebot, Wetter- bzw. Klimabedingungen) können Korrelationen berechnet, jedoch nur Vermutungen über Kausalitäten aufgestellt werden (Begon et al., 2017). Die einflussnehmenden Faktoren können allgemein in biotische Faktoren (belebte Faktoren; z. B. Symbiosen, Konkurrenz, Räuber-Beute-Beziehungen) und abiotische Faktoren (unbelebte Faktoren; z. B. Klima, Wetter, Temperatur, Wasser, Schadstoffe; Begon et al., 2017) kategorisiert werden.

Die in der Erklärung aufgegriffenen Zahlen für die Abundanz der Erdkröten stammen aus einer jährlichen Zählung an einem Schutzzaun entlang einer Straße in der Stadt Merklingen (L1182 zwischen Weil der Stadt und Merklingen) auf dem Weg zum Laichgewässer. Die Sammlung und die Erstellung der Statistik erfolgten durch den Verein der Vogel- und Naturfreunde Merklingen e. V. (Mettler, 2018).

Kommentar zu Umsetzung der UV:

Diese Erklärung weist einen höheren Grad an Komplexität als das inhaltliche Pendant im Themenbereich der Ökologie (Entstehung von Populationsschwankungen bei Schneeschuhhasen) auf. Exemplarisch wird der Einfluss von sechs Faktoren, die die Überlebens- und Reproduktionswahrscheinlichkeit und damit die Anzahl der Erdkröten beeinflussen, aufgezeigt. Dabei ist die Erklärung dennoch reduziert, da sich die Faktoren lediglich auf die Lebensphase der adulten Erdkröten beziehen. Die Multikausalität wird am Ende der Erklärung noch einmal gesondert hervorgehoben.

Im Anschluss daran wird auch der Aspekt Nichtwissen und unsicheres Wissen aufgegriffen. Dabei geht es in dieser Erklärung weniger um gesamtwissenschaftliches Nichtwissen über mögliche Mortalitätsfaktoren bei Erdkröten, sondern vielmehr um unsicheres Wissen, das bei der Aufklärung von Ursachen in diesem konkreten Fall bzw. im Übertragenen auch allgemein bei der Betrachtung von konkreten Fällen bestehen bleibt. Die Erklärung verliert daher an Eindeutigkeit und Abgeschlossenheit gegenüber der linearen Erklärung zur Entstehung von Populationsschwankungen bei Schneeschuhhasen.

e) Entstehung eines Vitamin-A-Mangels (P)

Mit dem nachstehenden Text wurden die Teilnehmenden vor der Videovignette über den Kontext der Erklärsituation informiert:

Die Klasse hat im Biologieunterricht gerade in einer Tabelle zusammengetragen, in welchen Lebensmitteln besonders viel Vitamin A vorkommt, welche Funktionen es im Körper hat und welche Folgen ein Mangel an Vitamin A haben kann.

Zum Ablauf der Erklärung:

Für den Einstieg in das Video zur Entstehung eines Vitamin-A-Mangels greift die erklärende Person ein Arbeitsergebnis von Schüler:innen auf, das in Form einer ausgefüllten Tabelle über eine Dokumentenkamera gezeigt wird, und fasst zusammen:

Okay, super. Also Vitamin-A-Mangel kann zu Wachstumsstörungen und Nachtblindheit führen, weil es eben wichtig für den Sehprozess ist und an Wachstum und Erneuerung der Haut beteiligt ist. (Zitat aus dem Video P)

Der Einstieg in die eigentliche Erklärsequenz erfolgt durch die erklärende Person:

Was wir uns jetzt noch nicht angeguckt haben, ist, wie denn überhaupt so ein Mangel zustande kommt. Und das möchte ich euch kurz an der Tafel zeigen. (Zitat aus dem Video P)

Abbildung 10 zeigt die während der Erklärung aufgebaute Visualisierung. Sie wurde von links nach rechts entwickelt.

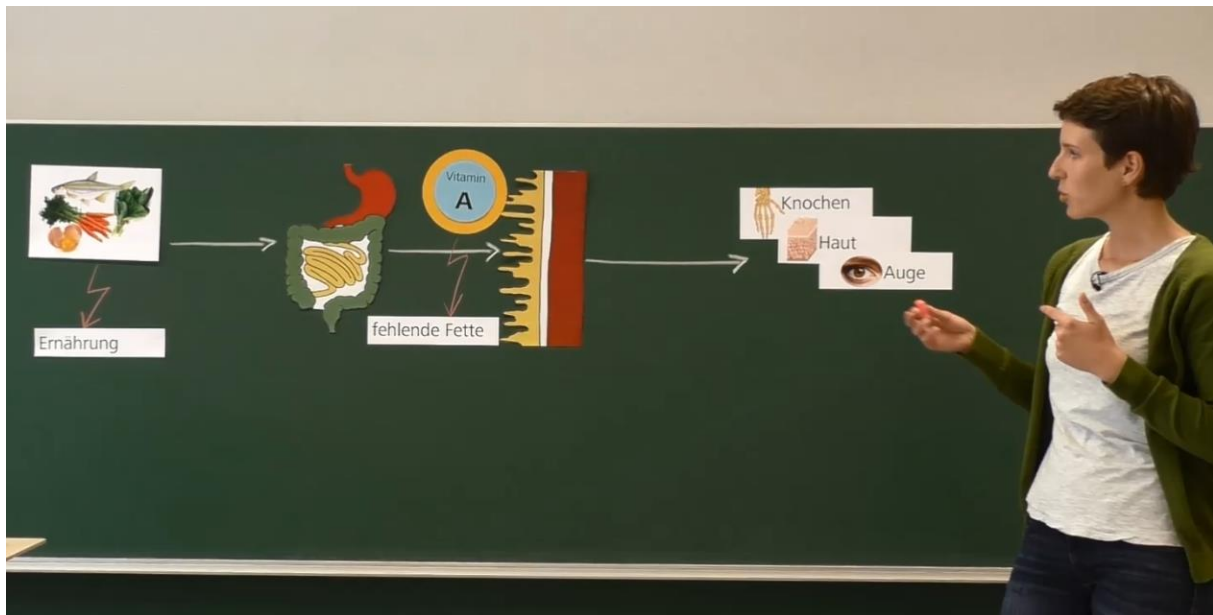


Abbildung 10: Screenshot aus dem Video zur Entstehung eines Vitamin-A-Mangels (Minute 2:55)

Nach der Formulierung des Erklärziels – die Entstehung eines Vitamin-A-Mangels – greift die erklärende Person auf, dass ein Vitaminmangel durch vitaminarme Ernährung zustande kommen kann. Diese unmittelbare Ursache wird weiter ausgeführt, indem eine Teildefinition von Vitaminen gegeben wird:

Weil wir eben Vitamine nicht selbst herstellen können, sondern eben mit der Nahrung aufnehmen müssen. (Zitat aus dem Video P₁)

Daraufhin werden – mit Bezug auf das zum Einstieg genutzte Arbeitsergebnis der Klasse – Lebensmittel genannt, die viel Vitamin A enthalten (Eigelb, Karotte, Spinat). Als Repräsentation für an Vitamin A reiche Ernährung heftet die Sprecherin eine Bildkarte mit entsprechenden Lebensmitteln an die Tafel, die **mit der Wortkarte „Ernährung“ unterschrieben** wird. Als Repräsentation dafür, dass eine vitaminarme Ernährung einen Vitamin-A-Mangel nach sich zieht, wird ein roter Blitz dazwischen gezeichnet. Die erklärende Person schließt den Punkt der Vitaminaufnahme durch die Ernährung mit einem zusammenfassenden Satz und leitet zu einer zweiten Möglichkeit über:

Es gibt allerdings noch eine zweite Möglichkeit. Nämlich, wenn wir diese Lebensmittel zwar essen, aber die Vitamine gar nicht aufnehmen können. (Zitat aus dem Video P₁)

Die Sprecherin leitet über, dass dazu zunächst die Verdauung anzusehen ist. Als Repräsentation für die Verdauung heftet sie eine Bildkarte an, auf der verschiedentlich eingefärbt der Magen, der Dünndarm sowie der Dickdarm zu sehen ist. Daraufhin wird eine kurze Erklärung zum Konzept der Verdauung gegeben, wobei die Zerkleinerung der Nahrung sowie die Aufnahme kleinerer Bestandteile im Dünndarm genannt werden. Die kleineren Bestandteile der Nahrung werden aufgrund des dafür notwendigen Vorwissens respektive der Kürze der Erklärung nicht näher bestimmt. Bei der Erwähnung des Dünndarms als Ort der Aufnahme der Nahrungsbestandteile wird auf den entsprechenden Abschnitt der Bildkarte gedeutet. Anschließend wird mit einer weiteren Bildkarte eine Vergrößerung der Dünndarmschleimhaut eingeführt. Sichtbarer Bestandteil der Repräsentation ist neben der stark ausgestülpten Dünndarmschleimhaut auch ein anliegendes Blutgefäß. Beide Bestandteile benennt die erklärende Person als solche und erklärt, dass

Nahrungsbestandteile durch die Dünndarmschleimhaut aufgenommen werden und anschließend in das Blutgefäß übergehen.

Vitamin A wird schließlich von der erklärenden Person als fettlösliches Vitamin eingeordnet. Das Fachwort „fettlöslich“ wird paraphrasiert, indem es in seine Wortbestandteile zerlegt wird:

Und bei Vitamin A gibt es nun eine Besonderheit, denn Vitamin A ist ein sogenanntes fettlösliches Vitamin. Das heißt, es löst sich wunderbar in Fett, aber sehr schlecht in Wasser. (Zitat aus dem Video P)

Anschließend wird der Darm als wässriges Milieu charakterisiert, in dem sich das Vitamin A aufgrund seiner Eigenschaft der Fettlöslichkeit wenig gut verteilen kann. Danach formuliert die erklärende Person die Kausalkette weiter aus: Aufgrund der schlechteren Verteilung im Dünndarm entsteht weniger häufig Kontakt zur Dünndarmschleimhaut, sodass folglich nur wenig Vitamin A aufgenommen werden kann und der Großteil des über die Nahrung Aufgenommenen wieder ausgeschieden wird.

Daran schließt sich der Vergleich zu einer effektiven Aufnahme des Vitamins an: Die Sprecherin leitet ein, indem sie informiert, dass Vitamin A in kleinen Fetttröpfchen transportiert wird. Die alltagssprachliche Formulierung „Fetttröpfchen“ wird in einem Einschub durch den Fachbegriff „Mizelle“ ergänzt. Anschließend formuliert sie auch hier die Kausalkette weiter aus: Durch den Einschluss in Mizellen kann Vitamin A besser im Dünndarm verteilt werden und es entsteht häufiger Kontakt mit der Dünndarmschleimhaut, sodass folglich mehr Vitamin A in die Blutbahn und damit in den Organismus gelangen kann. Anschließend wird Bezug auf die Funktionen des Vitamins im Körper genommen und angefügt, dass die in den Organismus aufgenommenen Vitamine sodann für Wachstumsprozesse oder den Sehprozess zur Verfügung steht.

Nachfolgend wird von der erklärenden Person darauf verwiesen, dass die fehlende Aufnahme von Fetten die zweite Möglichkeit ist, wie ein Vitaminmangel zustande kommen kann. Die Erklärsequenz wird mit einer kurzen Zusammenfassung geschlossen:

Und über eine ausgewogene Ernährung, das heißt vitaminreich, aber auch mit Fetten, können wir so einen Vitamin-A-Spiegel, der gesund ist, ganz gut erreichen. (Zitat aus dem Video P)

Zum fachlichen Hintergrund der Erklärung:

Vitamin A ist der Oberbegriff für mehrere Verbindungen, die eine ähnliche chemische Struktur und die biologische Aktivität von Retinol im Organismus aufweisen. Dazu gehören Retinol selbst sowie seine Ester (Retinylacetat, Retinylpalmitat, Retinylpropionat), teils Retinsäure sowie Retinaldehyd. Retinol ist den fettlöslichen Vitaminen zuzuordnen (Pietrzik, Golly & Loew, 2008). Während Retinol und Retinylester in tierischen Produkten enthalten sind, sind in Bestandteilen von Pflanzen Carotinoide (Provitamin A) zu finden, die im Organismus in Retinol bzw. dessen Derivate umgewandelt werden. Das in tierischen Produkten enthaltene Retinol stammt selbst wiederum aus der Verstoffwechslung von Carotinoiden (Pietrzik et al., 2008). Mit der Nahrung aufgenommen werden kann Vitamin A in besonders hoher Konzentration durch den Verzehr von beispielsweise Leber, Butter, Eigelb, diversen Fischarten, Karotten oder Spinat (Pietrzik et al., 2008).

Die Aufnahme von Vitamin A bzw. des Provitamins A in den Organismus ist unterschiedlich:

- i) Retinylester: Retinylester werden vor der Resorption im Darmlumen hydrolytisch zu Retinol gespalten und fast vollständig in Mizellen aufgenommen. Mizellen sind Aggregate, deren

Grundgerüst aus Gallsäuremolekülen besteht und in dessen Inneren die hydrophilen Abschnitte von Lipiden ragen und unter anderem fettlösliche Vitamine konzentriert sind (Jänig & Vaupel, 2019). Fette und Gallensäuren begünstigen die Resorption. Das Retinol wird anschließend in die Mukosazellen des Darms aufgenommen. Dort wird Retinol erneut verestert und anschließend in Lipoproteine integriert (Chylomikronen), die den Übertritt in die Blutbahn ermöglichen. Die Chylomikronen gelangen über die Blutbahn (dann als sogenannte Chylomikronen-Remnants) in die Leber. In den Parenchymzellen der Leber werden die Retinylester kurzzeitig, in den perisinusoidalen Stellatumzellen über einen langen Zeitraum gespeichert (Pietrzik et al., 2008). Von beiden Speicherorten können die Retinylester bei Bedarf wieder in das Blutplasma freigegeben werden.

- ii) Carotinoide: Carotinoide werden in der Darmmukosa unter Beteiligung von Galle zu Retinal gespalten, das anschließend enzymatisch reduziert wird, wobei Retinol entsteht. Retinol wird daraufhin verestert und gelangt über die Lymphe in den Blutkreislauf (Jänig & Vaupel, 2019; Pietrzik et al., 2008). Die Resorption und die Umwandlung von Carotinoiden werden durch gleichzeitig zugeführte Lipide, Proteine und Zink in der Nahrung sowie durch Gallensäure angeregt (Pietrzik et al., 2008).

Vitamin A kommt im Organismus keine einzelne Funktion, sondern eine Reihe von Wirkungsbereichen zu, unter anderem eine Funktion bei Zellwachstum und Zelldifferenzierung, bei der Reproduktion (Spermatogenese, Oogenese, Plazentaentwicklung, Embryonalentwicklung, Testosteronproduktion), bei der Regulation des Immunsystems sowie beim Sehvorgang (Pietrzik et al., 2008). (Pietrzik et al., 2008, S. 241). „Zum Mechanismus der Wachstumsförderung durch Vitamin A gibt es verschiedene Vorstellungen und wahrscheinlich treffen mehrere Faktoren zusammen“ (Pietrzik et al., 2008, S. 201). Wirksam sind dabei jeweils verschiedene Formen (Retinol, Retinal, Retinsäure).

Ein Vitamin-A-Mangel ist vor allem in Ländern der Dritten Welt verbreitet und auf eine Mangelernährung zurückzuführen (Pietrzik et al., 2008). In Industrienationen ist ein Vitamin-A-Mangel selten und häufig durch eine Störung der Resorption oder des Stoffwechsels von Retinol bedingt. Dazu können mehrere Faktoren beitragen: eine unzureichende Proteinaufnahme, hormonelle Störungen sowie eine Erkrankung von Darm, Leber oder Nieren (Zöliakie, Mukoviszidose, chronischer Alkoholismus; (Pietrzik et al., 2008).

Kommentar zur Umsetzung der UV:

Die Erklärung zur Entstehung eines Vitamin-A-Mangels nennt zwei Faktoren, die ursächlich zu einem Vitamin-A-Mangel führen können: zum einen die (direkte) Vitamin-A-arme Ernährung und zum anderen die Aufnahme einer unzureichenden Menge an Fetten mit der Nahrung, die für die Aufnahme des Vitamins in den Organismus essenziell sind. Die Erklärung ist dennoch als linear zu charakterisieren, da der Faktor der Aufnahme von Vitamin A mit der Nahrung nicht über Alltagswissen hinausgeht und eine einfache ‚wenn – dann‘ Verknüpfung darstellt. Eine Zurückführung eines Vitaminmangels ausschließlich auf eine vitaminarme Ernährung wäre darüber hinaus für die Schüler:innen vermutlich trivial und könnte als Übererklärung aufgefasst werden (Kotthoff, 2009).¹⁰² Weitere Faktoren wie die unzureichende Proteinaufnahme, hormonelle Störungen oder Darm-, Leber- oder Nierenerkrankungen werden nicht erwähnt, sodass die Erklärung bzw. die Faktoren des Systems insgesamt (sektoral) reduziert sind und allgemein die

¹⁰² Unter einer „Übererklärung“ versteht Kotthoff (2009, S. 122) unter Bezug auf Erickson und Shultz (1982) die deutliche Unterschätzung des Vorwissens von Rezipient:innen. Bei den Zuhörenden kommt es dabei zum Gefühl „für dumm verkauft“ (Kotthoff, 2009, S. 122) zu werden.

Aufnahme von fettlöslichen Vitaminen im Vordergrund steht. Weiterhin werden keine Verweise auf unsicheres Wissen und Nichtwissen angeführt.

f) Entstehung eines Folsäuremangels (P_k)

Vor der komplexer gestalteten Erklärung zur Entstehung eines Folsäuremangels erhielten alle Teilnehmenden den folgenden kontextualisierenden Text:

Die Klasse hat sich im Biologieunterricht gerade mit dem Vitamin „Folsäure“ beschäftigt und in einer Tabelle zusammengetragen, in welchen Lebensmitteln besonders viel davon enthalten ist, welche Funktionen sie im Körper hat und welche Folgen ein Folsäuremangel haben kann.

Zum Ablauf der Erklärung:

Analog zum Einstieg in das Video zur Entstehung eines Vitamin-A-Mangels greift die erklärende Person ein Arbeitsergebnis von Schüler:innen auf, das in Form einer ausgefüllten Tabelle über eine Dokumentenkamera gezeigt wird, und fasst zusammen:

Okay, super. Also ein Folsäuremangel kann zu Blutarmut und Fehlbildungen beim Baby im Mutterleib führen, weil es wichtig ist für die Bildung roter Blutkörperchen und die Herstellung von DNA. (Zitat aus dem Video P_k)

Der Einstieg in die Erklärung erfolgt auch hier durch die Sprecherin:

Was wir noch nicht besprochen haben, ist, wie es denn eigentlich zu so einem Mangel kommen kann. Und das möchte ich euch kurz an der Tafel zeigen. (Zitat aus dem Video P_k)

Abbildung 11 zeigt die während der Erklärung entstandene Visualisierung, die in der darunter stehenden Beschreibung der Vignette aufgegriffen wird.

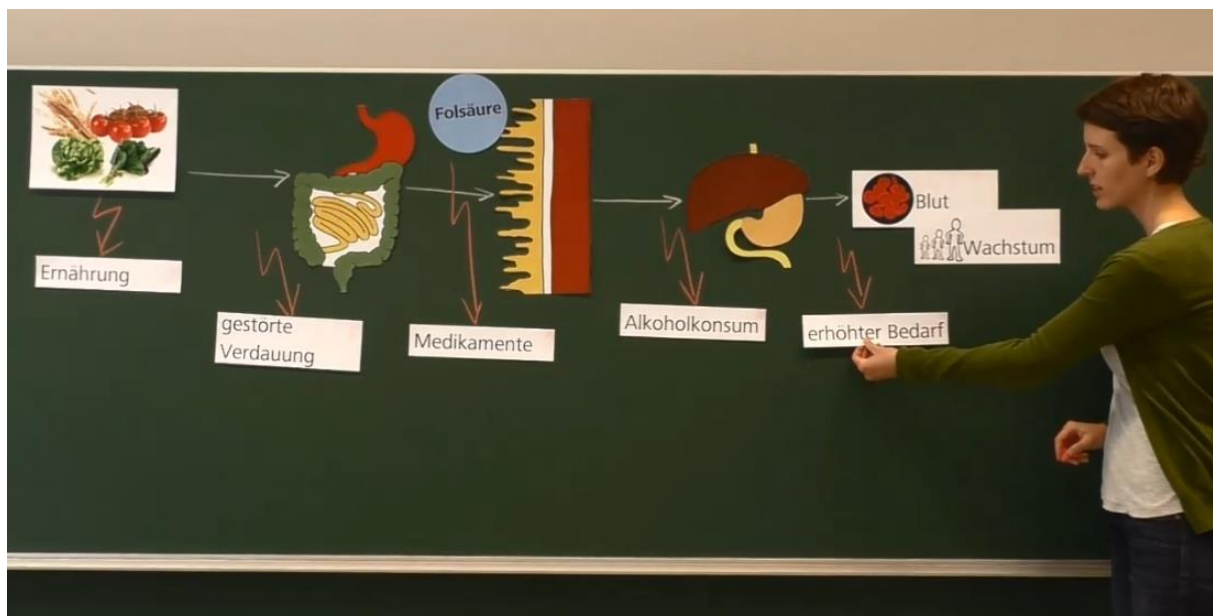


Abbildung 11: Screenshot aus dem Video zur Entstehung eines Folsäuremangels (Minute 2:51)

Auch bei dieser Erklärung wird als erster Faktor die Ernährung genannt, die die Quelle zur Aufnahme von Folsäure ist. Begründet wird dies ebenfalls mit der Teildefinition von Vitaminen, die nicht vom Körper selbst hergestellt werden können. Zeitgleich mit der Nennung des ersten Faktors heftet die erklärende Person als Repräsentation für die Ernährung eine Bildkarte mit folsäurereichen Lebensmitteln und sowie eine Wortkarte „Ernährung“, die diese Repräsentation bezeichnet, an. Die

Bedeutung als Risikofaktor für die Entstehung eines Mangels wird durch einen roten Blitz gekennzeichnet.

Die Sprecherin benennt nachfolgend die Ernährung als wichtigen Faktor und schränkt die Aussage ein, indem sie formuliert, dass die Ernährung nicht die einzige Ursache ist, die einen Mangel erklären kann. Den Schüler:innen wird benannt, dass für das Finden weiterer Ursachen eines Folsäuremangels der Weg der Folsäure im Körper zu betrachten ist. Dafür wird anschließend analog zur linearen Erklärung eines Vitamin-A-Mangels die Verdauung skizziert: die Zerkleinerung der Nahrung, die Aufnahme der kleineren Bestandteile, die auch hier nicht näher bestimmt werden, in den Körper, wobei der Dünndarm eine zentrale Rolle spielt, da hier unter anderem Folsäure in anliegende Blutgefäße übertreten und von dort durch den Blutkreislauf im Körper verteilt werden kann. Analog zum Vorgehen in der linearen Erklärung werden eine Repräsentation, die einen Ausschnitt des Verdauungstrakts zeigt (Magen, Dünndarm, Dickdarm), sowie eine Repräsentation, die eine Vergrößerung der Dünndarmschleimhaut mit anliegendem Blutgefäß zeigt, angeheftet. Nach dem Übertritt in den Blutkreislauf wird in der Erklärung die Leber als Speicherort für Folsäure genannt. Die Sprecherin erklärt, dass der Speicherort temporär ist und die Folsäure von dort aus bei Bedarf – als Beispiele nennt sie die Blutbildung und allgemein Wachstumsprozesse – wieder freigegeben wird. Die Station der Zwischenspeicherung in der Leber wird dabei wieder mit einer Bildkarte visualisiert, auf der der Umriss einer Leber zu sehen ist. Abschließend werden die Bedarfe mit einer Wortkarte, auf der „Blut“ und „Wachstum“ zu lesen ist, repräsentiert.

Im Anschluss an die Beschreibung von der Aufnahme bis zur Freisetzung bei Bedarf führt die erklärende Person ausgewählte Möglichkeiten für die Entstehung eines Mangels an. Sie ergänzt diese analog zur folsäurearmen Ernährung im Tafelbild mit einem roten Blitz und einer darunter stehenden Wortkarte. Zunächst wird dabei auf eine gestörte Verdauung als Faktor hingewiesen. Als Beispiele werden ein Magen-Darm-Infekt und Stress genannt. Die Ursache Stress als Auslöser für Verdauungsstörungen wird weiter ausgeführt und erläutert, dass Stresshormone Einfluss auf die Verdauung haben. Die erklärende Person macht anschließend explizit, dass eine gestörte Verdauung zur verminderten Aufnahme von Folsäure aus der Nahrung führen kann.

Beim Übergang der Folsäure in den Blutkreislauf werden als weitere mögliche Ursache für einen Mangel die Einnahme von Medikamenten und deren Wechselwirkungen genannt. Als Beispiel führt die Sprecherin Kontrazeptiva (im Video konkret „Pille zur Verhütung“) an.

Als vierte mögliche Ursache wird im Zusammenhang mit der Speicherung der Folsäure in der Leber angeführt, dass der Leber auch die Funktion der Ausscheidung von Giftstoffen zukommt und folglich die Speicherung sowie die Abgabe von Folsäure gehemmt sein kann, wenn das Organ im Übermaß Giftstoffen wie zum Beispiel Alkohol ausscheiden muss.

Als letzte mögliche Ursache wird ein erhöhter Bedarf genannt, für den die bisherige, normale Versorgung mit Folsäure nicht mehr ausreicht. Dabei verweist die erklärende Person auf die Beispiele, die bereits als Wortkarten an der Tafel angeheftet sind.

Anschließend fasst die Sprecherin zusammen, dass es neben der Ernährung selbst viele weitere Faktoren gibt, die zu einem Folsäuremangel führen können. Sie erweitert das Gesagte, indem sie darauf verweist, dass es neben den dargestellten Faktoren noch viele weitere vorliegen und meist nicht nur ein Faktor bei der Aufklärung eines Folsäuremangels auftritt:

[...] ihr seht schon. Neben der Ernährung gibt es noch ganz viele andere Faktoren, die einen Mangel erklären können. Und selbst diese vielen Faktoren – das war auch nur eine Auswahl von

mir. Das heißt, es gibt noch viele Weitere. Und was dazukommt ist, dass ein Mangel meist nicht nur durch einen isolierten Faktor erklärt werden kann, sondern es häufig eine Kombination von vielen ist. (Zitat aus dem Video P_k)

Zum fachlichen Hintergrund der Erklärung:

Folsäure, auch Pteroylglutaminsäure genannt, wurde 1930 als Antianämie- und Wachstumsfaktor entdeckt und 1946 in seiner chemischen Struktur beschrieben (Pietrzik et al., 2008). Die dabei untersuchte und beschriebene Form kommt so tatsächlich nicht in Zellen, sondern nur in Isolation vor. Natürliche Substrate hingegen enthalten eine Familie von teils hochlabilen Folsäureanaloga, die sich strukturell unterscheiden. Pietrzik et al. (2008) resümieren, dass die „[d]ie Kompliziertheit der Zusammensetzung der nativen Folsäuregemische [...] die Folsäureanalytik zu einem schwierigen Problem“ macht (Pietrzik et al., 2008, S. 82). Die verschiedenen Formen sind dabei unterschiedlich gut für die Aufnahme verfügbar. Sie kommt sowohl in pflanzlichen als auch tierischen Lebensmitteln vor und dabei insbesondere in Spinat, Tomaten, Blattsalaten und Leber. Aus tierischen Lebensmitteln werden Folate deutlich effektiver resorbiert (Pietrzik et al., 2008).

In der Nahrung enthalten ist meist Polyglutamylfolat. Dieses muss für die Aufnahme, die im Zwölffingerdarm (Duodenum) und im darauffolgenden Abschnitt des Dünndarms (Jejunum) stattfindet, zu Monoglutamat hydrolysiert werden. Daraufgehend findet der Großteil des Transports in die Mukosazelle carrier-vermittelt und energieabhängig statt, ein kleinerer Teil erfolgt durch Diffusion. In der Mukosazelle wird Monoglutamat reduziert und in das Blut abgegeben, wo dann vorrangig 5-Methyl-Tetrahydrofolsäure vorliegt. Letztere wird auch in Erythrozyten aufgenommen bzw. eingebaut, wobei die Aufnahme von Vitamin B12 abhängig ist (Pietrzik et al., 2008). Als Hauptspeicherorgan fungiert die Leber, wobei die Geschwindigkeit für die Freisetzung verschiedentlich diskutiert wird (Pietrzik et al., 2008). Funktionen kommen der Folsäure bzw. ihren Derivaten bei der Bildung von Erythrozyten, der Neubildung von Zellen sowie bei der Zellteilung zu (Pietrzik et al., 2008).

Im Gegensatz zu Vitamin A stellt sich bei Folsäure sehr leicht ein echter Mangel ein, da die speicherbaren Reserven gering sind (Pietrzik et al., 2008). Die Ernährung und damit die regelmäßige Zufuhr von Folaten spielen daher eine wichtige Rolle bei der Verhinderung eines Mangels.

Faktoren für eine eingeschränkte Verwertbarkeit der Folsäure aus Lebensmitteln sind für einen gleichzeitig auftretenden Eisenmangel sowie die Zusammensetzung der Nahrung nachgewiesen, da die Aufnahme aus pflanzlichen Lebensmitteln schwerer als aus tierischen Lebensmitteln erfolgen kann. Entzündliche Darmerkrankungen (z. B. chronisch auch Zöliakie) tragen darüber hinaus dazu bei, dass die Aufnahme sowie die Rückresorption von mit der Galle ausgeschiedenen Folaten nicht mehr funktioniert (Pietrzik et al., 2008).

Die wichtigste Ursache ist ein erhöhter Bedarf in der Schwangerschaft, der bei 30 bis 60 % der Schwangeren zu einem Mangel führt. Auch zwischen dem Alter und einem Folsäuremangel besteht eine Korrelation (Pietrzik et al., 2008). Des Weiteren sind Alkoholismus sowie die Einnahme von Medikamenten (z. B. Kontrazeptiva; Chemotherapeutika) Ursachen für eine verminderte Aufnahmefähigkeit, die zu einem Mangel führen können (Pietrzik et al., 2008).

Folgen eines Folsäuremangels sind schwerwiegend: unter anderem megaloblastäre Anämie (Verformung von Erythrozyten und damit das Fehlen von funktionierenden roten Blutkörperchen; Jelkmann, 2019) und megaloblastisches Knochenmark (Pietrzik et al., 2008).

Kommentar zur Umsetzung der UV:

Bei der Erklärung zur Entstehung eines Folsäuremangels ist der Grad der dargestellten Komplexität höher als bei der Erklärung zur Entstehung eines Vitamin-A-Mangels, da insgesamt fünf Faktoren angeführt werden. Diese wirken mitunter nicht isoliert auf die Entstehung des Mangels, sondern können additiv auftreten. Die Multikausalität wird auch in der Zusammenfassung noch einmal aufgegriffen und betont, dass die Konstellation von Faktoren variationsreich ist und im Einzelfall mitunter viele Faktoren zur Phänomenentstehung beitragen.

Der Aspekt ‚Nichtwissen und unsicheres Wissen‘ wird aufgegriffen, indem auf gesamtwissenschaftliches Nichtwissen in Bezug auf weitere Faktoren, die noch unbekannt sind, verwiesen wird. Die Wirkung der Abgeschlossenheit der Erklärung ist damit gegenüber der linearen Erklärung des inhaltlichen Pendantes als geringer einzuschätzen.

5.2.3 Triangulation als Möglichkeit, komplexen Forschungsgegenständen zu begegnen

Die Methode der Triangulation als Herangehensweise an einen Forschungsgegenstand beinhaltet dessen Betrachtung von mindestens zwei Punkten aus (Flick, 2011). Dem liegt die methodologische Annahme zugrunde, dass die Art der Erhebung die Gegenstände rekonstruiert (Flick spricht daher auch von der „Reaktivität der Methoden“ (Flick, 2011, S. 11)) und damit verschiedene Methoden und Herangehensweisen auch differierende Ergebnisse bringen können (Lamnek, 2010).

Die Standpunkte, die zur Betrachtung des Phänomens herangezogen und verschränkend in Verbindung gebracht werden können, sind beispielsweise verschiedene Datensorten, Forscher:innen, Methoden und Theorien (Einteilung der Konzeption Denzins (1970) folgend).¹⁰³ Als Datensorten werden von Denzin (1970) Datenquellen verstanden, die sich beispielsweise hinsichtlich des personellen (Befragung unterschiedlicher Personengruppen) oder zeitlichen Ursprungs (Befragung gleicher Personen zu unterschiedlichen Zeiten) unterscheiden. Die Triangulation von Forscher:innen bzw. Interviewer:innen und Beobachter:innen respektive deren Ergebnissen kann dazu dienen, die Zusammenhänge zwischen Personen und Ergebnissen zu analysieren, sodass diese in weiteren Schritten kontrolliert werden können. Bei der Methodentriangulation werden von Denzin (1970) zwei Versionen unterschieden. Zum einen nennt er die Triangulation innerhalb einer Methode (within-method), unter der er beispielsweise auch die Verwendung verschiedener Subskalen auf ein Phänomen bzw. einen Forschungsgegenstand versteht.¹⁰⁴ Zum anderen wird die Triangulation als Einsatz verschiedener Methoden (between-method) – zum Beispiel qualitative und quantitative Verfahren – beschrieben. Die Triangulation der Theorien dient unter anderem einer Ausweitung der Erkenntnismöglichkeiten anhand derselben Datengrundlage, wenn sie in Bezug auf verschiedene Theorien analysiert werden. Möglich werden aber auch das Loslösen von möglicherweise verfestigten Vorannahmen der Forscher:innen und ein Vergleich von Theorien über den gleichen Gegenstand.

¹⁰³ Eine alternative Einteilung erfolgt unter anderem durch Flick (2011). In seiner Konzeption zur systematischen Perspektiven-Triangulation beschreibt er die Verschränkung verschiedener qualitativer Forschungszugänge und -perspektiven, um unter Berücksichtigung deren Gegenstands- und Methodenverständnis den Forschungsgegenstand vertiefend und umfassend zu erforschen. Dieser Ansatz findet in der Anlage dieser Studie jedoch wenig Referenz, da ein qualitativer und ein quantitativer Zugang trianguliert werden soll. Deshalb wird auf weitere Ausführungen hierzu verzichtet.

¹⁰⁴ Der Forschungsgegenstand ‚Güte instruktionaler Erklärungen‘ wird in dieser Arbeit auch durch verschiedene Aspekte abgebildet. Diese Art der Triangulation wird im Folgenden theoretisch jedoch nicht berücksichtigt.

In den Beschreibungen Denzins (1970) zu den Formen der Triangulation wird deutlich, dass er die Vorzüge der Triangulation vor allem in einer Steigerung der Reliabilität durch die Verwendung verschiedener Methoden bzw. einer Art Validierung der Forschungsstrategien sieht.¹⁰⁵ Die Triangulation soll folglich vor allem dazu dienen, die Ergebnisse von Studien zu verbessern und ‚sicheres‘ Wissen zu generieren.

Diesem Verständnis stehen beispielsweise die Konzeptionen von Fielding und Fielding (1986) und Flick (2011) entgegen. Ebenso der Grundannahme folgend, dass unter anderem die jeweilige Forschungsmethode den Gegenstand mit konstruiert, sehen sie den Zugewinn der Triangulation in eben diesen divergierenden Ergebnissen, die verschiedene methodische Ansätze hervorbringen können. Die divergierenden Ergebnisse werden dabei nicht als Schwächen der jeweils eingesetzten Methoden, sondern als Folge der Rekonstruktion des Forschungsgegenstandes durch die Methoden bewertet, die in einer Verschränkung zu einem umfassenderen, tieferen Verständnis des Gegenstandes führen können (Flick, 2011).¹⁰⁶ Fielding und Fielding (1986) formulieren: „We should combine theories and methods carefully and purposefully with the intention of adding breadth or depth to our analysis, but not for the purpose of pursuing ‘objective truth’“ (Fielding & Fielding, 1986, S. 33). Flick (2011) fordert weiterhin, die Perspektiven der Triangulation soweit wie möglich gleichberechtigt zu behandeln. Als Folge daraus ergeben sich bei der Triangulation keine sich widersprechenden Ergebnisse über den Gegenstand, sondern diese sind als unterschiedliche Rekonstruktionen des Gegenstandes zu interpretieren (Flick, 2011).

Der Erkenntniszugewinn respektive sogar die Notwendigkeit einer Triangulation ergibt sich laut Flick (2011) insbesondere bei komplexen Forschungsgegenständen. Dieser Forderung soll in Anbetracht der Komplexität von Lehr-Lern-Prozessen und im Speziellen der Sprachhandlung des Erklärens – nachgekommen werden.

Dem Verständnis Flicks (2011) weiter folgend sollen die Datenquellen dieser Studie hinsichtlich ihrer Charakteristika reflektiert aufeinander bezogen und die jeweiligen Forschungsergebnisse wechselseitig ergänzt werden, sodass der Erkenntnisgewinn über den Forschungsgegenstand – Kriterien guter instruktionaler Erklärungen im Biologieunterricht – vergrößert wird. In der vorliegenden Studie findet sich die Triangulation auf zwei Ebenen wieder: zum einen aus methodischer Sicht, da qualitative Daten aus textbasierten, offenen Items und quantitative Daten mit geschlossenen Items gewonnen werden (Methodentriangulation als ‚between method triangulation‘; Denzin, 1970; siehe Kap. 5.2.3.1),¹⁰⁷ zum anderen aus inhaltlicher Sicht, da Perspektiven von verschiedenen Statusgruppen berücksichtigt und aufeinander bezogen werden, um den Gegenstand breiter zu erfassen (Datentriangulation; Denzin, 1970; siehe Kap. 5.2.3.2).

¹⁰⁵ Flick (2011) kommt zu dem Schluss, dass Denzin den Fokus seines Triangulationsverständnisses durch Kritik verschiedener Forscher:innen (er nennt Jane Fielding, Nigel Fielding und David Silverman) von einer Validierung und Objektivierung zu einem besseren Verständnis des Gegenstandes hin verschoben hat (Flick, 2011, S. 20).

¹⁰⁶ Lamnek (2010) verwendet in diesem Sinne für das Verständnis von Triangulation den Begriff der Konvergenz als Wortentlehnung der Biologie für die Ergänzung der Ergebnisse und Erkenntnisse auf einer Ebene. Er grenzt ihn dagegen von Kongruenz als mathematisch definierte Deckungsgleichheit ab.

¹⁰⁷ Die Methodentriangulation umspannt in diesem Fall weniger weit auseinanderliegende Pole, da die qualitativen Daten wenig rekonstruktiven Anteil haben bzw. die Forschungsfrage nicht darauf ausgerichtet ist, methodisch kontrolliertes Fremdverstehen zu erlangen (Bohnsack, 2005). Zusätzlich werden die Daten innerhalb der Auswertung quantifiziert werden.

5.2.3.1 Methodentriangulation – Verschränkung qualitativer und quantitativer Methoden

Die gewonnenen Daten unterscheiden sich hinsichtlich des Inhalts sowie der Erhebungssituation: Die Äußerungen der Teilnehmer:innen im offenen Textfeld geben zum einen in die Kriterien, die die Teilnehmer:innen an instruktionale Erklärungen stellen, und zum anderen diejenigen Kriterien, die diejenigen Aspekte der Erklärung, die für die Teilnehmer:innen in besonderer Weise auffällig sind, Einblick. Sie bilden demnach die individuellen Auffälligkeiten einer Erklärung für die Personen ab, die sich wiederum zwischen Personen und zwischen Personengruppen unterscheiden können. Es handelt sich um freie, ungerichtete Äußerungen, die ohne die Vorgabe von Themen bzw. Kategorien abgegeben werden, sodass die Daten einen Eindruck über die Diversität der Wahrnehmungs- und Bewertungsperspektiven erlauben. Die Angabe einer Begründung ist freiwillig, sodass der Motivation der Teilnehmer:innen, sich zur Frage zu äußern, auch eine Bedeutung zukommt. Im Gegensatz dazu erfassen die Items des quantitativen Zugangs eine obligatorische Beurteilung über die Ausprägung bzw. Qualität eines bestimmten Aspekts. Somit wird ein direkter Vergleich der Beurteilungen hinsichtlich definierter Merkmale der Erklärungen möglich. Damit kann der Forderung von Fielding und Fielding (1986) folgend durch die Kombination aus zwei Methoden respektive Forschungszugängen, die zum einen die individuellen Sichtweisen der Subjekte berücksichtigt und zum anderen die Struktur des Phänomens in den Blick nimmt, nachgekommen werden.

5.2.3.2 Datentriangulation – Erfassung des Messgegenstands aus unterschiedlichen Perspektiven

Sowohl im qualitativen als auch im quantitativen Zugang werden Schüler:innen, Studierende, Lehrkräfte und Didaktiker:innen zu ihrer Einschätzung der Qualität der Erklärungen befragt. Da davon ausgegangen werden kann, dass es sich um unterschiedliche Sichtweisen der Statusgruppen auf das Phänomen guter instruktionaler Erklärungen handelt (siehe Kap. 4.8), wird deren Verschränkung im Folgenden als Perspektiventriangulation bezeichnet (Denzin (1970) bezeichnet diese Form der Triangulation als eine Art der Datentriangulation).¹⁰⁸ Die erwarteten, differierenden Ergebnisse sollen konstruktiv genutzt werden, um gute instruktionale Erklärungen des Biologieunterrichts beschreiben zu können.¹⁰⁹

Durch die vier Statusgruppen werden die Perspektiven derer erfasst, die direkt oder indirekt an Unterricht und damit auch an schulischem Erklären beteiligt sind. Schüler:innen repräsentieren dabei die Perspektive potenzieller Rezipient:innen der Erklärungen. Sie können aus der Situation einer nicht wissenden Person in besonderer Weise bewerten, die wissende Personen – wie beispielsweise in der Regel Studierende, Lehrkräfte, Didaktiker:innen – nicht mehr einnehmen können. Studierende, Lehrkräfte und Didaktiker:innen nehmen hingegen eine Metaperspektive ein und bewerten die Erklärungen auf Grundlage ihres jeweiligen fachdidaktischen Wissens. Innerhalb der Metaperspektive sind damit auch verschiedene Professionalisierungsniveaus repräsentiert: Während Studierende noch über wenige Erfahrungen von Lehr-Lern-Prozesse verfügen und deshalb diesbezüglich eher als Noviz:innen bezeichnet werden können, sind Lehrkräfte aufgrund ihrer

¹⁰⁸ Mit dem Begriff der Perspektiventriangulation ist nicht der von Flick (2011) verwendete Begriff gemeint. Er versteht unter der systematischen Perspektiventriangulation die Kombination verschiedener qualitativer Methoden, die sich hinsichtlich deren theoretischen Grundannahmen unterscheiden, sodass durch die Kombination derer ein Zugewinn an Informationen entstehen kann.

¹⁰⁹ In ähnlicher Weise nutzt auch Gautschi (2011) in seiner explorativen, deskriptiv angelegten Untersuchung die Perspektiven von Lehrenden und Lernenden für die komplexe Frage nach gutem Geschichtsunterricht. Für die Erforschung von Unterrichtsqualität zieht er aus seinen empirischen Befunden den Schluss, dass Triangulation „eine angemessene und innovative Antwort auf den Umstand zu sein [scheint], dass Unterrichtsqualität systemisch bedingt und multipel determiniert ist“ (Gautschi, 2011, S. 256).

vielfältigen Erfahrungen im schulischen Kontext eher als Expert:innen zu sehen. Weiterhin können die Perspektiven durch verschiedene Professionalisierungsschwerpunkte charakterisiert werden: Während Lehrkräfte über einen berufserfahrenen Blick für alltägliche Unterrichtssituationen verfügen, kann angenommen werden, dass Didaktiker:innen aufgrund weniger Unterrichtserfahrung und größerer Forschungsnähe eher theoriegeleitete Unterrichtsreflexionen vornehmen. In Hinblick auf die damit einhergehenden Unterschiede in der Wahrnehmung und Bewertung der Erklärungen sei auf Kapitel 4.8. verwiesen.

Die beschriebenen Gruppen nehmen damit jeweils individuelle Sichtweisen ein und können als Expert:innen bezeichnet werden. Die Anerkennung der unterschiedlichen Perspektiven und aus der Konsequenz daraus deren gleichwertiger Einbezug in die Analysen finden Korrespondenz im Expert:innenbegriff der qualitativen Sozialforschung, wie ihn Gläser und Laudel (2010) verwenden.¹¹⁰

5.2.4 Konzeptualisierung und Operationalisierung des Messgegenstandes ‚Gutes instruktionales Erklären im Biologieunterricht‘

Instruktionale Erklärungen wurden in FALKE als vorbereitete¹¹¹ Erklärungen konzeptualisiert und operationalisiert, die auf einen Wissensaufbau und die Initiierung eines Verstehensprozesses bei Schüler:innen fokussieren. In FALKE-Biologie zeigen die instruktionalen Erklärungen jeweils kausale Ursache-Wirkungs-Zusammenhänge.

Für die Konzeptualisierung guten instruktionalen Erklärens wurden zunächst vier fachunspezifische – also fach- bzw. projektübergreifende – Aspekte ausgewählt: die Adressatenorientierung, die Strukturiertheit, die sprachliche Verständlichkeit sowie der Sprech- und Körperausdruck. Auf Grundlage der Literaturrecherche und fachübergreifender Diskussionen konnten diese als für alle Fächer in ähnlicher Weise relevante Aspekte guten Erklärens identifiziert werden. Die Auswahl dieser vier Aspekte für die Konzeptualisierung guten, instruktionalen Erklärens wird im Folgenden kurz begründet, die jeweiligen Konzeptualisierungen und Operationalisierungen der Aspekte werden in den folgenden Kapiteln (Kap. 5.2.4.1 bis 5.2.4.5) beschrieben.

Die Adressatenorientierung ist aus kognitionspsychologischer und lerntheoretischer Perspektive die Basis für einen gelingenden Wissenserwerb (z. B. Wellenreuther, 2013; Wittwer & Renkl, 2008) und wurde bereits mit verschiedenen Konzeptualisierungsschwerpunkten als wichtiger Faktor für effektive Erklärungen identifiziert (z. B. Kulgemeyer & Tomczyszyn, 2015; Sevia & Gonsalves, 2008). Nur durch die Berücksichtigung der Voraussetzungen der Rezipient:innen kann das Gesagte verstanden und gegebenenfalls in bestehende Wissensstrukturen eingebettet werden, sodass kein träges Wissen entsteht (Renkl, 2004).

Die Strukturiertheit einer Erklärung wird ebenso aus theoretischer (z. B. Wagner & Wörn, 2011; Wellenreuther, 2013) wie aus empirischer Sicht als relevanter Faktor für gute (z. B. Schopf & Zwischenbrugger, 2015) bzw. effektive (z. B. Leinhardt, 1987; Sevia & Gonsalves, 2008) Erklärungen diskutiert. Sie ist insofern ein interessanter Forschungsgegenstand, da die Struktur einer Erklärung in hohem Maße vom Erklärgegenstand und den Wissensstrukturen von Rezipient:innen

¹¹⁰ Gautschi (2011) bezeichnet lediglich die Gruppe der Didaktiker:innen als Expert:innen. Wie oben beschrieben, werden die Statusgruppen im Sinne eines gleichwertigen Einbezugs ihrer Perspektiven in dieser Arbeit nicht in generelle Expert:innen oder Laien eingeteilt.

¹¹¹ Im Gegensatz zu ad-hoc, also unvorbereitet gegebene Erklärungen.

und Erklärenden abhängt (Ehlich, 2009), sodass es keine prototypischen Vorschläge oder Muster für Themen oder Erklärgegenstände geben kann.

Die Sprache ist allgemein das zentrale Kommunikationsmedium, um Wissen auszutauschen (Morek & Heller, 2012; Wellenreuther, 2013). Die Rolle der Sprache ist daher für mündliche Erklärungen eine entscheidende. Für einen gelingenden Wissensaufbau und damit eine gute Erklärung ist die Verständlichkeit der Sprache von besonderem Interesse.

Der Sprech- und Körperausdruck der erklärenden Person stand bisher insgesamt wenig im Blickfeld didaktischer Forschung. Auch in Hinblick auf fachdidaktische Forschung zum Themenfeld des mündlichen Erklärens findet er nur wenig Beachtung: In der Untersuchung von Wragg (1993) und Odora (2014) werden einige Facetten dieses Aspekts berücksichtigt, in der breit angelegten Übersicht zu Kriterien guten Erklärens von Findeisen (2017) findet sich der Sprech- und Körperausdruck hingegen nicht. Ausgehend von der Bedeutung des Sprech- und Körperausdrucks für die Unterstützung des Kommunikationsprozesses und damit für das Verstehen, die beispielsweise von Neuber (2016) aus sprechwissenschaftlicher und von Odora (2014) aus didaktischer Sicht betont werden, ist davon auszugehen, dass der Aspekt einen relevanten Faktor für gutes, mündliches Erklären darstellt.

Der fünfte, biologiespezifische Aspekt fokussiert die Wahrnehmung des dargestellten Komplexitätsgrades. Da die Komplexität ein charakteristisches Element vieler biologischer Phänomene ist (Mainzer, 2008; Mitchell, 2008) und ein konzeptuelles Verständnis dessen in Anbetracht aktueller, hochkomplexer Problemlagen sowie deren Interpretation in naturwissenschaftlicher und sozialer Dimension eine Herausforderung darstellt (siehe Kap. 3.4; Jacobson, 2000; Jacobson & Wilensky, 2006; Grotzer, 2012; Resnick, 1995), lohnt sich die empirische Untersuchung.

Weitere Aspekte, die ebenfalls zur Güte der Erklärung beitragen, werden in dieser Studie nicht durch Items repräsentiert. Auch dies wird im Folgenden kurz begründet.

Hinsichtlich des fachspezifischen Aspekts ist die Facette der fachlichen Korrektheit von der Messung ausgenommen. Diese wird zwar häufig in der Literatur angeführt (z. B. Schopf & Zwischenbrugger, 2015; Wagner & Wörn, 2011), jedoch sind die in den Erklärungen aufgegriffenen Ursache-Wirkungs-Zusammenhänge stets fachlich korrekt, sodass eine Einschätzung durch die Teilnehmenden nicht notwendig ist. Beim Kriterium der Wahl der adäquaten fachlichen Erklärungsform muss zusätzlich davon ausgegangen werden, dass es aufgrund fehlenden Wissens bei den Teilnehmenden über Erklärungsformen in der Biologie nicht reliabel und valide gemessen werden kann.

Die Aspekte zum Einsatz und zur Gestaltung von Repräsentationen werden insgesamt nicht durch geschlossene Items untersucht. Da die Bedeutung und die Eigenschaften von Repräsentationen zwischen den im Projekt beteiligten Didaktiken sehr heterogen diskutiert werden, wurden keine Items implementiert bzw. keine Konzeptualisierung vorgenommen. Während beispielsweise in der Religionspädagogik eine Metapher ein bildhafter, individueller Denkanstoß ist, der einen hohen Interpretationsspielraum aufweisen kann, um eigene Sichtweisen und Lebensarten zu reflektieren (Halbfas, 2012), werden Metaphern in der Biologiedidaktik häufig als feststehende Ausdrücke mit unerwünschtem Interpretationsspielraum bzw. festgelegten Übertragungen vom Herkunftsbereich (erfahrungsnah, alltägliche Vorstellungen) in den Zielbereich (abstrakte, wissenschaftliche Vorstellungen) verstanden (z. B. Ohlhoff, 2002). Weiterhin sind beispielsweise unterschiedliche

Formen der Visualisierung notwendig, um Inhalte sprachlicher und naturwissenschaftlicher Unterrichtsfächer zu dokumentieren oder zu visualisieren, sodass die Generierung von konkreten, fachübergreifenden Items erschwert ist.

Obwohl das Erklären auf das Verstehen bei den Rezipient:innen oder zumindest das Initiieren eines Verstehensprozesses abzielt, ist die Überprüfung des Verstehens – also zum Beispiel in Form eines Leistungs- oder Wissenstests – nicht Teil der Studie. Diese Entscheidung erfolgte aufgrund theoretischer, aber auch empirischer Überlegungen. Leinhardt (2001) argumentiert, dass das Wirkungsgefüge, das schließlich zum Verstehen des Erklärinhalts führt, ein sehr komplexes ist: Zeitpunkt der Erklärung und Merkmale der Interaktionspartner:innen bzw. Rezipient:innen haben neben den Merkmalen der Erklärung selbst einen großen Anteil an der Verarbeitung des Gesagten und dem Verstehen dessen. Diese Merkmale können in der Studie nicht ausreichend berücksichtigt werden. Weiterhin sind die Aktivitäten, die nach einer lehrerzentrierten Erklärung im Unterricht folgen ebenso wesentlich, um das Verstehen bei Schüler:innen im Sinne eines eigenaktiven, konstruktiven Prozesses zu ermöglichen (Webb et al., 1995). Hinsichtlich der Messbarkeit stellt sich auch die Frage, welche Aufgabenstellung in einem begrenzten zeitlichen Rahmen die neuen Inhalte reliabel und valide messen kann, insbesondere im Fall der Erklärungen, die komplexer gestaltet sind, da es sich bei den Inhalten bzw. den zu erreichenden Lernzielen nicht ausschließlich um die Vermittlung von Fakten, sondern eben auch um Konzepte zu Nichtwissen und unsicherem Wissen handelt. Vielmehr soll in der Studie die lehrerzentrierte Sprechhandlung ‚Erklären‘ betrachtet werden und die Unterschiede bezüglich der wahrgenommenen Qualität sowohl zwischen den teilnehmenden Gruppen als auch der umgesetzten UV sollen im Fokus stehen, um aus den empirischen Ergebnissen neue Ansätze für ein strukturiertes Vorgehen bei Untersuchungen – in Anlehnung an die Begriffsunterscheidung von Berliner (2005) – zu effektivem oder qualitativem Erklären geben.

Da für die fünf gewählten Aspekte keine Items vorlagen, wurden diese selbst erstellt. Im Zuge der Operationalisierung wurde zunächst ein größerer Pool an möglichen Formulierungen für die ausgewählten Aspekte generiert. Diese wurden anschließend mit der Projektgruppe sowie fachinternen Kolleg:innen diskutiert, selektiert und gegebenenfalls hinsichtlich der Formulierungen angepasst. Die Items gingen dann in Pilotierungen durch kognitive Interviews ein und wurden danach auf Basis der Ergebnisse erneut selektiert und hinsichtlich der Formulierungen präzisiert (Vorgehen und Ergebnisse siehe Kap. 5.2.6). Insgesamt stand neben den inhaltlichen Kriterien zur Auswahl der Items auch der zeitliche Rahmen der Erhebung im Fokus. Um die Zumutbarkeit der Erhebung für die Teilnehmer:innen zu gewährleisten, wurde die Itemanzahl begrenzt und nur als zentral angenommene Indikatoren wurden in den Fragebogen aufgenommen.

Für die Erfassung der Merkmalsausprägung wurde eine sechsstufige, bipolare Ratingskala eingesetzt. Durch die fehlende mittlere Antwortkategorie werden die Teilnehmer:innen dazu gebracht, ihre Einschätzung zumindest tendenziell anzugeben, und dadurch die Schwierigkeiten der Interpretation einer neutralen Antwortkategorie (z. B. durch Indifferenz oder Ambivalenz bzw. die nicht gegebene Differenzierbarkeit) vermieden (Bortz & Döring, 2006). Die Ausprägungen des Merkmalskontinuums wurden stets mit den Bezeichnungen „stimme voll zu“, „stimme zu“, „stimme eher zu“, „stimme eher nicht zu“, „stimme nicht zu“ und „stimme gar nicht zu“ als verbale Marken überschrieben. Menold und Bogner (2015) geben bezüglich verbalisierter Ratingskalen einen Überblick zu Forschungsarbeiten, die empirisch die Bevorzugung dieser Form aufgrund erhöhter Reliabilität und Validität sowie Präferenzangaben von Teilnehmer:innen an entsprechenden Untersuchungen nahelegen.

Für die Aspekte Adressatenorientierung, Strukturiertheit, sprachliche Verständlichkeit sowie Sprech- und Körperausdruck wurde jeweils ein Item implementiert, das die globale Bewertung dieses Aspekts erfassen soll – konstruktbezogene Globalurteile. Die Globalurteile können anschließend auch zur Abschätzung der Validität der Skalen genutzt werden (siehe Kap. 6.1.2).

5.2.4.1 Adressatenorientierung

Unter Adressatenorientierung wird in dieser Studie die Anpassung der Erklärung und des Erklärprozesses an die Voraussetzungen der Rezipient:innen verstanden, um Verstehensprozesse zu ermöglichen und den Aufbau von Wissen zu unterstützen. In der Konzeptualisierung der Adressatenorientierung wurde in der Projektgruppe der Fokus auf die kognitiven Indikatoren gelegt. Dabei sind die kognitiven Aspekte zentral, wenn es um die Vermeidung trägen Wissens bzw. um den Aufbau von neuem Wissen geht (Renkl, 2004). Gleichzeitig stellt die Anpassung von Erklärungen an das Vorwissen eine zentrale Schwierigkeit für Lehrkräfte dar (Renkl et al., 2007).

Als zentrales Konzept bzw. als zentrale Methode für die Anpassung von Unterrichtsinhalten an das Vorwissen, Vorstellungen und kognitive Fähigkeiten der Schüler:innen wurde in der Konzeptualisierung der Adressatenorientierung weiterhin die didaktische Reduktion berücksichtigt (Item: *Ad_red*). Die didaktische Reduktion ist zentral, um den Umfang des Erklärinhalts zu reduzieren (sektorale Reduktion) und die Komplexität (auch im Sinne des Detailreichtums; Aussparung von Teilaspekten, die wiederum einer Erklärung bedürften) des Erklärinhalts zu reduzieren (strukturelle Reduktion). Das entsprechende Item wurde ausschließlich den Lehrenden (SLD-Gruppe) vorgelegt, da die didaktische Reduktion von Schüler:innen aufgrund fehlenden Professionswissens nicht beurteilt werden kann.

Insgesamt sollte eine angemessene didaktische Reduktion auch in einem angemessenen Schwierigkeitsgrad der Erklärung für die Adressat:innen münden. Der Schwierigkeitsgrad einer Erklärung wurde in der Konzeptualisierung gesondert aufgenommen, da die Anpassung des Schwierigkeitsgrads nicht nur durch die Anpassung an das Vorwissen, sondern auch durch die Anpassung an die kognitiven Fähigkeiten der Schüler:innen realisiert werden kann (Item: *Ad_sch*).

Ein weiteres Item nimmt direkt Bezug auf das Vorwissen der Schüler:innen bzw. die Angepasstheit der Erklärung an das Vorwissen der Rezipient:innen (Item: *Ad_vwi*; Pauli, 2015; Renkl, 2004; Schopf & Zwischenbrugger, 2015; Wellenreuther, 2013).

In der Konzeptualisierung wird weiterhin der Lebensweltbezug der Erklärung aufgegriffen (Item: *Ad_lbw*). Ein Lebenswelt- oder Alltagsbezug wird häufig gefordert. Angenommen wird ein positiver Effekt auf die Motivation sowie auf das Anbinden des neuen Wissens an bereits bestehendes Wissen (Pauli, 2015; Schopf & Zwischenbrugger, 2015; Wellenreuther, 2013).

Um trotz des Fokus auf die kognitiven Indikatoren die Breite des grundlegenden Aspekts der Adressatenorientierung erfassen zu können, wurden alle Teilnehmenden allgemein nach der Passung für die Adressat:innen befragt (Item: *Ad_pas*).

Innerhalb der kognitiven Indikatoren wurden die Anpassung der Erklärung an das Abstraktionsvermögen (Treagust & Harrison, 2000; Wagner & Wörn, 2011) sowie an die Kapazität des Arbeitsgedächtnisses der Rezipient:innen (Wellenreuther, 2013) aufgrund ihrer schwierigen Erfassung nicht in die Konzeptualisierung und Operationalisierung aufgenommen. Aufgrund der

monologischen Situation sind das Überprüfen des Verstehens des Inhalts im Nachgang durch die erklärende Person sowie das Eingehen der erklärenden Person auf wahrgenommene Rezipient:innensignale, die ein Nichtverstehen signalisieren können, ausgeklammert (Kennedy, 1996; Kulgemeyer & Tomczyszyn, 2015; Schopf & Zwischenbrugger, 2015; Sevia & Gonsalves, 2008; Treagust & Harrison, 2000). Keine Berücksichtigung in der Konzeptualisierung und Operationalisierung des Messgegenstands findet ebenfalls die Berücksichtigung der Motivation (Treagust & Harrison, 2000) und des Interesses (Kulgemeyer & Tomczyszyn, 2015; Treagust & Harrison, 2000).

Die Operationalisierung ist in Tabelle 6 einzusehen. Das Item *Ad_glo* dient als Globalurteil für die Adressatenorientierung der gesamten Erklärung. Dieses wurde aufgrund des didaktischen Fachwortes ‚adressatengerecht‘ lediglich der Gruppe der Lehrenden präsentiert.

Tabelle 6: Itemformulierungen zum Aspekt ‚Adressatenorientierung‘

Item-abkürzung	Items für SLD	Items für SuS
Ad_red	Die Erklärung war angemessen didaktisch reduziert.	/
Ad_sch	Die Erklärung war für Schüler_innen zu schwierig.	Die Erklärung war für mich zu schwierig.
Ad_vwi	Schüler_innen konnten der Erklärung mit ihrem Wissen gut folgen.	Ich konnte der Erklärung mit meinem Wissen gut folgen.
Ad_lbw ¹	Die Inhalte der Erklärung berücksichtigte die Lebenswelt der Schüler_innen.	Die Inhalte der Erklärung hatten etwas mit meinem Leben zu tun.
Ad_pas	Die Erklärung passte für Schüler_innen dieser Jahrgangsstufe.	Die Erklärung passte für mich.
Ad_glo	Die Erklärung war adressatengerecht.	/

Bem.: SLD: Gruppe der Studierenden, Lehrkräfte und Didaktiker:innen; SuS: Gruppe der Schüler:innen; /: das Item wurde dieser Gruppe nicht vorgelegt; ¹: Das Item wurde aufgrund negativer Trennschärfen ausgeschlossen.

5.2.4.2 Strukturiertheit

Für die Konzeptualisierung und Operationalisierung der Strukturiertheit einer Erklärung (siehe Tabelle 7) wurden von der Projektgruppe drei Facetten ausgewählt. Sie umfassen die Gegenstandsbestimmung, den logischen und schrittweisen Aufbau sowie die Fokussierung auf das Wesentliche. Die Facette der Transparenz des Aufbaus für die Rezipient:innen wurde dem Aspekt der verständlichen Sprache (Textkohärenz/Kohäsion) zugeordnet und ist dort durch ein Item repräsentiert.

Die Facette der Gegenstandsbestimmung beinhaltet, dass allen Adressat:innen klar ist, was der Erklärgegenstand ist (Item: *St_was*). Sie kann als zentral angesehen werden, da erst durch die Bestimmung dessen, was genau erklärt werden soll, aufseiten der Lehrkraft eine inhaltlich präzise Erklärung möglich wird und aufseiten der Schüler:innen beim Zuhören keine Gedächtniskapazitäten für die Erfassung des Erklärgegenstandes aufgewendet werden müssen.

Unter der Facette ‚Aufbau‘ werden hier die logische, kohärente Abfolge der Verstehenselemente (Item: *St_rot*) sowie deren schrittweise Einführung (Item: *St_bau*) in Anlehnung an Schopf und Zwischenbrugger (2015), Sevia und Gonsalves (2008) sowie Wragg (1993) verstanden. Ein gelungener Aufbau gewährleistet Stringenz und damit die Möglichkeit, dem Gesagten zu folgen

sowie die Teilschritte der Erklärung aufeinander zu beziehen, sodass er einen Kernaspekt von Strukturiertheit bildet.

Die Fokussierung auf das Wesentliche (Item: *St_ein*) adressieren in Übereinstimmung die Reduzierung von unwichtigen Exkursen oder Details in der Erklärung durch die Lehrkraft, sodass der Inhalt auf das Nötige und Wichtige konzentriert ist. So kann die Aufmerksamkeit der Adressat:innen auf relevante Aspekte gelenkt und mitunter die Effektivität der Erklärung gesteigert werden (Schopf & Zwischenbrugger, 2015).

Keine Berücksichtigung in der Konzeptualisierung und Operationalisierung findet hingegen die Wahrnehmung zur Herausarbeitung der relevanten Verstehenselemente durch die erklärende Person, wie sie beispielsweise bei Hargie (2013), Kennedy (1996) und Wragg (1993) in die Modellierung Eingang findet. Das Konzept der Schlüssel- oder Verstehenselemente ist zum einen abstrakt und zum anderen in der deutschsprachigen didaktischen Literatur nicht etabliert, sodass eine Einschätzung dessen sowohl für Schüler:innen als auch für die Teilnehmer:innen der Metaperspektive als schwer einzustufen ist. Des Weiteren wurden die Formulierung des Erklärziels nicht in der Konzeptualisierung und Operationalisierung repräsentiert. Sie gilt zwar als ein zentrales Qualitätskriterium (Kulgemeyer & Schecker, 2013; Leinhardt, 1987; Wagner & Wörn, 2011), wurde in den Vignetten der am Projekt beteiligten Fächer jedoch unterschiedlich umgesetzt, sodass eine Aufnahme in den fachunspezifischen Itemkatalog nicht möglich war. In den Videovignetten der hier vorgestellten biologiedidaktischen Studie wurde das Ziel der Erklärung in die Einleitung bzw. den Problemaufriss integriert (z. B. „und genau damit wollen wir uns jetzt beschäftigen, also warum es diese Schwankungen gibt, welche Ursache oder welche Gründe denn da dahinter stecken“ (Zitat aus dem Video Ö_k)). Das Erklärziel wird daher in den Erklärungen jeweils explizit verbalisiert. Die Relevanz des Themas bzw. die Einbettung des Erklärgegenstandes in einen größeren Zusammenhang (Kulgemeyer & Schecker, 2013; Schopf & Zwischenbrugger, 2015; Sevian & Gonsalves, 2008) sowie eine Zusammenfassung am Schluss (Kulgemeyer & Tomczyszyn, 2015; Odora, 2014; Wragg, 1993) wurden ebenso in den Videovignetten nicht realisiert, um die Erklärungen möglichst kurz zu halten. Für den Aspekt der Zusammenfassung wichtiger Punkte am Ende des Erklärprozesses zeigte sich darüber hinaus in Seminaren zum Themenbereich ‚Erklären‘, dass nach kurzen Erklärungen, wie sie hier als Stimuli realisiert und eingesetzt wurden, eine Zusammenfassung von den Studierenden eher als störend wahrgenommen wurde. Grund dafür könnte die fehlende Balance (Sevian & Gonsalves, 2008) zwischen den Teilen der Erklärung sein.

Die Operationalisierung ist in Tabelle 7 einzusehen. Die Formulierungen für die Gruppe der Lehrkräfte, Studierenden und Didaktiker:innen waren in diesem Fall die gleichen wie für die Gruppe der Schüler:innen. Das Item *St_glo* dient als Globalurteil für die Strukturiertheit der gesamten Erklärung.

Tabelle 7: Itemformulierungen zum Aspekt ‚Strukturiertheit‘

Item-abkürzung	Items für SLD und SuS
St_was	Es ist klar, was genau die Lehrerin erklären wollte.
St_rot	Die Erklärung hatte einen roten Faden.
St_bau	Die Lehrerin hat die Erklärung Schritt für Schritt aufgebaut.
St_ein	Die Lehrerin hat zu viele Einzelheiten erklärt.
St_glo	Die Erklärung war gut strukturiert.

Bem.: SLD: Gruppe der Studierenden, Lehrkräfte und Didaktiker:innen; SuS: Gruppe der Schüler:innen.

5.2.4.3 Sprachliche Verständlichkeit

Die Konzeptualisierung und Operationalisierung des Aspekts der sprachlichen Verständlichkeit wurden von Christiane Thim-Mabrey, Professorin für deutsche Sprachwissenschaft an der Universität Regensburg und FALKE-Projektbeteiligte, erarbeitet (Thim-Mabrey & Lindl, im Druck). Sie fasst die Konzeptualisierung des Aspekts ‚Verständliche Sprache‘ folgendermaßen zusammen:

Die sprachliche Verständlichkeit einer Erklärung erfordert, dass man einerseits Ausdrucksweisen verwendet, die das inhaltliche Verstehen für die Zuhörer unterstützen, und andererseits Ausdrucksweisen vermeidet, die das Verstehen erschweren. Hierfür zentrale sprachliche Mittel sind verständliche Wörter und Begriffe, eine angemessene Satzlänge sowie gezielte Hervorhebungen, um die Aufmerksamkeit der Zuhörer auf Wichtiges zu lenken. (Thim-Mabrey & Lindl, im Druck)

Insgesamt rekurren die drei Facetten auf die Verständlichkeit der Ebenen von Wort (Item: *Sw_bed*), Satz (Item: *Sw_zul*) und Text (Item: *Sw_wic*), die auch in der didaktischen Diskussion um den Einsatz bzw. die Verständlichkeit von naturwissenschaftlicher Fachsprache aufgegriffen werden (Leisen, 2010; Rincke, 2010; Zieprecht et al., 2017).

In Tabelle 8 sind die im Fragebogen enthaltenen Items zum Aspekt der sprachlichen Verständlichkeit eingetragen. Das Item *Sw_ver* repräsentiert den globalen Eindruck der sprachlichen Verständlichkeit der erklärenden Person.

Tabelle 8: Itemformulierungen für den Aspekt ‚Sprachliche Verständlichkeit‘

Item-abkürzung	Items für SLD	Items für SuS
<i>Sw_ver</i>	Die Lehrerin hat sich gut verständlich ausgedrückt.	
<i>Sw_wic</i> ¹	Manchmal hat die Lehrerin extra gesagt, dass etwas sehr wichtig ist.	
<i>Sw_bed</i>	Bei manchen Wörtern wissen die Schüler_innen eventuell nicht, was sie bedeuten.	Bei manchen Wörtern habe ich nicht gewusst, was sie bedeuten.
<i>Sw_zul</i>	Manche Sätze hat die Lehrerin zu lang gemacht.	

Bem.: SLD: Gruppe der Studierenden, Lehrkräfte und Didaktiker:innen; SuS: Gruppe der Schüler:innen; ¹: Das Item wurde aufgrund negativer Trennschärfen ausgeschlossen.

5.2.4.4 Sprech- und Körperausdruck

Die Konzeptualisierung und Operationalisierung des Aspekts zum Sprech- und Körperausdruck der erklärenden Person wurden von Eileen Gunga, Mitarbeiterin der Sprechwissenschaft an der Universität Regensburg im Projekt FALKE, erarbeitet. Die Argumente für die Auswahl der Aspekte lehnen sich daher an die von ihr in der gemeinsamen Projektpublikation genannten an (Lägel-Gunga et al., im Druck) Lägel-Gunga et al. (im Druck) fasst die Konzeptualisierung folgendermaßen zusammen:

Der Sprech- und Körperausdruck der erklärenden Lehrkraft umfasst zum einen akustisch-auditiv wahrnehmbare Merkmale gesprochener Sprache melodischer, dynamischer, temporaler und artikulatorischer Art sowie die Stimmqualität. Zum anderen werden darunter körperausdrucksbezogene Spezifika wie Mimik und Blickverhalten, Gestik, Körperhaltung, -spannung und -orientierung sowie die Bewegung und das Verhalten im Raum gefasst. (Lägel-Gunga et al., im Druck)

Die gewählten Facetten repräsentieren damit die Wahrnehmung der Prosodie (Items: *Se_sti*, *Se_ges*, *Se_pau*, *Se_abw*), der Artikulation (Item: *Se_aus*) sowie des Körperausdrucks (Items: *Se_kor*, *Se_zug*). Damit sind sowohl paraverbale (stimmlich-sprecherische) als auch extraverbale (körperausdrucksbezogene) Merkmale repräsentiert. Den genannten Merkmalen ist dabei gemein, dass sie das Verstehen des Gesagten beeinflussen können (siehe Kap. 4.5; Allhoff & Allhoff, 2010; Apel, 2009; Brown, 2006; Hargie, 2013; Morton & Watson, 2001; Neuber, 2016; Voigt-Zimmermann, 2011; Wragg, 1993) und damit als Variablen für gutes Erklären relevant sind.

Nicht berücksichtigt werden konnte in der Konzeptualisierung und Operationalisierung des Sprech- und Körperausdrucks für die vorliegende Erhebung der Blickkontakt der erklärenden Person, obwohl dieser eine Möglichkeit der Kontaktaufnahme zu den Zuhörer:innen mit extraverbalen Mitteln darstellt und darüber Aufmerksamkeit generiert werden kann (Allhoff & Allhoff, 2010). Aufgrund der Realisierung der mündlichen Erklärungen in Videovignetten ist jedoch davon auszugehen, dass dieser nicht in vergleichbarer Weise eingeschätzt werden würde wie in der Realsituation. Zudem wurde die Wahrnehmung der Mimik als extraverbales Mittel nicht abgefragt, da diese mitunter nicht in jeder der Videovignetten aufgrund der gewählten Kameraeinstellung erkennbar ist.

In Tabelle 9 sind die konkreten Formulierungen der Items eingetragen. Diese waren für Schüler:innen und Lehrende identisch formuliert. Das Item *Se_glo* stellt das skalenspezifische Globalurteil dar.

Tabelle 9: Itemformulierungen für den Aspekt ‚Sprech- und Körperausdruck‘

Item-abkürzung	Items für SLD und SuS
<i>Se_sti</i>	Die Lehrerin hatte eine angenehm klingende Stimme.
<i>Se_aus</i>	Die Lehrerin hatte eine deutliche Aussprache.
<i>Se_ges</i>	Die Lehrerin hatte eine angemessene Sprechgeschwindigkeit.
<i>Se_pau</i>	Die Lehrerin hat die Sprechpausen so gesetzt, dass ich gut folgen konnte.
<i>Se_abw</i>	Die Lehrerin hatte eine abwechslungsreiche Sprechweise.
<i>Se_kor</i>	Die Lehrerin hatte einen Körperausdruck, der die Erklärung unterstützt hat.
<i>Se_zug</i>	Die Lehrerin hat sich den Schülerinnen und Schülern zugewendet.
<i>Se_glo</i>	Die Art und Weise, wie die Lehrerin spricht, gefällt mir.

Bem.: SLD: Gruppe der Studierenden, Lehrkräfte und Didaktiker:innen; SuS: Gruppe der Schüler:innen.

5.2.4.5 Wahrnehmung des dargestellten Komplexitätsgrades

Der fachspezifische Schwerpunkt liegt auf der Wahrnehmung des dargestellten Komplexitätsgrades bzw. auf der Einschätzung, ob die Darstellung des biologischen Phänomens hinsichtlich des Komplexitätsgrades als angemessen beurteilt wird.¹¹² Die Wahrnehmung des dargestellten Komplexitätsgrades liegt damit am Schnittpunkt zwischen der Orientierung am Gegenstand und dessen adäquater Darstellung in Abhängigkeit von den Voraussetzungen der Adressat:innen.

¹¹² Die allgemeinen Facetten fachlich guter Erklärungen ‚Fachliche Korrektheit‘ und ‚Fachliche Vollständigkeit‘ wurden von der Messung ausgenommen. Eine Begründung findet sich bereits in Kapitel 5.2.4 bei der allgemeinen Konzeptualisierung und Operationalisierung guten Erklärens.

Dazu wurden zwei Facetten ausgewählt, die zwei zentrale Unterschiede zwischen linearen und komplexen Erklärungen betreffen: So wurden Kriterien zur Auswirkung des Gesagten berücksichtigt, das zwischen linearen und komplexen Erklärungen variieren kann, sowie Kriterien zur Wahrnehmung der Darstellung der Komplexitätsgrades implementiert, die die Wirkung der Orientierung am Phänomen im Gegensatz zur Wirkung der Orientierung an allgemeinen, übertragbaren Prinzipien erfassen.

Die Facette zur Auswirkung des Gesagten greift dabei zwei Kriterien auf. Zum einen das Gefühl der Unsicherheit (Item: *Fa_ver*), das insbesondere das Aufzeigen eines hohen Komplexitätsgrades verursachen kann, da die explizite Erwähnung von Nichtwissen und unsicherem Wissen zum Fehlen von Orientierung beiträgt (Böschchen et al., 2008; Kampourakis & McCain, 2020; Maier et al., 2018; Mehren et al., 2015; Varwig, 2020). Ein zu hohes Gefühl der Unsicherheit kann das Lernen behindern. Zum anderen die Auswirkung der Faszination und Interessiertheit (Items: *Fa_fas*, *Fa_erf*), die durch einen höheren Komplexitätsgrad mitunter sogar gesteigert werden kann (Retzbach & Maier, 2015) und die eine wichtige Voraussetzung für das Lernen ist.

Um die Wirkung der Orientierung am Phänomen oder an allgemeinen übertragbaren Prinzipien zu erfassen, wurden drei Indikatoren gewählt:

Die Angemessenheit der Anzahl der einbezogenen Faktoren und Zusammenhänge, die in der Erklärung herangezogen werden (Item: *Fa_zan*, *Fa_zvu*), wurde ausgewählt. Während komplexe Phänomene das Aufzeigen zahlreicher Faktoren verlangen, die über Wechselwirkungen miteinander verbunden sind (Kuhlmann, 2007; Mitchell, 2008), zeigen lineare Erklärungen eine Ursache auf, mit der das Phänomen erklärt wird. In Abhängigkeit des dargestellten Phänomens sollte die Anzahl der Faktoren und Zusammenhänge für Rezipient:innen angemessen sein. Für die Erklärung komplexer Phänomene stellt die Anzahl der Faktoren und Zusammenhänge eine Herausforderung dar, damit ein hoher ‚cognitive load‘ einhergehen kann (Grotzer et al., 2017).

Des Weiteren wurde die Adäquatheit der Abbildung des biologischen Phänomens (Items: *Fa_nat*, *Fa_aus*) aufgenommen, die die Angemessenheit der Darstellung des realen Phänomens in Abhängigkeit von der Perspektive der Teilnehmenden und des Ziels der Erklärung erfasst. Durch eine bewusste oder unbewusste Reduktion der Faktoren und Zusammenhänge für den unterrichtlichen Vermittlungszweck kann das reale Phänomen zunehmend weniger umfassend repräsentiert oder muss eher modellhaft und exemplarisch verstanden werden. Durch derartige Reduktionen gewinnt eine Erklärung insgesamt an Linearität und Eindeutigkeit, die realen Phänomenen nur in Teilen gerecht wird und daher ein begrenztes bzw. verzerrtes Verständnis vermittelt (Jacobson & Wilensky, 2006). Eine Bewusstheit dafür, dass die Erklärungen jeweils nur einen Ausschnitt des Phänomens darstellen bzw. modellhaft zu verstehen sind, ist für ein adäquates Verständnis der Phänomene relevant.

Zuletzt wurde die Übertragbarkeit der Inhalte auf andere Beispiele (Item: *Fa_hil*) ausgewählt. Renkl et al. (2007) formulieren als ein Qualitätskriterium von instruktionalen Erklärungen die Orientierung an grundlegenden, fachlichen Prinzipien, sodass Schüler:innen das Gelernte auf weitere Beispiele transferieren können und damit zum selbstständigen Problemlösen befähigt werden. Während bei komplexen Erklärungen die Komplexität selbst als Prinzip verstanden werden kann, das biologische Phänomene charakterisiert (Mitchell, 2008), sind bei linearen Erklärungen die teils modellhaften und exemplarischen Zusammenhänge prinzipienorientiert aufzufassen.

In Tabelle 10 sind die konkreten Formulierungen der Items des Konstrukts eingetragen.

Tabelle 10: Itemformulierungen zum Aspekt der Wahrnehmung des dargestellten Komplexitätsgrades

Item-abkürzung	Items für SLD	Items für SuS
Fa_zan	Die Anzahl der Zusammenhänge ist für Schüler_innen angemessen.	Die Anzahl der Zusammenhänge ist für mich angemessen.
Fa_zvu*	Für Schüler_innen zeigt die Erklärung zu viele Ursachen.	Für mich zeigt die Erklärung zu viele Ursachen.
Fa_ver*	Die Tatsache, dass nicht alle Dinge bekannt sind, verunsichert Schüler_innen.	Die Tatsache, dass nicht alle Dinge bekannt sind, verunsichert mich.
Fa_fas	Die Erklärung ist für Schüler_innen faszinierend.	Die Erklärung ist für mich faszinierend.
Fa_erf	Nach dieser Erklärung möchten Schüler_innen noch mehr darüber erfahren.	Nach dieser Erklärung möchte ich noch mehr darüber erfahren.
Fa_nat	Die Erklärung zeigt, wie es in der Natur wirklich ist.	Die Erklärung zeigt, wie es in der Natur wirklich ist.
Fa_aus ¹	Die Erklärung zeigt nur einen kleinen Ausschnitt auf die Frage („Thema der Erklärung“).	/
Fa_hil	Der Inhalt der Erklärung hilft Schüler_innen andere Beispiele zu verstehen.	/

Bem.: SLD: Gruppe der Studierenden, Lehrkräfte und Didaktiker:innen; SuS: Gruppe der Schüler:innen; *: Das Item wurde nur bei komplexen Erklärungen abgefragt und daher nicht in die Skala aufgenommen; /: Das Item wurde dieser Gruppe nicht vorgelegt; ¹: Das Item wurde aufgrund negativer Trennschärfen ausgeschlossen.

5.2.5 Kontrolle möglicher Störvariablen innerhalb des Fragebogens Kontrollvariablen

In den Fragebogen wurden zur Sicherung der internen Validität mögliche Ursachen für systematische Verzerrungen und Verfälschungen des wahren (Test-)Wertes der Items entsprechende Kontrollvariablen aufgenommen respektive die Reihenfolge der Videovignetten systematisch randomisiert (Bortz & Döhring, 2006). Neben der Videoreihenfolge betreffen mögliche Moderatoren auf die wahrgenommene Erklärqualität die Wirkung der erklärenden Person, das Vorwissen der Schüler:innen zum Erklärgegenstand und das Vorwissen der Studierenden in Hinblick auf erworbenes Wissen zu Kriterien guten Erklärens in universitären Seminaren. Die Kontrollvariablen werden im Folgenden kurz beschrieben und ihre Operationalisierung wird genannt. Die Ergebnisse der Analysen zu den jeweiligen Einflüssen und den Konsequenzen bei den Interpretationen der Ergebnisse finden sich in Kapitel 6.1.1.

Videoreihenfolge

Zur Kontrolle eines möglichen Reihenfolgeeffektes (Bortz & Döhring, 2006) bezüglich der Videoreihenfolge wurde diese systematisch randomisiert und zwei Versionen (A und B) des Fragebogens wurden erstellt. Dabei wechseln sich in beiden Versionen stets komplex und linear gestaltete Erklärungen ab. Weiterhin werden die thematisch ähnlichen Videovignetten nicht direkt hintereinander gezeigt, wodurch sich die in Abbildung 12 dargestellte Reihenfolge ergibt.

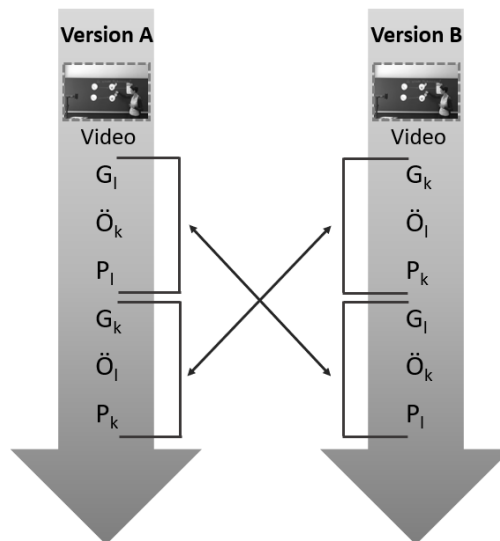


Abbildung 12: Systematisch randomisierte Videoreihenfolge in den zwei Versionen des Fragebogens

Die Versionen wurden den Teilnehmer:innen zufällig zugeteilt, wobei die Teilstichprobe der Schüler:innen je Klassenverbund einer Version zugewiesen wurden.

Wirkung der erklärenden Person

Als ein bekannter Beurteilungsfehler tritt der sogenannte Halo-Effekt auch in Kombination mit empfundener Sympathie gegenüber der zu beurteilenden Person auf. Der Halo-Effekt beschreibt allgemein eine Tendenz, bei der Beurteilung auf globale, holistische Wahrnehmungen zu rekurrieren und nicht das eigentlich zu beurteilende Merkmal differenziert zu betrachten (Bortz & Döhring, 2006; Krapp & Weidenmann, 2006). Übertragen auf die vorliegende Studie bedeutet das, dass sympathischen Erklärpersonen auch eine hohe Erklärkompetenz zugeschrieben bzw. davon abgeleitet die Erklärqualität als gut wahrgenommen wird – und invers dazu unsympathischen Erklärpersonen eine niedrige Erklärkompetenz zugeschrieben wird. Das entsprechende Item („Die Lehrerin wirkte in diesem Video auf mich sympathisch“ (Item: *Sy_sym*)) richtete sich an Schüler:innen und Lehrende und wurde auf der gleichen sechsstufigen Ratingskala eingeschätzt.

Aus dem Bereich der ‚teacher enthusiasm‘ Forschung liegen diverse Befunde zum Zusammenhang des gezeigten Enthusiasmus einer Lehrperson und der Wahrnehmung sowie dem Outcome von Unterricht vor (Babad, 2009; Keller, Hoy, Goetz & Frenzel, 2016).¹¹³ Bisher werden in empirischen Zugängen unterschiedliche Konzeptualisierung und Operationalisierung des Konstrukts zugrunde gelegt, wobei zudem viele Überschneidungen mit verwandten Konstrukten existieren (Babad, 2009; Bleck, 2019; Keller et al., 2016). Eine Konzeptualisierung rückt den nonverbalen Ausdruck der Lehrkraft in den Fokus (Keller et al., 2016). Indikatoren für die Einschätzung des latenten Konstrukts sind beispielsweise für die Ausprägung eines hohen Enthusiasmus eine deutliche Veränderung der Stimmmodulation, weit geöffnete Augen und Blickkontakt mit der Gruppe sowie dynamische und deutliche Bewegungen als Teil einer wechselnden, lebhaften Gestik und Mimik (Collins, 1978).

¹¹³ Die Darstellung wird an dieser Stelle eingegrenzt auf den gezeigten Enthusiasmus („displayed enthusiasm“; Keller et al., 2016, S. 745). Zu unterscheiden ist dieser laut den Autor:innen vom empfundenen Enthusiasmus („experienced enthusiasm“; Keller et al., 2016, S. 745), den Lehrkräfte für sich selbst einschätzen können und der sich in Enthusiasmus gegenüber dem Fach und gegenüber der Aktivität des Lehrens unterscheiden lässt.

Hinsichtlich des Einflusses des Enthusiasmus auf Schüler:innen resümieren Keller und Kolleg:innen (2016) aus einer Vielzahl der existierenden Studien, dass eine hohe wahrgenommene Begeisterung der Lehrperson zu positiven Effekten bezüglich der Motivation sowie der Interessiertheit von Schüler:innen führt. Weniger einheitliche Befunde existieren hingegen zur Entwicklung von Lernergebnissen: Kunter et al. (2013) gehen in ihrer Studie unter anderem dem Zusammenhang zwischen Merkmalen der Lehrkraft („pedagogical content knowledge“, „constructivist beliefs“, „teaching enthusiasm“ und „adaptive self-regulation“; Kunter et al., 2013, S. 808–809) und Merkmalen des Lernerfolgs („students' mathematics achievement“ und „students' mathematics enjoyment“; Kunter et al., 2013, S. 811) in der darauffolgenden Jahrgangsstufe nach und können einen signifikanten, positiven Zusammenhang zwischen dem Enthusiasmus der Lehrperson¹¹⁴ und dem Erfolg der Schüler:innen hinsichtlich beider Aspekte feststellen.¹¹⁵ Dementgegen können McKinney und Kolleg:innen (1983) in einem experimentellen Zugang mit trainierten Lehrkräften (diese zeigten in den darauffolgenden Unterrichtsstunden hohe, mittlere oder geringe Ausprägung an Enthusiasmus, wobei dieser an dem Kategorienvorschlag von Collins (1978) festgemacht wurde) keine systematischen Unterschiede in den Lernleistungen auf den gezeigten Enthusiasmus der Lehrperson zurückführen. Feldman (2007) untersuchte den Zusammenhang zwischen der Einschätzung von Studierenden zum Enthusiasmus der Lehrperson (gegenüber dem Fach und dem Lehren) und der insgesamt wahrgenommenen Lehrqualität. Die Ergebnisse zeigen einen moderaten Effekt auf die wahrgenommene Lehrqualität. Vor diesem Hintergrund erscheint die statistische Kontrolle dieses Aspektes insgesamt sinnvoll. Das implementierte Item war wie folgt formuliert: *Die Lehrerin wirkte in diesem Video auf mich begeistert* (Item: *Sy_beg*). Es wurde ebenfalls auf der sechsstufigen Ratingskala eingeschätzt. Die Paraphrasierung von ‚enthusiastisch‘ mit ‚begeistert‘ erfolgte aufgrund des antizipierten Wortverständnisses von Schüler:innen und der Bemühung, zentrale Begriffe in den Items von Lernenden und Lehrenden konstant zu halten.

In Pilotierungen verschiedener Videovignetten innerhalb der Projektgruppe zeigte sich zudem, dass die wahrgenommene Natürlichkeit der Person (Lägel-Gunga et al., im Druck)¹¹⁶ einen Einfluss auf die Einschätzung der Güte einer Erklärung nehmen kann, sodass das Item *„Die Lehrerin wirkte in diesem Video auf mich natürlich“* (Item: *Sy_nat*) implementiert wurde. Die Einschätzung erfolgte auch hier auf einer sechsstufigen Ratingskala.

Vorwissen der Schüler:innen zum Erkläregegenstand

Da das vorhandene Vorwissen der Schüler:innen das Verstehen der Erklärung und damit die Einschätzung der Qualität als abhängige Variable moderieren kann (Renkl et al., 2007), wurden die Schüler:innen zu einer Selbsteinschätzung bezüglich ihres themenbezogenen Vorwissens direkt nach der gesehenen Erklärung aufgefordert. Dabei handelt es sich um eine intuitive Einschätzung der

¹¹⁴ Konzeptualisiert und operationalisiert über zwei Aspekte, die sich in späteren Analysen auch als unterscheidbare Dimensionen erwiesen: Enthusiasmus für das Fach Mathematik und Enthusiasmus für das Lehren.

¹¹⁵ In weiteren Analysen erstellen die Autor:innen ein Mediationsmodell, bei dem die Merkmale der Lehrkraft über Variablen der Instruktion („cognitive activation“, „learning support“ und „classroom management“; Kunter et al., 2013, S. 809) die Merkmale des Lernerfolgs vorhersagen. Innerhalb des Modells mit der besten Passung hängt der Enthusiasmus (signifikant) positiv mit dem ‚learning support‘ zusammen, wobei dieser wiederum (signifikant) positiv auf die Variable ‚mathematics enjoyment‘ wirkt. Ebenso besteht ein (signifikanter) positiver Zusammenhang zwischen Enthusiasmus und dem ‚classroom management‘, das beide Merkmale des Lernerfolgs („mathematics achievement“ und ‚enjoyment‘) vorhersagen kann.

¹¹⁶ Diese war vor allem dann reduziert, wenn die erklärende Person sehr nahe an einem fest vorgegebenen Wortlaut bleiben musste und das Skript vermutlich noch nicht weit genug verinnerlicht wurde, um sich auf die Gestaltung der Situation zu konzentrieren.

eigenen Fähigkeiten, die zu einem bestimmten Zeitpunkt abgegeben wird (Bohl, 2016). Bei Selbsteinschätzungen des Fachwissens durch Schüler:innen ist zu beachten, dass die Schätzurteile von den jeweiligen Fähigkeitsselbstkonzepten der Lernenden abhängig sind (Stiensmeier-Pelster & Schöne, 2008). Des Weiteren neigen insbesondere leistungsschwächere Schüler:innen dazu, die eigene Leistung eher zu überschätzen (z. B. Dunlosky & Lipko, 2007). Trotz der beschriebenen Moderationen wird die Selbsteinschätzung für einen Richtwert der Kontrollvariable ‚Vorwissen‘ herangezogen. Die teilnehmenden Schüler:innen wurden dafür mit folgendem Item (Item: *Gl_wus*) und den drei Antwortalternativen befragt, wobei die dahinterstehenden Ziffern in Klammern die zugeordneten Werte auf der Ratingskala repräsentieren:

Was im Video erklärt wurde, ...

... wusste ich vorher schon. (= 1)

... wusste ich teilweise schon. (= 2)

... wusste ich vorher noch nicht. (= 3)

Vorwissen der Studierenden

Als mögliche Störvariable wurde auch der Seminarbesuch eines Tandemseminars (siehe Fußnote 3 in Kap. 1.3) des FALKE-Projektes zum Thema ‚Erklären‘ identifiziert. Das im Seminar erworbene Wissen zu Kriterien guten Erklärens kann sowohl die Einschätzung der Qualität als auch insbesondere die Begründung zur Notenvergabe beeinflussen. Daher wurde der Seminarbesuch als einfaches, bipolares Item mit den Antwortalternativen ‚Ja‘ und ‚Nein‘ aufgenommen.

Da am Projekt FALKE insgesamt elf Fachdidaktiken beteiligt waren und dort ein in weiten Teilen ähnlicher Fragebogen eingesetzt wurde, fragte ein einfaches, bipolares Item (Ja/Nein) auch die Teilnahme an weiteren Umfragen der FALKE Projektgruppe ab. Durch die Bearbeitung eines Fragebogens aus einer anderen Fachdidaktik könnte es ebenfalls zu Verzerrungen hinsichtlich des Globalurteils kommen, da dieses unvoreingenommen, also ohne die Vorgabe von Kriterien, vergeben werden soll. Davon kann nicht mehr ausgegangen werden, wenn bereits ein anderer Fragebogen und damit verbunden die kriteriengeleitete Einschätzung von Videovignetten bearbeitet wurden. Dies betraf lediglich $N = 3$ Personen, sodass aufgrund der geringen Anzahl keine weiteren Korrekturen vorgenommen werden.

5.2.6 Pilotierung des Messinstruments mittels kognitiver Interviews

Sowohl die Items als auch die Videovignetten wurden vor dem Einsatz hinsichtlich der Eignung für den Einsatz im Forschungsinstrument überprüft. Die fachübergreifenden Items wurden zunächst mit den Mitarbeiter:innen des Projekts hinsichtlich deren Verständlichkeit diskursiv überprüft. Dazu konnte insbesondere die Vertreterin der Sprachwissenschaft mit Ergebnissen der Verständlichkeitsforschung beitragen. Die fachlichen Items wurden hingegen fachintern diskutiert. Die Items wurden anschließend im Rahmen von Pilotierungsinterviews Vertreter:innen der vier Statusgruppen vorgelegt und es wurde überprüft, ob das gleiche semantische Verständnis vorliegt, das die Forscher:innen bei der Erstellung der Items intendiert hatten und damit einem Pretest zur Feststellung der Funktionsfähigkeit des Befragungsinstruments unterzogen. Aufgrund der hohen Anzahl an Items wurden die fachübergreifenden Aspekte der Adressatenorientierung und der Strukturiertheit innerhalb der Forschungsgruppe aufgeteilt. Die Pilotierungen der Skalen zur sprachlichen Verständlichkeit sowie zum Sprech- und Körperausdruck wurden von den

Projektmitarbeiterinnen der jeweiligen Fachwissenschaften selbst durchgeführt. Die Pilotierung der fachspezifischen Items wurden von den jeweiligen Mitarbeiter:innen der Fachdidaktiken durchgeführt. Die Pilotierungsinterviews der hier vorgestellten biologiedidaktischen Studie fokussierten deshalb die Items der Adressat:innenorientierung sowie des fachspezifischen Aspekts. Sie wurden mit Schüler:innen ($N = 2$), Studierenden ($N = 2$) und Biologiedidaktiker:innen ($N = 2$) geführt. Gezeigt wurden je zwei Vignetten.

Für diesen Schritt in der Fragebogenkonstruktion wurde die aus der Kognitionspsychologie stammende Methode des kognitiven Interviews bzw. genauer die Technik des ‚verbal probings‘ genutzt (Prüfer & Rexroth 2005). Sie stellt die kognitiven Prozesse der Teilnehmenden bei der Beantwortung der Items in den Fokus, indem der Forscher oder die Forscherin als Interviewer:in gezielte Nachfragen dazu stellt. Der Antwortprozess wird allgemein in einem Vierschritt beschrieben: Verstehen der Frage, Abruf relevanter Informationen aus dem Gedächtnis, Fällern eines Urteils sowie Abgleich mit der zum Urteil passenden Antwortkategorie im Fragebogen (Tourangeau, Rips & Rasinski, 2000). Aus dieser modellierten Abfolge ergeben sich für den Antwortprozess mögliche Quellen fehlerhaften Antwortverhaltens, denen in den Interviews nachgegangen wird. Die Fragen der kognitiven Interviews beziehen sich deshalb auf das Verständnis einer Frage, eines Aspekts oder eines Begriffs („comprehension probing“) sowie auf die Wahl der entsprechenden Antwortkategorie („category selection probing“), den Besitz relevanter Informationen zur Beantwortung dieses Items, also die Informationsbeschaffung zur Beantwortung der Frage („information retrieval probing“), und schließlich unspezifische Nachfragen zu Problemen („general probing“; Prüfer & Rexroth, 2005, S. 6–9).¹¹⁷

Bei der Pilotierung des Erhebungsinstrumentes wurde bei Schüler:innen eine Papierversion (im Gegensatz zur späteren Computerversion) eingesetzt und die Videovignetten wurden an entsprechender Stelle über ein Notebook abgespielt. Die für die Pilotierung wesentlichen Merkmale blieben jedoch erhalten. Das bedeutet, dass die generelle Abfolge der Fragebogenteile, die Strukturierung der Seiten, die Abfolge der Items sowie die Gestaltung der Ratingskalen bezüglich der Wortwahl sowie die Anordnung der Kategorien aus dem Computerformat identisch übernommen wurden. Studierende und Lehrkräfte wurden in Örtlichkeiten der Universität Regensburg befragt und konnten so die Pilotierung an einem Computer mit stabiler Internetverbindung durchführen.

Im Rahmen der Interviews wurden auch die instruktionalen Texte des Onlinefragebogens verwendet. Dabei wurde vor allem auf die Verständlichkeit der Items und Texte Wert gelegt, um die Reliabilität und Validität der Ergebnisse zu gewährleisten bzw. zu erhöhen.

Während des Interviewverlaufs internalisierten die Teilnehmer:innen aller Teilstichproben den Ablauf der Fragen zu den Items sehr schnell, sodass die Nachfragen nicht für jedes Item von der Interviewerin verbalisiert werden mussten. Als besonders informationsreich für den Abgleich des semantischen Verständnisses erwiesen sich die Nachfragen zum Verständnis eines bestimmten

¹¹⁷ Im Rahmen von Pretestverfahren bzw. kognitiven Interviews wird auch häufig die Technik des lauten Denkens eingesetzt. Sie birgt jedoch einige Nachteile gegenüber dem ‚verbal probing‘: Für die Beantwortung eines Items auf einer Ratingskala ist sie nur bedingt geeignet (Prüfer & Rexroth, 2005), sie fordert von den Teilnehmer:innen ein hohes Maß an Eigenaktivität, sodass vielen die Informationspreisgabe nicht gelingt und in jedem Fall vorab eine aufwendige Übungsphase vorausgehen muss und durch die Freiheit der Äußerungen gegebenenfalls der Output in Bezug auf den weiteren Konstruktionsprozess mitunter gering ist (Prüfer & Rexroth, 2005). Aus den genannten Gründen wurde die Technik des ‚verbal probings‘ gewählt.

Begriffs innerhalb eines Items, die Wahl der Antwortkategorie sowie die unspezifischen Nachfragen vor allem für die Formulierung der instruktionalen Texte.

Von allen Teilnehmer:innen wurden die instruktionalen Texte zur Erklärung des Vorhabens und der Aufgaben innerhalb des Fragebogens als hilfreich bzw. verständlich charakterisiert. Während des Interviewfortgangs (vor allem bei unspezifischen Nachfragen) zu den Items der Adressatenorientierung sowie der Wirkung der Vignetten (siehe unten) äußerten sich die befragten Studierenden und Didaktiker:innen jedoch zur fehlenden Involvierung von Schüler:innen in den Erklärprozess. Daher wurde die Einführung des Vorhabens um eine Instruktion ergänzt, in der hervorgehoben wird, dass die Vignetten unter der Vorgabe einer monologischen Gestaltung erstellt wurden („Die Lehrkraft sollte ohne Interaktion mit einzelnen Schülerinnen und Schülern erklären, da der Fokus auf die Erklärung der Lehrkraft gerichtet werden soll“), um eine reliable und valide Messung zu gewährleisten.

Die einführenden Texte zur den jeweiligen Videovignetten, die die vorhergehenden Aktivitäten schildern sowie notwendiges Vorwissen zur Erklärung entlasten, wurden als gut beschrieben. Die Texte wurden daraufhin in ihrer Konzeption und Formulierung so übernommen.

Hinsichtlich der Verständlichkeit der Items konnte sowohl bei den befragten Studierenden als auch bei den Didaktiker:innen nur in einem Fall ein Verständnisproblem aufgedeckt werden. Dies betraf das fachspezifische Item zur Entsprechung mit dem realen Phänomen. Die ursprüngliche Formulierung lautete: *Die Erklärung zeigt nur einen kleinen Ausschnitt des Phänomens*. Dabei traten in drei von vier Fällen Nachfragen seitens der Teilnehmer:innen auf, wie das Phänomen abzugrenzen sei. Daraufhin wurde das Item präzisiert und das konkrete Thema der Erklärung genannt. Die im Fragebogen aufgenommene Formulierung lautet: *Die Erklärung zeigt nur einen kleinen Ausschnitt der Antwort auf die Frage „Wie entsteht Mukoviszidose?“/„Wie entsteht ein Folsäuremangel?“ etc.*

In den Interviews mit Schüler:innen wurde diese Formulierung eingesetzt und wies keine Verständnisprobleme auf. Bei Schüler:innen wurden durch Nachfragen zur Wahl der Antwortkategorie weiterhin deutlich, dass Schwierigkeiten bestehen, die Items ausschließlich in Bezug auf die eigene Wahrnehmung hin einzuschätzen und die Mitschüler:innen nicht mit einzubeziehen. Eine Äußerung zum Item „*Die Erklärung war zu einfach*“¹¹⁸ lautete beispielsweise folgendermaßen:

- S: *(S. liest vor) Die Erklärung war zu einfach. (.) Also zu einfach gibt's ja eigentlich nicht, weil (.) wenn man's versteht, dann versteht man's. Es kommt immer drauf an, ob man es schwierig versteht oder man halt die einfache Formulierung braucht oder so.*
- I: *Okay.*
- S: *War zu einfach. Stimme nicht zu (.)*
- I: *Okay. (..) Und darf ich dich fragen, warum du da jetzt das fünfte und nicht das sechste angekreuzt hast?*
- S: *Ja, wegen diesen (.) Fachbegriffen. Weil eben manche Leute, die zum Beispiel jetzt eben nicht am Gymnasium sind oder jünger sind, die verstehen manche Fachbegriffe dann eben doch nicht, die ich dann eben schon verstehe.*

¹¹⁸ Das Item wurde in den Pilotierungskatalog mit aufgenommen, anschließend bei einer Kürzung jedoch wieder entfernt und ist damit nicht im endgültigen Messinstrument enthalten (für den Aspekt der Adressatenorientierung blieb das Item „*Die Erklärung war für mich zu schwierig*“ erhalten.)

- I: *Also du denkst dann beim Beantworten eigentlich auch an andere Kinder oder Schüler.*
- S: *Ja zum Vergleich so ein bisschen.*

Daher wurden die Items für diese Teilstichprobe um einen persönlichen Bezug erweitert (z. B. „Für mich zeigt die Erklärung zu viele Ursachen“ oder „Die Erklärung war für mich zu schwierig“). Analog dazu wurden die Items der Metaperspektive angepasst (z. B. „Für Schüler_innen zeigt die Erklärung zu viele Ursachen“ oder „Die Erklärung war für Schüler_innen zu schwierig“)

Die gezeigten Videovignetten wurden von Studierenden und Didaktiker:innen auf Nachfrage, überwiegend jedoch schon während der Befragung zum Verständnis der Items, mit Eigenschaften beschrieben, die auf die intendierte Umsetzung der UV in den Vignetten hinweisen. So lautet beispielsweise eine Äußerung einer Studierenden zum Item „Die Erklärung hilft, andere Beispiele zu verstehen“¹¹⁹:

- S: *Okay, hier würde ich jetzt sagen (.) dass, ähm (.), es wahrscheinlich schon helfen kann, aber dadurch, dass es halt eben (..) / also die Erklärung an sich (..) ähm (...) / Ich glaube bei der Aufgabe / Also die Erklärung ist schwieriger auf andere zu übertragen, weil es halt sehr / eher spezieller, oder halt ja oder ein bisschen komplexer einfach ist.*
- I: *Mhm.(...)*
- S: *[...] Also bestimmt hilft es (..) aber es ist vielleicht nicht so (.) geradlinig, wie jetzt das andere Video war.*

5.2.7 Stichprobe und Durchführung der Erhebung

Als Richtwert für die Größe der jeweiligen Teilstichproben wurden $N = 100$ Schüler:innen, $N = 50$ Biologielehramtsstudierende, $N = 30$ Biologielehrkräfte sowie $N = 30$ Biologiedidaktiker:innen angestrebt. Diese Verteilung dient aus theoretischer Sicht einem annähernd ausgewogenen Verhältnis von Lernenden und Lehrenden bzw. Rezipient:innenperspektive und Metaperspektive. Für die Studie mit explorativem Charakter ist die für eine quantitative Erhebung eher kleine Stichprobengröße ausreichend (Bortz & Döring, 2006).

Die zur Berechnung herangezogene Gesamtstichprobe umfasst insgesamt $N = 238$ Teilnehmer:innen, die sich zwischen Oktober 2017 und Januar 2019 freiwillig an der Studie beteiligten. Nach Ausschluss ungültiger Fälle (siehe Kap. 5.3) bestehen die Teilstichproben aus $N = 122$ Schüler:innen, $N = 51$ Biologielehramtsstudierenden, $N = 34$ Biologielehrkräften sowie $N = 31$ Biologiedidaktiker:innen. Alle Teilnehmer:innen wurden vor der Bearbeitung über die Ziele der Studie sowie den Umgang mit den entstehenden Daten aufgeklärt und nahmen anschließend freiwillig daran teil. Tabelle 11 gibt zunächst einen Überblick über die Teilstichproben. In den folgenden Abschnitten finden sich zudem Beschreibungen zur Durchführung der Erhebungen.

¹¹⁹ Nach der Pilotierungsphase wurde das Item im eingesetzten Fragebogen mit der geschilderten Erweiterung geführt („Die Erklärung hilft Schüler_innen, andere Beispiele zu verstehen“).

Tabelle 11: Überblick über die Zusammensetzung der Gesamtstichprobe

Statusgruppe	<i>N</i>	Geschlecht w/m/d	Schulart GY/RS	Jahrgangsstufe bzw. Erfahrung <i>M (SD)</i>	Funktion bzw. akademischer Grad
Schüler:innen	134	---	134/--	Jgst.: 10	---
Studierende	51	33/18/--	51/--	FS.: 5,33 (2,17)	---
Lehrkräfte	34	19/15/--	28/6/--	BJ.: 15,63 (9,48)	5 PL/10 BL/10 FB/10 SL
Didaktiker:innen	31	15/15/1	---	---	8 Dr./6 Prof.
Σ	250	67/48/1			

Bem.: *N*: Anzahl der teilnehmenden Personen zu mindestens einem Messzeitpunkt; GS: Grundschule, MS: Mittelschule, RS: Realschule, GY: Gymnasium; *M*: Mittelwert, *SD*: Standardabweichung; Jgst.: Jahrgangsstufe; FS: Fachsemester, BJ: Berufsjahre; PL: Praktikumslehrkraft, BL: Betreuungslernkraft, FB: Fachbetreuer, SL: Seminarlehrkraft.

Die untersuchten Schüler:innen ($N = 134$)¹²⁰ besuchten die 10. Jahrgangsstufe des Gymnasiums. Die 10. Jahrgangsstufe ist dabei für die Erhebung geeignet, da davon ausgegangen werden kann, dass die Schüler:innen grundlegendes Wissen für das Verstehen komplexer Phänomene erworben haben, das über deklaratives Wissen hinausgeht. Darunter sind ein Grundverständnis von Rückkopplungsschleifen in biologischen Systemen auf inhaltlicher Seite und ein Lesen von Mind-Maps respektive Ablaufdiagrammen auf lernmethodischer Seite zu fassen.

Erhoben wurde an drei bayerischen Schulen in fünf Klassen. Angaben zum Geschlecht wurden aufgrund datenschutzrechtlicher Vorgaben nicht abgefragt.

Schüler:innen benötigten zur vollständigen Bearbeitung des Fragebogens im Mittel 33:76 Minuten ($SD = 5:03$ Min.) zum ersten Messzeitpunkt und 32:00 Minuten ($SD = 3:93$ Min.) zum zweiten Messzeitpunkt. Zwischen den beiden Messzeitpunkten lag stets eine Woche.

Der Kontakt für die Teilnahme an der Studie wurde über die Schulleitung des jeweiligen Gymnasiums hergestellt und die Informationen zu den Zielen der Studie, organisatorischen Belangen und Umgang mit den erhobenen Schüler:innendaten wurden an Biologielehrkräfte weitergeleitet. Interessierte Lehrkräfte meldeten sich anschließend freiwillig für die Teilnahme mit ihrer Klasse. Die Bearbeitung des Fragebogens fand im Klassenverband in einem Computerraum der Schule statt. Die Schüler:innen wurden dabei zunächst über die Ziele der Studie, die Freiwilligkeit der Teilnahme sowie den Umgang mit den entstehenden Daten mithilfe eines standardisierten Textes unmittelbar vor der Bearbeitung durch eine eingewiesene studentische Hilfskraft informiert. Jede:r Teilnehmer:in hatte einen eigenen PC zur Verfügung. Die Bearbeitung fand durchgehend unter Aufsicht der Lehrkraft und einer eingewiesenen studentischen Hilfskraft statt.¹²¹ Während der Bearbeitung ergaben sich lediglich in zwei Fällen Nachfragen zur Bedeutung des Begriffs des roten Fadens (Item: *St_rot*: „Die Erklärung hatte einen roten Faden“). **Diese wurden von der** betreuenden Hilfskraft gemäß den vorab ausgearbeiteten Instruktionen bei Verständnisnachfragen geklärt. Zur Erhöhung der Repräsentativität der Gelegenheitsstichprobe trägt bei, dass im Klassenverband erhoben wurde und somit eher die Heterogenität der Lernenden innerhalb einer Jahrgangsstufe abgebildet werden kann.

¹²⁰ Die maximal in die Berechnung eingehende Anzahl reduziert sich auf $N_{max} = 122$, da zum ersten Messzeitpunkt und zum zweiten Messzeitpunkt Schüler:innen krankheitsbedingt in der Klasse fehlten, sodass jeweils nur der halbe Datensatz vorliegt. Für einzelne Items kann die maximal in die Berechnung eingehende Anzahl etwas geringer ausfallen, da vor allem gegen Ende des zweiten Messzeitpunkts einzelne Abbrüche zu verzeichnen sind.

¹²¹ Da es sich bei der in den Videos erklärenden Person um die Forscherin selbst handelte, wurde darauf geachtet, dass vor und während der Erhebung kein Kontakt zwischen Schüler:innen und Forscherin entsteht, der die Einschätzung verzerren könnte.

Die befragten Biologielehramtsstudierenden ($N = 51$; davon $N = 33$ weiblich und $N = 18$ männlich) besuchten im Mittel etwa das fünfte Semester ($M = 5,33$; $SD = 2,17$; $range = [2,9]$) und studierten Biologie als eines von zwei Hauptfächern für das Lehramt an Gymnasien. Im Mittel betrug die Bearbeitungszeit für den vollständigen Fragebogen nach Ausschluss von Ausreißern (insgesamt $N = 8$; z. B. durch längeres Pausieren der Bearbeitung) 40:47 Minuten ($SD = 18:14$ Min.) zum ersten Erhebungszeitpunkt und 40:27 Minuten ($SD = 10:60$ Min.) zum zweiten Erhebungszeitpunkt. Die Kontaktaufnahme mit der Bitte um Teilnahme an der Fragebogenerhebung, Informationen zur Studie und zum Umgang mit den entstehenden Daten sowie das Angebot einer geringen Vergütung (Gutschein für einen Kaffee an einer universitären Cafeteria) erhielten die Studierenden zunächst per E-Mail. Nach positiver Rückmeldung wurden die Links zur Umfrage versendet, sodass die Bearbeitung an einem beliebigen Computer erfolgen konnte. Es gab keine Vorgaben oder Regulationsmechanismen für die Bearbeitung der beiden Messzeitpunkte, sodass diese auch unmittelbar hintereinander bearbeitet werden konnten. Aufgrund der geringen Rückmeldungsquote ($N = 12$) wurde die Erhebung in verschiedene Kurse der Biologiedidaktik integriert. Dazu wurden die Links zur Umfrage, Informationen zur Studie sowie zum Umgang mit den entstehenden Daten vor der jeweiligen Seminarsitzung per E-Mail an die Kursteilnehmer:innen versendet, sodass die Bearbeitung auch in diesem Fall von einem beliebigen Computer selbstständig und zu frei gewählten Zeitpunkten erfolgen konnte. In einer Seminarsitzung wurde dann das instruktionale Erklären thematisiert und beispielsweise über Qualitätskriterien, die Angemessenheit lehrerzentrierten Erklärens aus lernpsychologischer Perspektive und die spezifischen Herausforderungen, die sich aus der Komplexität der biologischen Gegenstände ergeben, gesprochen und es wurden erste Ergebnisse der Studie präsentiert. Durch die Integration der Erhebung in universitäre Veranstaltungen konnte auch die positive Auswahl besonders interessierter und engagierter Studierender abgemildert werden, sodass eine höhere Repräsentativität der Teilstichprobe gegeben ist.

Die untersuchten Biologielehrkräfte ($N = 34$; davon $N = 19$ weiblich und $N = 15$ männlich) unterrichten an bayerischen Gymnasien ($N = 28$) und Realschulen ($N = 6$). Im Mittel beträgt die Berufserfahrung inklusive des Vorbereitungsdienstes knapp 16 Jahre ($M = 15,63$; $SD = 9,48$; $range = [2,38]$). Weiterhin zeichnen sich einige Lehrkräfte durch eine oder mehrere Funktionsstellen aus, bei der sie in besonderem Maße eine Metaperspektive einnehmen und Fachunterricht von Kolleg:innen oder Studierenden beurteilen: Von den befragten Lehrkräften gaben $N = 5$ an, als Betreuungslehrkraft für angehende Lehrkräfte im Vorbereitungsdienst zu fungieren. Als ausbildende Personen für angehende Biologielehrkräfte im Vorbereitungsdienst (in Bayern sogenannte Seminarlehrkräfte) sind $N = 10$ tätig. Weitere $N = 7$ gaben an, die Fachbetreuung an der jeweiligen Schule zu führen, wobei sie beispielsweise mit den Aufgaben betraut sind, eine fachspezifische Beratung bei der Anschaffung von Lehr- und Lernmaterial zu geben, hinsichtlich neuer didaktischer- und methodischer Veröffentlichungen zu informieren, im Fachkollegium beratend in Bezug auf didaktische Fragen zu wirken und gegebenenfalls Unterrichtsbesuche bei Kolleg:innen durchzuführen. Von den befragten Lehrkräften haben weitere $N = 12$ keine der genannten Funktionsstellen inne. Durch den insgesamt hohen Anteil derjenigen Lehrkräfte, die durch ihre Funktionsstelle regelmäßig eine Metaperspektive zu eigenem oder fremdem Unterricht einnehmen und diesen Unterricht beurteilen, kann der Problematik der geringen Erfahrung mit Unterrichtsbeurteilungen von Lehrkräften, die in Studien zur Validität von Unterrichtsbeurteilungen häufig als Nachteil für die Befragung von Lehrkräften angeführt wird, entgegengewirkt werden (Clausen, 2002). Bei Lehrkräften beträgt die mittlere Bearbeitungszeit für den vollständigen Fragebogen nach Ausschluss von Ausreißern (insgesamt $N = 5$; z. B. durch längeres Pausieren der Bearbeitung) 39:27 Minuten ($SD = 11:18$ Min.) zum ersten Messzeitpunkt und 41:93 Minuten

($SD = 11,98$ Min.) zum zweiten Messzeitpunkt. Die Bitte um Teilnahme richtete sich zunächst an Kontaktlehrkräfte an einigen bayerischen Gymnasien, die eine E-Mail mit den entsprechenden Informationen zur Studie und zum Datenschutz an das Kollegium weiterleiteten. Der Zugang zum Fragebogen wurde bei positiver Rückmeldung per E-Mail verschickt. Die Bearbeitung der beiden Teile des Fragebogens konnte zeitlich frei eingeteilt werden. Eine Vergütung für die Teilnahme wurde nicht angeboten. Das Finden von freiwilligen Teilnehmer:innen auf diesem Wege gestaltete sich jedoch insgesamt schwierig ($N = 3$), sodass die Teilnahme an der Studie in ein Fortbildungsangebot integriert wurde.

Die befragten Seminarlehrkräfte ($N = 10$) nahmen im Rahmen zweier vom Gesamtprojekt FALKE organisierter Kooperationstage für phasenübergreifende Lehrer:innenbildung an der Universität Regensburg teil. Dabei wurden unter anderem die Studie in den jeweiligen Fachgruppen vorgestellt, theoretische Fundierungen vor allem in Bezug auf die fachspezifischen Schwerpunkte erläutert und diskutiert sowie erste Ergebnisse präsentiert. Die Links zum Fragebogen wurden vor der Fortbildung per E-Mail versandt. Die Bearbeitung erfolgte im Vorfeld von zu Hause aus, wobei der zweite Messzeitpunkt von einigen Seminarlehrkräften am Tag der Fortbildung an einem Computer an der Universität Regensburg bearbeitet wurde. Des Weiteren wurden drei schulinterne Fortbildungen angeboten, an denen insgesamt $N = 21$ Biologielehrkräfte teilnahmen. Die Zugangslinks für beide Messzeitpunkte wurde vor der Fortbildung per E-Mail versandt, sodass die Bearbeitung individuell erfolgte. Die Inhalte der zweistündigen Fortbildungen waren thematisch an die Inhalte der Kooperationstage an der Universität Regensburg angelehnt. Durch die Integration der Erhebung in Fortbildungsangebote kann auch für die Teilstichprobe der Lehrkräfte die Repräsentativität erhöht werden, da nicht nur in besonderem Maße an Forschung interessierte Lehrkräfte an der Studie teilnahmen. Beide Fortbildungsangebote wurden als Weiterbildung für Lehrkräfte anerkannt.

Die Biologiedidaktiker:innen ($N = 31$; davon $N = 15$ weiblich, $N = 15$ männlich und $N = 1$ divers) forschen bzw. lehren an Universitäten aus Deutschland, Österreich und der Schweiz. Insgesamt gaben $N = 6$ Teilnehmer:innen an, eine Professur innezuhaben. $N = 8$ Personen sind promovierte wissenschaftliche Mitarbeiter:innen und die verbleibenden $N = 17$ Teilnehmer:innen sind Mitarbeiter:innen der Biologiedidaktik an Universitäten. Die mittlere Zeit für die vollständige Bearbeitung beträgt nach Ausschluss von Ausreißern ($N = 5$; z. B. durch längeres Pausieren der Bearbeitung) in dieser Teilstichprobe 46:34 Minuten ($SD = 14:18$ Min.) zum ersten Erhebungszeitpunkt und 38:81 Minuten ($SD = 11:11$ Min.) zum zweiten Erhebungszeitpunkt. Die Kontaktaufnahme erfolgte auch hier per E-Mail mit Informationen zur Studie und zum Umgang mit den Forschungsdaten. Nach positiver Rückmeldung wurden ebenfalls über die E-Mail die Zugänge für die zwei Teile des Fragebogens verschickt, sodass die Bearbeitung individuell stattfinden konnte. Eine Regulation bezüglich der Bearbeitungszeitpunkte gab es nicht.

Insgesamt handelt es sich bei den Teilstichproben stets um Ad-hoc-Stichproben (Bortz & Döhring, 2006), da die Auswahl aller Teilnehmenden vor allem nach den Kriterien der Verfügbarkeit und insbesondere bei Didaktiker:innen nach der Bereitschaft zur Teilnahme erfolgte. Die damit einhergehende reduzierte Repräsentativität im Vergleich zu exakten Zufallsstichproben wird bei der Dateninterpretation berücksichtigt.

5.3 Datenaufbereitungs- und Auswertungsmethoden

Vor der Auswertung wurden die erhobenen Daten zunächst aus dem Online-Tool (EFS-Survey) von Questback exportiert, in das Statistikprogramm SPSS überführt und die beiden Messzeitpunkte mit Hilfe der generierten Teilnehmer:innencodes sowie die Datensätze der Schüler:innen und Lehrenden zu einem Gesamtdatensatz zusammengeführt. Im Gesamtdatensatz wurden anschließend ungültige Fälle ausgeschlossen. Dies betraf $N = 3$ Schüler:innen, die ohne Berücksichtigung der positiven oder negativen Itemformulierungen den Ausprägungsgrad ‚*stimme gar nicht zu*‘ auf der Ratingskala wählten. Unvollständige Datensätze – beispielsweise erzeugt durch den Abbruch der Bearbeitung des Fragebogens oder durch die Bearbeitung nur einer der beiden Messzeitpunkte – wurden bei den Analysen dennoch berücksichtigt. Die fehlenden Werte wurden nicht durch computergestützte Verfahren ergänzt. Daher ergeben sich für die einzelnen Berechnungen mitunter Schwankungen in den Teilnehmer:innenzahlen, sodass jeweils ein N_{max} angegeben wird. Die Daten des offenen Items, also die Begründungen der Notenvergabe, wurden anschließend exportiert und zur weiteren Bearbeitung in das Programm MAXQDA importiert.

Im Folgenden werden zunächst die für die geschlossenen Items vorgenommenen Recodierungen, Transformationen und Skalenbildungen berichtet sowie die Umstrukturierung des Datensatzes für die hierarchischen Regressionsmodelle geschildert. Hiernach erfolgt eine Beschreibung der Vorgehensweise zur Auswertung des offenen Items.

5.3.1 Datenaufbereitung und Vorbereitungen für Auswertungen der geschlossenen Items

Vor den statistischen Analysen wurden im Fragebogen negativ formulierte Items recodiert, sodass alle Items in ihrer positiven Ausprägung in die Berechnungen eingehen. Dies betrifft die Items *„Bei manchen Wörtern wissen die Schüler_innen eventuell nicht, was sie bedeuten“* (Item: *Sw_bed*), *„Manche Sätze hat die Lehrerin zu lange gemacht“* (Item: *Sw_zul*), *„Die Erklärung war für Schüler_innen zu schwierig“* (Item: *Ad_sch*) sowie *„Die Lehrerin hat zu viele Einzelheiten erklärt“* (Item: *St_ein*).¹²²

Des Weiteren wurden die Daten des Globalurteils, die aus der Schulnote und der optionalen Tendenzangabe bestehen, transformiert respektive verrechnet. Dabei wurde ein vergebenes ‚Plus‘ als Verbesserung der Schulnote um 0,3 und ein ‚Minus‘ als Verschlechterung der Schulnote ebenfalls um 0,3 angerechnet. So ergeben sich Globalurteile mit den Werten von 0,7 bis 6,3.

Anschließend wurden skalenspezifische Mittelwerte gebildet, wobei Reliabilitätsindikatoren (Cronbachs Alpha; siehe Kap. 6.1.2) berücksichtigt wurden. Dabei gehen in die fachübergreifenden Skalen Adressatenorientierung, Strukturierung sowie Sprech- und Körperausdruck jeweils alle Items ausgenommen der skalenspezifischen Globalurteile (z. B. *„Die Erklärung war gut strukturiert“* (Item: *St_glo*)) ein. Eine Ausnahme bildet die Skala zur sprachlichen Verständlichkeit – hier musste aufgrund negativer Itemtrennschärfen das Item *„Manchmal hat die Lehrerin extra gesagt, dass etwas sehr wichtig ist“* (Item: *Sw_wic*) ausgeschlossen werden (siehe Kap. 6.1.2). Zugunsten der breiteren

¹²² Die Items *„Die Erklärung zeigt zu viele Ursachen“* (Item: *Fa_zuv*) und *„Die Tatsache, dass nicht alle Dinge bekannt sind, verunsichert Schüler_innen“* (Item: *Fa_ver*) sind ebenfalls negativ formuliert, gehen jedoch nicht in die fachspezifische Skala ein, sodass eine Recodierung an dieser Stelle nicht notwendig ist.

Erfassung des Aspekts wurde in die Skala das skalenspezifische Globalurteil daraufhin integriert (Item: *Sw_glo*; „Die Lehrerin hat sich verständlich ausgedrückt“).

Zur Bildung eines skalenspezifischen Mittelwertes für die fachspezifische Skala zur Wahrnehmung der dargestellten Komplexität wurden diejenigen Items herangezogen, die über alle Videovignetten hinweg abgefragt wurden. Das heißt, die Items „Die Erklärung zeigt zu viele Ursachen“ (Item: *Fa_zuv*) und „Die Tatsache, dass nicht alle Dinge bekannt sind, verunsichert Schüler_innen“ (Item: *Fa_ver*) wurden von der Bildung des Mittelwertes ausgenommen. Das Item „Die Erklärung zeigt nur einen Ausschnitt auf die Frage „...““ (Item: *Fa_aus*) wurde aufgrund negativer Itemtrennschärfen aus der Skala ausgeschlossen (siehe Kap. 6.1.2). Trotz der Ausschlüsse sind alle Facetten der Konzeptualisierung in der Skala repräsentiert. Eine genauere Darstellung mit entsprechenden Kennwerten zur Skalenbildung findet sich in Kapitel 6.1.2.

Zuletzt wurde eine Kopie des Datensatzes so umstrukturiert, dass die bisher nach Videovignetten strukturierten Daten nach Personen geordnet sind. Das heißt, dass pro Person abgegebene Einschätzungen nun nicht mehr in Spalten nebeneinander, sondern in Zeilen untereinander stehen. Damit einhergehend finden sich in entsprechenden Analysen hohe Fallzahlen, die dem Sechsfachen der Teilnehmer:innenzahlen entsprechen. Die Umstrukturierung ist für die Regressionsmodelle notwendig (siehe Kap. 6.2.3 und 6.2.4)

In diesen neuen, umstrukturierten Datensatz wurden weiterhin eine sechsstufige Variable für die Unterscheidung der Videovignetten sowie eine dichotome Variable für die Unterscheidung des dargestellten Komplexitätsgrades in den Videovignetten (linear/komplex) geschrieben, sodass die Unterscheidungen als Faktorstufen in Analysen eingehen und damit weitere Strukturierungsebenen in den Daten berücksichtigt werden können (siehe Kap. 6.2.4)

Die Voraussetzungsprüfungen für die Analysen finden sich im Abschnitt des jeweiligen Ergebnisberichts (siehe Kap. 6.2.2 bis 6.2.4).

5.3.2 Auswertungen der Begründungen des Globalurteils

Leitend für die Abfrage einer Begründung zur Notenvergabe war, die bereits bekannten respektive diskutierten Kriterien guten Erklärens unter Berücksichtigung der Perspektiven verschiedener Statusgruppen zu validieren und gegebenenfalls zu erweitern. Zudem wird damit die Möglichkeit genutzt, die Ergebnisse der standardisierten Erhebung mit den selbstgewählten Kriterien aus dem offenen Textfeld zu verschränken und Gründe für die Bewertungen zu eruieren. Insbesondere in Hinblick auf die Wahrnehmung von komplexen im Gegensatz zu linearen Erklärungen ist diese Möglichkeit von besonderer Relevanz. Für die Auswertung der offenen Antworten steht daher im Fokus, welche Kriterien genannt werden und welche Statusgruppen sich auf welche Kriterien beziehen.

Das zu analysierende Material umfasst 238 bearbeitete Fragebögen mit insgesamt 940 ausgefüllten Begründungsfeldern ($N = 533$ von Schüler:innen; $N = 158$ von Studierenden; $N = 108$ von Lehrkräften; $N = 141$ von Didaktiker:innen), die Äußerungen enthalten und für die Codierung herangezogen werden können. Daraus konnten 2186 Segmente gebildet werden. Die Äußerungen der Teilnehmer:innen unterscheiden sich dabei stark in Länge respektive Ausführlichkeit, sodass ein heterogener Textkorpus vorliegt. Das Spektrum reicht von einzelnen Stichworten ohne explizierte Wertung des genannten Aspekts (z. B. „Anschaulichkeit“ – Student:in/Ök, „Sprachduktus“ – Lehrer:in/Ök) bis hin zu ausführlichen Begründungen, die auch explizite

Wertungen der angesprochenen Aspekte beinhalten (z. B. *„Erklärung wird mit Abbildungen unterstützt, die nach und nach an Tafel angebracht werden. Dadurch kann der Erklärung sehr gut gefolgt werden. Lehrkraft spricht langsam, deutlich und dem Schüler zugewandt“* – Lehrer:in/ G). Insgesamt sind die Texte bzw. Textteile nicht als narrativ zu charakterisieren.

Da alle Daten bereits in digitalisierter Form vorlagen und den Teilnehmer:innen sowie Videovignetten eindeutig zugeordnet waren, mussten diese für die Datenanalysen nicht weiter aufbereitet werden.

5.3.2.1 Die inhaltlich-strukturierende qualitative Inhaltsanalyse

Für die Auswertung des Materials wurde die qualitative Inhaltsanalyse als übergeordneter Auswertungsansatz gewählt.¹²³ Das Spektrum der inhaltsanalytischen Verfahren reicht von einer Möglichkeit zur theoriebasierten, regelgeleiteten und quantifizierenden Beschreibung von Textinhalten (z. B. Mayring, 2010) bis zur Betonung des interpretativen Aspektes, der auf individuelles Verstehen von Texten bzw. durch Texte ausgedrückte nicht sprachliche Eigenschaften einer Person oder gesellschaftlichen Aggregaten fokussiert und die Ergebnisse überwiegend qualitativ beschreibt (Stamann, Jansse & Schreier, 2016). Der Kern der darunter subsumierten Verfahren liegt laut Schreier (2014), die auf die in der Forschungspraxis angewandten inhaltsanalytisch orientierten Auswertungsansätze rekurriert, in einer **„Beschreibung ausgewählter Textbedeutungen“** (Schreier, 2014, S. 3) anhand eines Kategoriensystems. Die Genese der Kategorien reicht von fast vollständiger Ableitung aus der Theorie bis hin zu weitestgehenden Abstraktion aus dem Material (Kuckartz, 2018). Entwicklung sowie Anwendung des Kategoriensystems sind auf Interpretationen des latenten Äußerungsgehalts durch die Forscher:innen angewiesen (Schreier, 2014).¹²⁴

In Bezug auf Merkmale des vorliegenden Datenmaterials (siehe Kap. 5.3.2) sowie die Ziele der Untersuchung, wurde die inhaltlich-strukturierende qualitative Inhaltsanalyse als Variante gewählt. Diese **„hat zum Ziel, eine bestimmte Struktur aus dem Material herauszufiltern“** (Mayring, 2010, S. 92), wobei sich die Struktur in diesem Fall auf extrahierte und zusammengefasste Inhaltsbereiche bezieht, die sich im Kategoriensystem widerspiegeln (Mayring, 2010).

Die Kategorien können anhand des Materials entwickelt oder bereits vor der Beschäftigung mit den Texten literaturbasiert erstellt werden. Während in älteren Arbeiten zu qualitativer Forschung die vorherige Beschäftigung mit thematischer Literatur als Verletzung der Unvoreingenommenheit der Forschenden und deren offenem Blick bei der Auswertung galt, betonen einige Autor:innen – vor allem für die (inhaltlich-)strukturierende qualitative Inhaltsanalyse – diese als Strukturierungshilfe und Sensibilisierung der Wahrnehmung (Hopf, 2016). Die Formulierung von Annahmen über Inhalte und Themen kann auch die Aufmerksamkeit auf fehlende oder dem theoretischen Vorverständnis widersprechende Aussagen lenken (Hopf, 2016).

¹²³ Der Begriff der qualitativen Inhaltsanalyse wird in vielen Disziplinen – auch der Biologiedidaktik – stark mit Phillip Mayring und seinen Ausführungen zur Methodik verknüpft. Schreier (z. B. 2014) nutzt ihn hingegen als allgemeinen Oberbegriff für qualitative Forschungszugänge, bei denen inhaltsanalytisch vorgegangen wird.

¹²⁴ Letztes Kriterium unterscheidet in Bezug auf die Untersuchungsziele und -inhalte die qualitative maßgeblich von der quantitativen Inhaltsanalyse (Lamnek, 2010; Schreier, 2014). Während in qualitativen Inhaltsanalysen der latente Äußerungsgehalt Berücksichtigung findet und die Ergebnisse auf Interpretationen des Forschers bzw. der Forscherin beruhen, werden bei quantitativen Inhaltsanalysen lediglich manifeste Äußerungsinhalte – also die direkt jedem Leser bzw. jeder Leserin zugänglichen Textinhalte – erfasst und quantifiziert (Lamnek, 2010).

Am Ende des regelgeleiteten Analyseprozesses steht ein Kategoriensystem, das die wesentlichen Inhalte des Textmaterials umfasst und zur systematischen Beschreibung der inhaltlichen Themen herangezogen werden kann. So können die Äußerungen der teilnehmenden Personen hinsichtlich der Kriterien untersucht werden, die für die Notenvergabe als Gründe angeführt werden und Rückschlüsse auf die Kriterien guten Erklärens derjenigen Person bzw. Statusgruppe ziehen lassen.

Varianten der inhaltlich-strukturierenden qualitativen Inhaltsanalyse finden sich mehrfach beschrieben bei verschiedenen Autor:innen (z. B. Mayring, 2010; Kuckartz, 2018), wobei sich diese in ihren Grundsätzen nicht unterscheiden (Schreier, 2014). Für die vorliegende Untersuchung lehnt sich die Auswertung vor allem an Verfahrensvorschläge von Kuckartz (2018) an. Das von ihm beschriebene Auswertungsvorgehen lässt dabei in unterschiedlichen Phasen des Prozesses explizit Raum für die Weiterentwicklung des Kategoriensystems anhand des vorliegenden Textmaterials. In Anlehnung an Schreier (2014) werden die Verfahrensvorschläge zudem erweitert und im Sinne eines Werkzeugkastenprinzips an das Untersuchungsanliegen angepasst. Dies betrifft vor allem die Ergänzung der inhaltlichen Strukturierung des Materials durch evaluative Aspekte (evaluative Inhaltsanalyse; ebenfalls in Anlehnung an Kuckartz (2018); Umsetzung in Kap. 5.3.2.3). So können die genannten Kriterien der teilnehmenden Personen auch dahingehend codiert und entsprechend ausgewertet werden, ob diese in Bezug auf die Videovignette positiv oder negativ erwähnt werden.

In methodologischer Hinsicht weist die (inhaltlich-)strukturierende qualitative Inhaltsanalyse insgesamt standardisierten Charakter¹²⁵ auf und ist theorieorientiert (Hopf, 2016; Lamnek, 2010). Sie ermöglicht dabei kein methodisch kontrolliertes Fremdverstehen, sodass sie nicht als rekonstruktiv sondern als offen charakterisiert werden kann (Bohnsack, 2005). Die Realisierung im Forschungsprojekt enthält in Übereinstimmung mit der geringen narrativen Charakter der gewonnenen Texte und den formulierten Untersuchungszielen entsprechend wenig rekonstruktive Anteile auf der Ebene der individuellen sozialen Wirklichkeiten einzelner Personen, sondern versucht, das Datenmaterial in Hinblick auf manifeste und latente Inhalte zu strukturieren und zu reduzieren (vgl. Lamnek, 2010).

5.3.2.2 Umsetzung der inhaltlich-strukturierenden qualitativen Inhaltsanalyse

Nach Kuckartz (2018) umfasst die inhaltlich-strukturierende Inhaltsanalyse die in der Abbildung 13 dargestellten Arbeitsphasen, die auch in mehrfachen Rückkopplungsschleifen durchlaufen werden können.

In der vorliegenden Studie wurde die vorgeschlagene Vorgehensweise für die inhaltliche Strukturierung des Textmaterials leicht variiert. Der konkrete Ablauf der Analyse ist in den Punkten eins bis sechs beschrieben und das entstandene Kategoriensystem in Kapitel 5.3.2.4 einsehbar.¹²⁶

¹²⁵ In Begrifflicher Anlehnung an Bohnsack (2005), der zwischen standardisierter (quantitativer) und nicht-standardisierter (qualitativer) Forschung unterscheidet.

¹²⁶ Das Kategorienhandbuch ist in Anhang B vollständig abgebildet.

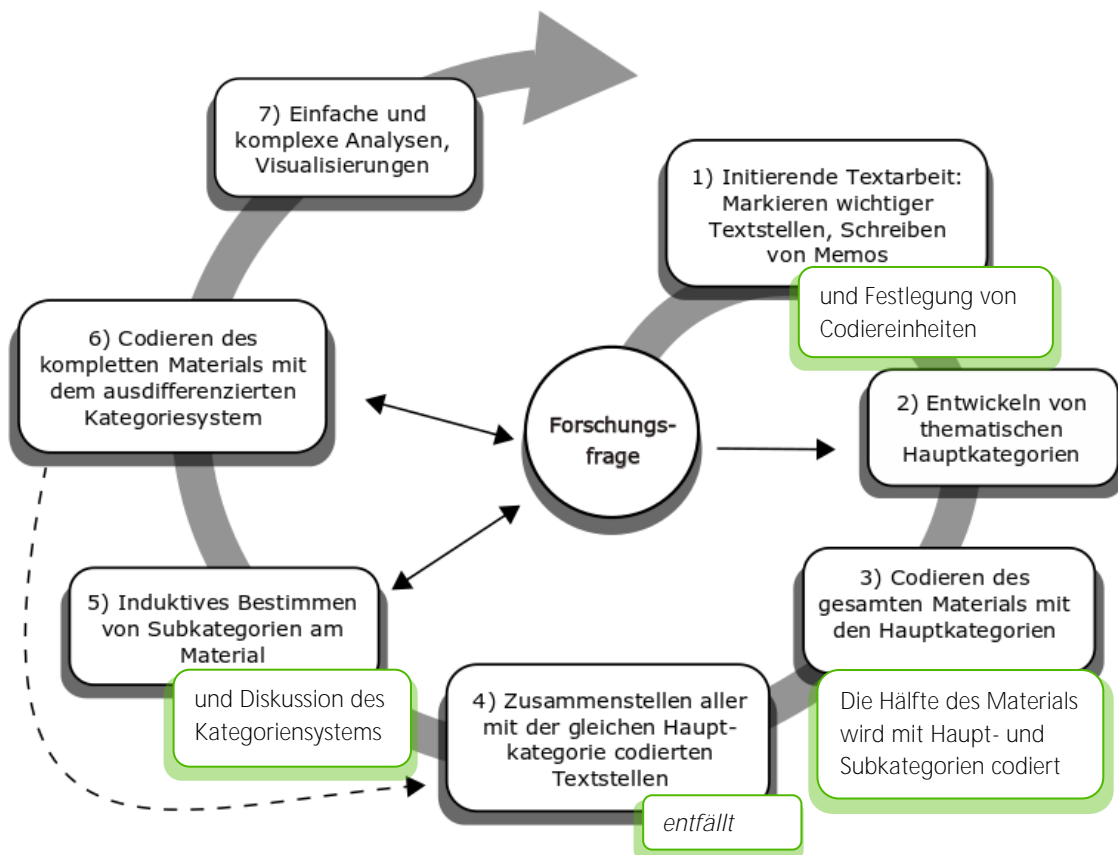


Abbildung 13: Ablaufschema einer inhaltlich-strukturierenden qualitativen Inhaltsanalyse nach Kuckartz (2018, S. 100); in grüner Umrandung sind die Modifikationen in Hinblick auf das vorliegende Dantematerial dargestellt

1. Initiierende Textarbeit und Festlegung der Codiereinheiten

Die initiierende Textarbeit dient dazu, einen ersten Überblick über die Art des Textmaterials auf sprachlicher sowie auf inhaltlicher Ebene zu gewinnen (Kuckartz, 2018). Über beide Ebenen hinweg war augenscheinlich, dass sowohl zwischen den Statusgruppen als auch innerhalb derer eine große Heterogenität vorzufinden ist.

Auf Unterschiede in der Ausführlichkeit der Äußerungen und damit auf Differenzen in der sprachlichen Darstellung wurde bereits in Kapitel 5.3.2 hingewiesen. Die gezeigten Kontraste lassen sich auch innerhalb einer Statusgruppe erkennen (z. B. „Oberflächlich, Exaktheit, zeitlicher Phasenversatz“ – Lehrer:in/Ö; „logisch aufgebaut, aber sie hat ein lineares Diagramm mit einem Zyklus weiter veranschaulicht, das verlangt ein Umdenken. Besser wäre eine Erläuterung direkt an dem bereits aufgeworfenen Diagramm bspw. Beide Kurven Hasen und Luchs übereinanderlegen [...]“ – Lehrer:in/Ö). Besonders erwähnenswert erscheinen hier auch einige engagierte und präzise Rückmeldungen von Schüler:innen (z. B. „Es fehlen Beschriftungen auf den Pfeilen. Z.B fettlöslich. Denn der Sachverhalt ist doch sehr komplex und wenn man sich zu hause den Eintrag noch ein mal durchliest kann man sich vielleicht an Machens nicht mehr so genau erinnert und versteht letztendlich das Ganze nicht mehr“ Schüler:in/P; „Sie hat alles genau begründet mit weil. Und die erkrankten und gesunden Vorgänge gegenübergestellt sodass man selber vergleich konnte, was wiederum zum nachdenken angeregt hat. Und die Reihenfolge hat gestimmt“ – Schüler:in/G). Auf inhaltlicher Ebene wurde weiterhin evident, dass eine große Zahl globaler Äußerungen enthalten ist. Beispielsweise beschreiben Teilnehmer:innen die Erklärungen häufig mit den Adjektiven

‚übersichtlich‘ oder ‚anschaulich‘. Das bedeutet auch, dass weniger Äußerungen explizit auf die Ursache der Einschätzung in der Erklärung verweisen.

Als Codiereinheiten, also diejenigen Elemente, die die Zuordnung zu einer Kategorie auslösen, wurden aufgrund des unterschiedlichen Umfangs der Antworten – von kurzen Prosatexten bis zu kurzen, sehr allgemein gehaltenen Stichworten – keine einheitlichen Definitionen formaler Art festgelegt. Bei ausformulierten Antworten wurden Sinneinheiten, mindestens jedoch ein vollständiger Satz codiert. Bei stichpunktartigen Antworten wurden nur die zusammenhängenden Wörter der einzelnen Stichpunkte als eine Codiereinheit aufgefasst.

2. Entwickeln von thematischen (Haupt-)Kategorien

Auf Grundlage der Konzeptualisierung und Operationalisierung der Aspekte Adressatenorientierung, Strukturierung, verständliche Sprache, Sprech- und Körperausdruck sowie der biologiedidaktischen Facette für den Fragebogen wurde ein a-priori Kategoriensystem erstellt. Der Begriff „a-priori“ wird von (Kuckartz, 2018, S. 64) verwendet, um diejenigen Kategorien zu beschreiben, die bereits vor der Beschäftigung mit dem Material auf Basis einer Theorie, einer Vorannahme oder eines eingesetzten Interviewleitfadens gebildet werden können. Er grenzt sich dadurch von dem üblichen Begriff der deduktiven Kategorienbildung ab, um die Assoziation eines ‚richtigen‘ Kategoriensystems zu umgehen, die durch das Schließen vom Allgemeinen auf das Besondere (im Sinne der Deduktion) gegeben sein müsste (Kuckartz, 2018). Da für die Untersuchung der strenge Theoriebegriff, aus dem Kriterien und damit Analysekategorien abgeleitet werden können, ohnehin nicht gegeben ist, wird der Begriff ‚a-priori‘ beibehalten. Für das a-priori Kategoriensystem bildeten die genannten Aspekte die Hauptkategorien sowie die dazugehörigen Kriterien die Subkategorien.

Nach Durchsicht der Memos aus der initiierten Textarbeit wurden außerdem Kategorien angelegt, unter die die globalen Äußerungen subsummiert werden können (z. B. sachlich, anschaulich, verständlich), da globale Aspekte nicht eindeutig einer einzelnen Kategorie zugeordnet werden. Sie können durch mehrere Merkmale der Erklärung respektive didaktische Entscheidungen und Handlungen der Lehrkraft realisiert werden, wodurch eine Zuordnung ohne hochinterpretative Abstraktion nicht möglich ist. Zum Beispiel könnte sich ‚gut verständlich‘ auf eine klare Aussprache, die Verwendung eines bekannten Wortschatzes oder die logische Abfolge von Schritten innerhalb der Erklärung beziehen und wäre damit dem Sprecherischen, der verständlichen Sprache oder der Strukturiertheit zuzuordnen. Zusätzlich wurde eine Kategorie ‚Sonstiges‘ gebildet, um diejenigen Segmente, die keiner bestehenden Kategorie zugeordnet werden können, gebündelt aufrufen zu können und so die induktive Kategorienbildung zu erleichtern.

Insgesamt entstand ein bereits ausdifferenziertes, hierarchisches Kategoriensystem mit thematischen Haupt- und Subkategorien.

In dieser Phase wurde auch die Arbeit an den Kategoriendefinitionen aufgenommen. Diese haben zum Ziel, die Kategorien inhaltlich zu beschreiben und die Abgrenzung zu anderen Kategorien zu ermöglichen. Dadurch wird die Zuverlässigkeit der Zuordnungen – auch zwischen verschiedenen Codierer:innen – erhöht (Kuckartz, 2018). Die Definitionen enthielten neben der inhaltlichen Beschreibung auch aussagekräftige Beispiele aus dem Datenmaterial (sogenannte Ankerbeispiele) und Abgrenzungen zu anderen inhaltlich oder alltagssprachlich nahen Kategorien (Kuckartz, 2018). Letztere wurden auf Grundlage von schwer zuzuordnenden Äußerungen erstellt. Innerhalb von fachinternen Gruppen wurden einige dieser schwer zuzuordnenden Äußerungen sowie die erstellten Kategoriendefinitionen anschließend diskutiert und damit argumentativ überprüft. Beispielsweise

bedurfte die Unterscheidung zwischen ‚Fokussierung auf Wesentliches‘ und ‚Didaktische Reduktion‘ einer Ausschärfung. Beide Kategorien beschäftigen sich mit dem (Nicht-)Vorhandensein bzw. (Nicht-)Geben von Informationen, wodurch die Kategorisierung von Äußerungen mitunter nicht eindeutig ist (z. B. *„sehr gelungene Reduktion auf das Wesentliche“*). Der folgende Auszug aus dem Codebuch illustriert den Aufbau am genannten Beispiel zur Kategorie ‚Didaktische Reduktion‘.

Tabelle 12: Auszug aus dem Codebuch zur exemplarischen Kategoriendefinition ‚Didaktische Reduktion‘

Inhaltliche Beschreibung:	<p>Der Begriff der didaktischen Reduktion wird in der Biologiedidaktik auf fünf Ebenen beschrieben:</p> <ol style="list-style-type: none"> auf sprachlicher Ebene, auf Ebene der Darstellungen, auf Ebene der naturwissenschaftlichen Arbeitsweisen auf sektoraler Ebene auf struktureller Ebene <p>Unter dieser Kategorie sollen globale Äußerungen mit direktem Bezug zur didaktischen Reduktion („Die Erklärung ist angemessen didaktisch reduziert“) sowie Aussagen in Bezug auf die inhaltliche Eingrenzung (also sektorale und strukturelle didaktische Reduktion) gefasst werden. Äußerungen, die unter dieser Kategorie subsumiert werden, können aufgrund von dafür notwendigem Professionswissen nur von Personen der Metaperspektive (Studierende, Lehrkräfte und Didaktiker:innen) kommen.</p>
Beispiele für die Anwendung des Codes:	<ul style="list-style-type: none"> - „sehr starke didaktische Reduktion“ (Did – MAPA898 – a) - „Informationsgehalt entsprechend eingegrenzt“ (Did – BEHE12 – c) - „m. M. nach unzulässige Vereinfachung Transkription / Translation“ (Leh – SOFR10 – a) - „sehr gelungene Reduktion auf das Wesentliche.“ (Leh – MAED33 – a)
Abgrenzung zu anderen Kategorien:	<p>Äußerungen, die sich explizit darauf beziehen, dass das natürliche Phänomen durch die sektorale Reduktion nicht mehr angemessen in der Erklärung repräsentiert wird, werden unter „Fach – Phänomenorientierung“ codiert.</p> <p>Äußerungen, bei denen sich Teilnehmer:innen darauf beziehen, dass nichts Unnötiges, sondern nur Wesentliches für die Sache in der Erklärung aufgegriffen wurde und damit nicht der Prozess des sektoralen und strukturellen Reduzierens im Vordergrund steht, werden unter „Struktur – Fokussierung“ codiert.</p>

3. Codieren eines Teils des Materials mit (Haupt-)Kategorien

Bei der ersten Erprobung des Kategoriensystems wurden entgegen der Empfehlungen von Kuckartz (2018) neben den Hauptkategorien auch die Subkategorien mit einbezogen, da bereits Beschreibungen der Kategorien durch die theoretische Vorarbeit der Facetten vorlagen.

Das entstandene Kategoriensystem wurde an einem Teil des bis zu diesem Zeitpunkt vorliegenden Datenmaterials erprobt und damit die Phase der induktiven Kategorienbildung eingeleitet. Insgesamt lagen zu diesem Zeitpunkt 217 bearbeitete Fragebögen vor. Für die Erprobung des Kategoriensystems wurde aufgrund der Fülle des gesamten Materials die Hälfte der Begründungen pro Statusgruppe über alle Videovignetten hinweg herangezogen. Dadurch konnten auch themenspezifisch auftretende Begründungen in einzelnen Statusgruppen bei der induktiven Kategorienbildung während dieser ersten Phase bereits berücksichtigt werden (in Anlehnung an die Empfehlung der maximalen Kontrastierung von Fällen von Rädiker und Kuckartz (2019)).

Sofern die Textstellen nicht einer bestehenden Subkategorie zugeordnet werden konnten, wurden die Segmente unter einer entsprechenden Hauptkategorie subsumiert. War dies ebenfalls nicht

möglich, da das Segment einen neuen inhaltlichen Aspekt aufgreift, wurde dieses mit ‚Sonstiges‘ codiert und gegebenenfalls mit einem Memo bzw. einer Notiz zu Ideen der induktiven Kategorienbildung versehen.

4. Zusammenstellen der (Haupt-)Kategorien

Durch das bisher beschriebene Vorgehen sind diejenigen Segmente, die noch nicht in Subkategorien eingeordnet werden konnten, in den entsprechenden Hauptkategorien oder der Kategorie ‚Sonstiges‘ bereits zusammengestellt.

5. Induktive Bestimmung von Subkategorien am Material

Der Fokus lag insgesamt auf der Bildung thematischer Kategorien (Kuckartz, 2018), die die quantitativ erfassten Einschätzungen der Güte einer Erklärung weiter beschreiben können und Hinweise darauf geben, wie komplexer gestaltete Erklärungen wahrgenommen werden. Der gesamte Prozess der induktiven Kategorienbildung war zyklisch geprägt. Insgesamt wurde das bestehende a-priori-Kategoriensystem dabei nicht umstrukturiert, sondern die induktiv gebildeten Kategorien in Hauptkategorien eingeordnet. Für ein erstes Zusammenführen von einzelnen Äußerungen zu neuen Kategorien wurden die Memos genutzt, die während des ersten Codierungsprozesses bei einigen Äußerungen angelegt wurden und als stark komprimierte Zusammenfassungen der Äußerungen mit Abstraktionen zum enthaltenen Thema betrachtet werden können (Mayring, 2010).¹²⁷ Bei zwei weiteren Durchgängen durch das Material wurden weitere derartige Memos erstellt, auf Basis derer neue Kategorien erstellt und die dazu passenden Äußerungen darunter subsummiert. Nachdem keine Gemeinsamkeiten zwischen den Äußerungen, die Grundlage für eine Kategorienbildung hätten sein können, mehr gefunden wurden, wurde der Prozess der Bestimmung von neuen Subkategorien abgeschlossen. Im Folgenden wurden die induktiven Kategorien auf ihre Notwendigkeit in Hinblick auf ihre Aussagekraft und das Differenzierungspotenzial sowie auf ihre Trennschärfe hin überprüft. Einige Kategorien wurden daraufhin (z. B. aufgrund zu geringer Nennungen, hoher Ähnlichkeit bzw. Überschneidungen mit einer weiteren Kategorie) aufgelöst bzw. zusammengelegt.

Die Benennung der neu gebildeten Kategorien orientierte sich zunächst an den Begrifflichkeiten der Teilnehmer:innen (in Anlehnung an In-Vivo-Codes), die einhergehend mit den Überarbeitungszyklen zunehmend vom konkreten Text abstrahiert wurden. Bei einigen Kategorien ist diese Nähe zum Text jedoch erhalten geblieben – insbesondere bei den globalen Äußerungen.

Das so entstandene Kategoriensystem wurde anschließend einer Fachgruppe präsentiert und das Resultat diskutiert. Dabei wurden sowohl die Benennungen von Kategorien in Abstimmung mit den darunter subsummierten Äußerungen als auch die Trennschärfe der Kategorien nochmals genauer betrachtet und gegebenenfalls überarbeitet bzw. präzisiert.¹²⁸

6. Codieren des kompletten Materials mit ausdifferenzierten Kategoriensystem

Mit dem so entstandenen Kategoriensystem wurde das bisher codierte Material von einer zweiten Codiererin unabhängig codiert, um die intersubjektive Übereinstimmung zu kontrollieren bzw. zu

¹²⁷ Aufgrund der Kürze der Segmente war ein Vorgehen anhand systematischer Zusammenfassungen nicht möglich bzw. erschien nicht sinnvoll, da die angesprochenen Themen (hier Kriterien) bereits ohne Abstraktion gut greifbar waren.

¹²⁸ Die induktive Kategorienbildung wird von Kuckartz (2018) explizit als kreativer Konstruktionsprozess beschrieben, bei dem verschiedene Personen nicht zu einem einheitlichen Ergebnis kommen können und müssen. Die Diskussion erfolgte daher nicht mit dem Ziel der Erreichung einer intersubjektiven Übereinstimmung, sondern mit dem Ziel der Erweiterung der Wissensbasis, auf deren Grundlage die Kategorienbildung stattfindet, was zur Qualität des Kategoriensystems beitragen kann.

erhöhen. Dies erfolgte auf Basis des erstellten Codierleitfadens – bestehend aus dem Codebuch sowie Anweisungen und Hilfen für das Codieren beispielsweise zur Segmentierung oder zur regelmäßigen Überprüfung der Stringenz der Codierungen innerhalb einer Kategorie – sowie einer literaturbasierten Einführung in die Methode und Instruktionen zur Handhabung von MAXQDA anhand eines Beispielprojektes.

Die Codierungen wurden anschließend der Vorgehensweise des konsensuellen Codierens folgend computergestützt innerhalb des MAXQDA Programms verglichen und die Abweichungen diskutiert.¹²⁹ Bezüglich aller Abweichenden Codierungen konnte argumentativ eine Einigung zur Zuordnung gefunden werden, sodass ein Hinzuziehen einer dritten Person aufgrund ausbleibender Konsensfindung in keinem Fall nötig war. Hohe Übereinstimmung konnte bei der Zuordnung zu globalen Äußerungen und der Subsummierung unter einer Hauptkategorie erreicht werden. Bei den Merkmalen zum Sprech- und Körperausdruck konnte auch bei den Subkategorien eine hohe Kongruenz festgestellt werden.

Häufige Differenzen gab es jedoch beispielsweise bezüglich der Zuordnung zu den Subkategorien ‚Fachlich vollständig‘ und ‚Fehlende Inhalte, die über das Erklärziel hinausgehen‘, was auf fehlende Anhaltspunkte in der Kategoriendefinition zurückzuführen war. Eine Ausschärfung der Kategoriendefinition durch Hervorhebung, dass die fachliche Vollständigkeit erst durch vertieftes, fachliches Wissen, also hauptsächlich von den Statusgruppen aus der Metaperspektive, eingeschätzt werden kann, sowie der Vergleichspunkt das in der Erklärung definierte Erklärziel (z. B. durch die Problemstellung oder Eingangsfrage) ist, wurde deshalb vorgenommen. Eine technische Quelle für eine Ausbleibende Übereinstimmung bei der Codierung war die freie Segmentierung der Texte durch die Codiererinnen. Vermutlich auch bedingt durch die Kürze der Äußerungen waren verschiedene Lesarten bzw. Schwerpunktsetzungen Grund für eine differierende Zuordnung. In offenen Diskussionen und der Darlegung von Anhaltspunkten für die Zuordnung im Segment selbst und vor- bzw. nachfolgenden Segmenten konnten diese Fälle geklärt werden. Aufgrund dessen wurde auch das restliche Material konsensuell codiert.

5.3.2.3 Umsetzung der Erweiterung zur evaluativen qualitativen Inhaltsanalyse

Da durch die inhaltlich strukturierende qualitative Inhaltsanalyse bereits ein Kategoriensystem vorliegt, mit dem das Material strukturiert beschrieben werden kann, beschränkt sich das Vorgehen zur evaluativen qualitativen Inhaltsanalyse lediglich auf einen Ausschnitt des von Kuckartz (2018) beschriebenen Prozesses. Es wurden die Ausprägungen der Valenzen bzw. Bewertungskategorien festgelegt, erprobt und im Anschluss das gesamte Material codiert.

Die Aussagen der Teilnehmer:innen in Bezug auf ihre Bewertung der gesehenen Videovignetten werden durch eine dreistufige Valenz (positiv, negativ oder neutral bzw. nicht eindeutig klassifizierbar) näher beschrieben. Nach der vollständigen, themenspezifischen Codierung wurde dafür zusätzlich zu jedem Code¹³⁰ ein Subcodesystem angelegt, mit dem die Valenz der Äußerung

¹²⁹ Als Richtwert für die Übereinstimmung der Codiererinnen wurde vor dem Beginn der Konsensfindung die prozentuale Übereinstimmung (Überlappung der codierten Segmente von mindestens 75 %) der Codiererinnen berechnet. Dies war möglich, da im vorliegenden Forschungsprojekt das freie Segmentieren aufgrund des stichwortartigen Datenmaterials eingeschränkt notwendig ist sowie kein freies Codieren stattfindet. Zwischen den Codiererinnen ließ sich eine prozentuale Inter-Coder-Übereinstimmung von 69,62 % feststellen. Die Berücksichtigung zufälliger Übereinstimmungen in der Codierung ist bei einer solch hohen Anzahl an Kategorien nicht zielführend, da die zufällige, korrekte Zuordnung zur entsprechenden Kategorie insgesamt gering ist.

¹³⁰ Für die Kategorien ‚Sonstiges‘ wurden keine Subcodes angelegt. Auch die Kategorie ‚Übererklärung‘ erhielt keine Subcodes für Valenzen, da die Kategorie ausschließlich negative Äußerungen zulässt.

codiert werden kann. So erhält beispielsweise der Code ‚Logischer Aufbau‘ die Subcodes ‚positiv‘ (z. B. *„die Reihenfolge hat gestimmt“* – Schüler:in/G_k), neutral bzw. nicht eindeutig klassifizierbar (z. B. *„Reihenfolge der Kärtchen“* – Schüler:in/G_k) und negativ (z. B. *„jedoch etwas verwirrend aufgebaut“* – Schüler:in/G_k). Die für die Valenz relevante Codiereinheit kann auch ein in der Äußerung enthaltenes „+“ oder „-“ sein. Die Kategorien ähneln damit einer Variable mit verschiedenen Ausprägungen in Anlehnung an quantitative Analysen.

Im Zuge der Codierung mit dem erweiterten, evaluativen Kategoriensystem wurde ein Teil des Materials (25 % pro Videovignette und Statusgruppe) ebenfalls durch zwei Codierinnen bearbeitet. Die prozentuale Übereinstimmung lag bei 89,56 % und ist damit als sehr hoch einzuschätzen. Für eine weitere Verbesserung der intersubjektiven Übereinstimmung wurden die divergierenden Fälle diskutiert und gegebenenfalls Notizen zur Kategoriendefinition oder zusätzliche Ankerbeispiele hinzugefügt. Aufgrund der bereits hohen Übereinstimmung der Codierungen vor den Präzisierungen wurde das weitere Material von nur einer Codiererin bearbeitet.

5.3.2.4 Kategoriensystem

Das entwickelte hierarchische Kategoriensystem ist in der nachfolgenden Abbildung (Abb. 14) einzusehen. Abgebildet sind die sechs Hauptkategorien (jeweils die linksbündigen, fettgedruckten Punkte; z. B. ‚Adressatenorientierung‘) sowie die darunter subsummierten inhaltlichen Subkategorien (eine Ebene eingerückt; z. B. ‚Didaktische Reduktion‘). Die Subkategorien wurden in einigen Fällen weiter untergliedert (zweite Ebene eingerückt; z. B. ‚Passung des Informationsgehalts‘). Beschreibungen der Kategorien inklusive der Ankerbeispiele finden sich im Kategorienhandbuch in Anhang B. Nicht abgebildet sind für eine bessere Übersichtlichkeit die jeweils drei weiteren Subkategorien der evaluativen qualitativen Inhaltsanalyse (positiv, neutral, negativ).

■ Adressatenorientierung

- Adressatenorientierung global
- Didaktische Reduktion
 - Passung des Informationsgehalts
 - Passung des Schwierigkeitsgrads
- Berücksichtigung des Vorwissens
 - Übererklärung
 - Erklärtempo
- Berücksichtigung von (Fehl-)Vorstellungen
- Interessante und motivierende Gestaltung
- Lebensweltbezug
- (Aktiver) Einbezug der Rezipient:innen

■ Strukturiertheit

- Strukturiertheit global
- Fokussierung auf Wesentliches
- Logischer Aufbau
- Schrittweise Erklärung
- Zusammenfassung zentraler Punkte
- Zielangabe
 - Übereinstimmung von Erklärziel und Erklärinhalt

■ Repräsentationen

- Repräsentationen global
- Einsatz bildlicher Visualisierung
 - Erklärbedürftigkeit der Grafik
 - Fachliche Angemessenheit
 - Sichtbarkeit der Visualisierung
 - Verständlichkeit der Visualisierung
 - Ästhetische Gestaltung
- Einsatz von Analogien, Metaphern u. Vergleichen
 - Themenspezifisch: Vergleich bei Mukoviszidose
- Einsatz von Beispielen
 - Ausreichende Bezugnahme auf das Beispiel
- Gelungene Verbindung der Repräsentationen

■ Sprech- und Körperausdruck

- Artikulation
- Prosodische Merkmale
- Stimmklang
- Sprechgeschwindigkeit
- Sprechpausen
- Variation der Stimme
- Sprechflüssigkeit
- Zugewandter Körperausdruck
- Unterstützender Körperausdruck
- Angemessene Bewegung im Raum

■ Verständliche Sprache

- Sprache global
- Wortschatz
- Umgang mit Fachworten
- Satzbau
- (Un-)präziser Ausdruck
- Hervorheben von Wichtigem

■ Fachliche Aspekte

- Fach global
- Korrektheit
- Fachspezifische Erklärungsform
- Berücksichtigung von Nichtwissen
- Vagheit/Eindeutigkeit
- Gefühl der (Un-)Sicherheit
- Faszination und Interessiertheit
- Phänomenorientierung
- Prinzipienorientierung
- Vollständigkeit
 - „Fehlende“ Inhalte, die über das Erklärziel hinausgehen
- Zusammenhänge werden deutlich

■ Globale Äußerungen

- Gut erklärt
- Verständlich
- Klar
- Nachvollziehbar
- Anschaulich

- Übersichtlich
- Sachlich
- Informativ
- Lehrperson
- Merkbarkeit und Lernbarkeit
- Sonstiges

Abbildung 14: Ergebnis der Kategorienbildung bzw. angewandtes Kategoriensystem

6 Untersuchungsergebnisse

Die Präsentation der Ergebnisse erfolgt entlang der Forschungsfragen, wobei zunächst die Gütekriterien des Messinstrumentes bzw. der Datenauswertung berichtet werden.

Im Folgenden werden die Videovignetten mit Abkürzungen benannt, die eine Zuordnung zum Themenbereich sowie zum Grad der dargestellten Komplexität (UV) erlauben. Die Abkürzungen der Videovignetten setzen sich folgendermaßen zusammen: Die Majuskeln kennzeichnen den Themenbereich der Erklärung (G: Genetik; Ö: Ökologie; P: Physiologie), die tiefgestellten Buchstaben den realisierten Grad der Komplexität (l: linear; k: komplex).

6.1 Gütekriterien und konfundierende Einflussgrößen

Für qualitative und quantitative Forschungsmethoden wird die Diskussion um die Güte einer Messung, Auswertung und Interpretation aufgrund divergierender methodologischer Grundannahmen unterschiedlich geführt (Bortz & Döring, 2006). Da in der vorliegenden Studie quantitative und qualitative Forschungsmethoden zur Anwendung kommen, wird die Güte getrennt in den Kapiteln 6.1.2 und 6.1.3 reflektiert. Zunächst werden in Kapitel 6.1.1 die Ergebnisse der erhobenen Kontrollvariablen und mögliche Konfundierungen berichtet.

6.1.1 Kontrollvariablen und mögliche Konfundierungen

Die in Kapitel 5.2.5 beschriebenen Kontrollvariablen werden nun hinsichtlich ihrer Effekte untersucht und Implikationen für die Interpretationen der Forschungsergebnisse abgeleitet.

Vidoreihenfolge

Die Vidoreihenfolge wurde in zwei Versionen des Fragebogens systematisch variiert (siehe Abb. 15; siehe auch Kap. 5.2.5), um mögliche Effekte auszumitteln.

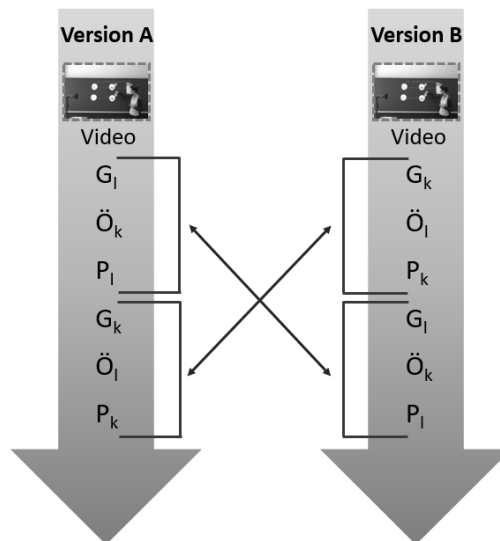


Abbildung 15: Grafik zur Reihenfolge der Videovignetten in den zwei Versionen des Fragebogens

Der Effekt der Videoreihenfolge wurde durch eine einfaktorische Varianzanalyse mit Messwiederholung geschätzt: Das Globalurteil wird als abhängige Variable, die sechs voneinander abhängigen Bewertungen der Personen werden als sechsstufiger Innersubjektfaktor mit Messwiederholung und die Version des Fragebogens wird als Zwischensubjektfaktor modelliert. Für die Abschätzung des Reihenfolgeeffekts ist der Interaktionseffekt relevant: Allgemein geben Interaktionseffekte an, ob bzw. wie stark der Effekt eines Faktors abhängig vom Effekt des anderen Faktors ist. Für die hier durchgeführte Analyse zum Einfluss der Reihenfolge auf die Bewertungen der Videovignetten bedeutet das, dass der Interaktionseffekt beschreibt, ob bzw. wie stark die Videobewertungen von der Reihenfolge (Version des Fragebogens) abhängen.¹³¹

Bei der Analyse über alle Statusgruppen ist der Test auf Sphärizität signifikant, sodass die Varianzen nicht homogen sind (Mauchly-W(14) = 0,88; $p = 0,01$). Entsprechend wird eine Korrektur der Freiheitsgrade nach Huynh-Feldt ($\epsilon = 0,98$) angewendet. Der Test auf Homogenität der Varianzen wurde mittels Levene-Tests überprüft. Lediglich für das Video G_k ist diese verletzt mit $p = 0,02$. Eine mögliche Überschätzung der Effekte (Bortz & Schuster, 2010) wird bei der Interpretation in Kauf genommen. Für alle weiteren Vignetten ist die Varianzhomogenität gegeben mit $0,08 \leq p \leq 0,71$. Die Voraussetzung der Normalverteilung wurde mittels Kolmogorov-Smirnov-Test überprüft. Sie ist bei beiden Versionen für alle Vignetten mit $p \leq 0,01$ verletzt. Aufgrund der Stichprobengröße ist das Verfahren der mixed ANOVA in diesem Fall dennoch anwendbar, da es gegenüber der Verletzung relativ robust ist (Bortz & Schuster, 2010).

Bei der statusgruppenübergreifenden Analyse zeigt sich eine statistisch signifikante Interaktion zwischen den Videobewertungen und der Version des Fragebogens ($F(4,89, 1095,61) = 3,13$; $p \leq 0,01$; $\eta^2 = 0,01$). Das bedeutet, dass die Videovignetten in Abhängigkeit von der Version des Fragebogens signifikant anders beurteilt werden. Mit $\eta^2 = 0,01$ entspricht das insgesamt jedoch einem kleinen Effekt (Cohen, 1988). Durch die Berechnung von Mittelwerten über beide Versionen schmälert sich der Effekt bzw. der Einfluss der Reihenfolge auf die globale Bewertung zudem, sodass

¹³¹ Es gibt zusätzlich zwei Haupteffekte bei der Analyse: Einen Haupteffekt ‚Version‘, der Auskunft darüber gibt, ob es einen Unterschied zwischen der gemittelten Bewertung aller Videos einer Version im Vergleich zur gemittelten Bewertung aller Videos der anderen Version gibt, und einen Haupteffekt ‚Videobewertung‘, der angibt, ob sich die Videobewertungen unabhängig von der Version voneinander unterscheiden. Dabei werden die Videobewertungen beider Versionen zusammengefasst. Beide Effekte sind in Hinblick auf die Ermittlung des Einflusses der Reihenfolge auf die Bewertung nicht dienlich, sodass die Ergebnisse hier nicht berichtet werden.

davon ausgegangen werden kann, dass der Effekt auf die Einschätzungen der Teilnehmer:innen zu vernachlässigen ist. Die systematische Randomisierung der Videovignetten war damit der Messgüte des Instruments zuträglich.

Nach Statusgruppen getrennt ergeben sich folgende Ergebnisse für die Überprüfung auf Reihenfolgeeffekte durch eine einfaktorielle Varianzanalyse mit Messwiederholung: Für die Teilstichprobe der Schüler:innen wurde zunächst ebenfalls die Voraussetzung der Varianzhomogenität mittels Levene-Tests überprüft. Sie ist für das Video G_i verletzt mit $p \leq 0,01$. Berücksichtigt werden muss daher eine mögliche Überschätzung der Effekte (Bortz & Schuster, 2010), die für die Fragestellung jedoch unproblematisch ist. Für alle weiteren Vignetten sind keine Verletzungen der Varianzhomogenität zu verzeichnen ($0,09 \leq p \leq 0,57$). Die durchgeführten Kolmogorov-Smirnov-Tests zeigen Verletzungen der Voraussetzung der Normalverteilung für alle Vignetten bis auf P_i in Version A. Auch hier ist aufgrund der Stichprobengröße von der Robustheit des Verfahrens gegenüber der Verletzung auszugehen (Bortz & Schuster, 2010). Insgesamt ergibt sich mit einer Korrektur der Freiheitsgrade nach Huynh-Feldt ($\epsilon = 0,97$; Sphärizität verletzt: Mauchly- $W(14) = 0,81$; $p \leq 0,05$) kein signifikanter Interaktionseffekt ($F(4,84, 547,03) = 0,75$; $p = 0,59$; $\eta^2 < 0,01$), sodass die Videobewertung unabhängig von der bearbeiteten Version des Fragebogens ist.

In der Teilstichprobe der Studierenden sind die Daten ebenfalls häufig nicht normalverteilt (Kolmogorov-Smirnov-Test signifikant für Video G_i in Version A, Video G_k in Version B, Video P_i in Version A und B, sowie Video \ddot{O}_i in Version A und B). Die Effekte können daher überschätzt werden (Bortz & Schuster, 2010). In Hinblick auf die Fragestellung ist eine Überschätzung jedoch akzeptabel und die Ergebnisse sind für eine Abschätzung hilfreich. Die Ergebnisse der mixed ANOVA zeigen bei angenommener Sphärizität (Mauchly- $W(14) = 0,67$; $p = 0,19$) und gegebener Varianzhomogenität (Levene-Tests mit $0,08 \leq p \leq 0,80$) einen signifikanten Interaktionseffekt ($F(5,240) = 3,03$; $p \leq 0,01$; $\eta^2 < 0,06$). Dieser Effekt ist nach Cohen (1988) als mittlerer Effekt einzuordnen. Es wird angenommen, dass sich der gefundene Effekt durch die systematische Randomisierung bzw. die Berechnungen der Mittelwerte über beide Versionen hinweg relativiert.

Innerhalb der Teilstichprobe der Lehrkräfte zeigt der Kolmogorov-Smirnov-Test in einigen Fällen eine Verletzung der Normalverteilung (Kolmogorov-Smirnov-Test signifikant für Video G_i in Version B, Video G_k in Version B, Video P_i in Version A und B sowie Video \ddot{O}_i in Version B). Ebenso wie bei Studierenden wird die Verletzung in Kauf genommen (Bortz & Schuster, 2010). Sphärizität wird angenommen (Mauchly- $W(14) = 0,53$; $p = 0,22$) und Varianzhomogenität der Residuen ist mit $0,06 \leq p \leq 0,93$ gegeben. In der Stichprobe besteht wiederum kein signifikanter Interaktionseffekt ($F(5,150) = 1,17$; $p = 0,33$; $\eta^2 = 0,04$), sodass die Reihenfolge keinen Einfluss auf die Bewertung der Videovignetten hat.

Auch bei der Teilstichprobe der Didaktiker:innen zeigen sich zunächst Verletzungen der Annahme der Normalverteilung (Kolmogorov-Smirnov-Test signifikant für Video \ddot{O}_k in Version A sowie Video G_k in Version A und B). Die Verletzung ist wie bei Studierenden und Lehrkräften in Hinblick auf die Fragestellung akzeptabel. Die Sphärizität als Voraussetzung ist angenommen (Mauchly- $W(14) = 0,64$; $p = 0,66$) sowie in allen Fällen die Varianzhomogenität mit $0,16 \leq p \leq 0,92$ gegeben. Es konnte auch bei Didaktiker:innen kein signifikanter Interaktionseffekt ($F(5,135) = 0,91$; $p = 0,48$; $\eta^2 = 0,03$) gefunden werden.

Wirkung der erklärenden Person

In den Daten konnten zwischen den drei Items zur wahrgenommenen Begeisterung, Natürlichkeit und Sympathie hohe Korrelationen gefunden werden. Die Korrelationskoeffizienten der bivariaten Korrelationen sind in der folgenden Tabelle (Tab. 13) eingetragen. Sie zeigt die jeweiligen Korrelationskoeffizienten pro Video getrennt nach Schüler:innen (oberhalb der Diagonalen) und Lehrenden (Studierende, Lehrkräfte und Didaktiker:innen; Teil unterhalb der Diagonalen). Pro Variable sind in den Zeilen die Ergebnisse der sechs Videos (Reihenfolge: G₁, G_k, Ö₁, Ö_k, P₁, P_k) angegeben.

Tabelle 13: Ergebnisse der Interkorrelationen der Items zur wahrgenommenen Begeisterung, Sympathie und Natürlichkeit der erklärenden Person bei Schüler:innen ($N_{max} = 70$) und Lehrenden (SLD; $N_{max} = 60$) sowie zugehörige Mittelwerte und Standardabweichungen

Item	begeistert (Video 1–6)	natürlich (Video 1–6)	sympathisch (Video 1–6)	M_{Sus} / M_{SLD} (Video 1–6)	SD_{Sus} / SD_{SLD} (Video 1–6)
begeistert		0,52**	0,65**	3,02 / 3,02	1,26 / 1,03
		0,64**	0,71**	2,86 / 2,80	1,27 / 1,00
		0,78**	0,67**	2,76 / 2,43	1,27 / 0,96
		0,34**	0,47**	3,44 / 3,52	1,25 / 1,23
		0,41**	0,60**	2,70 / 2,63	1,22 / 1,24
		0,49**	0,60**	3,00 / 3,02	1,12 / 1,21
natürlich	0,29*		0,73**	2,44 / 2,43	1,37 / 1,21
	0,53**		0,68**	2,58 / 2,39	1,28 / 1,13
	0,56**		0,76**	2,43 / 1,82	1,36 / 0,88
	0,31*		0,43**	2,71 / 2,52	1,26 / 1,23
	0,35**		0,43**	2,59 / 2,22	1,08 / 1,17
	0,54**		0,55**	2,60 / 2,28	1,07 / 1,18
sympathisch	0,45**	0,48**		2,23 / 1,63	1,07 / 0,78
	0,55**	0,58**		2,37 / 1,86	1,18 / 0,82
	0,57**	0,70**		2,21 / 1,57	1,16 / 0,65
	0,41**	0,46**		2,34 / 1,93	1,03 / 1,12
	0,42**	0,59**		2,26 / 1,73	0,97 / 0,92
	0,40**	0,67**		2,40 / 1,83	1,08 / 1,08

Bem.: Korrelationswerte oberhalb der Diagonale für Schüler:innen, unterhalb der Diagonale für Lehrende (SLD); Video 1–6: Videovignetten in der Reihenfolge G₁, G_k, Ö₁, Ö_k, P₁, P_k; M_{Sus} : Mittelwert Schüler:innen; M_{SLD} : Mittelwert Lehrende; SD_{Sus} : Standardabweichung Schüler:innen; SD_{SLD} : Standardabweichung Lehrende; *: $p \leq 0,05$; **: $p \leq 0,01$.

Da die Korrelationskoeffizienten stets als mittel oder hoch einzuordnen sind (Cohen, 1988), können die Items zu einem Aspekt ‚Persönlichkeitswirkung‘ zusammengefasst werden. Die Persönlichkeitswirkung dient als Kontrollvariable und trägt damit aus statistischer Sicht zur Reduktion von Fehlervarianzen bei. Sie wird nicht als Merkmal guter instruktionaler Erklärungen verstanden.¹³² Berücksichtigt wird sie bei der Frage nach Einflussfaktoren auf die Einschätzung der Güte einer instruktionalen Erklärung (siehe 6.2.3)

¹³² Auf eine inhaltliche Diskussion der Zusammenhänge des Teilkonstrukts zur Wirkung der Lehrerpersönlichkeit und der eingeschätzten Qualität der Erklärungen wird an dieser Stelle verzichtet. Es sei auf Krauss und Bruckmaier (2014) verwiesen, die resümieren, dass die im Rahmen des Persönlichkeitsparadigmas untersuchten Zusammenhänge zwischen Lehrerpersönlichkeitsmerkmalen und Unterrichtsqualität gering ausfallen.

Vorwissen der Schüler:innen zum Erklärgegenstand

Das Vorwissen der Schüler:innen wurde durch eine Selbsteinschätzung auf einer dreistufigen Ratingskala (1 = wusste ich vorher schon; 2 = wusste ich teilweise schon; 3 = wusste ich vorher noch nicht) erhoben. Tabelle 14 zeigt die deskriptiven Statistiken mit videospezifischen Mittelwerten und Standardabweichungen. Enthalten sind ebenfalls die inferenzstatistischen Ergebnisse der drei t-Tests mit verbundenen Stichproben zum Vergleich der Einschätzungen zwischen den inhaltsähnlichen Videopaaren, da sich weiterführende Berechnungen auch auf den Vergleich der Erklärvarianten auf Videoebene beziehen werden.

Tabelle 14: Ergebnisse des selbsteingeschätzten Vorwissens der Schüler:innen pro Video

Videovignette	G _i M (SD)	G _k M (SD)	Ö _i M (SD)	Ö _k M (SD)	P _i M (SD)	P _k M (SD)
M (SD)	2,54 (0,63)	2,34 (0,69)	2,06 (0,69)	2,05 (0,79)	2,12 (0,71)	2,48 (0,64)
N	121		119		115	
t	3,41**		-0,11		-4,86**	
d	-0,30		0,01		0,49	
df	120		118		114	

Bem.: N_{max} : aufgrund fehlender Werte bei einzelnen Videos den Berechnungen pro Statusgruppe maximal zugrunde liegende Fallanzahl; M: Mittelwert; SD: Standardabweichung; df: Freiheitsgrade; t: t-Wert; *: $p \leq 0,05$; **: $p \leq 0,01$; d = Effektstärke nach Cohens.

Die mittleren Einschätzungen der Schüler:innen liegen zwischen den Werten $M = 2,05$ (Ö_k) und $M = 2,54$ (G_i), sodass die Schlussfolgerung gezogen werden kann, dass keine Erklärung für alle Adressat:innen bereits vollständig bekannte Inhalte aufgreift. Die Standardabweichungen der Einschätzungen ($range = [0,63; 0,79]$) deuten auf die Heterogenität des Vorwissens innerhalb der Stichprobe hin. Insgesamt zeigen die angegebenen Verteilungsparameter, dass die Inhalte aller Themenbereiche überwiegend unbekannt sind. Tendenziell schätzen die befragten Schüler:innen ihr Vorwissen zu Inhalten der Themenbereiche Genetik und Physiologie geringer als zum Themenbereich Ökologie ein.

Wie die Mittelwerte bereits andeuten, lassen sich nur innerhalb des Themenbereichs der Ökologie keine signifikanten Unterschiede zwischen den Videopaaren feststellen. Innerhalb der Themenbereiche Genetik und Physiologie sind die Unterschiede auch inferenzstatistisch abzubilden, wobei Schüler:innen angeben, mehr Vorwissen bei der komplexeren Erklärungsvariante aus dem Themenbereich der Genetik und invers dazu weniger Vorwissen bei der komplexeren Erklärungsvariante aus dem Themenbereich der Physiologie zu besitzen. Die berechneten Effektstärken zeigen an, dass es sich um einen eher kleinen Effekt im Bereich der Genetik und einen mittleren Effekt im Bereich der Physiologie handelt (Cohen, 1988).

Über die Stärke des Zusammenhangs von niedrigem bzw. hohem Vorwissen auf die globale Bewertung von instruktionalen Erklärungen sowie auf die Aspekte instruktionaler Erklärungen können keine konkreten Hypothesen formuliert werden, sodass die Unterschiede bei der Betrachtung deskriptiver Statistiken nicht berücksichtigt werden. Da insgesamt keine systematischen Unterschiede zwischen linearen und komplexen Erklärungen vorliegen, kann auf die Implementation des Faktors ‚Vorwissen‘ bei den inferenzstatistischen Analysen zu Unterschieden zwischen komplexen und linearen Erklärungen verzichtet werden.

Vorwissen der Studierenden

Von den Studierenden besuchten $N = 6$ Personen bereits ein Seminar zum Thema ‚Erklären im Biologieunterricht‘ und weitere $N = 14$ ein Seminar zum Thema ‚Erklären‘ aus dem Kursangebot einer anderen Fachdidaktik, in dem sie die Möglichkeit zum Erwerb von Vorwissen zu Kriterien guten Erklärens hatten.

Unterschiede hinsichtlich der Einschätzung in Abhängigkeit eines Seminarbesuchs (Seminarbesuch vs. kein Seminarbesuch) wurden mit Varianzanalysen mit Messwiederholung überprüft. Als abhängige Variable wurde das Globalurteil verwendet. Die Bewertungen der sechs Videovignetten werden als sechsstufiger Innersubjektfaktor (Messwiederholung) und der Seminarbesuch (ja/nein) als zweistufiger Zwischensubjektfaktor modelliert. Die Voraussetzung der Normalverteilung wurde mittels Kolmogorov-Smirnov-Test überprüft und ist in acht Fällen nicht gegeben (Kolmogorov-Smirnov-Test signifikant für Video G_i , P_i und O_i in beiden Gruppen, für Video G_k und P_k in der Gruppe ohne Seminarbesuch). Gegebenenfalls sind die gefundenen Effekte überschätzt (Bortz & Schuster, 2010). Bei gegebener Sphärizität (Mauchly- $W(14) = 0,68$; $p = 0,20$) und Varianzhomogenität (Levene Tests mit $0,18 \leq p \leq 0,98$) ergibt sich keine signifikante Interaktion zwischen der Bewertung der Videovignetten und dem Seminarbesuch ($F(5, 240) = 0,74$; $p = 0,59$; $\eta^2 = 0,02$).

Hinsichtlich der Äußerungen der Studierenden im offenen Item zu Begründung der Notenvergabe können jedoch Unterschiede gefunden werden, die nahelegen, dass durch den Besuch eines Seminars (erwartungskonform) Kriterien guter Erklärungen erlernt werden, die zur Bewertung einer Erklärung herangezogen werden können. Tabelle 15 zeigt die Ergebnisse des Vergleichs der in den Äußerungen über alle Videovignetten angesprochenen Hauptkategorien unter Angabe der relativen und absoluten Häufigkeiten pro Gruppe.

Tabelle 15: Relative und absolute Häufigkeiten der Äußerungen innerhalb der Hauptkategorien von Studierenden mit und ohne Seminarbesuch über alle Videovignetten

	Ohne Seminarbesuch ($N_{max} = 31$)		mit Seminarbesuch ($N_{max} = 20$)		Gesamt ($N_{max} = 51$)	
	R.H.	A.H.	R.H.	A.H.	R.H.	A.H.
Adressatenorientierung	16,97	47	21,93	25	18,41	72
Strukturiertheit	8,66	24	4,39	5	7,42	29
Repräsentationen	27,44	76	15,79	18	24,04	94
Verständliche Sprache	6,86	19	10,53	12	7,93	31
Sprech- und Körperausdruck	8,30	23	15,79	18	10,49	41
Fach	12,64	35	15,79	18	13,55	53
Globale Aussagen	12,99	36	13,16	15	13,04	51
Lehrperson	0,00	0	0,87	1	0,26	1
Merkbarkeit u. Lernbarkeit	2,53	7	0,00	0	1,79	7
Sonstiges	3,61	10	1,75	2	3,07	12
Σ	100	277	100	114	100	391

Bem.: N_{max} : maximal zugrunde liegende Anzahl der Teilnehmer:innen; R.H.: relative Häufigkeit der Nennungen; A.H.: absolute Häufigkeit der Nennungen.

Betrachtet man die Verteilungen der Äußerungen im Vergleich, so verwenden Studierende, die kein Seminar zum Thema ‚Erklären‘ besucht haben, mit Abstand am häufigsten Kriterien zur Bewertung, die sich auf die Repräsentationen beziehen. Ebenfalls oft werden Kriterien genutzt, die sich den Kategorien ‚Adressatenorientierung‘, ‚Globale Aussagen‘ oder den fachlichen Aspekten zuordnen lassen. Bei der Gruppe der Studierenden, die ein Seminar besucht haben, ergeben sich weniger deutliche Differenzen zwischen den Häufigkeiten der Hauptkategorien. Sie beziehen sich bei ihren Begründungen zur Notenvergabe ebenfalls am häufigsten auf die eingesetzten Repräsentationen, sprechen jedoch alle weiteren Aspekte guten Erklärens – mit Ausnahme der Strukturiertheit – ebenso oft an.¹³³ Neben der Auffälligkeit in der generellen Verteilung der Äußerungen auf die Hauptkategorien ergeben sich insbesondere in den Kategorien ‚Verständliche Sprache‘ sowie ‚Sprech- und Körperausdruck‘ bemerkenswerte Differenzen in den Häufigkeiten. So nennen die Studierenden, die ein Seminar besucht haben, deutlich häufiger Kriterien, die sich diesen beiden Hauptkategorien zuordnen lassen, als Studierende, die an keinem Seminar teilgenommen haben. Ein Grund dafür kann die überwiegend erstmalige, intensivere Beschäftigung mit Sprechen und Sprache innerhalb des Lehramtsstudiums an der Universität Regensburg in den Seminaren zum Erklären sein, die die Studierenden dazu befähigt, diese Aspekte der erklärenden Person wahrzunehmen und für eine Einschätzung von Erklärungen heranzuziehen. Den theoretischen Rahmen hierfür liefern die Ansätze ‚professional vision‘ (Sherin, 2001) und ‚perception-interpretation-decision-making‘ (Blömeke et al., 2015).

Die Ergebnisse legen nahe, dass die relativen Häufigkeiten in der gesamten Teilstichprobe der Studierenden in den Kategorien Sprech- und Körperausdruck sowie Sprache insgesamt eher überschätzt werden bzw. höher ausfallen.

6.1.2 Gütekriterien der quantitativen Erhebung

Die Testgüte respektive die Güte der gewonnenen Daten und Ergebnisse für den quantitativen Forschungszugang wird im Folgenden hinsichtlich der drei genannten, zentralen Hauptgütekriterien Objektivität, Reliabilität und Validität abgeschätzt (Bortz & Döhring, 2006; Bühner, 2011).

Die Objektivität der Ergebnisse als Grad ihrer Unabhängigkeit vom Untersucher kann für dieses Untersuchungsvorhaben die Durchführungs- und Auswertungsobjektivität betrachtet werden. Dabei konnte die Durchführungsobjektivität erhöht werden, weil die mündlich gegebenen Instruktionen bei der Erhebung der Schüler:innen zu Bearbeitungsanweisungen und Nachfragen bei Verständnisproblemen vorab formuliert und dadurch standardisiert wurden. Bei den weiteren Teilstichproben wurden die Informationen zur Erhebung schriftlich und in wesentlichen Teilen identisch formuliert per E-Mail übermittelt. Die Bearbeitungsanweisungen wurden in den ersten Seiten des Onlinefragebogens ebenfalls schriftlich mitgeteilt. Nicht vorgegeben war hingegen bei den Teilstichproben der Studierenden, Lehrkräfte und Didaktiker:innen der Abstand zwischen den zwei Messzeitpunkten. Insgesamt kann jedoch von einer hohen Durchführungsobjektivität ausgegangen werden. Die Auswertungsobjektivität konnte durch das geschlossene Antwortformat im quantitativen Teil der Untersuchung, die computergestützte Datenerhebung und den direkten Export der Daten in die entsprechenden Programme gewährleistet werden.

¹³³ Die Kategorien ‚Lehrperson‘, ‚Merkbarkeit und Lernbarkeit‘ sowie ‚Sonstiges‘ sind aus diesem Vergleich ausgenommen.

Als empirischen Indikator für diese Messgenauigkeit wurde Cronbachs Alpha als Maß für die interne Konsistenz berechnet. Die Berechnungen erfolgten pro Vignette. Da unterschiedliche Werte für die verschiedenen Perspektiven zu erwarten waren, wurde Cronbachs Alpha nach Statusgruppen getrennt berechnet. Die Ergebnisse sind in den Tabellen 16 bis 20 einsehbar.¹³⁴

Das fachspezifische Konstrukt wird mit denjenigen Items gebildet, die sowohl bei linearen als auch bei komplexen Erklärungen abgefragt werden. Das Item „*Die Erklärung zeigt nur einen Ausschnitt auf die Frage „...“.*“ (Item: *Fa_aus*) wurde aufgrund negativer Itemtrennschärfen aus der Skala ausgeschlossen.

In den Aspekt der sprachlichen Verständlichkeit gehen drei der vier im Fragebogen enthaltenen Items ein, da sich für das Item „*Manchmal hat die Lehrerin extra gesagt, dass etwas sehr wichtig ist*“ (Item: *Sw_wic*) negative Trennschärfen ergeben haben.

Tabelle 16: Interne Konsistenzen der sechs Skalen für Schüler:innen – Cronbachs Alpha und mittlere part-whole korrigierte Itemtrennschärfen inkl. Standardabweichungen

Video	Video G _i	Video G _k	Video Ö _i	Video Ö _k	Video P _i	Video P _k
	α	α	α	α	α	α
Konstrukt	$r_{it}: M (SD)$	$r_{it}: M (SD)$	$r_{it}: M (SD)$	$r_{it}: M (SD)$	$r_{it}: M (SD)$	$r_{it}: M (SD)$
Adressatenorientierung (4 Items)	0,54 0,34 (0,29)	0,38 0,20 (0,26)	0,29 0,15 (0,29)	0,16 0,06 (0,21)	0,53 0,31 (0,20)	0,55 0,35 (0,23)
Strukturiertheit (4 Items)	0,74 0,54 (0,23)	0,69 0,47 (0,27)	0,63 0,40 (0,26)	0,71 0,51 (0,30)	0,75 0,53 (0,17)	0,74 0,53 (0,19)
Sprech- u. Körperausdruck (7 Items)	0,78 0,50 (0,09)	0,69 0,41 (0,10)	0,77 0,49 (0,10)	0,88 0,66 (0,08)	0,85 0,61 (0,12)	0,80 0,53 (0,11)
Sprachliche Verständlichkeit (3 Items)	0,43 0,29 (0,19)	0,45 0,30 (0,17)	0,48 0,35 (0,21)	0,50 0,33 (0,11)	0,37 0,25 (0,18)	0,57 0,39 (0,11)
Persönlichkeitswirkung (3 Items)	0,84 0,70 (0,09)	0,68 0,49 (0,05)	0,74 0,56 (0,08)	0,86 0,74 (0,03)	0,89 0,79 (0,05)	0,78 0,62 (0,05)
Fachspezifische Facette (4 Items)	0,76 0,56 (0,09)	0,72 0,51 (0,13)	0,69 0,48 (0,07)	0,72 0,51 (0,12)	0,80 0,62 (0,06)	0,77 0,57 (0,07)

Bem.: N_{max} : maximale Stichprobengröße; α : Cronbachs Alpha; r_{it} : mittlere part-whole korrigierte Trennschärfe; M : Mittelwert; SD : Standardabweichung.

¹³⁴ Die vollständige Tabelle mit allen deskriptiven Werten pro Item und die jeweiligen Itemtrennschärfen pro Videovignette sind im Anhang D einzusehen.

Tabelle 17: Interne Konsistenzen der sechs Skalen für Studierende – Cronbachs Alpha und mittlere part-whole korrigierte Itemtrennschärfen inkl. Standardabweichungen

Video	Video G _i	Video G _k	Video Ö _i	Video Ö _k	Video P _i	Video P _k
Konstrukt	α $r_{it}: M (SD)$	α $r_{it}: M (SD)$	α $r_{it}: M (SD)$	α $r_{it}: M (SD)$	α $r_{it}: M (SD)$	α $r_{it}: M (SD)$
Adressatenorientierung (5 Items)	0,70 0,46 (0,18)	0,81 0,60 (0,11)	0,62 0,38 (0,22)	0,38 0,20 (0,19)	0,86 0,68 (0,11)	0,84 0,64 (0,05)
Strukturiertheit (4 Items)	0,80 0,59 (0,17)	0,85 0,68 (0,17)	0,72 0,43 (0,22)	0,80 0,60 (0,18)	0,77 0,55 (0,22)	0,85 0,66 (0,11)
Sprech- u. Körperausdruck (7 Items)	0,84 0,60 (0,14)	0,59 0,32 (0,17)	0,80 0,53 (0,17)	0,84 0,60 (0,12)	0,91 0,71 (0,08)	0,78 0,52 (0,19)
Sprachliche Verständlichkeit (3 Items)	0,62 0,42 (0,02)	0,69 0,51 (0,05)	0,67 0,48 (0,07)	0,53 0,35 (0,12)	0,63 0,46 (0,10)	0,73 0,55 (0,07)
Persönlichkeitswirkung (3 Items)	0,73 0,54 (0,10)	0,69 0,50 (0,7)	0,76 0,59 (0,09)	0,82 0,67 (0,08)	0,89 0,74 (0,09)	0,82 0,67 (0,14)
Fachspezifische Facette (5 Items)	0,78 0,56 (0,09)	0,83 0,62 (0,04)	0,78 0,55 (0,04)	0,72 0,49 (0,17)	0,83 0,63 (0,16)	0,87 0,70 (0,10)

Bem.: N_{max} : maximale Stichprobengröße; α : Cronbachs Alpha; r_{it} : mittlere part-whole korrigierte Trennschärfe; M : Mittelwert; SD : Standardabweichung.

Tabelle 18: Interne Konsistenzen der sechs Skalen für Lehrkräfte – Cronbachs Alpha und mittlere part-whole korrigierte Itemtrennschärfen inkl. Standardabweichungen

Video	Video G _i	Video G _k	Video Ö _i	Video Ö _k	Video P _i	Video P _k
Konstrukt	α $r_{it}: M (SD)$	α $r_{it}: M (SD)$	α $r_{it}: M (SD)$	α $r_{it}: M (SD)$	α $r_{it}: M (SD)$	α $r_{it}: M (SD)$
Adressatenorientierung (5 Items)	0,64 0,40 (0,14)	0,60 0,36 (0,11)	-0,19 0,00 (0,35)	0,48 0,26 (0,18)	0,84 0,63 (0,04)	0,80 0,58 (0,09)
Strukturiertheit (4 Items)	0,74 0,51 (0,24)	0,78 0,57 (0,14)	0,83 0,67 (0,21)	0,92 0,82 (0,10)	0,90 0,78 (0,09)	0,92 0,78 (0,13)
Sprech- u. Körperausdruck (7 Items)	0,72 0,43 (0,18)	0,90 0,71 (0,09)	0,89 0,68 (0,11)	0,90 0,72 (0,09)	0,87 0,66 (0,10)	0,94 0,80 (0,05)
Sprachliche Verständlichkeit (3 Items)	0,57 0,45 (0,24)	0,38 0,20 (0,13)	0,67 0,48 (0,04)	0,64 0,45 (0,05)	0,74 0,56 (0,04)	0,73 0,58 (0,16)
Persönlichkeitswirkung (3 Items)	0,67 0,45 (0,08)	0,74 0,57 (0,07)	0,76 0,58 (0,21)	0,84 0,68 (0,16)	0,84 0,70 (0,03)	0,84 0,70 (0,14)
Fachspezifische Facette (5 Items)	0,72 0,49 (0,22)	0,75 0,52 (0,18)	0,59 0,35 (0,15)	0,58 0,35 (0,14)	0,81 0,61 (0,13)	0,74 0,51 (0,12)

Bem.: N_{max} : maximale Stichprobengröße; α : Cronbachs Alpha; r_{it} : mittlere part-whole korrigierte Trennschärfe; M : Mittelwert; SD : Standardabweichung.

Obwohl sich bei Lehrkräften für das Video Öl ein negatives Cronbachs α in Bezug auf die Skala der Adressatenorientierung ergibt, wird diese aus Gründen der Einheitlich- und Vollständigkeit wie bei allen anderen Vignetten und Statusgruppen dennoch mit allen fünf Items gebildet und ausgewiesen.

Die nur hier vorliegende Inkonsistenz der Skala ist insbesondere auf das Item „Die Erklärung war für Schüler_innen zu schwierig“ zurückzuführen. Lehrkräfte empfinden die Erklärung für Schüler:innen eher als zu einfach, was aus inhaltlicher Sicht insofern nachvollziehbar ist, als darin der einfache Zusammenhang zwischen der Größe der Population von Räufern und der Population der Beutetiere monokausal erklärt wird. Der hohe repetitive Charakter könnte zu dieser Einschätzung führen.

Tabelle 19: Interne Konsistenzen der sechs Skalen für Didaktiker:innen – Cronbachs Alpha und mittlere part-whole korrigierte Itemtrennschärfen inkl. Standardabweichungen

Video	Video G _i	Video G _k	Video Ö _i	Video Ö _k	Video P _i	Video P _k
	α	α	α	α	α	α
Konstrukt	$r_{it}: M (SD)$	$r_{it}: M (SD)$	$r_{it}: M (SD)$	$r_{it}: M (SD)$	$r_{it}: M (SD)$	$r_{it}: M (SD)$
Adressatenorientierung (5 Items)	0,76 0,53 (0,18)	0,64 0,40 (0,11)	0,69 0,46 (0,22)	0,84 0,65 (0,22)	0,78 0,57 (0,25)	0,57 0,34 (0,31)
Strukturiertheit (4 Items)	0,89 0,76 (0,15)	0,84 0,67 (0,17)	0,84 0,67 (0,14)	0,86 0,71 (0,19)	0,88 0,73 (0,11)	0,88 0,73 (0,14)
Sprech- und Körperausdruck (7 Items)	0,87 0,66 (0,16)	0,89 0,68 (0,11)	0,87 0,65 (0,13)	0,80 0,52 (0,13)	0,89 0,69 (0,16)	0,91 0,74 (0,07)
Sprachliche Verständlichkeit (3 Items)	0,62 0,44 (0,08)	0,44 0,30 (0,18)	0,53 0,41 (0,32)	0,68 0,48 (0,12)	0,70 0,52 (0,07)	0,63 0,54 (0,33)
Persönlichkeitswirkung (3 Items)	0,61 0,42 (0,13)	0,56 0,37 (0,13)	0,65 0,45 (0,04)	0,68 0,48 (0,14)	0,67 0,49 (0,09)	0,63 0,44 (0,06)
Fachspezifische Facette (5 Items)	0,72 0,49 (0,18)	0,74 0,51 (0,16)	0,69 0,44 (0,10)	0,74 0,50 (0,24)	0,80 0,58 (0,09)	0,56 0,34 (0,18)

Bem.: N_{max} : maximale Stichprobengröße; α : Cronbachs Alpha; r_{it} : mittlere part-whole korrigierte Trennschärfe; M : Mittelwert; SD : Standardabweichung.

Field (2018) gibt an, dass bei neu erstellten Skalen Werte ab einem Cronbachs Alpha = 0,5 als akzeptabel betrachtet werden können. Mit diesem Richtwert bedeutet das für die Statusgruppe der Didaktiker:innen, dass alle angegebenen internen Konsistenzen deutlich darüber liegen und somit reliabile Skalen vorliegen. Für die Gruppe der Studierenden kann dies mit einer Ausnahme ebenso formuliert werden (Skala zur Adressatenorientierung für das Video Ö_k). Die reduzierte Reliabilität wird bei der Interpretation der Ergebnisse berücksichtigt. Für Lehrkräfte ergeben sich – neben dem bereits erwähnten negativen Cronbachs Alpha – bei den Skalen zur sprachlichen Verständlichkeit teilweise niedrigere Werte als 0,5. Aufgrund theoretischer Überlegungen wurden die Items trotz eher niedriger Werte in der Skala belassen. Die Werte der Teilstichprobe der Schüler:innen für die Aspekte des Sprech- und Körperausdrucks, der Strukturiertheit, der Persönlichkeitswirkung sowie der fachspezifischen Facette sind in allen Videos größer als 0,63 und damit als gut einzuordnen. Für die sprachliche Verständlichkeit sowie die Adressatenorientierung ergeben sich bei einzelnen Vignetten hingegen eher niedrige Werte. Aufgrund theoretischer Überlegungen wurden die Items mit geringer Trennschärfe dennoch für in den jeweiligen Aspekte beibehalten und nicht zugunsten ihrer Homogenisierung verworfen (Bühner, 2011). Die teils geringen Werte von Cronbachs Alpha sind als Ausdruck für die Breite der Aspekte zu sehen. Das ermöglicht insgesamt, die entsprechenden Konstrukte zu bilden bzw. statistische Aussagen über diese zu treffen – mit Ausnahme des Falles der linearen Erklärung im Themenbereich Ökologie bei Lehrkräften.

Die Validität kann – wie die Objektivität – nicht unmittelbar anhand von spezifischen Kennwerten empirisch überprüft werden (Bühner, 2011). Das Kriterium bezieht sich auf die Gültigkeit und damit darauf, dass der Test bzw. der Fragebogen auch das misst, was er zu messen vorgibt.¹³⁵ Der Zielaspekt der Inhaltsvalidität als Angabe darüber, dass die Items den entsprechenden Aspekt hinsichtlich der wichtigsten Facetten erfassen (Bortz & Döhring, 2006) und dies präzise tun, ohne einen Überschneidungsbereich mit anderen Aspekten bzw. Skalen aufzuweisen (Bühner, 2011), wurde in Bezug auf die Items zum einen durch die Konstruktion auf Basis von entsprechender Literatur erhöht. Zum anderen wurden innerhalb der Fachgruppe der Biologiedidaktik sowie der fachübergreifenden Projektgruppe intensive Diskussionen über die Aspekte geführt. Weiterhin können die Pilotierungsinterviews mit zwei Schüler:innen, zwei Studierenden und zwei Didaktikerinnen zum Verständnis der Items zur Sicherstellung der Inhaltsvalidität beitragen (siehe Kap. 5.2.6).

Als empirischer Indikator für die Validität wurden für die Aspekte der Strukturiertheit, der Adressatenorientierung, des Sprech- und Körperausdrucks sowie der sprachlichen Verständlichkeit skalenbezogene Globalurteile abgefragt. Für die Statusgruppe der Schüler:innen erfolgte keine Abfrage eines solchen Globalurteils für Adressatenorientierung. Die jeweiligen skalenspezifische Globalurteile können mit den Skalenmittelwerten korreliert werden. Eine hohe Korrelation spricht dann für eine hohe Abdeckung des Aspekts durch die im Fragebogen enthaltenen Items. Tabelle 20 zeigt die Ergebnisse der Korrelationen.

Für die aufgeführten Aspekte der Adressatenorientierung, der Strukturiertheit sowie des Sprech- und Körperausdrucks kann mit Blick auf die insgesamt hohen Korrelationskoeffizienten, die nach Cohen (1988) einen großen Effekt beschreiben, von einer hohen Abdeckung der Aspekte und damit einhergehend von einer hohen Validität ausgegangen werden. Für die Skala zur sprachlichen Verständlichkeit fällt der Korrelationskoeffizient insbesondere bei Schüler:innen sehr niedrig aus. Wie in Kapitel 5.3.1 bereits angeführt wurde, wird in die Skala deshalb das skalenspezifische Globalurteil mit aufgenommen, sodass der Aspekt breiter erfasst werden kann.

¹³⁵ Das Kriterium der Validität bezieht sich klassischerweise auf die Inhaltsvalidität (im folgenden Abschnitt erläutert) sowie die Kriteriums- und Konstruktvalidität. Die Kriteriumsvalidität kann aufgrund einer fehlenden Norm bzw. eines manifesten Merkmals, das entsprechende Kriteriumswerte liefert, die wiederum mit den Werten der Testpersonen verglichen werden können, nicht ohne Weiteres angegeben werden. Einen Anhaltspunkt liefern die Ergebnisse des hierarchischen, linearen Regressionsmodells mit der abhängigen Variable Globalurteil (siehe Kap. 6.2.3.). Die Konstruktvalidität kann hingegen nicht berechnet werden, da vergleichbare, konstruktverwandte oder konstruktferne Tests fehlen, die für einen Vergleich herangezogen werden könnten.

Tabelle 20: Korrelationskoeffizienten r nach Pearson für die vier skalenbezogenen Globalurteile der Aspekte und den implementierten Items der Aspekte (ohne die aufgrund negativer Trennschärfen ausgeschlossenen Items) je Videovignette getrennt nach den Statusgruppen der Schüler:innen ($N_{max} = 122$) und der Lehrenden (Studierende, Lehrkräfte, Fachdidaktiker:innen ($N_{max} = 97$))

Aspekt	Adressatenorientierung		Strukturiertheit		Sprech- und Körperausdruck		Verständliche Sprache		
	SuS (/)	SLD (4 Items)	SuS (4 Items)	SLD (4 Items)	SuS (7 Items)	SLD (7 Items)	SuS (2 Items)	SLD (2 Items)	
G _i	r	0,72**	0,65**	0,73**	0,72**	0,86**	0,08	0,33**	
	p	/	< 0,01	< 0,01	< 0,01	< 0,01	0,37	< 0,01	
	N		97	122	97	43	49	122	96
G _k	r	0,79**	0,60**	0,73**	0,67**	0,76**	0,22*	0,40**	
	p	/	< 0,01	< 0,01	< 0,01	< 0,01	< 0,01	0,02	< 0,01
	N		96	121	96	70	60	121	94
Ö _i	r	0,72**	0,49**	0,72**	0,68**	0,71**	0,05	0,44**	
	p	/	< 0,01	< 0,01	< 0,01	< 0,01	< 0,01	0,57	< 0,01
	N		95	120	95	70	60	121	93
Ö _k	r	0,65**	0,75**	0,72**	0,87**	0,82**	0,11	0,38**	
	p	/	< 0,01	< 0,01	< 0,01	< 0,01	< 0,01	0,24	< 0,01
	N		95	121	95	43	49	121	95
P _i	r	0,78**	0,66**	0,73**	0,61**	0,85**	0,11	0,37**	
	p	/	< 0,01	< 0,01	< 0,01	< 0,01	< 0,01	0,23	< 0,01
	N		94	121	94	42	49	120	94
P _k	r	0,73**	0,74**	0,69**	0,78**	0,82**	0,27**	0,39**	
	p	/	< 0,01	< 0,01	< 0,01	< 0,01	< 0,01	< 0,01	< 0,01
	N		94	120	94	68	60	120	93

Bem.: SuS: Schüler:innen; SLD: Studierende, Lehrkräfte und Hochschuldidaktiker:innen; r : Korrelationskoeffizient nach Pearson; p : p-Wert; N : Stichprobengröße nach paarweisem Fallausschluss; *: $p \leq 0,05$; **: $p \leq 0,01$.

Zur Gewährleistung der Inhaltsvalidität in Bezug auf die interessierende Variable (UV) in den Videovignetten – also die Variation im dargestellten Komplexitätsgrad – tragen die bereits erwähnten Diskussionen mit Kolleg:innen sowie die Rückfragen in den Pilotierungsinterviews zu den Unterschieden der zwei gezeigten Vignetten ebenfalls bei (siehe Kap. 5.2.6).

In Hinblick auf die Abschätzung der Konstruktvalidität kann den Ergebnissen vorgegriffen werden: In den Begründungen zur Notenvergabe in den offenen Textfeldern lassen sich die im standardisierten Zugang aufgenommenen Aspekte guten Erklärens finden (siehe Kap. 6.2.1). Auch die Ergebnisse der hierarchischen linearen Regressionsmodelle sprechen für eine hohe Konstruktvalidität (siehe Kap. 6.2.3).

Zur Absicherung der Augenscheinvalidität bzw. der Trennschärfe der fächerübergreifenden Aspekte Adressatenorientierung, Strukturiertheit, sprachliche Verständlichkeit, Sprech- und Körperausdruck sowie der Wirkung der erklärenden Person wurden Lehramtsstudierende, Promovierende und Professor:innen der empirischen Bildungsforschung ($N = 98$) der Universität Regensburg befragt (Schilcher, Krauss, Lindl & Hilbert, im Druck). Dabei wurden die Items ungeordnet präsentiert und sollten den fünf Aspekten zugeordnet werden. Die Aspekte wurden vorab kurz skizziert, sodass eine reliable Messung bzw. Zuordnung ermöglicht wird. Die Ergebnisse der Zuordnung der Items zeigen, dass in 24 von 26 Fällen die höchsten relativen Häufigkeiten bei dem intendierten Aspekt auftreten (die vollständige Tabelle ist im Anhang C einzusehen). **Das Item „Die Lehrkraft hat die Sprechpausen**

so gesetzt, dass man gut folgen konnte“ (Item: *Se_pau*) wird hingegen in 42 % der Fälle der sprachlichen Verständlichkeit und in 40 % der Fälle dem Sprech- und Körperausdruck zugeordnet. Diese etwa gleichmäßige Verteilung auf die Aspekte der sprachlichen Verständlichkeit und des Sprech- und Körperausdrucks spiegelt die auch bei der Codierung der offenen Antworten hervortretende Schwierigkeit der Differenzierung zwischen ‚Sprache‘ und ‚Sprechen‘ wider, die sich erfahrungsgemäß auch bei der Thematisierung von Kriterien guten Erklärens mit Studierenden und dem Geben von Feedback in Seminarangeboten zeigt. Trotz der hohen Schwierigkeit der Zuordnung für die befragten Personen wird das Item in der Facette des Sprech- und Körperausdrucks belassen. Zum einen sprechen die Korrelationen des Items mit den übrigen Items der Facette aus statistischer Sicht nicht gegen einen Einbezug und zum anderen spricht die theoretische Fundierung für einen Einbezug (Lägel-Gunga et al., im Druck). Ein eindeutiges Bild zeigt sich bei der Zuordnung des Items „Manchmal hat die Lehrkraft gesagt, dass etwas sehr wichtig ist“ (Item: *Sw_wic*), die in 61 % der Fälle zur Strukturiertheit und nicht wie intendiert zum Aspekt der sprachlichen Verständlichkeit (13 %) erfolgt. Damit liefern die Ergebnisse ein weiteres Argument für das Entfernen dieses Items aus dem Aspekt der sprachlichen Verständlichkeit.

6.1.3 Gütekriterien der qualitativen Erhebung

Die Frage, ob in qualitativen Forschungszugängen die Gütekriterien Objektivität, Reliabilität und Validität ebenfalls Gültigkeit haben bzw. haben können, wird vielseitig diskutiert, wobei mehrere Grundhaltungen differenzierbar sind (Bortz & Döhring, 2006; Kuckartz, 2018; Steinke, 2007): die Annahme der Generalisierbarkeit der Kriterien und damit einhergehend die Anwendung derer auf qualitative Forschungsvorhaben¹³⁶, die Grundhaltung zur Spezifität der Gütekriterien in Abhängigkeit des Forschungszugangs und damit wiederum einhergehend die Reformulierung und Erweiterung der Gütekriterien standardisierter Forschung oder die Entwicklung eigener Gütekriterien zur Abgrenzung der Forschungszugänge.

Da die gewählte Auswertungsmethode der inhaltlich-strukturierenden qualitativen Inhaltsanalyse in der Forschungspraxis etabliert ist, sind hier bereits Empfehlungen für die interne Studiengüte expliziert worden (Kuckartz, 2018).¹³⁷ Die Empfehlungen folgen dem Ansatz, die aus den quantitativen respektive standardisierten Forschungszugängen stammenden Gütekriterien zu reformulieren, auf die Bedingungen der qualitativen Forschungsmethoden zu übertragen und so zu spezifischen Standards weiterzuentwickeln. In weiten Teilen lassen sich die von Kuckartz (2018) aufgelisteten Punkte zur Reflexion bzw. zur Sicherung der internen Studiengüte, die sehr konkret formuliert sind („Ist die gewählte inhaltsanalytische Methode der Fragestellung angemessen?“, „Wurde mit Originalzitate gearbeitet und nach welchen Kriterien wurden diese ausgewählt?“, „Wurde eine vollständige Transkription des Interviews vorgenommen?“ (Kuckartz, 2018, S. 204–205)), allgemeinen bzw. nach Prinzipien strukturierte Kriterien zuordnen, wie sie beispielsweise von Steinke (2012) vorgeschlagen werden. Sie nennt für die interne Studiengüte Folgendes:

¹³⁶ Dieser Ansatz scheint der problematischste zu sein. Bortz und Döhring (2006) führen dafür zwei Gründe an: „Zum einen werden bei diesen Übertragungsversuchen nicht selten die Systematik sowie die Definitionen quantitativer Gütekriterien unscharf verwendet: Oft findet eine Vermischung der Testgütekriterien Objektivität, Reliabilität und Validität (die sich auf Messinstrumente beziehen) mit den Validitätskriterien wissenschaftlicher Aussagen statt [...]. Zum anderen ist zu hinterfragen, ob es sinnvoll ist, Kriterien für Studien, die einem ganz anderen wissenschaftstheoretischen Paradigma folgen, in das qualitative Paradigma zu importieren“ (Bortz & Döhring, 2006, S. 107).

¹³⁷ Bohnsack (2005) beschreibt als Charakteristikum qualitativer bzw. rekonstruktiver Forschung, dass die Explikation und Festlegung von Standards zur Studiengüte erst nach der Etablierung der spezifischen Methode in der Forschungspraxis geschehen können, da die Standards aus der konkreten Forschungspraxis abgeleitet werden.

- Die Indikation des Forschungsprozesses unter anderem als Reflexion über die Angemessenheit des qualitativen Vorgehens, der Methodenauswahl und insbesondere der Passung von Erhebungs- und Auswertungsmethode (siehe Kap. 5.3.2 und 5.3.2) sowie der ausgewählten Untersuchungsfälle (hier Statusgruppen; siehe Kap. 4.8 und 5.2.7) für den gewählten Forschungsgegenstand. Auch sukzessive Anpassung der Kategoriensystems im Sinne der induktiven Kategorienbildung und die damit einhergehende Öffnung des Forschungsprozesses (siehe Kap. 5.2.3.3) erhöhen die Gegenstandsangemessenheit in konkreter inhaltlicher Dimension. Aufgeführt wird unter dem Stichwort der Indikation auch die Indikation von Transkriptionsregeln. Diese entfällt bei der hier durchgeführten Studie, da das zu analysierende Textmaterial von den Teilnehmer:innen selbst produziert wird und der gesamte Textkorpus in die Analyse eingeht.
- Die empirische Verankerung, also die Begründung der Hypothesen oder Theorien in den Daten, die beispielsweise durch den Beleg anhand konkreter Textstellen bzw. anhand des vorliegenden empirischen Materials sichergestellt werden kann (siehe z. B. Kap. 5.3.2.2).
- Die intersubjektive Nachvollziehbarkeit, die häufig als zentrales Kriterium für qualitative Forschungsarbeiten angeführt wird und sich durch durchgehende Transparenz und Kommunikation auszeichnet. Sie wird unter anderem durch eine möglichst genaue Darstellung des gesamten Forschungsprozesses gewährleistet. Die Dokumentation der Auswertungsmethoden (Kap. 5.3.2) und der Informationsquellen (Kap. 5.2.7) erhöhen die Nachvollziehbarkeit und damit die Überprüfbarkeit der Vorgehensweisen sowie der daraus resultierenden Forschungsergebnisse bzw. deren Zuverlässigkeit und Glaubwürdigkeit. Weiterhin wird die Anwendung kodifizierter Verfahren **als Anlehnung „an herkömmliche Wege zur Herstellung von Intersubjektivität“** (Steinke, 2012, S. 326) durch die Verwendung einheitlicher, etablierter Methodiken angeführt. Dieser Aspekt steht in Verbindung mit dem von Kuckartz (2018) geforderten korrekten, regelgeleiteten Anwendung der Methode. Zuletzt nennt Steinke (2012) **die Interpretation in Gruppen „als eine diskursive Form der Herstellung von Intersubjektivität und Nachvollziehbarkeit durch expliziten Umgang mit Daten und deren Interpretation“** (Steinke, 2012, S. 326). Dazu dienten insbesondere die fachinternen Diskussionen im Zuge der Kategorienbildung bzw. der Ausschärfung der Kategoriendefinitionen sowie der Codierungsprozess, der durch das konsensuelle Codieren durch einen hohen Anteil an Kommunikation über die Daten und deren Interpretation geprägt war.

In Bezug auf die Methode der qualitativen Inhaltsanalyse wird als Kriterium der Zuverlässigkeit auch die Güte des Codierungsprozesses diskutiert (Kuckartz, 2018; Mayring, 2010). Hierzu wurde das gesamte Material von zwei Codiererinnen codiert und die Ergebnisse wurden computergestützt mit der MAXQDA Software verglichen. Wie bereits in Abschnitt 5.3.2.2 erläutert wurde für eine Orientierung ebenfalls die prozentuale Inter-Coder-Übereinstimmung berechnet. Sie betrug für die erste Hälfte des Materials 69,62 %. Die Abweichungen wurden diskutiert und die Segmente entsprechend dem Konsens der jeweiligen Kategorie zugeordnet. Gleichzeitig wurde das Kategoriensystem ausgeschärft und das Codebuch präzisiert. Insgesamt entstanden im Verlauf des Codierungsprozesses ein Kategoriensystem und ein Codebuch als Interpretationsbasis mit hoher Materialnähe und ein intersubjektives Kategorienverständnis der an der Auswertung beteiligten Personen. Für die anschließende Codierung der Valenzen im Zuge der evaluativen qualitativen Inhaltsanalyse wurde nach der Bearbeitung von etwa 25 % des Materials ebenfalls die prozentuale Übereinstimmung der Codierungen berechnet. Die Berechnung über MAXQDA ergibt einen Wert von 89,72 % und ist damit als sehr hoch einzustufen. Die weitere Codierung erfolgte daher durch eine Person.

6.2 Ergebnisse

Die Analysen sind in den folgenden Abschnitten nach Forschungsfragen gegliedert. Die untenstehende Tabelle (Tab. 21) zeigt für einen besseren Überblick dazu noch einmal die Forschungsfragen, die bereits in Kapitel 5.1 genannt wurden, sowie die zur Beantwortung herangezogene Datengrundlage und Auswertungsmethode.

Tabelle 21: Überblick zur Struktur des Ergebnistells anhand der Kapitelnummerierung, der dort adressierten Forschungsfragen und der dafür verwendeten Daten bzw. Auswertungsmethoden

Kapitel	Forschungsfrage	Datengrundlage und Auswertungsmethode
	1. Welche Kriterien werden zur Bewertung instruktionaler Erklärungen herangezogen?	
6.2.1.1	1.1 Welche Kriterien werden statusgruppenübergreifend zur Bewertung instruktionaler Erklärungen herangezogen?	Begründungen zur Notenvergabe aus dem offenen Textfeld Strukturierende QIA Über alle Statusgruppen Über alle Vignetten
6.2.1.3	1.2 Welche Kriterien werden von verschiedenen Statusgruppen (Schüler:innen, Studierende, Lehrkräfte, Didaktiker:innen) zur Bewertung instruktionaler Erklärungen herangezogen?	Begründungen zur Notenvergabe aus dem offenen Textfeld Inhaltlich strukturierende QIA Getrennt nach Statusgruppen Über alle Vignetten
6.2.2.1	2. Wie unterscheiden sich die Bewertungen instruktionaler Erklärungen durch verschiedene Statusgruppen (Schüler:innen, Studierende, Lehrkräfte, Didaktiker:innen) in Hinblick auf vorgegebene Kriterien?	Bewertungen der geschlossenen Items Einfaktorielle ANOVA Getrennt nach Statusgruppen Getrennt nach Vignetten
	3. Welche Aspekte sind für die globale Bewertung von instruktionalen Erklärungen große Einflussfaktoren?	
6.2.3.1	3.1 Welche Aspekte sind statusgruppenübergreifend für die globale Bewertung von instruktionalen Erklärungen große Einflussfaktoren?	Bewertungen der geschlossenen Items Hierarchisches Regressionsmodell Über alle Statusgruppen Über alle Vignetten
6.2.3.2	3.2 Welche Aspekte sind aus der Perspektive der Schüler:innen im Unterschied zur Perspektive der Lehrenden für die globale Bewertung von instruktionalen Erklärungen große Einflussfaktoren?	Bewertungen der geschlossenen Items Hierarchisches Regressionsmodell Getrennt Statusgruppen Über alle Vignetten
	4. Wie unterscheiden sich die globalen Bewertungen und deren Begründungen zwischen linearen und komplexen instruktionalen Erklärungen in Abhängigkeit von der Statusgruppe?	
6.2.4.1	4.1 Wie unterscheiden sich die globalen Bewertungen zwischen linearen und komplexen instruktionalen Erklärungen in Abhängigkeit von der Statusgruppe?	Bewertungen der geschlossenen Items; Mixed ANOVA Getrennt nach Statusgruppen Getrennt nach Vignetten
6.2.4.3	4.2 Wie unterscheiden sich die Begründungen der globalen Bewertungen zwischen linearen und komplexen instruktionalen Erklärungen in Abhängigkeit von der Statusgruppe?	Begründungen zur Notenvergabe aus dem offenen Textfeld Evaluative QIA Getrennt nach Statusgruppen Getrennt nach Vignetten

Bem.: QIA: Qualitative Inhaltsanalyse.

Ein Überblick zu den berichteten Ergebnissen sowie deren Einordnung finden sich am Ende nach den jeweiligen Unterkapitel (für die Unterkapitel 6.2.3.1 und 6.2.3.2 nicht getrennt, sondern gemeinsam). Die zentralen Ergebnisse werden anschließend in Kapitel 7.1 zusammengefasst und diskutiert.

6.2.1 Ergebnisse zur Nutzung von Kriterien für die Bewertung instruktionaler Erklärungen

Die Analyse der Begründungen aus dem offenen Textfeld wird im Folgenden dazu genutzt, die in der Literatur diskutierten Kriterien guter instruktionaler Erklärungen zu validieren und gegebenenfalls zu erweitern oder einzugrenzen. Während die statusgruppenübergreifende Betrachtung einen Überblick zu Kriterien gibt (Kap. 6.2.1.1), vertieft und differenziert die statusgruppenspezifische Betrachtung die Nutzung von Kriterien (Kap. 6.2.1.1). Es resultieren folglich nicht nur Erkenntnisse über gute instruktionaler Erklärungen, sondern auch Erkenntnisse über die Statusgruppen bzw. deren Perspektiven.

Berücksichtigt werden muss, dass die Äußerungen zum einen Einblicke geben, welche Kriterien den Teilnehmer:innen bekannt sind, und zum anderen zeigen, welche Kriterien für die Bewertung der gezeigten Videovignetten wichtig sind. Darüber hinaus ist anzunehmen, dass es einen moderierenden Effekt gibt: Je auffälliger die Umsetzung eines bestimmten Aspekts ist, desto mehr rückt dieser in den Vordergrund. Die Umsetzung kann dabei sowohl besonders gut als auch besonders schlecht gelungen sein, wobei eine störende oder für das Verstehen hinderliche Umsetzung mitunter auffälliger als eine gefällige oder für das Verstehen nicht hinderliche Umsetzung ist. Entsprechend häufiger wird auf das dazugehörige Kriterium referiert, da Teilnehmende in der begrenzten Zeit die ihnen am auffälligsten erscheinenden Kriterien nennen. Ist beispielsweise der logische Aufbau durch auffällig viele Brüche gekennzeichnet, wird der unlogische Aufbau in den Äußerungen häufig genannt werden. Pabst-Weinschenk (2011) erläutert in diesem Kontext in Bezug auf sprachliche Aspekte, dass erst ein gehäuftes Auftreten von ‚Ähs‘ die Aufmerksamkeit negativ beeinflusst, gelegentliche ‚Ähs‘ haben hingegen meist keinen Einfluss auf die Aufmerksamkeit und auf das Hörverstehen. Damit spiegeln die Äußerungen auch Merkmale der Videovignetten wider.

6.2.1.1 Ergebnisse zur statusgruppenübergreifenden Nutzung von Kriterien für die Bewertung instruktionaler Erklärungen

In der folgenden Abbildung (Abb. 16) sind die Hauptkategorien mit ihren jeweiligen Subkategorien abgebildet. Obwohl die Fragestellung (*Welche* Kriterien werden statusgruppenübergreifend zur Bewertung instruktionaler Erklärungen herangezogen?) nicht zwingend auf die Häufigkeit der Nennung einzelner Kriterien ausgerichtet ist, werden die relativen Häufigkeiten in der Abbildung aufgeführt. Dies geschieht, da nur so die Aussage des Vorkommens angemessen beantwortet bzw. gegebenenfalls relativiert werden kann. Das würde beispielsweise notwendig werden, wenn eine Subkategorie nur wenige Äußerungen enthält. Die folgende Abbildung (Abb. 16) enthält in der letzten Spalte die kumulierten relativen Häufigkeiten über die Subkategorien pro Hauptkategorie. Eine vollständige Tabelle mit der Ausgabe der absoluten und relativen Häufigkeiten je Kategorie findet sich in Anhang G.

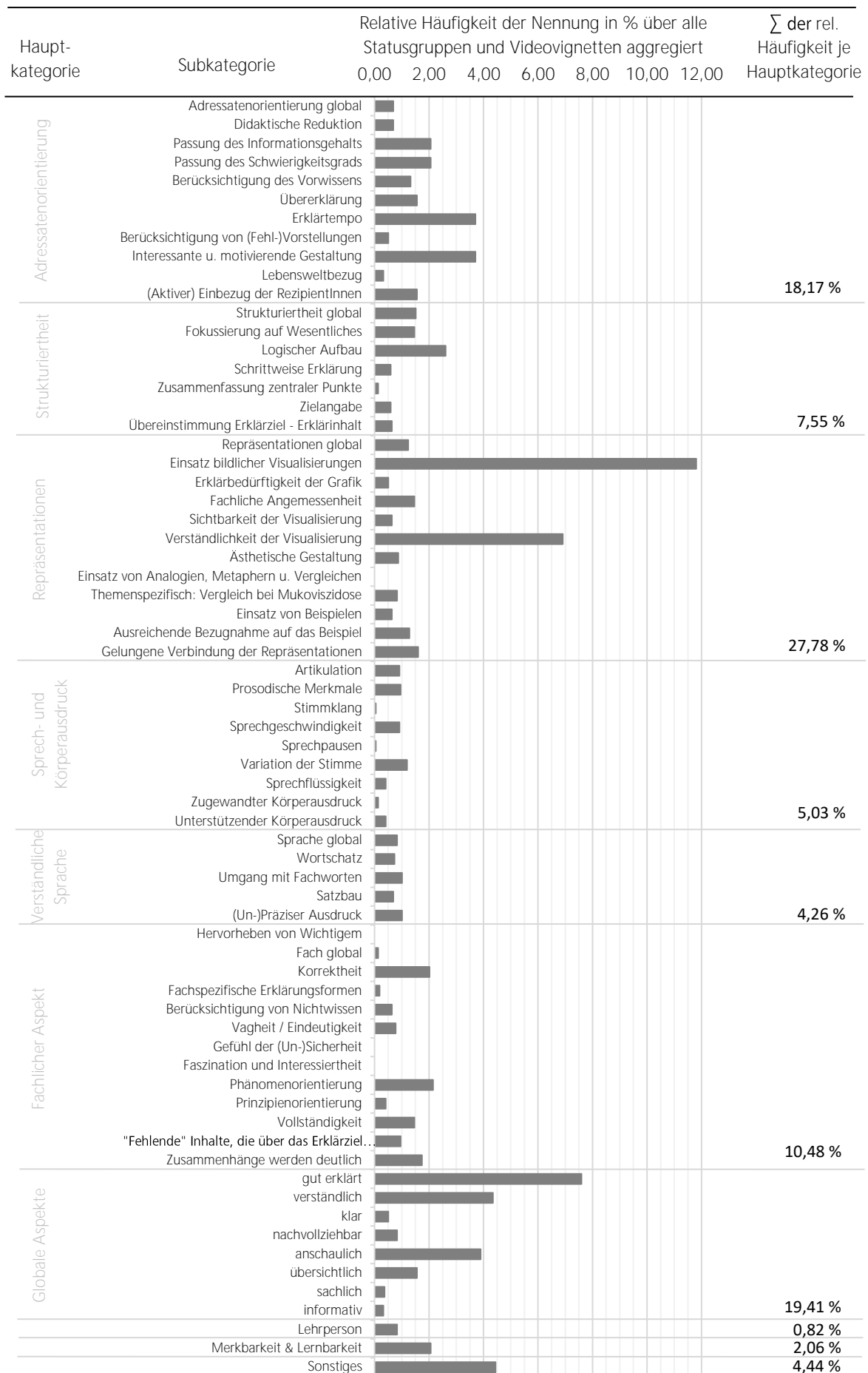


Abbildung 16: Relative Häufigkeiten der Nennungen aller Subkategorien über alle Statusgruppen ($N_{max} = 238$) und Videovignetten ($n = 6$) hinweg

Aus der Abbildung wird ersichtlich, dass die Begründungen zur Notenvergabe aus dem offenen Textfeld aller Teilnehmer:innen in acht thematische Hauptkategorien strukturiert werden können: die Adressatenorientierung, die Strukturiertheit, Repräsentationen, den Sprech- und Körperausdruck, eine verständliche Sprache, den fachlichen Aspekt mit Bezug zum biologischen Gegenstand sowie Aussagen zur Lehrperson und Aussagen zur Merkbarkeit und Lernbarkeit des Erklärten. Hinzu kommen Äußerungen zu globalen Aspekten, die keinem der einzelnen Aspekte zugeordnet werden können. Vergleicht man dieses Ergebnis mit den in der Literatursynopse rekonstruierten Aspekten guten Erklärens (siehe Tab. 4 in Kap. 4.7), lässt sich feststellen, dass die Aspekte Adressatenorientierung, Strukturiertheit, Nutzung von externen Repräsentationen, verständliche Sprache, Sprech- und Körperausdruck sowie fachliche Aspekte auch im Material zu finden sind. Damit können die genannten Aspekte validiert und folglich für die Beschreibung guter Erklärungen im Biologieunterricht herangezogen werden.

In den Äußerungen der Teilnehmenden kommen darüber hinaus vereinzelt Äußerungen zur Lehrperson vor (18 Äußerungen; 0,82% aller Äußerungen). Die Kategorie umfasst verschiedene Merkmale der Lehrperson wie Humor, wahrgenommene Kompetenz, Autorität, sicheres Auftreten sowie den Unterrichtsstil (die Kategorie wurde nicht weiter in Subkategorien unterteilt, da die Anzahl zu gering ist). Folgende Beispiele können die Kategorie illustrieren:

Humor: *„Zu wenig Humor“* (Schüler:in/Ö)

Kompetenz: *„fachlich informiert“* (Didaktikerin /G); *„Die Lehrerin hat die Thematik kompetent mithilfe bildlicher Darstellung erklärt“* (Schüler:in/G_k)

Autorität: *„Außerdem erweckt die Lehrerin nicht den Eindruck, eine hohe Autorität zu besitzen, was ihr vor allem in älteren Klassen sehr zum Verhängnis werden kann“* (Schüler:in/G_k)

Sicheres Auftreten: *„jedoch kam sie beim reden nicht wirklich sicher rüber“* (Schüler:in/Ö)

Stil: *„Ich mag den Unterrichtsstil nicht“* (Schüler:in/Ö_k)

In der rezipierten Literatur wurde von den genannten Charakteristika der Lehrperson der Humor als Merkmal von Lehrkräften, die qualitätsvolle Erklärungen geben, durch empirischen Daten von Geelan (2013) sowie individuelle Erklärstile und ästhetische Präferenzen von Lehrpersonen als Einflussfaktor auf die Gestaltung von Erklärungen durch Treagust und Harrison (2000) beschrieben. Insgesamt ist die Kategorie facettenreich und wird nicht auffällig häufig angesprochen.

Als weiteres neues Kriterium, konnten die Merkbarkeit und Lernbarkeit des Gesagten in den Äußerungen der Teilnehmenden gefunden werden (45 Äußerungen, 2,06 % aller Äußerungen). Die folgenden Zitate aus dem Material können die Kategorie verdeutlichen:

„kann man gut ins Heft übertragen und dann zuhause noch nachvollziehen warum das so ist“ (Schüler:in/Ö_k)

„wieder wenig Text für die Schüler zum Mitschreiben. Wie sollen sich die Schüler das ganze mündlich Gesagte alles bis zur nächsten Stunde merken?“ (Schüler:in/P)

„Lernen von abgeschriebenem Tafelbild zuhause schwer durchzuführen“ (Schüler:in/G_k)

„nachhaltig einprägsam“ (Student:in/Ö)

Das Kriterium ‚Merkbarkeit und Lernbarkeit‘ bezieht sich nicht auf eine einzelne Facette der Erklärung, sondern ist eher (wie die Verständlichkeit oder die Klarheit) auf die insgesamt Gestaltung zu beziehen. Das Kriterium ‚Merkbarkeit und Lernbarkeit‘ konnte in den rezipierten Texten nicht als explizit formuliertes Kriterium guten Erklärens identifiziert werden. Abgeleitet werden kann es jedoch aus dem Vermittlungsziel instruktionaler Erklärungen – der Wissensvermittlung bzw. der Initiierung eines Verstehensprozesses oder Fähigkeitserwerbs (siehe Kap. 2.4). Die explizite Äußerung des Kriteriums wurde nicht erwartet, sodass es auch nicht als a-priori Kategorie in das Kategoriensystem aufgenommen wurde. Die Nennung gibt jedoch Anlass, zu reflektieren, warum das Kriterium genannt wird und welche Bedeutung ihm beigemessen werden kann. Einen Anhaltspunkt dafür kann die schulische Lehr-Lern-Kultur sein, die beispielsweise von Grotzer (2012) in Bezug auf den Unterricht in den USA als stark von Outputorientierung und Leistungsprüfungen geprägt beschrieben wird. Dittmer (2010) arbeitet anhand von Interviewdaten mit deutschen Biologielehrkräften heraus, dass das Schulfach Biologie das Images eines Lernfachs hat und **„prüfungsorientierte[s] Lernen, mit [...] einer starken Fokussierung auf deklaratives Wissen“** (Dittmer, 2010, S. 99) charakteristisch ist. Vor dem Hintergrund eines nachfolgenden Leistungsnachweises wird der Aspekt nachvollziehbar und gewinnt für instruktionale Erklärungen – beispielsweise im Gegensatz zu Erklärungen von Spielanleitungen – eine besondere Bedeutung.

Im Material finden sich weiterhin zahlreiche Äußerungen von Teilnehmer:innen, die sich keinem einzelnen Aspekt bzw. keiner einzelnen Kategorie guter Erklärungen zuordnen lassen, sondern stets auf verschiedene Merkmale zurückzuführen sind. So geht die Äußerung ‚klar‘ beispielsweise auf die Klarheit im sprachlichen Ausdruck oder möglicherweise auch die Klarheit im Aufbau der Erklärung, also deren Struktur, zurück. Äußerungen dieser Art sind in der Hauptkategorie ‚Globale Aspekte‘ zusammengefasst. Die Äußerungen in dieser Hauptkategorie machen insgesamt 19,41 % aller codierten Segmente aus (424 von 2185 Äußerungen). Um die Güte einer Erklärung zu beschreiben, nutzen die Teilnehmer:innen – geordnet nach deren Häufigkeit im Material im Vergleich zu allen weiteren Subkategorien – folgende Adjektive: verständlich (4,35 %), anschaulich (3,89 %), übersichtlich (1,56 %), nachvollziehbar (0,82 %), klar (0,50 %), sachlich (0,37 %) und informativ (0,32 %).

Betrachtet man die Abbildung mit Fokus auf die Subkategorien, so kann zunächst festgehalten werden, dass die Äußerungen der Teilnehmenden zur Begründung der Bewertung in insgesamt 60 Subkategorien eingeordnet werden können (zusätzlich zu den Kategorien ‚Lehrperson‘ und ‚Merkbarkeit und Lernbarkeit‘).¹³⁸ Eine spezifischere Betrachtung und ein Abgleich mit den in der Literatur diskutierten Kriterien erfolgen nun pro Hauptkategorie bzw. Aspekt. Dazu wird die Tabelle aus Kapitel 4.7 mit den literaturbasiert erarbeiteten Kriterien erweitert: Zum einen erfolgt dies durch eine Spalte, in der die entsprechende Subkategorie eingetragen ist, sodass ein direkter Vergleich zwischen den Ergebnissen der Literaturzusammenschau und den Ergebnissen der strukturellen qualitativen Inhaltsanalyse, also den im Material gefundenen Facetten guten Erklärens, durchgeführt werden kann. Zum anderen werden ergänzend dazu auch die im quantitativen Zugang der empirischen Studie eingesetzten Items aufgeführt. Den Tabellen schließen sich jeweils genauere Betrachtungen von Auffälligkeiten oder Abweichungen von den aufgrund der Literaturrecherche erwarteten Facetten an.

¹³⁸ Die a-priori aufgenommene Kategorie zum Hervorheben von Wichtigem (sprachliche Verständlichkeit) und zum Gefühl der Unsicherheit von Schüler:innen sowie zur Faszination und Interessiertheit am Fachgegenstand (Fachlicher Aspekt/Wirkung des dargestellten Komplexitätsgrads) wurde von keinem/keiner Teilnehmer:in gewählt. Sie ist demnach nicht in der genannten Anzahl berücksichtigt. Die Nichtnennung wird im Folgenden noch einmal aufgegriffen. Auch die Kategorie ‚Sonstiges‘ sowie die Subkategorie ‚gut erklärt‘ wurden in der Anzahl nicht berücksichtigt.

Adressatenorientierung

Im gesamten Material können 397 von insgesamt 2185 Äußerungen der Adressatenorientierung zugeordnet werden (18,17 % aller Äußerungen). Die nachstehende Tabelle (Tab. 122) zeigt den Vergleich zwischen den in der Literatur und den im Material gefundenen Kriterien.

Tabelle 22: Übersicht zum Aspekt ‚Adressatenorientierung‘ mit den aus der Literatur rekonstruierten Facetten und Kriterien, den Items des standardisierten Zugangs und den entsprechenden Subkategorien aus dem Kategoriensystem

Facette	Literaturbasiertes Kriterium	Item im standardisierten Zugang (mit Itemkürzel)	Subkategorie der Adressatenorientierung
Anpassung durch didaktische Reduktion	Didaktische Reduktion	Die Erklärung war angemessen didaktisch reduziert. (Ad_red)	Didaktische Reduktion
	Anpassung des Informationsgehalts durch sektorale didaktische Reduktion		Passung des Informationsgehalts
	Anpassung des Schwierigkeitsgrads durch strukturelle didaktische Reduktion	Die Erklärung war für Schüler_innen zu schwierig. (Ad_sch)	Passung des Schwierigkeitsgrads
Vorwissen und Vorstellungen	Berücksichtigung des Vorwissens über den Erklärgegenstand	Schüler_innen konnten der Erklärung mit ihrem Wissen gut folgen. (Ad_vwi)	Berücksichtigung des Vorwissens Übererklärung Erklärtempo
	Berücksichtigung der Vorstellungen über den Erklärgegenstand		Berücksichtigung von (Fehl-)Vorstellungen
	Berücksichtigung des Vorwissens über (fachspezifische) Repräsentationen		*
	Berücksichtigung des Vorwissens über Sprache		*
Kognitive Fähigkeiten	Berücksichtigung des Abstraktionsvermögens		
	Berücksichtigung der Kapazität des Arbeitsgedächtnisses		
Affektiv-motivationale Voraussetzungen	Berücksichtigung des Interesses z. B. durch Auswahl des Beispiels oder Lebensweltbezug	Die Inhalte der Erklärung berücksichtigten die Lebenswelt der Schüler_innen. (Ad_lbw)	Interessante und motivierende Gestaltung Lebensweltbezug
	Berücksichtigung der Motivation z. B. durch Auswahl des Beispiels oder Lebensweltbezug		
			(Aktiver) Einbezug der Rezipient:innen

Bem.: die zwei mit * gekennzeichneten Facetten sind in den Hauptkategorien ‚Repräsentationen‘ und ‚Verständliche Sprache‘ enthalten.

Die Teilnehmer:innen beziehen sich bei der Begründung zur Notenvergabe auf zehn Kriterien. Auf die didaktische Reduktion als grundlegende Methode der Anpassung an die Voraussetzungen der Adressat:innen nehmen die Teilnehmenden kaum direkten Bezug (15 Äußerungen, 0,69 % aller Äußerungen). Jedoch werden die Kriterien der Passung des Informationsgehalts (45 Äußerungen, 2,06 % aller Äußerungen) sowie der Passung des Schwierigkeitsgrads (ebenfalls 45 Äußerungen, 2,06 % aller Äußerungen), die die didaktische Reduktion näher beschreiben, relativ häufig als Begründungen angeführt.

Auch das Kriterium der Anpassung an das Vorwissen wird von Teilnehmenden herangezogen, um die Notenvergabe zu begründen (29 Äußerungen, 1,33 % aller Äußerungen). Im Material findet sich weiterhin das Kriterium der „Übererklärung“, das von Kotthoff (2009, S. 122) unter Bezugnahme auf Erickson und Shultz (1982) beschrieben wurde. Das Kriterium ist als Ausprägung der Berücksichtigung des Vorwissens zu verstehen, wobei dieses in der Erklärung unterschätzt wird. Sie stellt somit kein eigenständiges Kriterium dar. Entscheidend für die Zuordnung zur Subkategorie ‚Übererklärung‘ im Gegensatz zur Subkategorie ‚Berücksichtigung des Vorwissens‘ war eine erkennbare emotionale Komponente, die auf ein sich „für dumm verkauft“ (Kotthoff, 2009, S. 122) hindeutet. Folgende Beispiele können die Subkategorie illustrieren:

„Und nur wenige Jahre später, im Jahre 1870, gibt es sehr sehr wenige Hasen. JA DANKE FÜR DAS HERAUSSTELLEN DES OFFENSICHTLICHEN“ (Schüler:in/Ö)

„Da muss nicht so viel erklärt werden. z. B. das Fressfeinde die Anzahl verringern sollte jeder selbst draufkommen“ (Student:in/Ö)

Die meisten Äußerungen innerhalb der Kategorie ‚Adressatenorientierung‘ entfallen auf die Kriterien ‚Erklärtempo‘ (81 Äußerungen, 3,71 % aller Äußerungen) sowie auf ‚Interessante und motivierende Gestaltung‘ (ebenfalls 81 Äußerungen, 3,71 % aller Äußerungen).

Auf die Berücksichtigung von fachlich nicht angemessenen Vorstellungen von Schüler:innen wird kaum Bezug genommen (11 Äußerungen, 0,50 % aller Äußerungen). Vor dem Hintergrund der zahlreichen Literatur zu möglichen Schülervorstellungen, die das fachliche Lernen behindern können (z. B. Hammann & Asshof, 2014; Kattmann, 2015), ist das Ergebnis bemerkenswert, wobei die niedrige Häufigkeit auch aufgrund der fehlenden Referenzpunkte in den Vignetten zustande kommen kann.

Im Vergleich zu den aus der Literatur rekonstruierten Kriterien guten Erklärens ist auffällig, dass sich keine Äußerung direkt auf die kognitiven Fähigkeiten der Adressat:innen (Abstraktionsvermögen, Kapazität des Arbeitsgedächtnisses) bezieht. Dies könnte darin begründet liegen, dass die Facette nicht direkt beobachtbar ist und für ein psychologisches Konstrukt steht. In der Bewertung muss dadurch auf mehrere beobachtbare Merkmale und mitunter anderen Aspekten zugeordnete Kriterien zurückgegriffen werden. Für das latente Konstrukt ‚Arbeitszeitgedächtnis‘ könnten das beispielsweise der Informationsgehalt (dem Aspekt der Adressatenorientierung und der Facette ‚Anpassung durch didaktische Reduktion‘ zugeordnet) oder das Erklärtempo (dem Aspekt der Adressatenorientierung und der Facette ‚Vorwissen und Vorstellungen‘ zugeordnet) sein. Die Berücksichtigung der Kapazität des Arbeitszeitgedächtnisses ist für die Studienteilnehmer:innen möglicherweise deshalb als konkretes Kriterium nicht verfügbar.

Auf den Lebensweltbezug der Erklärung als Möglichkeit, die Interessen der Rezipient:innen zu berücksichtigen, wird unerwartet selten – lediglich siebenmal (0,32 % aller Äußerungen) – rekurriert, um die Erklärung zu beurteilen. Dieses Resultat ist interessant, da insbesondere in praxisnahen Literaturquellen zu gutem Erklären der Lebensweltbezug als Kriterium angeführt wird (Pauli, 2015; Schopf & Zwischenbrugger, 2015; Wellenreuther, 2013).

Auffällig ist weiterhin, dass neben der Berücksichtigung des Vorwissens in den Erklärsituationen ein aktiver Einbezug der Schüler:innen für die Begründung der Notenvergabe und damit als Kriterium für die Bewertung der Erklärung herangezogen wird – insgesamt 34 Äußerungen, was 1,56 % aller

Äußerungen entspricht. Folgende Beispiele geben einen Einblick in die entsprechende Subkategorie ‚Aktiver Einbezug der Rezipient:innen‘:

„Allerdings finde ich dass dieses Thema ein gutes Thema gewesen wäre wo sich die Schüler die Gründe/ Ursachen mehr selber erarbeiten“ (Schüler:in/Ö_k)

„durch Fragen denkt man mit, die zwischendurch gestellt wurden“ (Schüler:in/Ö_i)

„das Thema eignet sich eher für ein Unterrichtsgespräch als für einen Lehrervortrag“ (Lehrer:in/Ö_i)

„Würde das fragend-entwickelnd erarbeiten“ (Didaktiker:in/Ö_k)

Die Äußerungen verweisen auf die in Kapitel 1 aufgegriffene Debatte um die Angemessenheit instruktionaler Erklärungen (u.a. Brüning & Saum, 2019; Hattie, 2015; Helmke, 2007). Sie sind dabei nicht direkt als Kriterien für instruktionaler Erklärungen im Biologieunterricht zu verstehen, sondern stellen die Angemessenheit der instruktionalen Erklärung des konkreten Gegenstands in dieser Art und Weise (also als Instruktion) infrage. Interessant ist dabei der Hinweis einer Schülerin/ eines Schülers, dass der aktive Einbezug auch durch die rhetorischen Fragen der Lehrkraft gelingen kann.

Zusammenfassend lässt sich festhalten, dass die Anpassung an das Vorwissen und Vorstellungen der Lernenden differenziert in die Bewertung von instruktionalen Erklärungen eingeht. Maßgeblich werden die Kriterien zur Passung des Informationsgehalts und des Schwierigkeitsgrads, der Berücksichtigung des Vorwissens und des Erklärtempos genannt, seltener wird allgemein auf das Kriterium der didaktischen Reduktion verwiesen. Die Anpassung an kognitive Fähigkeiten wird von den Teilnehmer:innen nicht direkt angesprochen. Die affektiv-motivationalen Voraussetzungen der Rezipient:innen werden bei der Bewertung von Erklärungen ebenfalls mit einbezogen, wobei vorrangig die interessante und motivierende Gestaltung angesprochen und nur selten der Lebensweltbezug als Variante dieser Anpassung explizit als Kriterium herangezogen wird.¹³⁹

¹³⁹ Die Häufigkeit der Kategorien, die sich auf die Anpassung an die affektiv-motivationalen Voraussetzungen beziehen, beträgt insgesamt 4,03 % aller Äußerungen im Material.

Strukturiertheit

Der Hauptkategorie ‚Strukturiertheit‘ können 165 von insgesamt 2185 codierten Segmenten zugeordnet werden (7,55 % aller Äußerungen). Die folgende Tabelle (Tab. 23) gibt auch für den Aspekt der Strukturiertheit einen Überblick zu den in der Literatur und im Material gefundenen Kriterien.

Tabelle 23: Übersicht zum Aspekt ‚Strukturiertheit‘ mit den aus der Literatur rekonstruierten Facetten und Kriterien, den Items des standardisierten Zugangs und den entsprechenden Subkategorien aus dem Kategoriensystem

Facette	Literaturbasiertes Kriterium	Item im standardisierten Zugang (mit Itemkürzel)	Subkategorien der Strukturiertheit
Inhaltliche Fokussierung	Fokussierung auf Wesentliches	Die Lehrerin hat zu viele Einzelheiten erklärt. (St_ein)	Fokussierung auf Wesentliches
	Klare Herausarbeitung von Verstehenselementen		
	Fokussierung auf Konzepte und Prinzipien		*
Gliederung und Aufbau	Logische Anordnung der Verstehenselemente und Teilerklärungen; Herstellung eines roten Fadens	Die Erklärung hatte einen roten Faden. (St_rot)	Logischer Aufbau
	Schrittweise Einführung der Verstehenselemente	Die Lehrerin hat die Erklärung Schritt für Schritt aufgebaut. (St_bau)	Schrittweise Erklärung
	Zusammenfassung des Inhalts bei längeren Erklärungen		Zusammenfassung zentraler Punkte
Orientierung und Transparenz	Klare Zielformulierung bzw. Gegenstandsbestimmung	Es ist klar, was genau die Lehrerin erklären wollte. (St_was)	Zielangabe Übereinstimmung Erklärziel – Erklärinhalt
	Offenlegung der Struktur		
	Einbettung in einen größeren Zusammenhang		

Bem.: die zwei mit * gekennzeichneten Facetten sind in der Hauptkategorie ‚Fachlicher Aspekt‘ verortet.

Insgesamt zeigt die Übersicht, dass sich die für den standardisierten Zugang gewählten und operationalisierten Kriterien der Strukturiertheit auch in den Äußerungen der Teilnehmenden finden lassen.

In der Literatur wird darüber hinaus auch die klare Herausarbeitung der Verstehenselemente als Kriterium guten Erklärens aufgegriffen (Brown, 2006; Hargie, 2013; Wagner & Wörn, 2011; Wragg, 1993). Diese kann am Material nicht rekonstruiert werden, mitunter, weil das Konzept der Verstehenselemente als solche nicht bekannt ist. Weiterhin zeigt sich die Umsetzung dessen, dass die erklärende Person relevante Verstehenselemente identifiziert und abgegrenzt hat, auch in einer schrittweisen Einführung der Verstehenselemente (Facette ‚Gliederung und Aufbau‘) – bzw. nur dann, wenn die Verstehenselemente klar abgegrenzt sind, können sie auch nacheinander eingeführt werden.

Darüber hinaus werden in der Literatur die Offenlegung der Struktur (Brown, 2006; Sevan & Gonsalves, 2008) sowie die Einbettung in einen größeren Zusammenhang diskutiert (Schopf &

Zwischenbrugger, 2015; Sevia & Gonsalves, 2008). Auf beide Kriterien nehmen die Teilnehmenden in ihren Begründungen keinen Bezug. Dabei ist zu berücksichtigen, dass für kurze instruktionale Erklärungen mitunter keine explizite Offenlegung der Struktur (z. B. in Form eines ‚advanced organizers‘) notwendig wird, um den Zuhörer:innen Orientierung zu geben, da die zu überblickende Einheit relativ kurz ist.

Dass die Einbettung in einen größeren Zusammenhang nicht in den Äußerungen zu finden ist, könnte ebenso am Studiendesign sowie darüber hinaus auch an den gewählten Erklärgegenständen liegen: Vor dem Zeigen der Erklärungen werden den Teilnehmenden Informationen über das vorhergehende Unterrichtsgeschehen gegeben und dabei wird der bereits erarbeitete Inhalt im Unterricht aufgegriffen. Die nachfolgenden Erklärungen ergeben sich aus dem Erarbeiteten, sodass eine Einbettung in das unmittelbare Unterrichtsgeschehen und das Thema gegeben ist. Insgesamt sind die Erklärgegenstände so gewählt, dass sie als abgeschlossene Einheiten gesehen werden können, die mitunter keine Einbettung in größere Zusammenhänge benötigen.

Ergänzend zur Angabe der Gliederung und des Aufbaus (13 Äußerungen, 0,59 % aller Äußerungen) nehmen die Teilnehmenden darüber hinaus vereinzelt auch auf die Übereinstimmung des zu Beginn der Erklärung genannten Ziels und des daraufhin folgenden Erklärinhalts Bezug (14 Äußerungen, 0,64 % aller Äußerungen). Folgende Beispiele geben Einblick in die konkreten Formulierungen der Teilnehmenden:

„es wurde nicht erklärt, was genau zu dem Anstieg und Abstieg der Anzahl der Kröten im Diagramm führte“ (Schüler:in/Ö_k)

„Aber: Irgendwie war das zu viel versprochen, weil letztlich eben nicht erklärt wurde wie MS entsteht“ (Didaktiker:in/G_k)

Dieses Kriterium differenziert die Zielangabe, die lediglich auf das Vorhandensein einer solchen verweist. In Zusammenhang mit der implementierten UV ist interessant, dass sich Äußerungen zur Übereinstimmung des Erklärziels mit dem Erklärinhalt bis auf eine Ausnahme ausschließlich auf die komplexen Erklärvarianten beziehen (siehe auch Kap. 6.2.4.3).

Das Kriterium ‚Zusammenfassung zentraler Punkte‘ wird lediglich dreimal im gesamten Material erwähnt. Es kann davon ausgegangen werden, dass unter anderem aufgrund der Kürze der Erklärungen eine Zusammenfassung nicht notwendig wird (siehe auch Kap 5.3.4.2).

Resümiert werden kann, dass sich der Aspekt der Strukturiertheit für die Teilnehmenden maßgeblich aus fünf differenzierbaren Kriterien zusammensetzt (Fokussierung auf Wesentliches, logischer Aufbau, schrittweise Erklärung, Zielangabe und Übereinstimmung von Erklärziel und Erklärinhalt). Dabei eignet sich das Kriterium ‚Klare Herausarbeitung von Verstehenselementen‘ mitunter nicht zur Bewertung von Erklärungen. Die Kriterien ‚Offenlegung der Struktur‘ und ‚Einbettung in einen größeren Zusammenhang‘ werden gar nicht und ‚Zusammenfassung zentraler Punkte‘ wird nur selten genutzt, sodass sie bei kurzen instruktionalen Erklärungen nicht unmittelbar für eine Bewertung der Güte notwendig sind.

Repräsentationen

Auf den Einsatz von Repräsentationen wird mit 607 von insgesamt 2185 Äußerungen (27,78 % aller Äußerungen) am häufigsten Bezug genommen. Da keine Items zu Repräsentationen im standardisierten Zugang implementiert wurden, fehlt die entsprechende Spalte in der folgenden Übersichtstabelle (Tab. 24).

Tabelle 24: Übersicht zum Aspekt ‚Repräsentationen‘ mit den aus der Literatur rekonstruierten Facetten und Kriterien sowie den entsprechenden Subkategorien aus dem Kategoriensystem

Facette	Literaturbasiertes Kriterium	Subkategorien zu Repräsentationen
Auswahl passender externer Repräsentationen	Berücksichtigung der Passung zum Gegenstand	Fachliche Angemessenheit
	Berücksichtigung des Vorwissens der Adressat:innen	
	Explikation des Ziels der Betrachtung	
Veranschaulichung durch reale Objekte	Berücksichtigung der Komplexität der realen Objekte	
	Berücksichtigung von Empfindungen gegenüber den realen Objekten (z. B. Ekel, Angst)	
Nutzung von realistischen und logischen Bildern	Nutzung von Bildern	Einsatz bildlicher Visualisierungen
	Anleitung der Adressat:innen bei der Identifikation relevanter Elemente und der Interpretation	Erklärbedürftigkeit der Grafik
	Zeitgleiche Präsentation von Text und Bild	
	Integration von Text und Bild	
	Einhaltung von Gestaltungsprämissen zur Verständlichkeit von Visualisierungen/ logischen Bildern	Verständlichkeit der Visualisierung
		Ästhetische Gestaltung
	Sichtbarkeit der Visualisierung	Sichtbarkeit der Visualisierung
Nutzung von Analogien, Metaphern und Vergleichen	Nutzung Analogien, Metaphern und Vergleichen	Themenspezifisch: Vergleich bei Mukoviszidose
	Herausarbeitung des Herkunftsbereichs und des Übertragungsbereichs	
	Aufzeigen der Grenzen der Übertragung	
Nutzung von Beispielen	Nutzung von Beispielen	Einsatz von Beispielen
	Auswahl oder Gestaltung der Beispiele nach den Eigenschaften: passend, typisch, deckend, konkret, lebensweltnah, interessant und relevant	
		Ausreichende Bezugnahme auf das Beispiel
Herstellung von Verbindungen zwischen verschiedenen Repräsentationen		Gelungene Verbindung der Repräsentationen

Der Aspekt zum Einsatz von Repräsentationen ist breit angelegt, da unter externen Repräsentationen sowohl natürliche Objekte, realistische und logische Bilder, Analogien, Metaphern und Vergleiche als auch Beispiele gefasst werden können, die jeweils wiederum durch spezifische Kriterien bewertet werden können.

Aus der Übersicht geht hervor, dass die Teilnehmenden die fachliche Angemessenheit als ein für alle externen Repräsentationen geltendes Kriterium bei ihren Bewertungen der Erklärungen einbeziehen (32 Äußerungen, 1,46 % aller Äußerungen). Die weiteren Kriterien unter der Facette der Auswahl passender externer Repräsentationen werden hingegen nicht aufgegriffen.

Die Facette ‚Veranschaulichung durch reale Objekte‘ wird von den Teilnehmenden ebenfalls nicht aufgegriffen. In den Vignetten werden jedoch auch keine realen Objekte eingesetzt, sodass die Annahme nahe liegt, dass das Ausbleiben von Bezügen darauf hierin begründet liegt.

In den drei Facetten ‚Nutzung von realistischen und logischen Bildern‘, ‚Nutzung von Analogien, Metaphern und Vergleichen‘ und ‚Nutzung von Beispielen‘ zeigt sich ein wiederkehrendes Muster: Es wird häufig auf das Vorhandensein der externen Repräsentationen als Begründung für die Bewertung der Erklärung Bezug genommen, bei den Facetten ‚Nutzung von realistischen und logischen Bildern‘ sowie ‚Nutzung von Analogien, Metaphern und Vergleichen‘ sogar häufiger als auf spezifische Gestaltungskriterien.

Detaillierter betrachtet ergibt sich für die Facette ‚Nutzung von realistischen und logischen Bildern‘ folgendes Bild: Die meisten Äußerungen entfallen auf den Einsatz bildlicher Visualisierungen (258 Äußerungen, 11,81 % aller Äußerungen). Die Äußerungen beziehen sich allgemein auf die Anwesenheit bildlicher Visualisierungen – häufig in Kombination mit ihrer veranschaulichenden Wirkung. Folgende Zitate können einen exemplarischen Einblick in die Inhalte der Subkategorie geben:

„Durch das Schaubild war die Erklärung besser verständlich“ (Schüler:in/G_k)

„durch die Bilder kommt man gut mit“ (Student:in/G_i)

„gute Veranschaulichung durch Applikationen“ (Lehrer:in/G_i)

Daneben wird bei der Begründung der Notenvergabe häufig darauf Bezug genommen, ob die Visualisierung verständlich war (151 Äußerungen, 6,91 % aller Äußerungen). Damit werden überwiegend Beschriftungen von Pfeilen und Elementen in der Visualisierung angesprochen:

„dadurch, dass die Pfeile beschriftet sind und es nur Textfelder gibt, kann man sich's besser merken“ (Schüler:in/Ö_k)

„Bei der Visualisierung könnte man noch etwas genauer angeben was z.B die Pfeile bedeuten (Bsp.: Immunsystem -----> Nervenzelle auf den Pfeil schreiben greift eigene NZ an)“ (Student:in/G_k)

„Gen nicht beschriftet“ (Didaktiker:in/G_i)

Diese Äußerungen sind dahingehend interessant, da sie gezielt den Lernaspekt in den Blick nehmen (in Kap. 6.2.1.3 wird untersucht, ob die Bezugnahme abhängig von der Statusgruppe ist). Selten wird hingegen darauf rekuriert, ob die Anleitung der Adressat:innen bei der Identifikation

relevanter Elemente der Grafik gelungen oder nicht gelungen ist, das bedeutet, inwiefern die Grafik erklärbedürftig war oder ist (11 Äußerungen, 0,50 % aller Äußerungen).

Als ein weiteres Merkmal bildlicher Visualisierungen, das die Teilnehmenden als Grund für ihre Bewertung der Erklärung angeben, ist die Ästhetik der eingesetzten Bilder (19 Äußerungen, 0,87 % aller Äußerungen). Diese Äußerungen beziehen sich im Gegensatz zu denen der Subkategorie ‚Verständlichkeit der Visualisierung‘ nicht direkt auf lernrelevante Merkmale, sondern auf eher oberflächlich wahrnehmbare Aspekte der Visualisierungen. So äußern sich Teilnehmende zum Beispiel folgendermaßen:

„weniger laminiertes Papier Zeichnungen an die Tafel mit bunter Kreide“ (Schüler:in/G_k)

„die Visualisierung war noch attraktiver“ (Didaktiker:in/G_k)

Auf die Nutzung von Analogien, Metaphern und Vergleichen als Facette einer guten Erklärung wird von den Teilnehmer:innen kaum eingegangen. Analogien und insbesondere Metaphern sind mitunter im gesprochenen Text nicht bewusst wahrnehmbar, wenn nicht direkt darauf Bezug genommen wird, indem beispielsweise die Grenzen der Übertragung kenntlich gemacht werden. So lassen sich in den sechs Erklärungen sprachliche Bilder finden: beispielsweise bei der Erklärung zur Entstehung eines Folsäuremangels die „gestörte Verdauung“ (dahinterliegende Metapher des Gleichgewichts), bei der Erklärung zur Entstehung von Multipler Sklerose das „Angreifen“ des Immunsystems (dahinterliegende Metapher des Krieges) oder die Analogie der Flimmerhärchen und deren Transportfunktionsweise mit einem Fließband in der Erklärung zur Entstehung von Mukoviszidose. Eine Ausnahme bildet die Erklärung zur Entstehung von Mukoviszidose, da hierbei der Vergleich zwischen gesunden und erkrankten Personen als grundlegende Struktur für die Erklärung genutzt wird und damit stark im Wahrnehmungsfokus der Rezipient:innen bzw. Teilnehmer:innen liegt. Dieser wird sodann in den Bewertungen aufgenommen (18 Äußerungen; 0,82 % aller Äußerungen).

Innerhalb der Facette zum Einsatz von Beispielen verweisen die Teilnehmenden zwar auch auf die bloße Nutzung von Beispielen (14 Äußerungen; 0,64 % aller Äußerungen), jedoch wird häufiger in den Begründungen erwähnt, dass in ausreichendem Maß auf das Beispiel Bezug genommen bzw. dies nicht erfolgt (28 Äußerungen; 1,28 % aller Äußerungen). Damit verhält es sich in dieser Facette anders als bei denen zur ‚Nutzung von realistischen und logischen Bildern‘ sowie zur ‚Nutzung von Analogien, Metaphern und Vergleichen‘. Das Kriterium zum ausreichenden Bezug auf Beispiele konnte in der rezipierten Literatur nicht gefunden werden. Die folgenden Beispiele geben Einblick in die Äußerungen der Teilnehmenden, die die Grundlage für die Subkategorie bzw. das neue Kriterium bilden:

„aber man hätte das Diagramm mehr einbeziehen können“ (Student:in/Ö_i)

„Schwankungen nicht direkt am Beispiel erklärt“ (Didaktiker:in/Ö_k)

„Diagramm kann allerdings nicht vollständig erklärt werden, mir fehlen Informationen zur Situation in Merklingen, um das Diagramm mit dem Lebensraum in Verbindung zubringen“
(Didaktiker:in/ Ö_k)

Alle Äußerungen innerhalb der Subkategorie beziehen sich auf die zwei Erklärungen im Themenbereich der Ökologie. Dabei wurde ein Diagramm zur Populationsdichte als Einstieg in die Erklärung und Ansatzpunkt für die Ableitung der Fragestellung genutzt.

Zuletzt lässt sich im Material auch die theoretisch herausgearbeitete Facette der Verbindung der verschiedenen Repräsentationen innerhalb der Erklärung finden (35 Äußerungen; 1,60 % aller Äußerungen).

Über alle Subkategorien hinweg lässt sich festhalten, dass häufig das Vorhandensein von externen Repräsentationen als Begründung für die Notenvergabe angeführt wird (317 Äußerungen; 14,51 % aller Äußerungen)¹⁴⁰ und etwas seltener deren tatsächliche Gestaltung und die Einbettung in die Erklärung, die für das Lernen relevant sind, einbezogen werden (290 Äußerungen; 13,27 % aller Äußerungen).¹⁴¹ Insgesamt werden die Visualisierungen hauptsächlich anhand der Kriterien der fachlichen Angemessenheit der externen Repräsentationen, des Einsatzes von Bildern, der Verständlichkeit der Visualisierung, der ästhetischen Gestaltung, der Sichtbarkeit der Visualisierung, des Einsatzes von Beispielen, der ausreichenden Bezugnahme auf die Beispiele als neues Kriterium und der gelungenen Verbindung von Repräsentationen bewertet. Selten wird die Erklärbedürftigkeit der Grafik herangezogen.

Sprech- und Körperausdruck

Der Sprech- und Körperausdruck der erklärenden Person wird mit 110 von insgesamt 2185 Äußerungen adressiert (5,03 % aller Äußerungen). Damit entfallen auf den Sprech- und Körperausdruck neben der Verständlichkeit der Sprache (4,26 % aller Äußerungen) die wenigsten Äußerungen. Die geringe Häufigkeit wird durch einen Blick auf zwei theoretische Bezüge nachvollziehbar: Zum einen können erst durch entsprechendes Wissen Merkmale bewusst wahrgenommen sowie schließlich benannt und beurteilt werden. Für den Bereich der Lehrkräftebildung ist ein fachliches Wissen über den Sprech- und Körperausdruck nicht anzunehmen, da die Inhalte bisher in der Regel kaum Gegenstand der Ausbildung sind (Lägel-Gunga et al., im Druck). Berücksichtigt man, dass ein Teil der teilnehmenden Studierenden ein Seminar zum Erklären besuchte, wodurch Wissen innerhalb einer Seminareinheit aufgebaut und Kriterien zum Sprech- und Körperausdrucks thematisiert wurden (siehe Kap. 6.1.1), könnte die relative Häufigkeit dieses Aspekts bei der Beurteilung von instruktionalen Erklärungen in einer anderen Stichprobe sogar noch etwas geringer ausfallen. Zum anderen rücken negative bzw. störende Aspekte im Gegensatz zu positiven Aspekten mitunter verstärkt in den Wahrnehmungsfokus (Pabst-Weinschenk, 2011). Dieser Effekt könnte moderierend auf die Nennung von Facetten des Sprech- und Körperausdrucks wirken, wenn dieser insgesamt als wenig auffällig bzw. negativ erlebt wird. Die folgende Zusammenschau (Tab. 25) verdeutlicht, welche Kriterien in den Begründungen zur Notenvergabe gefunden und welche darüber hinaus in der Literatur diskutiert werden.

¹⁴⁰ Die Anzahl ergibt sich aus folgenden Subkategorien: Repräsentationen global, Einsatz bildlicher Visualisierungen, Einsatz von Analogien, Metaphern und Vergleichen bzw. Themenspezifisch: Vergleich bei Mukoviszidose sowie Einsatz von Beispielen.

¹⁴¹ Die Anzahl ergibt sich aus folgenden Subkategorien: Erklärbedürftigkeit der Grafik, fachliche Angemessenheit, Sichtbarkeit der Visualisierung, Verständlichkeit der Visualisierung, ästhetische Gestaltung der Visualisierung, ausreichende Bezugnahme auf das Beispiel, gelungene Verbindung der Repräsentationen.

Tabelle 25: Übersicht zum Aspekt ‚Sprech- und Körperausdruck‘ mit den aus der Literatur rekonstruierten Facetten und Kriterien, den Items des standardisierten Zugangs und den entsprechenden Subkategorien aus dem Kategoriensystem

Facette	Literaturbasiertes Kriterium	Item im standardisierten Zugang (mit Itemkürzel)	Entsprechende Subkategorie
Artikulation	Deutliche Artikulation	Die Lehrerin hatte eine deutliche Aussprache. (Se_aus)	Artikulation
Prosodie			Prosodische Merkmale
	Gesunder Stimmklang	Die Lehrerin hatte eine angenehm klingende Stimme. (Se_sti)	Stimmklang
	Angemessene Sprechgeschwindigkeit	Die Lehrerin hatte eine angemessene Sprechgeschwindigkeit. (Se_ges)	Sprechgeschwindigkeit
	Bedeutungsstützende Sprechpausen	Die Lehrerin hat die Sprechpausen so gesetzt, dass ich gut folgen konnte. (Se_pau)	Sprechpausen
	Variationsreiche Stimme	Die Lehrerin hatte eine abwechslungsreiche Sprechweise. (Se_abw)	Variation der Stimme
			Sprechflüssigkeit
Kinesik	Blickkontakt mit den Adressat:innen	Die Lehrerin hat sich den Schülerinnen und Schülern zugewendet. (Se_zug)	Zugewandter Körperausdruck
	Einladende Mimik		
	Offene und zugewandte Körperhaltung		
	Gestik und Mimik, die das Gesagte unterstützt	Die Lehrerin hatte einen Körperausdruck, der die Erklärung unterstützt hat. (Se_kor)	Unterstützender Körperausdruck
	Entspannte Körperhaltung		
Proxemie	Angemessene Bewegung im Raum		

Von den im standardisierten Zugang durch Items repräsentierten Facetten nutzen die Teilnehmenden einige Kriterien zur freien Bewertung der Erklärungen, wobei die Sprechpausen sowie der Stimmklang (beides Kriterien der Facette der Prosodie) jeweils nur einmal angesprochen werden. Auch der zugewandte Körperausdruck wird lediglich dreimal zur Begründung angeführt. Daher wird nicht mehr zwischen dem Blickkontakt mit den Adressat:innen, der einladenden Mimik sowie der offenen und zugewandten Körperhaltung differenziert, sondern eine Zusammenfassung zum zugewandten Körperausdruck vorgenommen. Folgende Äußerungen finden sich im Material:

„*evtl mehr Augenkontakt*“ (Lehrer:in/G_{KI})

„*zugewandt*“ (Didaktiker:in/G_I)

Während die Kriterien ‚Sprechpausen‘ und ‚Stimmklang‘ mitunter aufgrund der oben angeführten Ursachen nicht in die Begründung einfließen, ist anzunehmen, dass die Zugewandtheit des Körperausdrucks auch durch die Situation der Videografie nicht in den Wahrnehmungsfokus rückt. Die Ausprägung des Körperausdrucks, die das Gesagte unterstützen kann, wird von den Teilnehmenden insgesamt selten angesprochen (9 Äußerungen; 0,41 % aller Äußerungen). Sie beziehen sich dabei allgemein auf Gestik („*Schlecht gestikuliert*“ (Schüler:in/G_I)) und Mimik („*Mimik*“

und Gestik ist ausgeblieben“ (Student:in/G_i)) der erklärenden Person. Keine Äußerung entfällt auf das Kriterium der Körperhaltung, die bei Sprecher:innen entspannt sein sollte.

Vergleicht man darüber hinaus die am Material gebildeten Subkategorien mit den in der Literatur genannten Kriterien, kann weiterhin festgestellt werden, dass die Teilnehmenden in einigen wenigen Fällen auch auf die Sprechflüssigkeit (9 von 2185 Äußerungen; 0,41 % aller Äußerungen) Bezug nehmen. Sie formulieren beispielsweise:

„[f]lüssig gesprochen“ (Schüler:in/G_k)

„Zusätzlich stocke der Redefluss zwischendurch, wodurch das Zuhören schwerer fiel“
(Student:in/Ö_k)

Auffällig ist, dass innerhalb der Hauptkategorie Sprech- und Körperausdruck am häufigsten auf die Variation der Stimme Bezug genommen wird (26 von insgesamt 2185 Äußerungen; 1,19 % aller Äußerungen). Dabei rekurren die Teilnehmenden sowohl auf die Sprechmelodie (häufig in der negativen Formulierung mit *„Monotone Sprechweise beim Erklären“* (Schüler:in/G_k)) als auch auf die Lautstärke (*„keine Akzente bei der Lautstärke“* (Student:in/G_i)).

Zusammenfassend wird der Sprech- und Körperausdruck von den Teilnehmenden nicht vorrangig global, sondern mit neun unterscheidbaren Subkategorien beschrieben, wobei zwei Subkategorien (Stimmklang und Sprechpausen) mit nur einer Nennung enthalten sind. Am häufigsten werden die Kriterien der Artikulation, der allgemeinen Prosodie, der Sprechgeschwindigkeit und der Variation der Stimme herangezogen. Insgesamt lässt sich für den Aspekt des Sprech- und Körperausdrucks festhalten, dass die Teilnehmenden insbesondere Kriterien äußern, die sich auf die Stimme der erklärenden Person (98 Äußerungen; 4,49 % aller Äußerungen) und hierbei insbesondere auf allgemeine prosodische Merkmale beziehen.¹⁴² Der Körperausdruck mit den Kriterien zugewandter Körperausdruck und unterstützender Körperausdruck wird hingegen kaum als Begründung für die Einschätzung der Erklärung angeführt (12 Äußerungen; 0,55 % aller Äußerungen), die Proxemik überhaupt nicht.¹⁴³

Verständliche Sprache

Auf die Verständlichkeit der Sprache wird in 93 von insgesamt 2185 Äußerungen Bezug genommen (4,26 % aller Äußerungen). Von den im standardisierten Zugang durch Items repräsentierten Kriterien nutzen die Teilnehmenden lediglich die Hervorhebung von Wichtigem nicht zur freien Bewertung der Erklärungen. Dies ist nachvollziehbar, da es in den Erklärungen keine expliziten Verweise auf Wichtiges gibt. Jedoch ist in einer Nichtnennung auch enthalten, dass die Teilnehmenden derartige Verweise in den gezeigten instruktionalen Erklärungen für nicht notwendig halten.¹⁴⁴

¹⁴² Die Anzahl setzt sich aus den folgenden Subkategorien zusammen: Artikulation, prosodische Merkmale, Stimmklang, Sprechgeschwindigkeit, Sprechpausen, Variation der Stimme, Sprechflüssigkeit.

¹⁴³ Die Anzahl setzt sich aus den Subkategorien ‚Zugewandter Körperausdruck‘ und ‚Unterstützender Körperausdruck‘ zusammen.

¹⁴⁴ Im standardisierten Zugang ist das entsprechende Item nach Prüfung der internen Konsistenz der Skala auch aus dieser entfernt worden (siehe Kap. 6.1.2).

Die Gegenüberstellung der am Material gebildeten Subkategorien mit den in der Literatur genannten Kriterien kann der folgenden Tabelle (Tab. 26) entnommen werden. Demnach werden alle weiteren in der Literatur diskutierten Kriterien auch von Teilnehmenden bei der Begründung der Notenvergabe angesprochen, also bei der Bewertung von instruktionalen Erklärungen im Biologieunterricht angelegt.

Tabelle 26: Übersicht zum Aspekt ‚Verständliche Sprache‘ mit den aus der Literatur rekonstruierten Facetten und Kriterien, den Items des standardisierten Zugangs und den entsprechenden Subkategorien aus dem Kategoriensystem

Facette	Literaturbasiertes Kriterium	Item im standardisierten Zugang (mit Itemkürzel)	Entsprechende Subkategorie
Wortebene	Nutzung geläufiger Wörter	Bei manchen Wörtern wissen die Schüler_innen eventuell nicht, was sie bedeuten. (Sw_bed)	Wortschatz
	Aussparung, Paraphrasierung oder Erklärung nicht geläufiger Wörter		Umgang mit Fachworten
Satzebene	Anpassung der Komplexität der Syntax	Manche Sätze hat die Lehrerin zu lang gemacht. (Sw_zul)	Satzbau
	Eindeutige und präzise Formulierung der Aussagen		(Un-)Präziser Ausdruck
Textebene	Herstellung lokaler und globaler Kohärenz (z. B. durch Kohärenzbildungshilfen)	Manchmal hat die Lehrerin gesagt, dass etwas sehr wichtig ist. (Sw_wic)	Hervorheben von Wichtigem ¹

Bem.: ¹ Die Subkategorie enthält keine codierten Segmente, das Kriterium wird entsprechend nicht von den Teilnehmenden angeführt.

Insgesamt befinden sich in allen Subkategorien ähnlich viele bzw. ähnlich wenige codierte Segmente. Die wenigsten Äußerungen entfallen auf den Satzbau (15 Äußerungen; 0,73 % aller Äußerungen), die meisten auf den Umgang mit Fachworten sowie den (un-)präzisen Ausdruck (je 22 Äußerungen; 1,01 % aller Äußerungen). Allgemein auf den Wortschatz entfallen 16 Äußerungen (0,82 % aller Äußerungen). Wie beim Sprech- und Körperausdruck muss bei der sprachlichen Verständlichkeit berücksichtigt werden, dass ein Teil der Studierenden an einem Seminar zum guten Erklären teilnahm und Kriterien zur sprachlichen Verständlichkeit erlernte. In einer Stichprobe, in der keine teilnehmende Person ein Seminar besucht, könnte die Anzahl der Nennung leicht niedriger ausfallen (siehe Kap. 6.1.1).

Da die Verwendung von Fachbegriffen in den Naturwissenschaften vielfach diskutiert wird (Rincke, 2010; Schmiemann, 2011; Kulgemeyer & Tomczyszyn, 2015; Ziepprecht et al., 2017) und die entsprechende Subkategorie ‚Umgang mit Fachworten‘ von den Teilnehmenden innerhalb der Kategorie ‚Verständliche Sprache‘ mit am häufigsten angesprochen wird (22 Äußerungen; 1,01 % aller Äußerungen), soll anhand von Beispielen an dieser Stelle Einblick in die Äußerungen der Teilnehmer:innen gegeben werden:

„und ich fand auch gut dass immer bei Wörtern die manche Schüler eventuell nicht kennen (z. B. Insektizide) auch erklärt wurden“ (Schüler:in/Ök)

„nutzt zunächst einen alltäglichen, dann einen fachlichen Begriff oder paraphrasiert das Gesagte“ (Student:in/G_i)

„viele Fachbegriffe ohne zusätzliche schriftl. Fixierung (Bsp. Micelle)“ (Lehrer:in/P)

„positiv: Erklärung von wichtigen Begriffen (Verdauung)“ (Didaktiker:in/P)

Zusammenfassend kann festgehalten werden, dass die Verständlichkeit der Sprache zwar selten in den Begründungen zur Notenvergabe angeführt wird. Die Äußerungen sind jedoch differenziert, sodass mit Ausnahme des Hervorhebens von Wichtigem die in der Literatur diskutierten Kriterien zur Verständlichkeit der Sprache (bezogen auf den Wortschatz, den Umgang mit Fachworten und den Satzbau, den (un-)präzisen Ausdruck) von den Teilnehmenden an instruktionalen Erklärungen angelegt werden.

Fachlicher Aspekt

229 von insgesamt 2185 Äußerungen lassen sich dem fachlichen Aspekt guten Erklärens zuordnen (10,48 % aller Äußerungen). Die nachstehende Tabelle (Tab. 27) gibt einen Überblick zum Vergleich der im Material und der in der Literatur gefundenen Kriterien.

Die Teilnehmenden beziehen sich bei der Bewertung der instruktionalen Erklärungen nicht auf alle Kriterien, die auf Basis der Literatur rekonstruiert wurden. So nehmen sie keinen Bezug auf ein entstehendes Gefühl der Unsicherheit oder Sicherheit (Kategorie ‚Gefühl der (Un-)Sicherheit‘), das insbesondere bei Erklärungen der komplexen Variante durch das Aufgreifen von Nichtwissen und unsicherem Wissen sowie von Multikausalität entstehen kann (Kampourakis & McCain, 2020; Mitchell, 2008). Darüber hinaus lassen sich auch keine Äußerungen finden, die auf die Interessiertheit oder Faszination gegenüber dem fachlichen Gegenstand hinweisen, die durch die instruktionalen Erklärungen hervorgerufen wurden. Betrachtet man die weiteren Kriterien, so wird deutlich, dass über diese zwei Ausnahmen hinaus alle weiteren literaturbasiert erarbeiteten Facetten auch im Material gefunden werden.

Innerhalb der Facette der fachlichen Korrektheit äußern sich die Teilnehmenden neben dem Kriterium der fachlichen Korrektheit von Fakten und Zusammenhängen (44 Äußerungen; 2,01 % aller Äußerungen) auch zur Wahl der fachspezifischen Erklärungsform. Die Häufigkeit der Äußerungen zur Erklärungsform ist gering (4 Äußerungen; 0,18 % aller Äußerungen), jedoch konform mit den Annahmen, da eine Thematisierung der wissenschaftstheoretischen Grundlagen in der Regel nicht Teil der universitären Lehramtsausbildung ist.

Tabelle 27: Übersicht zum fachlichen Aspekt mit den aus der Literatur rekonstruierten Facetten und Kriterien, den Items des standardisierten Zugangs und den entsprechenden Subkategorien aus dem Kategoriensystem

Facette	Literaturbasiertes Kriterium	Item im standardisierten Zugang (mit Itemkürzel)	Entsprechende Subkategorie
Fachliche Korrektheit	Richtigkeit der Fakten, Zusammenhänge und Prinzipien		Korrektheit
	Wahl der adäquaten fachlichen Erklärungsform		Fachspezifische Erklärungsformen
	Angemessene Darstellung des aktuellen Wissensstandes		Berücksichtigung von Nichtwissen
			Vagheit/Eindeutigkeit
Berücksichtigung der Auswirkung des Gesagten	Gefühl der (Un-)Sicherheit	Die Tatsache, dass nicht alle Dinge bekannt sind, verunsichert Schüler_innen. (Fa_ver*)	Gefühl der (Un-)Sicherheit ¹
	Faszination und Interessiertheit	Die Erklärung ist für Schüler_innen faszinierend. (Fa_fas) Nach dieser Erklärung möchten Schüler_innen noch mehr darüber erfahren. (Fa_eref)	Faszination und Interessiertheit ¹
Auswahl repräsentativer und geeigneter Phänomene (Phänomen- vs. Prinzipienorientierung)	Orientierung am Phänomen und Darstellung dessen Komplexität	Die Anzahl der Zusammenhänge ist für Schüler_innen angemessen. (Fa_zan)	Phänomenorientierung
	Orientierung an allgemeinen, übertragbaren Prinzipien der Biologie	Für Schüler_innen zeigt die Erklärung zu viele Ursachen. (Fa_zvu) Die Erklärung zeigt, wie es in der Natur wirklich ist. (Fa_nat) Die Erklärung zeigt nur einen Ausschnitt auf die Frage („Thema der Erklärung“). (Fa_aus) Der Inhalt der Erklärung hilft Schüler_innen, andere Beispiele zu verstehen. (Fa_hil)	Prinzipienorientierung
Fachliche Vollständigkeit	Berücksichtigung aller notwendiger Verstehenselemente		Vollständigkeit
			„Fehlende“ Inhalte, die über das Erklärziel hinausgehen Zusammenhänge werden deutlich

Bem.: ¹ Die Subkategorien enthalten keine codierten Segmente.

Weiterhin berücksichtigen die Teilnehmenden bei ihrer Notenvergabe zum Teil, wie bzw. ob Nichtwissen in der Erklärung thematisiert wird (14 Äußerungen; 0,64 % aller Äußerungen). Auffällig ist in diesem Kontext, dass sie darüber hinaus auch eine Bewertung vornehmen, inwiefern die Erklärung für sie vage oder eindeutig erscheint (17 Äußerungen; 0,78 % aller Äußerungen). Dabei wird meist auf die durch die Berücksichtigung von Nichtwissen und Multikausalität entstehende Vagheit Bezug genommen. Lediglich zweimal wird die Eindeutigkeit – aufgefasst als das Gegenteil

von Vagheit – der Erklärung als Begründung für die Bewertung geäußert. Folgende Beispiele können die Bewertung der Erklärung durch die Subkategorie verdeutlichen:

„kein klares ergebnis“ (Schüler:in/Ö_k)

„klares Ergebnis“ (Schüler:in/P_i)

„Gesamteindruck der Erklärung ist ja etwas vage“ (Didaktiker:in/G_k)

„Irgendwie multifaktoriell / weiß man nicht genau“ (Didaktiker:in/G_k)

Diese Subkategorie als Kriterium zu formulieren, bedarf neben einer didaktischen auch einer fachlich orientierten Reflexion. Während aus didaktischer bzw. lernorientierten Perspektive die Entstehung von Vagheit nicht Ziel einer Erklärung ist, ist die Vermeidung von Vagheit aus fachlicher Perspektive mitunter nicht möglich, da sie dem Erklärgegenstand – im weitesten Sinne komplexe biologische Phänomene – und dessen Eigenschaften (Multikausalität, Nichtwissen, unsicheres Wissen) inhärent ist. Die Äußerungen leiten daher eher zu Fragen, wie mit der Vagheit bzw. der fehlenden Eindeutigkeit im Kontext von instruktionalen Erklärungen umzugehen ist. Das Kriterium bzw. die Äußerungen hierzu werden im Kapitel 6.2.1.3 bei der statusgruppenspezifischen Analyse näher betrachtet, um Gründe für die Vagheit zu eruieren.

Innerhalb der Facette zur Auswahl repräsentativer und geeigneter Phänomene beziehen sich die Teilnehmenden auch auf die auf Literaturbasis abgeleiteten Kriterien der Phänomenorientierung (47 Äußerungen; 2,15 % der Äußerungen) und seltener der Prinzipienorientierung (9 Äußerungen; 0,41% aller Äußerungen). Unter der Phänomenorientierung sind dabei Äußerungen subsummiert, die daraufhin deuten, dass das natürliche Phänomen mit seiner Komplexität gut bzw. nicht gut abgebildet wurde. Folgende Beispiele geben Einblick in die Einschätzungen der Teilnehmenden:

„Es fehlt ob sich die Faktoren gegenseitig beeinflussen“ (Student:in/G_k)

„Zusammenspiel Faktoren wird deutlich“ (Lehrer:in/Ö_k)

„Hinweis auf Problem der Erklärung eines Phänomens wenn viele Einflussfaktoren beteiligt sind, ist wichtig -> fördert eine realistische Sichtweise“ (Didaktiker:in/Ö_k)

In der Subkategorie der Prinzipienorientierung wurden hingegen Äußerungen subsummiert, die daraufhin deuten, dass in besonderer Weise Prinzipien in den Fokus der Erklärung gestellt wurden bzw. nicht gestellt wurden, also von einem geeigneten Phänomen abstrahiert wurde bzw. nicht abstrahiert wurde, um Allgemeines und Regelhaftes in der Erklärung herauszuarbeiten. Folgende Beispiele können die wenigen Äußerungen mit Bezug auf das Kriterium exemplarisch illustrieren:

„Je - Desto - Beziehungen kommen wenig zum tragen“ (Student:in/Ö_k)

„möglich auch Hinweis auf Basiskonzepte bzw. Erschließungsfelder“ (Didaktiker:in/G_i)

„Also expliziter Bezug zur Regelmäßigkeit natürlicher Phänomene möglich aber nicht genutzt“ (Didaktiker:in/Ö_i)

Die letzte Facette ‚Fachliche Vollständigkeit‘ wurde auf Basis der Literaturrecherche lediglich mit einem Kriterium – Berücksichtigung aller notwendiger Verstehenselemente bzw. ‚Vollständigkeit‘ – angelegt. Das Kriterium wird von den Teilnehmenden auch geäußert (32 Äußerungen; 1,46 % aller

Äußerungen), jedoch konnten im Material weitere davon unterscheidbare Kriterien gefunden werden, die sich ebenfalls auf die fachliche Vollständigkeit der Erklärung beziehen: Zum einen konnte die Subkategorie „**Fehlende**“ Inhalte, die über das Erklärziel hinaus gehen‘ gebildet werden. Darunter wurden Äußerungen gefasst, bei denen Teilnehmende weitere Erklärgegenstände bzw. Teilerklärungen zur gegebenen Erklärung einfordern, die jedoch nicht zwingend zur Beantwortung der Fragestellung notwendig sind bzw. zum Teil auch durch die Tätigkeiten der Klasse vor der instruktionalen Erklärung erarbeitet wurden (diese wurden in der Kontextualisierung vor dem Zeigen der Erklärung eigentlich als bekannte Fakten benannt). Zum Beispiel:

„fehlt: Folgen des Mangels“ (Schüler:in/P_k)

„Das Protein kann nicht mehr in die Membran eingebaut werden, aber was geschieht damit dann?“ (Student:in/G_i)

„Evtl. noch Folgen des Vitaminmangels im Schaubild klar machen“ (Lehrerin/P_i)

Die Subkategorie kann somit nicht direkt in ein Kriterium für gute instruktionale Erklärungen überführt werden, sondern bezieht sich eher darauf, im Anschluss an die instruktionalen Erklärungen Raum für Fragen zu geben.

Weiterhin konnte die Subkategorie ‚Zusammenhänge werden deutlich‘ am Material gebildet werden (38 Äußerungen; 1,74 % aller Äußerungen), die auch der Facette der fachlichen Vollständigkeit zuzuordnen ist. Die Äußerungen verweisen – meist negativ formuliert – darauf, dass bei kausalen Erklärungen die Zusammenhänge deutlich herausgearbeitet werden oder deutlicher herausgearbeitet werden müssten. Sie eröffnen eine interessante Perspektive, um über die fachliche Vollständigkeit nachzudenken, und wurden deshalb getrennt in einer eigenen Subkategorie zusammengefasst. Folgende Beispiele können das illustrieren:

„man versteht die Zusammenhänge“ (Schüler:in/G_k)

„Manche Zusammenhänge wurden nicht ganz deutlich gemacht. Warum genau braucht Amphibien gemäßigtes Klima?“ (Student:in/Ö_k)

„+ Ursache-Wirkungs-Zusammenhänge werden erklärt“ (Didaktiker:in/Ö)

Resümiert kann werden, dass instruktionale Erklärungen hinsichtlich biologiedidaktischer Facetten insbesondere auf die fachliche Korrektheit hin beurteilt werden. Selten wird hingegen auf die Nutzung der korrekten fachspezifischen Erklärungsform verwiesen. Im Kontext von Wissenslücken wird das Kriterium der Berücksichtigung von Nichtwissen von den Teilnehmenden zur Bewertung herangezogen. Darüber werden Erklärungen anhand deren Vagheit bewertet, die sich in einem interessanten Spannungsverhältnis zwischen fachlicher Vagheit und wünschenswerter didaktischer bzw. lernorientierter Eindeutigkeit befindet. Die Auswirkungen des Gesagten spielen für die Teilnehmenden scheinbar keine prominente Rolle beim Erklären biologischer Inhalte, da es keine Äußerungen im Material gibt, die darunter subsumiert werden können. Die Prinzipienorientierung wird selten und die Phänomenorientierung häufiger als Kriterium herangezogen, um die Erklärungen zu bewerten. Darüber hinaus wird die fachliche Vollständigkeit als Kriterium der Facette zur Berücksichtigung aller notwendiger Verstehenselemente häufig mit in die Bewertungen einbezogen. Als neues Kriterium dieser Facette wurde die deutliche Herausarbeitung von Zusammenhängen evident.

6.2.1.2 Überblick und Einordnung der Ergebnisse zur statusgruppenübergreifenden Nutzung von Kriterien für die Bewertung instruktionaler Erklärungen

Zentrale Befunde in Bezug auf die Aspekte

Die Güte von instruktionalen Erklärungen wird von Teilnehmenden mit Kriterien bewertet, die in acht thematischen Hauptkategorien (Adressatenorientierung, Strukturiertheit, Repräsentationen, Sprech- und Körperausdruck, verständliche Sprache, fachlicher Aspekt, Merkmale der Lehrperson, Merkbarkeit und Lernbarkeit des Gesagten) sowie globalen Aspekten der Güte einer instruktionalen Erklärung zusammengefasst werden können. Die Vorannahmen können dahingehend bestätigt werden, dass die Aspekte der Adressatenorientierung, Strukturiertheit, Repräsentationen, Sprech- und Körperausdruck, verständliche Sprache und fachliche Aspekte zur Bewertung herangezogen bzw. Kriterien, die diesen Aspekten zuordenbar sind, zur Bewertung der Erklärungen genutzt werden. In einigen Fällen wird auch auf verschiedene Merkmale der Lehrperson als Bewertungskriterien zurückgegriffen. Insgesamt lassen sich die in der Literatur diskutierten Aspekte damit auch im Material finden (z. B. Kiel, 1999; Odora, 2014; Schopf & Zwischenbrugger, 2015; Sevian & Gonsalves, 2008; Treagust & Harrison, 2000; Wagner & Wörn, 2011). Darüber hinaus spielen die Merkbarkeit und Lernbarkeit des Gesagten in einigen wenigen Begründungen eine Rolle, wodurch die Bedeutung von instruktionalen Erklärungen im Unterricht deutlich gemacht bzw. implizit auf die Outputorientierung verwiesen wird (Dittmer, 2010; Grotzer, 2012).

Die Häufigkeit der Bezüge auf die acht Hauptkategorien unterscheiden sich deutlich: Am häufigsten werden Äußerungen zu den Aspekten der Repräsentationen (27,78 %), der globalen Aspekte (19,41 %), der Adressatenorientierung (18,17 %) sowie zum fachlichen Aspekt (10,48 %) getroffen. Weniger häufig kommen Äußerungen zu den Aspekten der Strukturiertheit (7,55 %), des Sprech- und Körperausdrucks (5,03 %), der verständlichen Sprache (4,26 %), der Merkbarkeit und Lernbarkeit (2,03 %) sowie den Merkmalen der Lehrperson (0,82 %) vor. Die Vorannahme, dass der Sprech- und Körperausdruck sowie die Verständlichkeit der Sprache seltener als die weiteren Aspekte genannt werden, kann bestätigt werden. Die Ergebnisse können nur bedingt Hinweise in Bezug auf die Relevanz der Aspekte für die Güte der instruktionalen Erklärung geben (siehe dazu auch Kap. 6.2.3). Mit einbezogen werden muss, dass die Kategorien bzw. Aspekte unterschiedlich breit angelegt sind und deshalb auch unterschiedlich viele Möglichkeiten für eine Äußerung bestehen, die die kumulierte Häufigkeit für die Hauptkategorie bzw. den Aspekt moderieren. Weiterhin kann angenommen werden, dass die Häufigkeit der Nennung auch vom Wissen über den jeweiligen Aspekt und die Eigenschaften der Vignetten moderiert wird.

Zentrale Befunde in Bezug auf die Kriterien

Neben Kriterien, die sich ganz allgemein auf die Adressatenorientierung, die Strukturiertheit, die Repräsentationen, die Sprache sowie das Fach beziehen,¹⁴⁵ werden 55 inhaltlich unterscheidbare

¹⁴⁵ Nennungen, die bspw. allgemein auf den Aspekt der Adressatenorientierung referieren und daher unter „Adressatenorientierung global“ subsummiert wurden.

Kriterien genannt.¹⁴⁶ Dabei werden auf der einen Seite nicht alle Kriterien, die in der Literatur diskutiert werden, aufgegriffen und auf der anderen Seite weitere Kriterien angelegt.

Die Gegenüberstellung der selbst herangezogenen Kriterien mit denjenigen Kriterien, die in der Literatur genannt bzw. diskutiert werden zeigt, dass von den Teilnehmenden nicht alle Kriterien zur Bewertung herangezogen werden. Darüber hinaus werden Kriterien teilweise nur selten¹⁴⁷ genutzt. Der Umfang differiert dabei je nach Aspekt (beim Aspekt der Repräsentationen ist dabei zu beachten, dass zum Beispiel keine realen Objekte eingesetzt werden, die Angst und Ekel auslösen könnten). Die nicht bzw. selten genannten Kriterien verteilen sich wie folgt auf die sechs inhaltlich differenzierbaren Aspekte guten Erklärens:

- Adressatenorientierung: Zwei von zehn literaturbasierten Kriterien werden gar nicht¹⁴⁸ und zwei werden selten genannt.¹⁴⁹
- Strukturiertheit: Drei von neun literaturbasierten Kriterien werden gar nicht¹⁵⁰ und ein weiteres Kriterium wird selten genannt.¹⁵¹
- Repräsentationen: Neun von 17 literaturbasierten Kriterien werden gar nicht¹⁵² genannt und ein weiteres Kriterium wird selten genannt.¹⁵³
- Sprech- und Körperausdruck: Drei von elf literaturbasierten Kriterien werden gar nicht¹⁵⁴ und sechs werden selten genannt.¹⁵⁵
- Verständliche Sprache: Eins von fünf literaturbasierten Kriterien wird gar nicht¹⁵⁶ und keines der literaturbasierten selten genannt.
- Fachspezifische Aspekte: Zwei von acht literaturbasierten Kriterien werden gar nicht¹⁵⁷ und zwei werden selten genannt.¹⁵⁸

Darüber hinaus lassen sich sieben globale Aspekte im Material finden: Die instruktionale Erklärung sollte verständlich, anschaulich, übersichtlich, nachvollziehbar, klar, sachlich und informativ sein.

¹⁴⁶ In diese Zahl wurden die Subkategorien ‚Übererklärung‘, ‚Fehlende Inhalte, die über das Erklärziel hinausgehen‘ sowie ‚gut erklärt‘ nicht mit einbezogen, da sie zwar von den Teilnehmenden als Begründung herangezogen werden, jedoch nicht als Kriterien guten Erklärens aufgenommen werden können.

¹⁴⁷ Kriterium für die Bezeichnung selten: Kriterien, auf die weniger als 0,50 % der Äußerungen entfallen; entspricht 11 von 2185 Äußerungen.

¹⁴⁸ Berücksichtigung des Abstraktionsvermögens und Berücksichtigung der Kapazität des Arbeitszeitgedächtnisses.

¹⁴⁹ Berücksichtigung von (Fehl-)Vorstellungen, Lebensweltbezug

¹⁵⁰ Klare Herausarbeitung von Verstehenselementen, Offenlegung der Struktur, Einbettung in einen größeren Zusammenhang

¹⁵¹ Zusammenfassung zentraler Punkte.

¹⁵² Berücksichtigung des Vorwissens der Adressat:innen in Bezug auf die Darstellung, Explikation des Ziels der Betrachtung, Berücksichtigung der Komplexität der realen Objekte, Berücksichtigung von Empfindungen gegenüber den realen Objekten (z. B. Ekel/Angst), zeitgleiche Präsentation von Bild und Text, Integration von Bild und Text, Herausarbeitung des Herkunftsbereichs und des Übertragungsbereichs bei Metaphern und Vergleichen sowie Aufzeigen der Grenzen der Übertragung bei Metaphern und Vergleichen, Auswahl oder Gestaltung der Beispiele nach den Eigenschaften passend, typisch, deckend, konkret, lebensweltnah, interessant und relevant.

¹⁵³ Erklärbedürftigkeit der Grafik.

¹⁵⁴ Entspannte Körperhaltung, angemessene Bewegung im Raum und einladende Mimik (im zugewandten Körperausdruck enthalten, der nur einmal erwähnt wird; keine der Äußerungen bezieht sich auf eine einladende Mimik).

¹⁵⁵ Stimmklang, Sprechpausen, Sprechflüssigkeit, offene und zugewandte Körperhaltung inklusive Blickkontakt mit den Adressat:innen (Zugewandter Körperausdruck), unterstützender Körperausdruck.

¹⁵⁶ Hervorheben von Wichtigem.

¹⁵⁷ Gefühl der (Un-)Sicherheit, Faszination und Interessiertheit.

¹⁵⁸ Fachspezifische Erklärungsformen und Prinzipienorientierung.

Die Schlussfolgerung, dass die nicht oder kaum im Material enthaltenen Kriterien allgemein keine Relevanz für die Güte instruktionaler Erklärungen haben, ist jedoch unzulässig. Betrachtet man die Nichtnennungen und die selten genannten Kriterien, können verschiedene Gründe dafür vermutet werden, wobei mehrere Gründe auf ein Kriterium zutreffen können:

- Die gezeigten Erklärungen geben keinen Anlass, das Kriterium in die Bewertung mit einzubeziehen (Berücksichtigung der Komplexität der realen Objekte (Repräsentationen), Berücksichtigung von Empfindungen gegenüber den realen Objekten (Repräsentationen)).
- Die Kürze der Erklärungen macht es nicht notwendig, das Kriterium anzulegen (Zusammenfassung zentraler Punkte (Strukturiertheit), Offenlegung der Struktur (Strukturiertheit), Einbettung in einen größeren Zusammenhang (Strukturiertheit)).
- Durch die Situation des Videos bzw. des Studiendesigns ist eine Bewertung anhand des Kriteriums nicht sinnvoll (angemessene Bewegung im Raum (Sprech- und Körperausdruck), Einbettung in einen größeren Zusammenhang (Strukturiertheit)).
- Es fehlt Wissen und Expertise bzw. das Merkmal ist in einer Theorie verankert, die den bewertenden Personen nicht ausreichend bekannt ist, um den Aspekt wahrzunehmen und bewerten zu können (Berücksichtigung des Abstraktionsvermögens (Adressatenorientierung), Berücksichtigung der Kapazität des Arbeitsgedächtnisses (Adressatenorientierung), Nutzung der korrekten fachspezifischen Erklärungsform (fachspezifische Aspekte), Prinzipienorientierung (fachspezifische Aspekte)). Insbesondere beim Aspekt des Sprech- und Körperausdrucks kann vermutet werden, dass das Wissen bzw. das fehlende Wissen ausschlaggebend für die eher niedrige Anzahl an selbst herangezogenen Kriterien ist.
- Im Zusammenhang mit fehlendem Wissen ist auch die Nichtnennung der Kriterien ‚Gefühl der (Un-)Sicherheit‘ und ‚Faszination und Interessiertheit‘ unter der Facette ‚Berücksichtigung der Auswirkungen des Gesagten‘ zu sehen – angenommen werden kann jedoch auch, dass nicht nur fehlendes Wissen darüber, sondern insbesondere bei Gefühlen der Unsicherheit, die biologische Erklärgegenstände auslösen können, auch die schulische Lehr-Lern-Kultur einen Einfluss hat, wobei die Bewertung von Erklärungen eher auf den mit der Erklärung möglichen Wissenserwerb und nicht die emotionalen Aspekte des Gesagten ausgerichtet sind.
- Bereits hervorgehoben wurde (siehe Kap. 6.2.1.1), dass die Nichtnennung des Lebensweltbezugs (Adressatenorientierung) interessant ist, da dieser in praxisnahen Quellen häufig zum Erklären sowie zum Unterrichten allgemein genannt wird (Pauli, 2015; Schopf & Zwischenbrugger, 2015; Wellenreuther, 2013), sodass eigentlich davon ausgegangen werden könnte, dass er als Kriterium genannt wird. Eine vermutungsbasierte Zuordnung zu einem der bereits genannten Gründe für eine Nichtnennung ist unplausibel.
- Insgesamt muss auch davon ausgegangen werden, dass nicht alle Kriterien, die die Teilnehmenden kennen bzw. die zur Entscheidung für eine Notenvergabe herangezogen wurden, genannt werden, sondern in Relation zur angegebenen Zeit für die Bearbeitung des Fragebogens eine Auswahl an Kriterien niedergeschrieben wird. Insofern ist es auch interessant, die am häufigsten genannten Kriterien anzusehen.

Die Aspekte der Verständlichkeit der Sprache und des Sprech- und Körperausdrucks werden entgegen den Vorannahmen nicht häufiger als andere Aspekte global in die Bewertung der Güte der instruktionalen Erklärung mit einbezogen. Die Verständlichkeit der Sprache wird differenziert anhand von vier Kriterien, die auch in der Literatur Erwähnung finden, bewertet. Der Sprech- und Körperausdruck hingegen wird zwar ebenfalls nicht global bewertet, jedoch weniger differenziert und maßgeblich nur anhand von vier Kriterien bewertet, die der Stimme zuzuordnen sind (Artikulation, allgemeine prosodische Merkmale, Sprechgeschwindigkeit, Variation der Stimme).

Datenbasiert kommen sechs Kriterien hinzu, die sich den sechs Aspekten zuordnen lassen: Aktiver Einbezug der Rezipient:innen (Adressatenorientierung), Übereinstimmung des Erklärziels mit dem Erklärinhalt (Strukturiertheit), ästhetische Gestaltung der realistischen Bilder (Repräsentationen), ausreichende Bezugnahme auf das Beispiel (Repräsentationen), Sprechflüssigkeit (Sprech- und Körperausdruck) und Zusammenhänge zwischen den Faktoren innerhalb der Erklärung werden deutlich (fachspezifischer Aspekt). Auch die Lehrperson sowie die Merkbarkeit und Lernbarkeit werden in die Bewertungen einbezogen.

Insgesamt wird anhand der Vielzahl genannter Kriterien die Komplexität der Sprachhandlung ‚Erklären‘ evident. Die Vielzahl genannter Kriterien unterstützt auch die Frage nach der Relevanz von Kriterien für die Güte der instruktionalen Erklärung (siehe Forschungsfrage 3; Kap. 6.2.3), um zum einen Lehr-Lern-Konzepte für instruktionales Erklären im Rahmen der hochschuldidaktischen Ausbildung von angehenden Lehrkräften oder der Weiterbildung von Lehrkräften zu erarbeiten (z. B. Asen-Molz, Knott & Schilcher, 2022; Ehras et al., 2021) und zum anderen Anhaltspunkte für die Überprüfung bzw. Studien zu effektiven instruktionalen Erklärens zu konzipieren.

6.2.1.3 Ergebnisse zur statusgruppenspezifischen Nutzung von Kriterien für die Bewertung instruktionaler Erklärungen

Um Erkenntnisse darüber zu erhalten, welche Kriterien aus der Sicht von verschiedenen Statusgruppen zur Bewertung der Güte einer instruktionalen Erklärung herangezogen werden, werden die Begründungen der Notenvergabe bzw. die Ergebnisse der strukturierenden qualitativen Inhaltsanalyse getrennt nach Statusgruppen betrachtet. Für eine bessere Übersichtlichkeit wird die statusgruppenspezifische Betrachtung der selbst herangezogenen Kriterien getrennt nach Hauptkategorien bzw. Aspekten durchgeführt. Für einen gesamten Überblick zu den Häufigkeiten dient jedoch zunächst Tabelle 28. Die Ergebnisse werden an dieser Stelle nicht näher ausgeführt, sondern in den entsprechenden Abschnitten zu den Hauptkategorien bzw. Aspekten erläutert.

Tabelle 28: Übersicht zu den statusgruppenspezifischen (Spalte 2-5) und statusgruppenübergreifenden (Spalte 6) relativen Häufigkeiten der angesprochenen Hauptkategorie bzw. der angesprochenen Aspekte guten Erklärens über alle Videovignetten hinweg

Statusgruppe	Schüler:innen (N _{max} = 122)	Studierende (N _{max} = 51)	Lehrkräfte (N _{max} = 34)	Didaktiker:innen (N _{max} = 30)	über alle Statusgruppen (N _{max} = 237)
Aspekt					
Adressatenorientierung	20,35	18,41	19,05	11,30	18,17
Strukturiertheit	6,78	7,42	11,11	7,62	7,55
Repräsentationen	28,37	24,04	23,81	32,19	27,79
Verständliche Sprache	1,41	7,93	7,54	6,63	4,26
Sprech- u. Körperausdruck	4,05	10,49	4,37	2,95	5,03
Fachliche Aspekte	3,26	13,55	15,87	24,32	10,48
Globale Aspekte	27,14	13,04	11,11	9,09	19,41
Lehrperson	0,97	0,26	0,40	1,23	0,82
Lernbarkeit und Merkbarkeit	2,91	1,79	1,98	0,00	2,06
Sonstiges	4,76	3,07	4,76	4,67	4,44

Bem.: N_{max}: aufgrund fehlender Werte bei einzelnen Videos den Berechnungen maximal zugrunde liegende Personenanzahl.

Eine Tabelle mit Angaben der absoluten und relativen Häufigkeiten getrennt nach Statusgruppe je Subkategorie bzw. Kriterium findet sich in Anhang G.

Adressatenorientierung

Die folgende Tabelle (Tab. 29) gibt einen ersten Überblick zu den absoluten und relativen Häufigkeiten der Äußerungen je Statusgruppe, die zur Hauptkategorie der Adressatenorientierung zuordenbar sind.

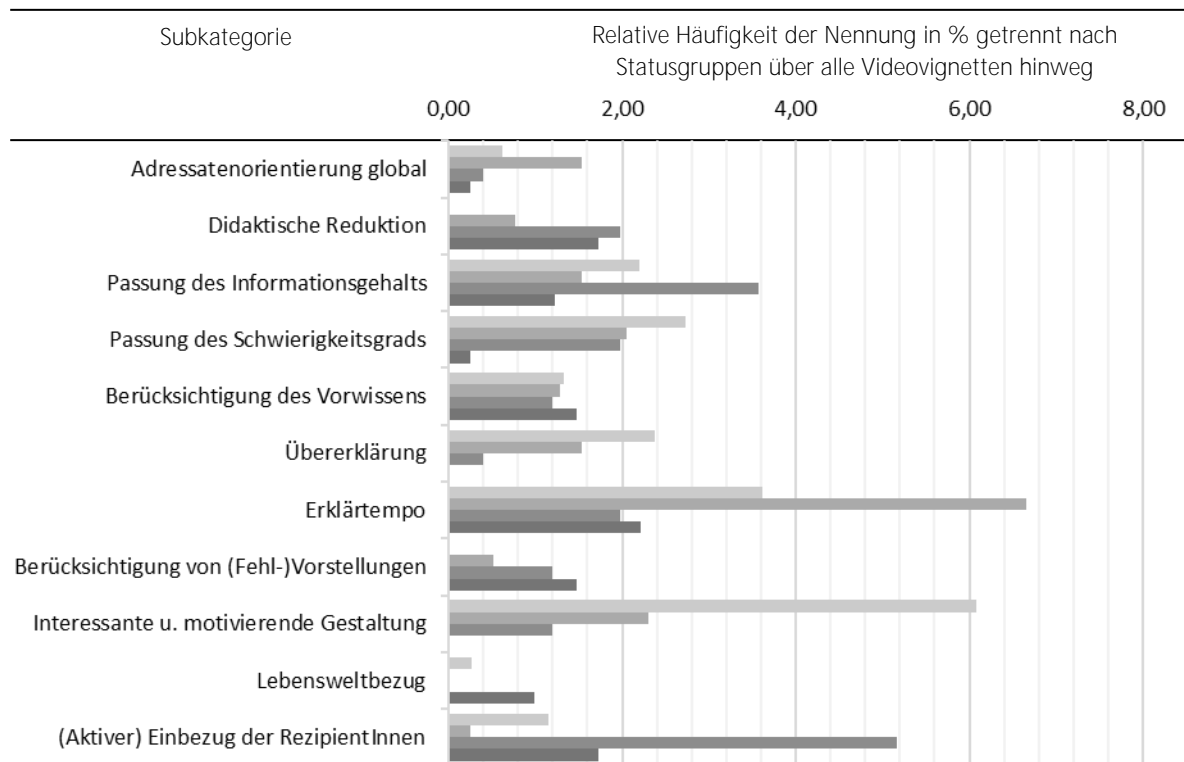
Tabelle 29: Häufigkeiten der Äußerungen, die der Hauptkategorie ‚Adressatenorientierung‘ zugeordnet werden können getrennt nach Statusgruppe über alle Videovignetten hinweg

Statusgruppe Häufigkeit	Schüler:innen ($N_{max} = 122$)	Studierende ($N_{max} = 51$)	Lehrkräfte ($N_{max} = 34$)	Didaktiker:innen ($N_{max} = 30$)	Summe über alle Statusgruppen ($N_{max} = 237$)
Absolute Häufigkeiten	231	72	48	46	397
Relative Häufigkeiten in %	20,35	18,41	19,05	11,30	18,17

Bem.: N_{max} : aufgrund fehlender Werte bei einzelnen Videos den Berechnungen maximal zugrunde liegende Personenanzahl.

Der Vergleich der relativen Häufigkeiten macht deutlich, dass Didaktiker:innen seltener Bezug auf die Passung der Erklärung für die Adressat:innen nehmen als die drei weiteren Statusgruppen. Es kann vermutet werden, dass Schüler:innen besonders häufig Kriterien der Adressatenorientierung anführen, da sie selbst als Adressat:innen die Passung in besonderer Weise wahrnehmen. Für sie sind die gezeigten Videovignetten echte Erklärungen, die neue Inhalte thematisieren. Die unterschiedlichen Häufigkeiten zwischen den Gruppen der Metaperspektive – Studierende, Lehrkräfte auf der einen Seite und Didaktiker:innen auf der anderen Seiten – sind mitunter durch die relative Praxisnähe begründet: Während Studierende mitunter anhand ihrer eigenen noch präsenten Lernerfahrungen sowie mit den Praxiserfahrungen im Praktikum und Lehrkräfte täglich mit Schüler:innen arbeiten und ihnen deren Blick auf Erklärungen näher liegt bzw. präsent ist, sind die Didaktiker:innen der Schulpraxis etwas ferner.

Abbildung 17 gibt Aufschluss über die konkreten Kriterien der Adressatenorientierung, die die Statusgruppen an die Erklärung anlegen.



Bemerkung: Relative Häufigkeiten der Nennungen in den Statusgruppen:

■ Schüler:innen ■ Studierende ■ Lehrkräfte ■ Didaktiker:innen

Abbildung 17: Relative Häufigkeiten der Nennungen der Subkategorien aus der Hauptkategorie ‚Adressatenorientierung‘ getrennt nach Statusgruppen (Schüler:innen $N_{max} = 122$; Studierende $N_{max} = 51$; Lehrkräfte $N_{max} = 34$; Didaktiker:innen $N_{max} = 30$) über alle Videovignetten hinweg

Bei der Betrachtung der statusgruppenspezifischen Häufigkeiten der Nennungen fällt zunächst auf, dass Schüler:innen nicht mehr globale Begründungen zur Passung der Erklärung abgeben als die weiteren drei Statusgruppen. Die folgenden Beispiele geben einen Einblick in die Inhalte der Kategorie ‚Adressatenorientierung global‘:

„*altersgemäße Erklärung*“ (Studentin/P)

„*Unterrichtsweise nicht für höhere Klassen geeignet*“ (Schüler:in/G)

Die Statusgruppe der Schüler:innen kann durch folgende Muster in der Bewertung charakterisiert werden: Am häufigsten werden die Kriterien der interessanten und motivierenden Gestaltung (69 von 231 Äußerungen innerhalb der Kategorie ‚Adressatenorientierung‘; 6,08 % aller Äußerungen), des Erklärtempos (41 von 231 Äußerungen innerhalb der Kategorie ‚Adressatenorientierung‘; 3,61 % aller Äußerungen) und der Passung des Schwierigkeitsgrads (31 von 231 Äußerungen innerhalb der Kategorie ‚Adressatenorientierung‘; 2,73 % aller Äußerungen) für die Begründung der Notenvergabe herangezogen. Dass Schüler:innen die Subkategorien ‚Berücksichtigung von (Fehl-)Vorstellungen‘ und ‚didaktische Reduktion‘ nicht ansprechen, ist naheliegend, da für eine derartige Einschätzung professionelle Wissensbestände bzw. fachdidaktisches Wissen notwendig sind.

Die Begründungen von Studierenden weisen ein recht ähnliches Muster zu demjenigen der Schüler:innen auf, wobei die Rangfolge der drei am meisten angesprochenen Kriterien variiert: Am häufigsten werden Kriterien des Erklärtempos (26 von 72 Äußerungen innerhalb der Kategorie ‚Adressatenorientierung‘; 6,65 % aller Äußerungen), der interessanten und motivierenden Gestaltung (9 von 72 Äußerungen innerhalb der Kategorie ‚Adressatenorientierung‘; 2,30 % aller Äußerungen) und der Passung des Schwierigkeitsgrads (8 von 72 Äußerungen innerhalb der Kategorie ‚Adressatenorientierung‘; 2,05 % aller Äußerungen) für die Begründung der Notenvergabe herangezogen. Die Häufigkeit des Kriteriums ‚Erklärtempo‘ ist dabei im Vergleich zu den anderen Kriterien innerhalb der Studierenden sowie im Vergleich zu den Häufigkeiten dieses Kriteriums bei den anderen Statusgruppen besonders hoch. Im Gegensatz zu den Schüler:innen sprechen die Studierenden – wenn auch deutlich seltener als Lehrkräfte und Didaktiker:innen – Kriterien an, für die professionelles Wissen bzw. fachdidaktisches Wissen notenwendig ist: Das Kriterium ‚Didaktische Reduktion‘ wird dreimal (3 von 72 Äußerungen innerhalb der Kategorie ‚Adressatenorientierung‘; 0,77 % aller Äußerungen) und das Kriterium ‚Berücksichtigung von (Fehl-)Vorstellungen‘ lediglich zweimal (2 von 72 Äußerungen innerhalb der Kategorie ‚Adressatenorientierung‘; 0,51 % aller Äußerungen) genannt. Lehrkräfte und Didaktiker:innen tun dies verhältnismäßig häufiger, was die Annahme des Zusammenhangs zwischen dem verfügbaren professionellen Wissen und der Art der Bewertung instruktionaler Erklärungen für dieses Kriterium stützt. Auffällig ist weiterhin, dass Studierende in keinem Fall den Lebensweltbezug als Kriterium für die Bewertung der Erklärung anführen.

Die Begründungen von Lehrkräften weisen ein von Schüler:innen und Studierenden differenzierbares Muster auf: Am häufigsten wird das Kriterium ‚(Aktiver) Einbezug der Rezipient:innen‘ (13 von 48 Äußerungen innerhalb der Kategorie ‚Adressatenorientierung‘; 5,16 % aller Äußerungen) genannt. Der Aktive Einbezug von Rezipient:innen, beispielsweise durch ein fragend-entwickelndes Unterrichtsgespräch oder eigenständige Erarbeitungsphasen der Schüler:innen, wird also vorrangig von Lehrkräften gefordert. Darüber hinaus wird etwas häufiger auf die Kriterien, die hier der didaktischen Reduktion zugeordnet werden, rekurriert: die Passung des Informationsgehalts (9 von 48 Äußerungen innerhalb der Kategorie ‚Adressatenorientierung‘; 3,57 % aller Äußerungen), die Passung des Schwierigkeitsgrads (5 von 48 innerhalb der Kategorie ‚Adressatenorientierung‘; 1,98 % aller Äußerungen) und allgemein die didaktische Reduktion (ebenfalls 5 von 48 innerhalb der Kategorie ‚Adressatenorientierung‘; 1,98 % aller Äußerungen). Auffällig ist weiterhin, dass Lehrkräfte wie Studierende in keinem Fall den Lebensweltbezug als Kriterium für die Bewertung der Erklärung anführen.

Didaktiker:innen nehmen bei ihren Begründungen insgesamt seltener Bezug auf die Adressatenorientierung (11,30 %) als alle weiteren Statusgruppen. Dabei führen sie in ihren Begründungen nicht an, dass die Erklärung interessant und motivierend gestaltet ist oder eine fehlerhafte Anpassung an die Adressat:innen im Sinne einer Übererklärung vorliegt. Am häufigsten beziehen sie sich auf das Erklärtempo (9 von 46 Äußerungen innerhalb der Kategorie ‚Adressatenorientierung‘; 2,21 % aller Äußerungen), was insofern überraschend ist, als die weiteren Kriterien, die inhaltlich nahe am Kriterium des Erklärtempos sind (Passung des Informationsgehalts, Passung des Schwierigkeitsgrads, Übererklärung), selten genannt werden. Die weiteren drei Statusgruppen führen das Erklärtempo wie auch die inhaltlich näheren Kriterien deutlich häufiger an. Auffällig ist außerdem, dass Didaktiker:innen den Lebensweltbezug anführen (7 von 46 Äußerungen innerhalb der Kategorie ‚Adressatenorientierung‘; 1,72 % aller Äußerungen).

Insgesamt ist auffällig, dass das Kriterium des Lebensweltbezugs nicht von den Statusgruppen, denen eher eine größere Praxisnähe zugeschrieben werden kann – also Schüler:innen und Lehrkräfte –, sondern nur selten von Schüler:innen (3 von 231 Äußerungen innerhalb der Kategorie ‚Adressatenorientierung‘; 0,26 % aller Äußerungen) und vor allem von Didaktiker:innen (4 von 46 Äußerungen innerhalb der Kategorie ‚Adressatenorientierung‘; 0,98 % aller Äußerungen) in den Begründungen zur Bewertung der instruktionalen Erklärungen angeführt wird. Im Spiegel des bereits in Kapitel 6.2.1.1 angemerkten Ergebnisses der Literaturrecherche, in der dieses Kriterium eher in praxisnahen Literaturquellen zu gutem Erklären gefunden werden kann (Pauli, 2015; Schopf & Zwischenbrugger, 2015; Wellenreuther, 2013), ist dieses Ergebnis überraschend.

Auf der Ebene der Kriterien kann resümiert werden, dass Schüler:innen und Studierende ihre Bewertungen hauptsächlich auf drei Kriterien stützen (Passung des Schwierigkeitsgrads, Erklärtempo, interessante und motivierende Gestaltung). Aufgrund der Ähnlichkeit der Verteilung der Häufigkeiten der Kriterien kann vermutet werden, dass sich Studierende zum Teil als Adressat:innen wahrnehmen bzw. sich eher in diese Position hineinversetzen, um die Adressatenorientierung zu beurteilen. Lehrkräfte und Didaktiker:innen ziehen jeweils fünf, nur teilweise identische Kriterien am häufigsten zur Begründung heran, sodass sich der Aspekt Adressatenorientierung für diese Perspektiven als breiteres Konstrukt darstellt. Das am Material gebildete Kriterium ‚(Aktiver) Einbezug der Adressat:innen‘ wird vorrangig von Lehrkräften genannt.

Um die hohe Anzahl an Kategorien besser zu überblicken, werden die Kategorien noch entlang der in Kapitel 6.2.1.1 aufgeführten Facetten zusammengefasst. Dabei ergibt sich folgendes Bild:

Tabelle 30: Absolute und relative Häufigkeiten der Äußerungen, die den verschiedenen Facetten des Aspekts der ‚Adressatenorientierung‘ zugeordnet werden können, getrennt nach Statusgruppe über alle Videovignetten hinweg

Statusgruppe	Schüler:innen ($N_{max} = 122$)		Studierende ($N_{max} = 51$)		Lehrkräfte ($N_{max} = 34$)		Didaktiker:innen ($N_{max} = 30$)		Summe über alle Statusgruppen ($N_{max} = 237$)	
	A.H.	R.H.	A.H.	R.H.	A.H.	R.H.	A.H.	R.H.	A.H.	R.H.
Didaktische Reduktion	56	4,93	17	4,35	19	7,54	13	3,19	105	4,81
Vorwissen und Vorstellungen	83	7,31	39	12,02	12	4,76	21	5,16	155	7,09
Kognitive Fähigkeiten	0	0,00	0	0,00	0	0,00	0	0,00	0	0,00
Affektiv motivationale Voraussetzungen	72	6,34	9	2,30	3	1,19	4	0,98	88	4,03
(Aktiver) Einbezug der Rezipient:innen	13	1,15	1	0,26	13	5,16	7	1,72	34	1,56

Bem.: N_{max} : aufgrund fehlender Werte bei einzelnen Videos den Berechnungen maximal zugrunde liegende Personenanzahl; A.H.: absolute Häufigkeit; R.H.: relative Häufigkeit.

Vergleicht man die Häufigkeiten der Nennungen innerhalb der Statusgruppen, so ziehen Schüler:innen vorrangig Kriterien zur Bewertung heran, die die Anpassung an ihr Vorwissen und ihre Vorstellungen sowie an ihre affektiv-motivationalen Voraussetzungen betreffen. Auch die aus ihrer Sicht gelungene oder weniger gelungene didaktische Reduktion (die Anpassung des Informationsgehalts und des Schwierigkeitsgrads) spielt bei der Bewertung eine größere Rolle. Studierende hingegen beziehen sich vor allem auf die Anpassung an das Vorwissen und die Vorstellungen von Schüler:innen. Die Umsetzung der didaktischen Reduktion wird ebenfalls mit einbezogen, jedoch deutlich seltener als die Anpassung an das Vorwissen und die Vorstellungen. Lehrkräfte bewerten die Adressatenorientierung überwiegend anhand der Kriterien ‚Didaktische

Reduktion', ‚(Aktiver) Einbezug der Rezipient:innen‘ und der Berücksichtigung von ‚Vorwissen und Vorstellungen‘. Didaktiker:innen bewerten die Adressatenorientierung wie Studierende vor allem anhand der Anpassung der Erklärung an das Vorwissen und die Vorstellungen von Schüler:innen, wobei auch die Umsetzung der didaktischen Reduktion als Kriterium herangezogen wird.

Zusammenfassend ergibt sich das Bild, dass auf Ebene der Facetten alle Statusgruppen die Reduktion sowie die Anpassung an das Vorwissen und die Vorstellungen häufig zur Bewertung der Erklärungen heranziehen. Schüler:innen, Studierende und Didaktiker:innen nennen dabei häufiger Kriterien mit Bezug zum Vorwissen und Vorstellungen als Grund für die Notenvergabe, Lehrkräfte Kriterien, die die Umsetzung der didaktischen Reduktion betreffen. Für Schüler:innen kommt als wesentlicher Bezugspunkt für die Bewertung der Erklärung die Berücksichtigung der affektiv-motivationalen Voraussetzungen hinzu. Für Lehrkräfte ist auch der aktive Einbezug der Rezipient:innen relevant.

Strukturiertheit

Im Vergleich zwischen den Statusgruppen kann zunächst mit Blick auf die Tabelle 31 festgehalten werden, dass Lehrkräfte bei ihren Begründungen zur Notenvergabe häufiger Bezug auf die Strukturiertheit nehmen als Schüler:innen, Studierende und Didaktiker:innen. Zwischen den Statusgruppen der Schüler:innen, Studierenden und Didaktiker:innen sind dabei keine auffälligen Unterschiede in der relativen Häufigkeit des Bezugs auf Kriterien der Strukturiertheit festzustellen.

Tabelle 31: Häufigkeiten der Äußerungen, die der Hauptkategorie ‚Strukturiertheit‘ zugeordnet werden können, getrennt nach Statusgruppe über alle Videovignetten hinweg

Statusgruppe Häufigkeit	Schüler:innen ($N_{max} = 122$)	Studierende ($N_{max} = 51$)	Lehrkräfte ($N_{max} = 34$)	Didaktiker:innen ($N_{max} = 30$)	Summe über alle Statusgruppen ($N_{max} = 237$)
Absolute Häufigkeiten	77	29	28	31	165
Relative Häufigkeiten in %	6,78	7,42	11,11	7,62	7,55

Bem.: N_{max} : aufgrund fehlender Werte bei einzelnen Videos den Berechnungen maximal zugrunde liegende Personenanzahl.

Abbildung (Abb. 18) zur Verteilung der Äußerungen über die Subkategorien bzw. die differenzierbaren Kriterien zur Strukturiertheit gibt unter anderem Aufschluss darüber, welche Kriterien Lehrkräfte besonders häufig für die Bewertung der Erklärungen heranziehen.

Ersichtlich wird, dass auch hier Schüler:innen nicht häufiger als andere Teilnehmende Äußerungen abgeben, die sich auf die globale Strukturiertheit der instruktionalen Erklärung beziehen. Im Vergleich tun dies hingegen Lehrkräfte (3,17 % aller Äußerungen) und Didaktiker:innen (2,21 % aller Äußerungen) häufiger.

Schüler:innen beziehen sich bei ihren Begründungen am häufigsten auf das Kriterium des logischen Aufbaus (29 von 77 Äußerungen innerhalb der Kategorie ‚Strukturiertheit‘; 2,56 % aller Äußerungen) sowie auf das Kriterium der Fokussierung auf Wesentliches (24 von 77 Äußerungen innerhalb der Kategorie ‚Strukturiertheit‘; 2,11 % aller Äußerungen). Nicht genannt wird die Zusammenfassung zentraler Punkte.

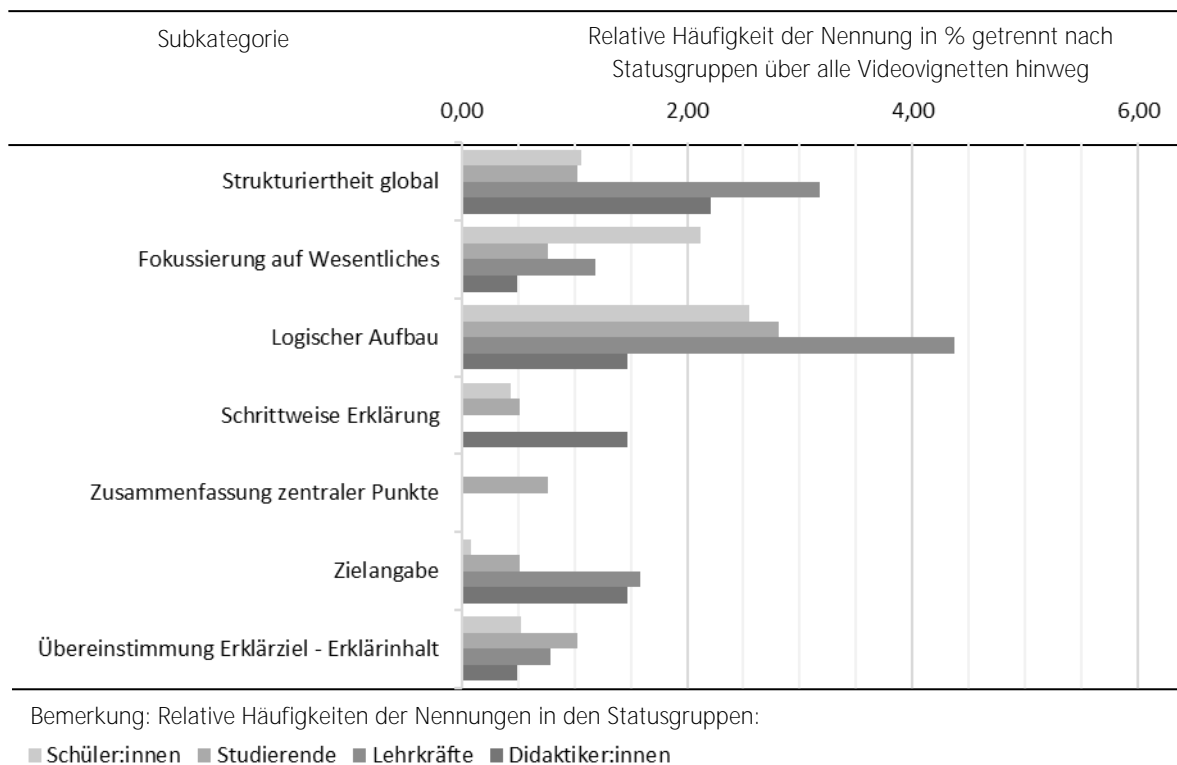


Abbildung 18: Relative Häufigkeiten der Nennungen der Subkategorien aus der Hauptkategorie ‚Strukturiertheit‘ getrennt nach Statusgruppe (Schüler:innen $N_{max} = 122$; Studierende $N_{max} = 51$; Lehrkräfte $N_{max} = 34$; Didaktiker:innen $N_{max} = 30$) über alle Videovignetten hinweg

Studierende begründen ihre Notenvergabe ebenfalls am häufigsten mit dem Kriterium des logischen Aufbaus (11 von 29 Äußerungen innerhalb der Kategorie ‚Strukturiertheit‘; 2,81 % aller Äußerungen). Auch die Übereinstimmung von Erklärziel und Erklärinhalt wird in den Begründungen angeführt (4 von 29 Äußerungen innerhalb der Kategorie ‚Strukturiertheit‘; 1,02 % aller Äußerungen) – häufiger als das Vorhandensein der Zielangabe selbst (2 von 29 Äußerungen innerhalb der Kategorie ‚Strukturiertheit‘; 0,51 % aller Äußerungen). Sie sind weiterhin die einzige Statusgruppe, die sich bei der Bewertung der Erklärung auf die Zusammenfassung zentraler Punkte als Kriterium bezieht (lediglich 3 von 29 Äußerungen innerhalb der Kategorie ‚Strukturiertheit‘; 0,77 % aller Äußerungen).

Lehrkräfte beziehen sich wie bereits Schüler:innen und Studierende vorrangig auf den logischen Aufbau (11 von 28 Äußerungen innerhalb der Kategorie ‚Strukturiertheit‘; 4,37 % aller Äußerungen). Auffällig hoch ist in dieser Statusgruppe in Vergleich zu den weiteren Statusgruppen die Anzahl globaler Nennungen des Kriteriums ‚Strukturiertheit‘ (8 von 28 Äußerungen innerhalb der Kategorie ‚Strukturiertheit‘; 3,17 % aller Äußerungen). Lehrkräfte beziehen sich bei ihrer Bewertung nicht auf das Kriterium der schrittweisen Erklärung sowie die Zusammenfassung zentraler Punkte.

Didaktiker:innen nennen in ihren Begründungen am häufigsten das globale Kriterium ‚Strukturiertheit global‘ (9 von 31 Äußerungen innerhalb der Kategorie ‚Strukturiertheit‘; 2,21 % aller Äußerungen). Auch der logische Aufbau wird relativ häufig genannt (6 von 31 Äußerungen innerhalb der Kategorie ‚Strukturiertheit‘; 1,47 % aller Äußerungen). Darüber hinaus werden im Gegensatz zu den weiteren drei Statusgruppen aber auch die Kriterien ‚Schrittweise Erklärung‘ und ‚Zielangabe‘ ebenso häufig genannt (jeweils 6 von 31 Äußerungen innerhalb der Kategorie ‚Strukturiertheit‘; 1,47 % aller Äußerungen). Zur Bewertung des Aspekts der Strukturiertheit einer

Erklärung werden von Didaktiker:innen daher durchschnittlich mehr Kriterien herangezogen. Keine Nennung erhält das Kriterium ‚Zusammenfassung zentraler Punkte‘.

Das neue Kriterium ‚Übereinstimmung von Erklärziel und Erklärinhalt‘ wird selten angeführt, davon verhältnismäßig häufig von Studierenden (4 von 29 Äußerungen innerhalb der Kategorie ‚Strukturiertheit‘; 1,02 % aller Äußerungen).

Zusammenfassend ist das Kriterium des logischen Aufbaus damit für alle Statusgruppen zentral für die Strukturiertheit einer Erklärung bzw. die Bewertung der Erklärung insgesamt. Schüler:innen, Lehrkräfte und Didaktiker:innen führen das Kriterium der Zusammenfassung zentraler Punkte nicht an und Lehrkräfte nennen die schrittweise Erklärung nicht als Kriterium für ihre Notenvergabe. Für Schüler:innen scheint weiterhin die Zielangabe weniger ausschlaggebend zu sein als die Übereinstimmung zwischen dem Erklärziel und dem Erklärinhalt, ebenso wie für Studierende.

Fasst man die Kriterien zu den in Kapitel 6.2.1.1 aufgeführten Facetten zusammen, ergibt sich folgendes Bild (siehe Tab. 32):

Tabelle 32: Absolute und relative Häufigkeiten der Äußerungen, die den verschiedenen Facetten des Aspekts der ‚Strukturiertheit‘ zugeordnet werden können, getrennt nach Statusgruppe über alle Videovignetten hinweg

Facette	Schüler:innen ($N_{max} = 122$)		Studierende ($N_{max} = 51$)		Lehrkräfte ($N_{max} = 34$)		Didaktiker:innen ($N_{max} = 30$)		Summe über alle Statusgruppen ($N_{max} = 237$)	
	A.H.	R.H.	A.H.	R.H.	A.H.	R.H.	A.H.	R.H.	A.H.	R.H.
Inhaltliche Fokussierung	24	2,11	3	0,77	3	1,19	2	0,49	32	1,46
Gliederung und Aufbau	34	3,00	16	4,09	11	4,37	12	2,95	73	3,34
Orientierung und Transparenz	7	0,62	6	1,53	8	2,38	8	1,97	27	1,24

Bem.: N_{max} : aufgrund fehlender Werte bei einzelnen Videos den Berechnungen maximal zugrunde liegende Personenanzahl; A.H.: absolute Häufigkeit; R.H.: relative Häufigkeit.

Alle Statusgruppen beziehen sich am häufigsten auf die Facette ‚Gliederung und Aufbau‘. Während jedoch Schüler:innen am zweithäufigsten Kriterien nennen, die der Facette ‚Inhaltliche Fokussierung‘ zugeordnet werden können, nennen Studierende, Lehrkräfte und Didaktiker:innen Kriterien der Facette ‚Orientierung und Transparenz‘ am zweithäufigsten.

Repräsentationen

Aus der untenstehenden Tabelle (Tab. 33) geht hervor, dass insbesondere Didaktiker:innen bei der Bewertung der Erklärungen häufig Kriterien heranziehen, die sich auf die eingesetzten Repräsentationen beziehen. Weiterhin zeigt sich, dass Schüler:innen häufiger als Studierende und Lehrkräfte auf die Repräsentationen Bezug nehmen.

Tabelle 33: Häufigkeiten der Äußerungen, die der Hauptkategorie ‚Repräsentationen‘ zugeordnet werden können, getrennt nach Statusgruppe über alle Videovignetten hinweg

Statusgruppe	Schüler:innen ($N_{max} = 122$)	Studierende ($N_{max} = 51$)	Lehrkräfte ($N_{max} = 34$)	Didaktiker:innen ($N_{max} = 30$)	Summe über alle Statusgruppen ($N_{max} = 237$)
Absolute Häufigkeiten	322	94	60	131	607
Relative Häufigkeiten in %	28,37	24,04	23,81	32,19	27,79

Bem.: N_{max} : aufgrund fehlender Werte bei einzelnen Videos den Berechnungen maximal zugrunde liegende Personenanzahl.

Um zu klären, welche spezifischen Kriterien dabei in den Blick genommen werden, ist in der folgenden Abbildung (Abb. 19) wiederum angegeben, wie sich die Nennungen pro Statusgruppe innerhalb der Kategorie ‚Repräsentationen‘ auf die Subkategorien verteilen.

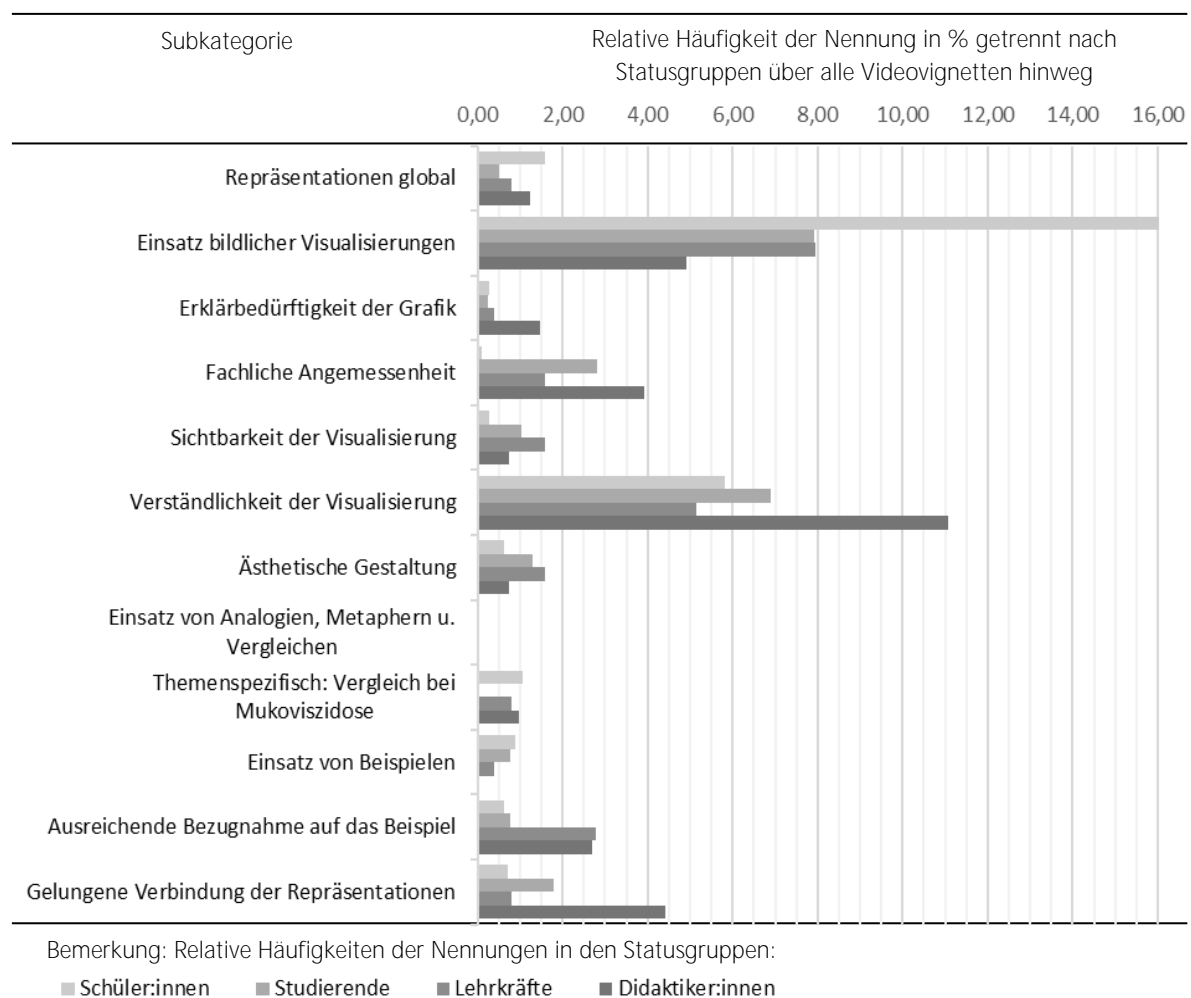


Abbildung 19: Relative Häufigkeiten der Nennungen der Subkategorien aus der Hauptkategorie ‚Repräsentationen‘ getrennt nach Statusgruppe (Schüler:innen $N_{max} = 122$; Studierende $N_{max} = 51$; Lehrkräfte $N_{max} = 34$; Didaktiker:innen $N_{max} = 30$) über alle Videovignetten hinweg

Deutlich wird, dass Schüler:innen vor allem das Kriterium des Vorhandenseins bildlicher Visualisierungen anführen (187 von 322 Äußerungen innerhalb der Kategorie ‚Repräsentationen‘; 16,48 % aller Äußerungen). Deutlich seltener wird das Kriterium der Verständlichkeit der Visualisierung genutzt, um die Notenvergabe zu begründen (66 von 322 Äußerungen innerhalb der Kategorie ‚Repräsentationen‘; 5,81 % aller Äußerungen). Im Vergleich dazu werden alle weiteren

Kriterien, die sich auf die eingesetzten Repräsentationen beziehen, kaum genutzt. Dass Schüler:innen nicht anhand der fachlichen Angemessenheit beurteilen, kann mit fehlendem Professionswissen begründet werden.

Die befragten Studierenden beziehen sich ebenfalls am häufigsten auf das Vorhandensein bildlicher Visualisierungen (31 von 94 Äußerungen innerhalb der Kategorie ‚Repräsentationen‘; 7,93 % aller Äußerungen). Am zweithäufigsten wird – wie in den Begründungen der Schüler:innen – das Kriterium der Verständlichkeit der Visualisierungen angeführt (27 von 94 Äußerungen innerhalb der Kategorie ‚Repräsentationen‘; 6,91 % aller Äußerungen). Im Vergleich zu den weiteren Kriterien entfallen auch auf das Kriterium der fachlichen Angemessenheit noch einige Nennungen (11 von 94 Äußerungen innerhalb der Kategorie ‚Repräsentationen‘; 2,81 % aller Äußerungen). Die Erklärbedürftigkeit der Grafik wird von Studierenden hingegen nicht mit in die Bewertung mit einbezogen (1 von 94 Äußerungen innerhalb der Kategorie ‚Repräsentationen‘; 0,26 % aller Äußerungen).

Auch bei Lehrkräften ist eine ähnliche Verteilung zu finden: Am häufigsten wird das Kriterium des Vorhandenseins bildlicher Visualisierungen (20 von 60 Äußerungen innerhalb der Kategorie ‚Repräsentationen‘; 7,94 % aller Äußerungen) und am zweithäufigsten das Kriterium der Verständlichkeit der Visualisierung (13 von 60 Äußerungen innerhalb der Kategorie ‚Repräsentationen‘; 5,16 % aller Äußerungen) genannt. Einige Äußerungen entfallen – nicht wie bei Studierenden – noch auf die ausreichende Bezugnahme auf das Beispiel (7 von 60 Äußerungen innerhalb der Kategorie ‚Repräsentationen‘; 2,78 % aller Äußerungen). Auch sie nutzen die Erklärbedürftigkeit der Grafik nicht für die Begründung der Globalbeurteilung (1 von 60 Äußerungen innerhalb der Kategorie ‚Repräsentationen‘; 0,40 % aller Äußerungen). Hinzu kommt, dass der Einsatz von Beispielen nicht mit in die Bewertung einfließt (1 von 60 Äußerungen innerhalb der Kategorie ‚Repräsentationen‘; 0,40 % aller Äußerungen).

Die Begründungen von Didaktiker:innen lassen sich nicht mit einem Muster beschreiben, das dem von Schüler:innen, Studierende und Lehrkräfte ähnelt: Didaktiker:innen nennen am häufigsten die Verständlichkeit der Visualisierung (45 von 131 Äußerungen innerhalb der Kategorie ‚Repräsentationen‘; 11,06 % aller Äußerungen). Des Weiteren führen sie vier weitere Kriterien in etwa vergleichbar häufig an – nutzen also eine Breite von Kriterien für die Beschreibung der Güte der instruktionalen Erklärung, die sich auf die Repräsentationen beziehen: den Einsatz bildlicher Visualisierungen (20 von 131 Äußerungen innerhalb der Kategorie ‚Repräsentationen‘; 4,91 % aller Äußerungen), die gelungene Verbindung der Repräsentationen (18 von 131 Äußerungen innerhalb der Kategorie ‚Repräsentationen‘; 4,42 % aller Äußerungen), die fachliche Angemessenheit der bildlichen Visualisierungen (16 von 131 Äußerungen innerhalb der Kategorie ‚Repräsentationen‘; 3,93 % aller Äußerungen) sowie die ausreichende Bezugnahme auf das Beispiel (11 von 131 Äußerungen innerhalb der Kategorie ‚Repräsentationen‘; 2,70 % aller Äußerungen).

Das neu aufgenommene Kriterium ‚Ästhetische Gestaltung‘ wird bei statusgruppenspezifischer Betrachtung nicht sehr häufig genutzt. Entgegen den Vermutungen, dass Schüler:innen dieses Kriterium im Vergleich zu den anderen Statusgruppen häufiger nutzen, verwenden sie es am seltensten (7 von 322 Äußerungen innerhalb der Kategorie ‚Repräsentationen‘; 0,62 % aller Äußerungen). Studierende (5 von 94 Äußerungen innerhalb der Kategorie ‚Repräsentationen‘; 1,28 % aller Äußerungen) und Lehrkräfte beziehen sich im Vergleich (4 von 60 Äußerungen innerhalb der Kategorie ‚Repräsentationen‘; 1,59 % aller Äußerungen) häufiger auf das eher oberflächliche Merkmal.

Die ebenfalls am Material gebildete Kategorie ‚Ausreichende Bezugnahme auf das Beispiel‘ wird von Lehrkräften (7 von 60 Äußerungen innerhalb der Kategorie ‚Repräsentationen‘; 2,78 % aller Äußerungen) und Didaktiker:innen (11 von 131 Äußerungen innerhalb der Kategorie ‚Repräsentationen‘; 2,70 % aller Äußerungen) relativ häufig angesprochen.

Durch die statusgruppenspezifische Analyse kann das Ergebnis aus Kapitel 6.2.1.1 weiter differenziert werden: Es sind nicht ausschließlich Schüler:innen, die sich auf das bloße Vorhandensein von externen Repräsentationen beziehen. Auch alle weiteren Statusgruppen führen diese Kriterien häufig an. Für Didaktiker:innen gilt jedoch auch, dass sie häufiger auf die Verständlichkeit der Visualisierungen achten als alle anderen Statusgruppen. Studierende und Didaktiker:innen führen überdies die fachliche Angemessenheit an. Lehrkräfte und Didaktiker:innen beziehen sich weiterhin auf die ausreichende Bezugnahme auf das Beispiel, Didaktiker:innen zusätzlich auf die gelungene Verbindung zwischen den Repräsentationen.

Die folgende Tabelle (Tab. 34) gibt noch einmal die in 6.2.1.1 getroffene Unterteilung in Kriterien des bloßen Vorhandenseins externer Repräsentationen¹⁵⁹, Kriterien, die sich auf die Gestaltung von oberflächlich wahrnehmbaren Aspekten beziehen,¹⁶⁰ und Kriterien, die sich auf die Gestaltung von lernrelevanten Aspekten der externen Repräsentationen beziehen,¹⁶¹ je Statusgruppe wieder.

Tabelle 34: Absolute und relative Häufigkeiten der Äußerungen zum Aspekt der ‚Repräsentationen‘ getrennt nach Verweisen auf das Vorhandensein von Repräsentationen, oberflächliche Gestaltungsmerkmale und lernrelevante Gestaltungsmerkmale getrennt nach Statusgruppe über alle Videovignetten hinweg

Statusgruppe	Schüler:innen ($N_{max} = 122$)		Studierende ($N_{max} = 51$)		Lehrkräfte ($N_{max} = 34$)		Didaktiker:innen ($N_{max} = 30$)		Summe über alle Statusgruppen ($N_{max} = 237$)	
	A.H.	R.H.	A.H.	R.H.	A.H.	R.H.	A.H.	R.H.	A.H.	R.H.
Vorhandensein von externen Repräsentationen	227	20,00	36	9,21	25	9,92	29	7,13	317	14,51
Oberflächliche Merkmale der externen Repräsentationen	10	0,88	9	2,30	8	3,17	6	1,47	33	1,51
Lernrelevante Merkmale der externen Repräsentationen	85	7,49	49	12,53	27	10,71	96	23,59	257	11,76

Bem.: N_{max} : aufgrund fehlender Werte bei einzelnen Videos den Berechnungen maximal zugrunde liegende Personenanzahl; A.H.: absolute Häufigkeit; R.H.: relative Häufigkeit.

Wie bereits angedeutet zeigt sich, dass Schüler:innen deutlich häufiger in ihren Begründungen auf das Vorhandensein rekurrieren als auf lernrelevante Merkmale der Repräsentationen. Für die Statusgruppen der Metaperspektive sind stets mehr Äußerungen mit Bezug auf lernrelevante Merkmale der externen Repräsentationen als solche mit Bezug auf das bloße Vorhandensein zu verzeichnen. Jedoch ist bei Studierenden und Lehrkräften das Verhältnis zwischen den Bezügen auf das Vorhandensein und auf die lernrelevanten Merkmale deutlich von dem bei Didaktiker:innen zu

¹⁵⁹ Die Anzahl ergibt sich aus folgenden Subkategorien: Repräsentationen global, Einsatz bildlicher Visualisierungen, Einsatz von Analogien, Metaphern und Vergleichen bzw. Themenspezifisch: Vergleich bei Mukoviszidose sowie Einsatz von Beispielen.

¹⁶⁰ Die Anzahl ergibt sich aus den Subkategorien ‚Sichtbarkeit der Visualisierung‘ und ‚Ästhetische Gestaltung der Visualisierung‘.

¹⁶¹ Die Anzahl ergibt sich aus den Subkategorien Erklärbedürftigkeit der Grafik, fachliche Angemessenheit, Verständlichkeit der Visualisierung, ausreichende Bezugnahme auf das Beispiel, gelungene Verbindung der Repräsentationen.

unterscheiden. Nur bei Didaktiker:innen ist eine sehr deutliche Verschiebung zugunsten der Bewertung anhand lernrelevanter Merkmale von Repräsentationen erkennbar.

Zusammenfassend zeigen die Daten zwei besonders auffällige Werte: Schüler:innen nehmen mit Abstand am häufigsten Bezug auf das Vorhandensein von externen Repräsentationen. Sie beurteilen die Erklärung im Falle der Repräsentationen damit überwiegend anhand einer eher oberflächlichen Sichtstruktur. Didaktiker:innen beziehen sich hingegen auffällig häufig auf Merkmale der externen Repräsentationen, die das Lernen mit diesen fördern oder hemmen können. Studierende und Lehrkräfte nehmen jeweils auch am häufigsten auf lernrelevante Merkmale der Repräsentationen Bezug, jedoch mit einer deutlich geringeren Differenz zum Vorhandensein externer Repräsentationen.

Sprech- und Körperausdruck

In der Übersicht zu den Häufigkeiten getrennt nach Statusgruppe (Tab. 35) wird sichtbar, dass Studierende mehr als doppelt so häufig Bezug auf die Gestaltung des Sprech- und Körperausdrucks nehmen als die weiteren drei Statusgruppen. Wie in Kapitel 6.1.1 im Zuge der Konfundierungen erörtert wurde, muss hierbei berücksichtigt werden, dass durch die Teilnahme an Seminaren zum guten Erklären auch Wissen zum Sprech- und Körperausdruck vermittelt wurde und daher Studierende mitunter auch deshalb auffällig häufig bei der Begründung auf den Sprech- und Körperausdruck rekurren. Didaktiker:innen nehmen am seltensten Bezug auf den Sprech- und Körperausdruck der erklärenden Person.

Tabelle 35: Häufigkeiten der Äußerungen, die der Hauptkategorie ‚Sprech- und Körperausdruck‘ zugeordnet werden können, getrennt nach Statusgruppe über alle Videovignetten hinweg

Statusgruppe Häufigkeit	Schüler:innen ($N_{max} = 122$)	Studierende ($N_{max} = 51$)	Lehrkräfte ($N_{max} = 34$)	Didaktiker:innen ($N_{max} = 30$)	Summe über alle Statusgruppen ($N_{max} = 237$)
Absolute Häufigkeiten	46	41	11	12	110
Relative Häufigkeiten in %	4,05	10,49	4,37	2,95	5,03

Bem.: N_{max} : aufgrund fehlender Werte bei einzelnen Videos den Berechnungen maximal zugrunde liegende Personenanzahl.

Die folgende Abbildung (Abb. 20) gibt Aufschluss über die Verteilung der Äußerungen auf Ebene der einzelnen Kriterien.

Schüler:innen nennen häufig prosodische Merkmale bei der Begründung der Notenvergabe (13 von 46 Äußerungen innerhalb der Kategorie ‚Sprech- und Körperausdruck‘; 1,15 % aller Äußerungen), wobei sie häufig direkt auf die Wirkung der Sprechweise eingehen. So äußern Schüler:innen beispielsweise:

„redet zu unsicher“ (Schüler:in/Ö_k)

„Das hat bei der Frau auch eben etwas auswenig gelernt gewirkt, deshalb das Minus“
(Schüler:in/Ö_k)

Auch auf die Sprechgeschwindigkeit (10 von 46 Äußerungen innerhalb der Kategorie ‚Sprech- und Körperausdruck‘; 0,88 % aller Äußerungen) und die Variation der Stimme (9 von 46 Äußerungen

innerhalb der Kategorie ‚Sprech- und Körperausdruck‘; 0,79 % aller Äußerungen) entfallen noch verhältnismäßig viele Äußerungen. Nicht zur Begründung genutzt werden die Kategorien ‚Stimmklang‘, ‚Sprechpausen‘ sowie ‚Zugewandter Körperausdruck‘. Sehr selten werden die Sprechflüssigkeit (4 von 46 Äußerungen innerhalb der Kategorie ‚Sprech- und Körperausdruck‘; 0,35 % aller Äußerungen) und der unterstützende Körperausdruck (3 von 46 Äußerungen innerhalb der Kategorie ‚Sprech- und Körperausdruck‘; 0,26 % aller Äußerungen) genannt.

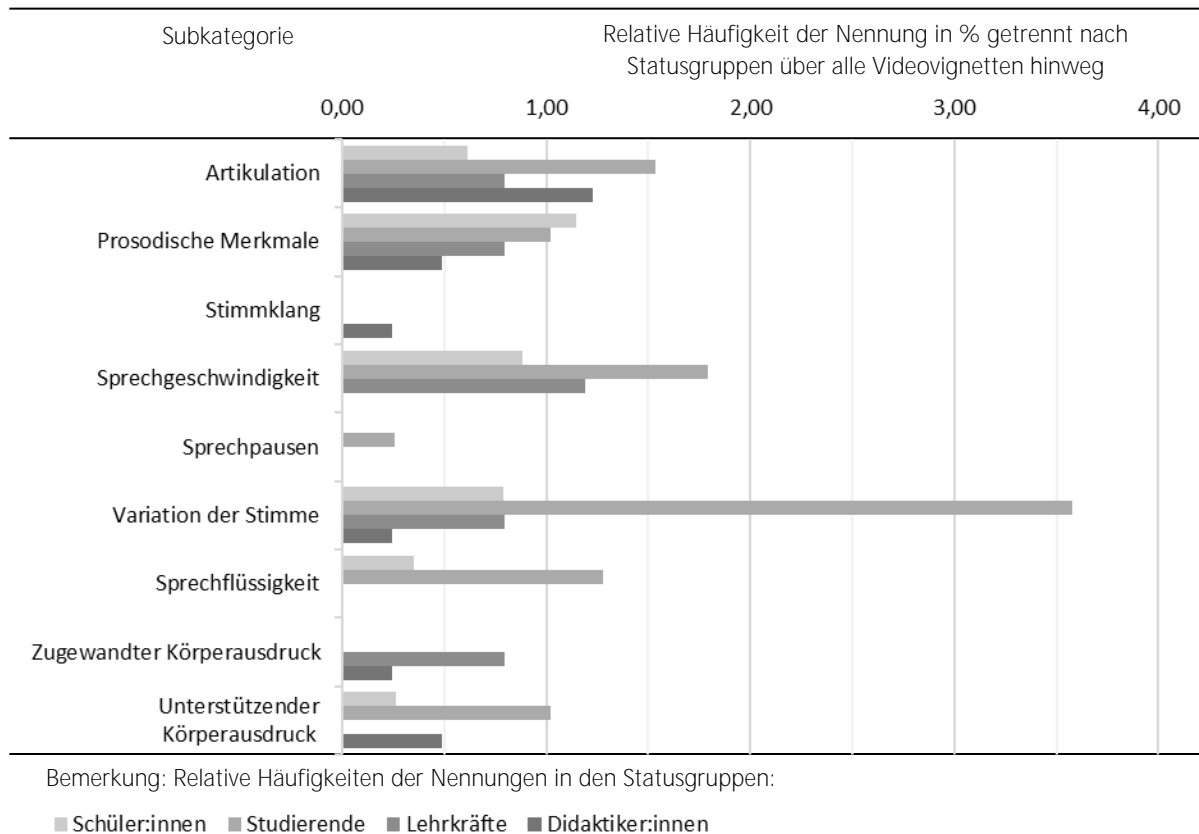


Abbildung 20: Relative Häufigkeiten der Nennungen der Subkategorien aus der Hauptkategorie ‚Sprech- und Körperausdruck‘ getrennt nach Statusgruppe (Schüler:innen $N_{max} = 122$; Studierende $N_{max} = 51$; Lehrkräfte $N_{max} = 34$; Didaktiker:innen $N_{max} = 30$) über alle Videovignetten hinweg

Studierende beziehen sich am häufigsten auf die Variation der Stimme (14 von 41 Äußerungen innerhalb der Kategorie ‚Sprech- und Körperausdruck‘; 3,58 % aller Äußerungen). Deutlich seltener – aber im Vergleich zu den weiteren Statusgruppen immer noch häufiger – werden die Kriterien ‚Sprechgeschwindigkeit‘ (7 von 41 Äußerungen; 1,79 % aller Äußerungen), ‚Artikulation‘ (6 von 41 Äußerungen; 1,53 % aller Äußerungen) und ‚Sprechflüssigkeit‘ (5 von 41 Äußerungen; 1,28 % aller Äußerungen) angeführt. Die prosodischen Merkmale sowie der unterstützende Körperausdruck werden gleich häufig herangezogen (je 4 von 41 Äußerungen; 1,02 % aller Äußerungen). Auch Studierende nutzen die Kategorie Sprechflüssigkeit nicht zur Begründung des Globalurteils. Nur eine Person nimmt in einem Fall Bezug auf die Sprechpausen (1 von 41 Äußerungen; 0,26 % aller Äußerungen).

Das am Material gebildete Kriterium ‚Sprechflüssigkeit‘ wird von Schüler:innen (4 von 46 Äußerungen innerhalb der Kategorie ‚Sprech- und Körperausdruck‘; 0,35 % aller Äußerungen) und Studierenden (5 von 45 Äußerungen innerhalb der Kategorie ‚Sprech- und Körperausdruck‘; 1,28 %

aller Äußerungen) moderat genutzt. Aufgrund der wenigen Nennung zum Aspekt des Sprech- und Körperausdrucks erhält es innerhalb der Kategorie jedoch anteilig relativ viele Äußerungen.

Lehrkräfte nehmen insgesamt sehr selten Bezug auf den Sprech- und Körperausdruck. Genutzt werden die Kategorien der Artikulation, der prosodischen Merkmale, der Variation der Stimme und des zugewandten Körperausdrucks je zweimal (0,79 % aller Äußerungen) und die Kategorie der Sprechgeschwindigkeit dreimal (1,19% aller Äußerungen). Nicht genutzt werden die Kriterien des Stimmklangs, der Sprechpausen, der Sprechflüssigkeit sowie der unterstützende Körperausdruck.

Bei Didaktiker:innen verhält es sich ähnlich, wobei es eine leichte Häufung in der Kategorie zur Artikulation gibt (5 von 12 Äußerungen innerhalb der Kategorie ‚Sprech- und Körperausdruck‘; 1,23 % aller Äußerungen). Didaktiker:innen nutzen ebenfalls viele Kriterien nicht bzw. nur selten: Keine Begründung enthält Äußerungen zu der Sprechgeschwindigkeit, den Sprechpausen sowie zur Sprechflüssigkeit. Bis auf die Artikulation werden alle weiteren Kriterien nur selten angesprochen und sind im Material jeweils nur ein oder zwei Mal zu finden (was 0,25 % bzw. 0,49 % aller Äußerungen entspricht).

Insgesamt kann formuliert werden, dass auf den Sprech- und Körperausdruck von den verschiedenen Statusgruppen deutlich unterschiedlich häufig rekurriert wird. Studierende nehmen – mitunter auch bedingt durch den Seminarbesuch – am häufigsten darauf Bezug. Schüler:innen und Lehrkräfte beziehen sich etwa gleich häufig auf den Sprech- und Körperausdruck und Didaktiker:innen noch seltener. Die Statusgruppen nutzen dabei eine vergleichbare Breite von Kriterien, jedoch zeichnet sich ab, dass Studierende vor allem die Variation der Stimme und Didaktiker:innen die Artikulation in den Blick nehmen, um die gesehenen Erklärungen zu bewerten.

Die folgende Tabelle (Tab. 36) stellt die Ergebnisse noch einmal getrennt nach Facetten des Sprech- und Körperausdrucks zusammen, die auf Basis der Theorie auch in Kapitel 6.2.1.1 aufgegriffen wurden.

Tabelle 36: Absolute und relative Häufigkeiten der Äußerungen, die den verschiedenen Facetten des Aspekts der ‚Sprech- und Körperausdruck‘ zugeordnet werden können, getrennt nach Statusgruppe über alle Videovignetten hinweg

Statusgruppe	Schüler:innen ($N_{max} = 122$)		Studierende ($N_{max} = 51$)		Lehrkräfte ($N_{max} = 34$)		Didaktiker:innen ($N_{max} = 30$)		Summe über alle Statusgruppen ($N_{max} = 237$)	
	A.H.	R.H.	A.H.	R.H.	A.H.	R.H.	A.H.	R.H.	A.H.	R.H.
Artikulation	7	0,62	6	1,53	2	0,79	5	1,23	20	0,92
Prosodie	32	2,82	26	6,65	7	2,78	4	0,98	69	3,16
Sprechflüssigkeit	4	0,35	5	1,28	0	0,00	0	0,00	9	0,41
Kinesik	3	0,26	4	1,02	2	0,79	3	0,74	12	0,55
Proxemik	0	0,00	0	0,00	0	0,00	0	0,00	0	0,00

Bem.: N_{max} : aufgrund fehlender Werte bei einzelnen Videos den Berechnungen maximal zugrunde liegende Personenanzahl; A.H.: absolute Häufigkeit; R.H.: relative Häufigkeit.

Insgesamt beziehen sich damit Schüler:innen, Studierende und Lehrkräfte am häufigsten auf Facetten der Prosodie und Didaktiker:innen am häufigsten auf die Artikulation.

Verständliche Sprache

Die Daten in Tabelle 37 zeigen, dass Schüler:innen bei ihren Bewertungen der instruktionalen Erklärungen seltener als Studierende, Lehrkräfte und Didaktiker:innen auf die sprachlichen Merkmale der Erklärung rekurrieren. Dieses Ergebnis ist insofern plausibel, als eine Bewertung der sprachlichen Verständlichkeit sehr spezifisches Wissen über diese voraussetzt und die Sprache durch ihre Flüchtigkeit zudem weniger gut zugänglich ist. In Hinblick auf die Ergebnisse zu möglichen Konfundierungen (siehe Kap. 6.1.1) muss beachtet werden, dass die relative Häufigkeit der Nennung von Kriterien, die sich auf die sprachliche Verständlichkeit beziehen, für die Statusgruppe der Studierenden mitunter etwas überschätzt ist, da diejenigen Studierenden, die ein Seminar zum Thema ‚Erklären‘ besuchten, häufiger in ihren Begründungen darauf Bezug nahmen.

Tabelle 37: Häufigkeiten der Äußerungen, die der Hauptkategorie ‚Verständliche Sprache‘ zugeordnet werden können, getrennt nach Statusgruppe über alle Videovignetten hinweg

Statusgruppe	Schüler:innen ($N_{max} = 122$)	Studierende ($N_{max} = 51$)	Lehrkräfte ($N_{max} = 34$)	Didaktiker:innen ($N_{max} = 30$)	Summe über alle Statusgruppen ($N_{max} = 237$)
Absolute Häufigkeiten	16	31	19	27	93
Relative Häufigkeiten in %	1,41	7,93	7,54	6,63	4,26

Bem.: N_{max} : aufgrund fehlender Werte bei einzelnen Videos den Berechnungen maximal zugrunde liegende Personenanzahl.

Für eine differenzierte Betrachtung zeigt die nachstehende Abbildung 21 die Häufigkeiten pro Kriterium.

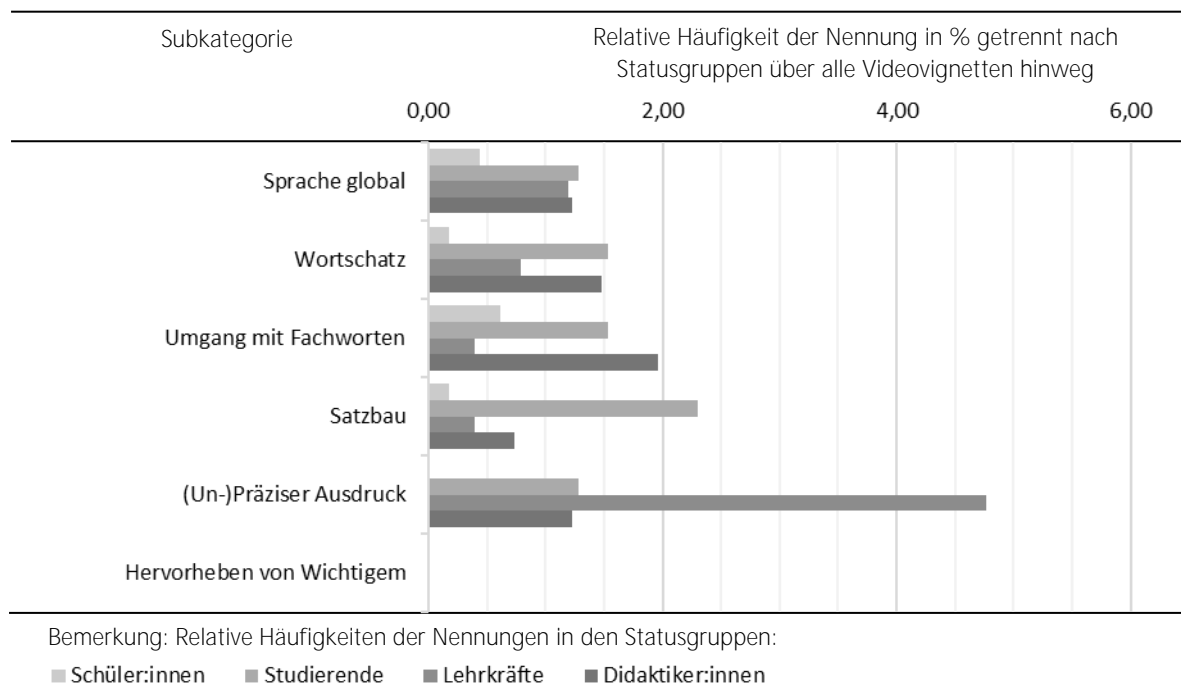


Abbildung 21: Relative Häufigkeiten der Nennungen der Subkategorien aus der Hauptkategorie ‚Verständliche Sprache‘ getrennt nach Statusgruppe (Schüler:innen $N_{max} = 122$; Studierende $N_{max} = 51$; Lehrkräfte $N_{max} = 34$; Didaktiker:innen $N_{max} = 30$) über alle Videovignetten hinweg

Ersichtlich wird, dass Schüler:innen neben einigen wenigen globalen Äußerungen zur Sprache (5 von 16 Äußerungen innerhalb der Kategorie ‚Sprachliche Verständlichkeit‘; 0,44 % aller Äußerungen) vor allem Kriterien nutzen, die die Wortebene der sprachlichen Verständlichkeit betreffen. Dabei wird von Schüler:innen vor allem der Umgang mit Fachworten in den Blick genommen (7 von 16 Äußerungen innerhalb der Kategorie ‚Verständliche Sprache‘; 0,62 % aller Äußerungen). Auf alle weiteren Kriterien entfallen kaum Nennungen bzw. auf den (un-)präzisen Ausdruck keine Nennungen.

Studierende nutzen hingegen Kriterien, die sich auf verschiedene Ebenen der sprachlichen Verständlichkeit beziehen – alle Kriterien werden etwa gleich häufig zur Begründung der Notenvergabe genutzt, wobei der Satzbau im Vergleich etwas häufiger auftaucht (9 von 19 Äußerungen innerhalb der Kategorie ‚Verständliche Sprache‘; 2,30 % aller Äußerungen).

Bei Lehrkräften entfallen die meisten Äußerungen auf das Kriterium des präzisen Ausdrucks – 12 von insgesamt 19 Äußerungen innerhalb der Kategorie ‚Verständliche Sprache‘ beziehen sich darauf, was 4,76 % aller Äußerungen entspricht. Das Kriterium macht für die befragten Lehrkräfte damit einen Großteil der Äußerungen aus, während Schüler:innen dieses Kriterium gar nicht nutzen. Selten werden von Lehrkräften hingegen der Umgang mit Fachworten sowie der Satzbau zur Begründung des Globalurteils herangezogen (jeweils 1 von 19 Äußerungen innerhalb der Kategorie ‚Verständliche Sprache‘; 0,40 % aller Äußerungen).

Didaktiker:innen beziehen sich bei ihren Begründungen – eher vergleichbar mit Studierenden – etwa ähnlich häufig auf die unterscheidbaren Kriterien verständlicher Sprache. Lediglich auf den Satzbau (3 von 27 Äußerungen innerhalb der Kategorie ‚Verständliche Sprache‘; 0,74 % aller Äußerungen) entfallen im Verhältnis etwas weniger Äußerungen als auf die anderen Kriterien.

Zusammenfassend wird die sprachliche Verständlichkeit von allen Statusgruppen eher selten in den selbst herangezogenen Kriterien gefunden, sodass es aufgrund der wenigen Nennungen innerhalb der Kategorie ‚Sprachliche Verständlichkeit‘ kaum Auffälligkeiten gibt. Studierende und Didaktiker:innen nutzen deutlich mehr Kriterien als Schüler:innen, die sich hauptsächlich auf den Umgang mit Fachworten beziehen, und Lehrkräfte, die sich meist auf den präzisen Ausdruck der erklärenden Person Bezug nehmen.

Auf Ebene der Facetten ergibt sich ein Bild, das sich bereits auf Ebene der Kriterien abgezeichnet hat (siehe Tab. 38).

Tabelle 38: Absolute und relative Häufigkeiten der Äußerungen, die den verschiedenen Facetten des Aspekts ‚Verständliche Sprache‘ zugeordnet werden können, getrennt nach Statusgruppe über alle Videovignetten hinweg

Statusgruppe	Schüler:innen ($N_{max} = 122$)		Studierende ($N_{max} = 51$)		Lehrkräfte ($N_{max} = 34$)		Didaktiker:innen ($N_{max} = 30$)		Summe über alle Statusgruppen ($N_{max} = 237$)	
	A.H.	R.H.	A.H.	R.H.	A.H.	R.H.	A.H.	R.H.	A.H.	R.H.
Wortebene	9	0,80	12	3,07	3	1,19	14	3,44	38	1,74
Satzebene	2	0,18	14	3,84	13	5,16	8	1,97	37	1,69
Textebene	0	0,00	0	0,00	0	0,00	0	0,00	0	0,00

Bem.: N_{max} : aufgrund fehlender Werte bei einzelnen Videos den Berechnungen maximal zugrunde liegende Personenanzahl; A.H.: absolute Häufigkeit; R.H.: relative Häufigkeit.

Schüler:innen führen in den wenigen Nennungen hauptsächlich Kriterien an, die der Wortebene zuzuordnen sind. Studierende beziehen sich in etwa ausgewogen auf Kriterien der Wort- und Satzebene. Lehrkräfte nehmen insbesondere Kriterien auf Satzebene und weniger häufig solche auf Wortebene in den Blick. Didaktiker:innen nutzen wie Schüler:innen vermehrt jene Kriterien, die der Wortebene zuzuordnen sind.

Fachlicher Aspekt

Die nachstehende Tabelle (Tab. 39) zeigt, dass innerhalb der Hauptkategorie des fachlichen Aspekts die relativen Häufigkeiten die größten Differenzen zwischen den befragten Statusgruppen aufweisen.

Tabelle 39: Häufigkeiten der Äußerungen, die der Hauptkategorie ‚Fachlicher Aspekte‘ zugeordnet werden können, getrennt nach Statusgruppe über alle Videovignetten hinweg

Statusgruppe	Schüler:innen ($N_{max} = 122$)	Studierende ($N_{max} = 51$)	Lehrkräfte ($N_{max} = 34$)	Didaktiker:innen ($N_{max} = 30$)	Summe über alle Statusgruppen ($N_{max} = 237$)
Häufigkeit					
Absolute Häufigkeiten	37	53	40	99	229
Relative Häufigkeiten in %	3,26	13,55	15,87	24,32	10,48

Bem.: N_{max} : aufgrund fehlender Werte bei einzelnen Videos den Berechnungen maximal zugrunde liegende Personenanzahl.

So beziehen sich Schüler:innen erwartungsgemäß selten auf fachliche Aspekte bei der Begründung der Notenvergabe, da ihnen für viele Kriterien notwendiges Fachwissen fehlt. Innerhalb der Metaperspektive, also bei den Statusgruppen der Lehrenden, die insbesondere innerhalb des Studiums das entsprechende Fachwissen erwerben oder erworben haben, begründen vor allem Didaktiker:innen die Bewertung der Erklärung über den fachlichen Aspekt. Die Betrachtung auf Ebene der einzelnen Subkategorien kann auch hier Aufschluss über die angelegten Kriterien für Erklärungen geben.

Abbildung 22 macht deutlich, dass Schüler:innen lediglich nennenswerte Häufigkeiten in den Kategorien ‚Fehlende“ Inhalte, die über das Erklärziel hinausgehen‘ (16 von 37 Äußerungen innerhalb der Kategorie ‚Fachlicher Aspekt‘; 1,41 % aller Äußerungen) und ‚Zusammenhänge werden deutlich‘ (16 von 37 Äußerungen innerhalb der Kategorie ‚Fachlicher Aspekt‘; 1,41 % aller Äußerungen) aufweisen. Diese Ergebnisse sind insofern plausibel, als beide Kategorien auch in Zusammenhang mit der Adressatenorientierung stehen. Die Kriterien ‚Fach global‘, ‚Fachspezifische Erklärungsformen‘, ‚Berücksichtigung von Nichtwissen‘, ‚Phänomenorientierung‘ und ‚Prinzipienorientierung‘ werden nicht genannt. Selten nutzen Schüler:innen die fachliche Korrektheit (3 von 37 Äußerungen innerhalb der Kategorie ‚Fachlicher Aspekt‘; 0,26 % aller Äußerungen), die Vollständigkeit (1 von 37 Äußerungen innerhalb der Kategorie ‚Fachlicher Aspekt‘; 0,09 % aller Äußerungen) sowie das Kriterium zur Vagheit (4 von 37 Äußerungen innerhalb der Kategorie ‚Fachlicher Aspekt‘; 0,35 % aller Äußerungen) für die Begründung der Notenvergabe.

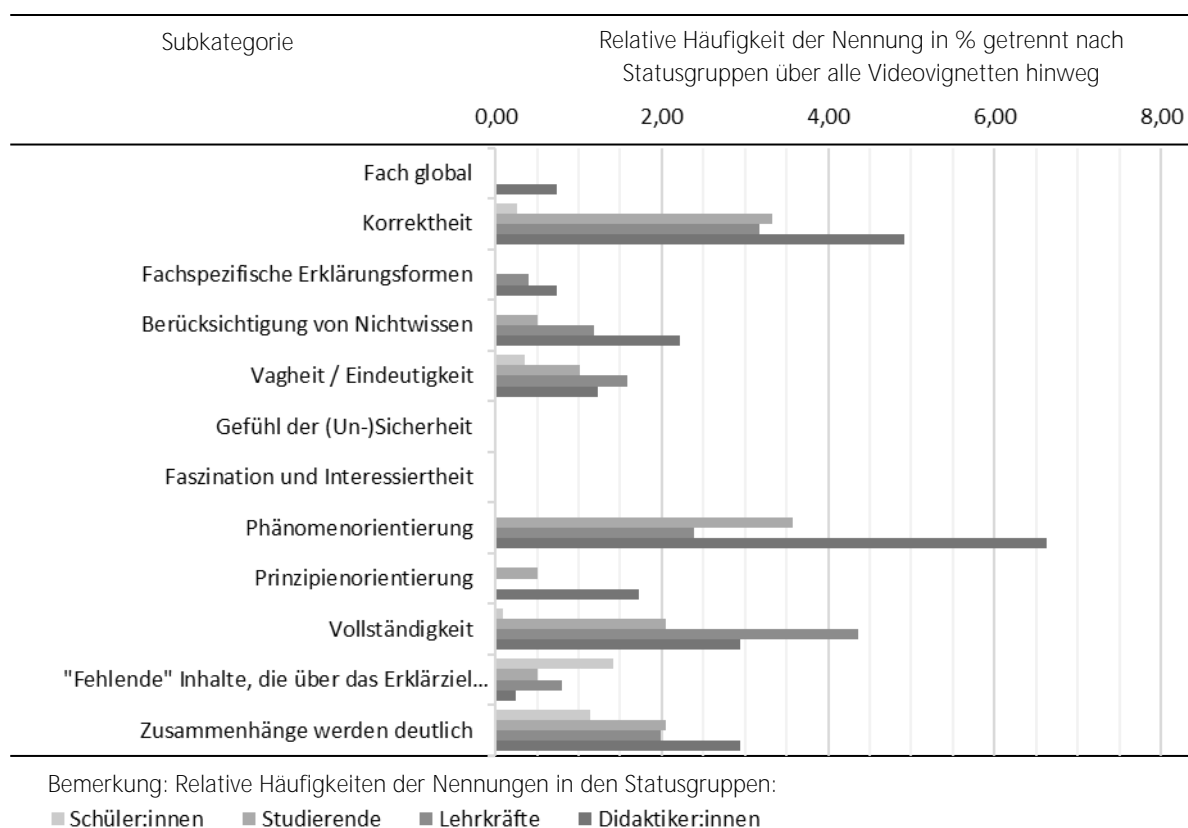


Abbildung 22: Relative Häufigkeiten der Nennungen der Subkategorien aus der Hauptkategorie ‚Fachlicher Aspekt‘ getrennt nach Statusgruppe (Schüler:innen $N_{max} = 122$; Studierende $N_{max} = 51$; Lehrkräfte $N_{max} = 34$; Didaktiker:innen $N_{max} = 30$) über alle Videovignetten hinweg

Studierende nennen am häufigsten die Kriterien ‚Korrektheit‘ (13 von 53 Äußerungen innerhalb der Kategorie ‚Fachlicher Aspekt‘; 3,32 % aller Äußerungen) und ‚Phänomenorientierung‘ (14 von 37 Äußerungen innerhalb der Kategorie ‚Fachlicher Aspekt‘; 3,58 % aller Äußerungen). Auch ihre Begründungen enthalten keine Äußerungen, die den Kategorien ‚Fach global‘ und ‚Fachspezifische Erklärungsformen‘ zugeordnet werden können. Selten nutzen sie darüber hinaus die Berücksichtigung des Nichtwissens (2 von 37 Äußerungen innerhalb der Kategorie ‚Fachliche Aspekte‘; 0,51 % aller Äußerungen) sowie die Prinzipienorientierung (2 von 37 Äußerungen innerhalb der Kategorie ‚Fachlicher Aspekt‘; 0,51 % aller Äußerungen) zur Begründung.

Lehrkräfte bedienen sich ähnlicher Kriterien wie Studierende: Neben der ‚Korrektheit‘ (8 von 40 Äußerungen innerhalb der Kategorie ‚Fachliche Aspekte‘; 3,17 % aller Äußerungen) und der ‚Phänomenorientierung‘ (6 von 40 Äußerungen innerhalb der Kategorie ‚Fachlicher Aspekt‘; 2,38 % aller Äußerungen) nennen sie jedoch die Vollständigkeit am häufigsten von allen Kriterien (11 von 40 Äußerungen innerhalb der Kategorie ‚Fachlicher Aspekt‘; 4,37 % aller Äußerungen), die dem fachlichen Aspekt zugeordnet werden können. Die Prinzipienorientierung wird von Lehrkräften – wie auch von Schüler:innen – nicht zur Begründung der Notenvergabe herangezogen, Gleiches gilt für Äußerungen, die sich global auf das Fach beziehen. Eine Person führt zur Begründung die fachspezifischen Erklärungsform an.

Didaktiker:innen führen wie Studierende am häufigsten die ‚Phänomenorientierung‘ (27 von 99 Äußerungen innerhalb der Kategorie ‚Fachlicher Aspekt‘; 6,63 % aller Äußerungen) an. Die Kategorie markiert damit einen interessanten Unterschied zwischen den Perspektiven. Auf die ‚Korrektheit‘ wird ebenfalls häufig rekuriert (20 von 99 Äußerungen innerhalb der Kategorie ‚Fachlicher Aspekt‘; 4,91 % aller Äußerungen) an. Darüber hinaus werden auch die Vollständigkeit

(12 von 99 Äußerungen innerhalb der Kategorie ‚Fachlicher Aspekt‘; 2,95 % aller Äußerungen), ‚Zusammenhänge werden deutlich‘ (ebenfalls 12 von 99 Äußerungen innerhalb der Kategorie ‚Fachlicher Aspekt‘; 2,95 % aller Äußerungen) und die Berücksichtigung von Nichtwissen (9 von 99 Äußerungen innerhalb der Kategorie ‚Fachlicher Aspekt‘; 2,21 % aller Äußerungen) als Kriterien für die Bewertung herangezogen.

Auffällig ist weiterhin das Ergebnis, dass Didaktiker:innen weitaus häufiger als alle weiteren Statusgruppen – bzw. häufiger als die Studierenden, die als einzige weitere Statusgruppe darauf Bezug nehmen – die Prinzipienorientierung anführen. Auch die fachspezifischen Erklärungsformen werden als Kriterium hauptsächlich von Didaktiker:innen genannt, wobei die Anzahl der Nennungen hier insgesamt gering ausfällt (3 von 99 Äußerungen innerhalb der Kategorie ‚Fachlicher Aspekt‘; 0,74 % aller Äußerungen).

Das am Material gebildete Kriterium ‚Zusammenhänge werden deutlich‘ wird von allen vier Statusgruppen genutzt: von Schüler:innen (13 von 37 Äußerungen innerhalb der Kategorie ‚Fachlicher Aspekt‘; 1,15 % aller Äußerungen), Studierenden (8 von 53 Äußerungen innerhalb der Kategorie ‚Fachlicher Aspekt‘; 2,05 % aller Äußerungen), Lehrkräften (5 von 40 Äußerungen innerhalb der Kategorie ‚Fachlicher Aspekt‘; 1,98 % aller Äußerungen) und Didaktiker:innen (12 von 99 Äußerungen innerhalb der Kategorie ‚Fachlicher Aspekt‘; 2,95 % aller Äußerungen).

Zusammenfassend wird auf der Ebene der Kriterien deutlich, dass Schüler:innen lediglich auf zwei Kriterien Bezug nehmen, die sich auch anhand einer Perspektive einschätzen lassen, die kein unmittelbares professionelles Wissen voraussetzt. Die Statusgruppen der Metaperspektive beziehen sich in ihren Begründungen alle auf die Korrektheit der Erklärung sowie auf die Phänomenorientierung, wobei sich die Häufigkeitsverhältnisse unterscheiden und Didaktiker:innen deutlich häufiger auf die Phänomenorientierung eingehen. Lehrkräfte beziehen sich häufig noch auf die Vollständigkeit. Auffällig ist, dass Didaktiker:innen nicht nur häufiger auf fachliche Aspekte rekurren, sondern gemessen an der Anzahl häufig genannter Kriterien, die Bewertungen auch breiter ausfallen.

Betrachtet man die Ergebnisse noch in Hinblick auf die unterscheidbaren Facetten innerhalb des fachlichen Aspekts ergibt sich folgendes Bild (Tab. 40):

Tabelle 40: Absolute und relative Häufigkeiten der Äußerungen, die den verschiedenen Facetten des fachlichen Aspekts zugeordnet werden können, getrennt nach Statusgruppe über alle Videovignetten hinweg

Statusgruppe	Schüler:innen ($N_{max} = 122$)		Studierende ($N_{max} = 51$)		Lehrkräfte ($N_{max} = 34$)		Didaktiker:innen ($N_{max} = 30$)		Summe über alle Status- gruppen ($N_{max} = 237$)	
	A.H.	R.H.	A.H.	R.H.	A.H.	R.H.	A.H.	R.H.	A.H.	R.H.
Facette										
Fachliche Korrektheit	7	0,62	19	4,86	16	6,35	37	9,09	79	3,62
Berücksichtigung der Auswirkung des Gesagten	0	0,00	0	0,00	0	0,00	0	0,00	0	0,00
Phänomen- vs. Prinzipienorientierung	0	0,00	16	4,09	6	2,38	34	8,35	56	2,56
Fachliche Vollständigkeit	30	2,64	18	4,60	18	7,14	25	6,14	91	4,16

Bem.: N_{max} : aufgrund fehlender Werte bei einzelnen Videos den Berechnungen maximal zugrunde liegende Personenanzahl; A.H.: absolute Häufigkeit; R.H.: relative Häufigkeit.

Während die Kriterien der Studierenden über alle Facetten in etwa gleich verteilt sind, begründen Lehrkräfte die Notenvergabe etwas seltener anhand der Auswahl repräsentativer und geeigneter Phänomene sowie Didaktiker:innen etwas seltener an der Vollständigkeit der Erklärung. Schüler:innen begründen die Notenvergabe meist mit der fachlichen Vollständigkeit, wobei diese aus ihrer subjektiven Sicht heraus einbezogen und nicht anhand von fachlichen Referenzpunkten beurteilt wird.

Globale Äußerungen

Die Übersicht (Tab. 41) verdeutlicht, dass Schüler:innen erheblich häufiger als alle weiteren Statusgruppen globale Aspekte zur Begründung der Notenvergabe anführen (27,14 % aller Äußerungen). In den Bewertungen der Schüler:innen ist damit etwas mehr als jedes vierte Kriterium, das genannt wird, als global zu bezeichnen. Studierende tun dies nur etwa halb so häufig (13,04 % aller Äußerungen) – Lehrkräfte (11,11 % aller Äußerungen) und Didaktiker:innen (9,09 % aller Äußerungen) noch seltener.

Tabelle 41: Häufigkeiten der Äußerungen, die der Hauptkategorie ‚Globale Aspekte‘ zugeordnet werden können, getrennt nach Statusgruppe über alle Videovignetten hinweg

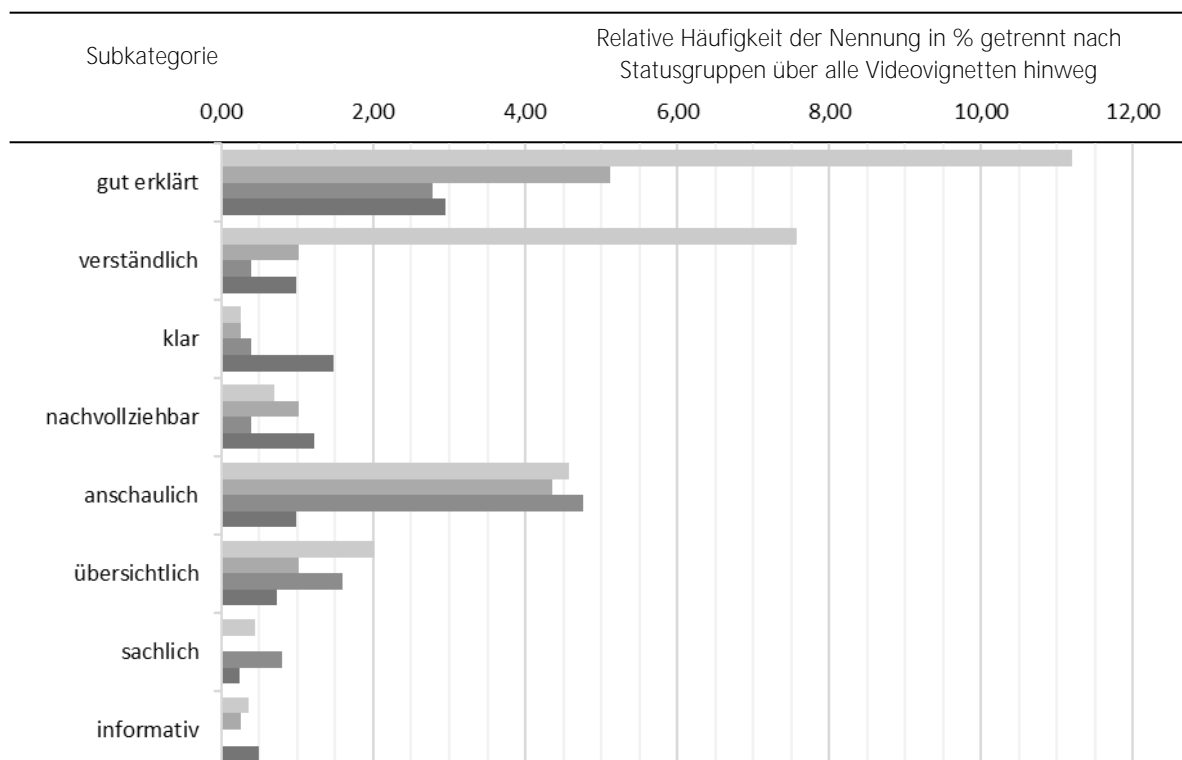
Statusgruppe	Schüler:innen ($N_{max} = 122$)	Studierende ($N_{max} = 51$)	Lehrkräfte ($N_{max} = 34$)	Didaktiker:innen ($N_{max} = 30$)	Summe über alle Statusgruppen ($N_{max} = 237$)
Häufigkeit					
Absolute Häufigkeiten	308	51	28	37	424
Relative Häufigkeiten in %	27,14	13,04	11,11	9,09	19,41

Bem.: N_{max} : aufgrund fehlender Werte bei einzelnen Videos den Berechnungen maximal zugrunde liegende Personenanzahl.

Abbildung (Abb. 23) differenziert den Blick auf die Verteilung der globalen Aussagen auf die acht Subkategorien innerhalb der Hauptkategorie.

Dabei wird deutlich, dass Schüler:innen häufig die am wenigsten differenzierte Äußerung ‚gut erklärt‘ als Begründung für die Bewertung der Notenvergabe heranziehen (127 von 303 Äußerungen in ‚Globale Aspekte‘; 11,19 % aller Äußerungen) – die Subkategorie, die unter den globalen Aspekten die globalste ist. Ebenfalls häufig im Material zu finden sind in den Bewertungen der Schüler:innen die Äußerung ‚verständlich‘ (86 von 303 Äußerungen in ‚Globale Aspekte‘; 7,58 % aller Äußerungen) und ‚anschaulich‘ (52 von 303 Äußerungen in ‚Globale Aspekte‘; 4,58 % aller Äußerungen). Insbesondere die Äußerung ‚verständlich‘ rückt dabei die Rezipient:innenperspektive der Schüler:innen in den Vordergrund: Nur sie können beurteilen, ob der Inhalt der (echten) instruktionalen Erklärung tatsächlich für sie verständlich ist. Selten werden ‚klar‘, ‚sachlich‘ und ‚informativ‘ genannt.

Die Statusgruppe der Studierenden nennt bei den Begründungen etwa vergleichbar häufig ‚gut erklärt‘ (20 von 51 Äußerungen in ‚Globale Aspekte‘; 5,12 % aller Äußerungen) und ‚anschaulich‘ (17 von 51 Äußerungen in ‚Globale Aspekte‘; 4,35 % aller Äußerungen). Keine Äußerung enthält das Kriterium ‚sachlich‘ als Begründung für die globale Bewertung der Erklärung und selten nutzen Studierende ‚klar‘ und ‚informativ‘.



Bemerkung: Relative Häufigkeiten der Nennungen in den Statusgruppen:

■ Schüler:innen ■ Studierende ■ Lehrkräfte ■ Didaktiker:innen

Abbildung 23: Relative Häufigkeiten der Nennungen der Subkategorien aus der Hauptkategorie ‚Globale Aspekte‘ getrennt nach Statusgruppe (Schüler:innen $N_{max} = 122$; Studierende $N_{max} = 51$; Lehrkräfte $N_{max} = 34$; Didaktiker:innen $N_{max} = 30$) über alle Videovignetten hinweg

Ähnlich zu diesem Begründungsverhalten nennt die Statusgruppe der Lehrkräfte ‚anschaulich‘ (12 von 28 Äußerungen in ‚Globale Aspekte‘; 4,76 % aller Äußerungen) und ‚gut erklärt‘ (7 von 28 Äußerungen in ‚Globale Aspekte‘; 2,78 % aller Äußerungen) innerhalb der globalen Aussagen in ihren Bewertungen. Keine Nennungen entfallen nur auf ‚informativ‘, selten werden ‚verständlich‘, ‚klar‘ sowie ‚nachvollziehbar‘ genutzt.

Auch die Statusgruppe der Didaktiker:innen nennt am häufigsten innerhalb der globalen Aussagen ‚gut erklärt‘ (12 von 37 Äußerungen in ‚Globale Aspekte‘; 2,95 % aller Äußerungen). Selten wird ‚sachlich‘ als Begründung angeführt, wobei alle weiteren globalen Aussagen im Material etwa ähnlich häufig vorkommen, was die Verteilung dieser Statusgruppe von den weiteren drei Gruppen differenziert, bei denen es jeweils zwei bzw. bei Schüler:innen drei sehr häufig vorkommende globale Aussagen gibt.

Zusammenfassend kann festgehalten werden, dass alle Statusgruppen Erklärungen pauschal mit ‚gut erklärt‘ bewerten, Schüler:innen dies jedoch mit deutlichem Abstand am häufigsten tun. Für Schüler:innen ist die Verständlichkeit weiterhin ein wichtiges überspannendes Kriterium. In der Zusammenschau der Verteilungen ist auffällig, dass die Anschaulichkeit der Erklärung von den Statusgruppen, die der schulischen Praxis näher sind – Schüler:innen, Studierende, Lehrkräfte –, häufiger als Kriterium herangezogen wird als von Didaktiker:innen.

Lehrperson, Merkbarkeit und Lernbarkeit, Sonstiges:

Die drei letzten Kategorien wurden in einer Abbildung (Abb. 24) zusammengefasst.

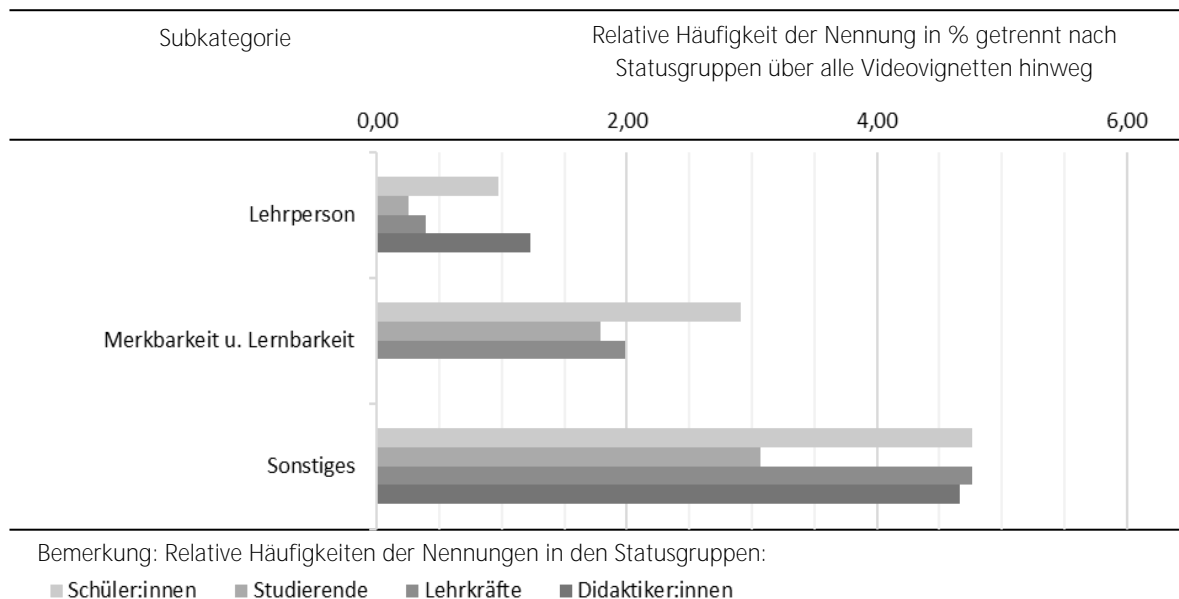


Abbildung 24: Relative Häufigkeiten der Nennungen in den Kategorien ‚Lehrperson‘, ‚Merkbarkeit und Lernbarkeit‘ sowie ‚Sonstiges‘ getrennt nach Statusgruppe (Schüler:innen $N_{max} = 122$; Studierende $N_{max} = 51$; Lehrkräfte $N_{max} = 34$; Didaktiker:innen $N_{max} = 30$) über alle Videovignetten hinweg

Bei der Verteilung innerhalb der Kategorie ‚Lehrperson‘ ist auffällig, dass Schüler:innen (0,97 % aller Äußerungen) und Didaktiker:innen (1,23 % aller Äußerungen) häufiger auf diese Bezug nehmen als Studierende (0,26 % aller Äußerungen) oder Lehrkräfte (0,40 % aller Äußerungen). Für Schüler:innen ist ein Einbezug von Persönlichkeitsmerkmalen in die Bewertung von instruktionalen Erklärungen eher plausibel, da ihnen beispielsweise von Clausen (2002) affektiv geprägte Beurteilungen von Unterricht konstatiert werden. Für Didaktiker:innen ist der Einbezug hingegen aufgrund des professionellen Wissens eher unerwartet, wobei Gautschi (2011) in seiner sogenannten Expertengruppe (lehrerfahrene Geschichtsdidaktiker:innen) ebenso Beurteilungen findet, die sich auf die Persönlichkeitsmerkmale der unterrichtenden Person beziehen.

Die Merkbarkeit und Lernbarkeit des Inhalts der instruktionalen Erklärung werden von Schüler:innen (2,91 % aller Äußerungen), Studierenden (1,79 % aller Äußerungen) und Lehrkräften (1,98 % aller Äußerungen) in die Bewertung der Erklärung mit einbezogen. Didaktiker:innen beziehen sich nicht auf dieses Kriterium. Damit differenzieren sich auch bei diesem Kriterium die Perspektiven zwischen den Professionalisierungsschwerpunkten bzw. der schulischen Praxisnähe. Die Verteilung der Häufigkeiten zwischen den Statusgruppen bzw. die Nichtnennung durch Didaktiker:innen kann auch als weiterer Hinweis darauf gedeutet werden, dass das Kriterium Ausdruck der schulischen Lehr-Lern-Kultur und der damit verbundenen Outputorientierung sowie der durchgeführten Leistungsnachweise ist (siehe 6.2.1.1).

Auf die Kategorie ‚Sonstiges‘ entfallen von allen Statusgruppen in etwa gleich häufig Äußerungen. Damit kann auch resümiert werden, dass die Bewertungen von Schüler:innen ebenso gut in differenzierbare Subkategorien eingeordnet werden können bzw. ebenso häufig differenzierbare Kriterien enthalten wie die Bewertungen der Teilnehmenden der Metaperspektive (Studierende, Lehrkräfte, Didaktiker:innen).

6.2.1.4 Exkurs: Personenspezifische Nutzung von Kriterien für die Bewertung instruktionaler Erklärungen

Die Frage, welche Kriterien zur Bewertung instruktionaler Erklärungen herangezogen werden, kann auch auf der kleinsten Analyseebene der Daten – pro Person und Videovignette – betrachtet werden. Die personenspezifischen Bewertungen tragen nicht wie die vorhergehenden Analysen zum Erkenntnisgewinn über Kriterien guter instruktionaler Erklärungen bei. Die Einblicke im Rahmen des Exkurses können hingegen das Bewertungsverhalten innerhalb der Statusgruppen näher beschreiben. Die folgende Tabelle (Tab. 42)¹⁶² gibt dazu einen Einblick in die angesprochenen Hauptkategorien bzw. Aspekte je Videovignette (Zeilen) für eine Zufallsstichprobe von 16 Schüler:innen (Spalten).¹⁶³ Dabei sind die Schüler:innen jeweils durch ein Pseudonym (erste Zeile) benannt. Die farblich codierten Rechtecke repräsentieren die Hauptkategorien des Kategoriensystems bzw. die Aspekte guten Erklärens (siehe Bemerkung unterhalb der Tabelle). Da die Begründung der Notenvergabe optional ist, ergeben sich in der Tabelle auch leere, weiße Felder.

Tabelle 42: Tabellarisch dargestelltes visuelles Dokumentenvergleichsdiagramm für eine Zufallsstichprobe von $N = 16$ Schüler:innen pro Videovignette

Schüler:in	URCH20	EDUD31	ANIN13	KALU21	BERO16	KEST2	ANJE31	ANGE20	KAYÜ3	NAWO9	CLTH04	ELST6	CARO9	UTAN26	COJO231	ANST15
G _i	■	■	■			■			■							
G _k								■	■							
Ö _i		■	■				■	■	■	■					■	■
Ö _k	■	■	■													
P _i	■	■	■		■						■					
P _k		■	■				■		■	■			■	■		■

Bem.: Farbzuordnungen der Kategorien: **Fach**, **Adressatenorientierung**, **Strukturiertheit**, **Repräsentationen**, **Verständliche Sprache**, **Sprech- und Körperausdruck**, Globale Aussagen, Lehrperson, Sonstiges, Merkbareit und Lernbarkeit; URCH20, EDUD31 usw.: Pseudonym eines Schülers bzw. einer Schülerin.

Die Zufallsstichprobe macht deutlich, was sich auch in der gesamten Teilstichprobe zeigt: Die Möglichkeit zur Abgabe einer Begründung wird häufig genutzt. Insgesamt werden so in der gesamten Teilstichprobe 533 von 719 (entspricht 76,91 %) möglichen offenen Textfeldern zur Begründung genutzt. Wie in Kapitel 6.2.1.3 bereits gezeigt wurde, bedienen sich Schüler:innen neben globalen Aussagen auch differenzierbarer Kriterien, um ihre Notenvergabe zu begründen. Überwiegend rekurren sie auf einen oder zwei Aspekte, häufig in Kombination mit einer globalen Aussage. Die angesprochenen Aspekte guten Erklärens variieren dabei sowohl innerhalb der Personen in Abhängigkeit von der jeweiligen Videovignette als auch zwischen den Personen zu einer Videovignette. Die reduzierte Erklärung zum Themenbereich Genetik kommentieren drei der Schüler:innen aus der Zufallsstichprobe beispielsweise wie folgt:

„Da es ab und zu etwas durcheinander war und man sich sehr konzentrieren musste um mitzukommen aber man es schlussendlich verstanden hat“ (EDUD31/Video G)

¹⁶² Die Darstellungsform entstand auf Grundlage eines visuellen Dokumentenvergleichsdiagramm in MAXODA18.

¹⁶³ Die Zufallsstichprobe wurde über das Programm SPSS erzeugt und aus der Gesamtstichprobe der Schüler:innen ($N = 122$) gezogen.

„ich fand das Video gut verständlich erklärt. Das minus ist da, weil ich finde, dass sie ein bisschen langsam und unsicher gesprochen hat. Ist aber nur eine Kleinigkeit. Ich habe aber alles verstanden.“ (ANIN13/Video G)

„Es war sehr gut das der notwendige Stoff noch einmal wiederholt wurde so kann man besser an den neuen Stoff anknüpfen und der neue Stoff wurde gut erklärt“ (CARO09/Video G)

Schüler:in EDUD31 spricht damit die Verständlichkeit (globaler Aspekt) und die Strukturiertheit der Erklärung an, während Schüler:in ANIN13 neben der Verständlichkeit (globaler Aspekt) den Sprech- und Körperausdruck als Kriterium heranzieht. Schüler:in CARO09 verweist in der Begründung neben der globalen Aussage, dass gut erklärt wurde, mit der Stoffwiederholung auf die Adressatenorientierung innerhalb der Erklärung. Gleichzeitig scheinen bei einzelnen Personen bestimmte Aspekte der Erklärung stärker im Fokus der Aufmerksamkeit zu stehen. Dies wird beispielsweise bei Schüler:in ANIN13 deutlich, deren/dessen Begründungen sich neben globalen Aspekten auf den Sprech- und Körperausdruck sowie die Gestaltung der Adressatenorientierung in der Erklärung konzentrieren.

Wie für die Statusgruppe der Schüler:innen illustriert die untenstehende Tabelle (Tab. 43) diejenigen Aspekte, die Studierende als Begründung zur Vergabe des Globalurteils anführen. In den Spalten sind unter den Pseudonymen der Studierenden die angesprochenen Hauptkategorien pro Videovignette farblich gekennzeichnet.

Tabelle 43: Tabellarisch dargestelltes visuelles Dokumentenvergleichsdiagramm für eine Zufallsstichprobe von N = 16 Studierenden pro Videovignette

Studierende:r	BRHE95	MOAL18	JUMA08	ERAL15	GEER18	CHCH10	YYPE04	EHKO04	KATH05	HEJO14	ANAN85	BASI02	MATH01	CAAN02	INST04	KAMA21
Gi	■	■	■	■					■	■	■	■	■		■	■
Gk	■	■	■					■	■	■	■	■		■	■	■
Öl		■						■	■	■	■	■	■		■	■
Ök	■	■	■						■	■	■	■	■		■	■
Pi	■	■	■	■							■	■	■		■	■
Pk	■	■	■	■				■	■	■	■	■	■		■	■

Bem.: Farbzuordnungen der Kategorien: Fach, Adressatenorientierung, Strukturiertheit, Repräsentationen, Verständliche Sprache, Sprech- und Körperausdruck, Globale Aussagen, Lehrperson, Sonstiges, Merkbareit und Lernbarkeit; BRHE95, MOAL18 usw.: Pseudonym eines Studenten bzw. deiner Studentin.

Das Diagramm zeigt anhand der Zufallsstichprobe, dass Studierende die Möglichkeit zur Begründung seltener als Schüler:innen nutzen. In der gesamten Teilstichprobe verwenden sie diese in 158 von 302 möglichen Fällen (entspricht 52,32 %). An den Studierenden der Zufallsstichprobe wird weiterhin deutlich, dass mindestens eine, meist jedoch zwei oder mehr differenzierbare Hauptkategorien angesprochen werden. Die Aspekte, auf die bei den Begründungen rekurriert wird, unterscheiden sich dabei in einigen Fällen zwischen den Studierenden stärker als zwischen den Videovignetten.¹⁶⁴ Besonders deutlich wird das in den Fällen BASI02 und KAMA21: Während BASI02

¹⁶⁴ Bei Schüler:innen zeigt sich diese Systematik noch weniger stark, da insgesamt weniger Kategorien angesprochen werden und in den Äußerungen viele globale Aussagen enthalten sind.

die fachliche Dimension der Erklärung bei fünf der Vignetten als Begründung heranzieht, beschreibt KAMA21 die Güte von fünf Erklärungen mit Merkmalen der Sprache. Auch zu den Repräsentationen (Begründungen in vier Vignetten) und der Adressatenorientierung (Begründungen in drei Vignetten) wird angegeben, diese seien bei der Vergabe des Globalurteils mit einbezogen worden. In diesen Fällen werden einige Hauptkategorien stringent bei einem Großteil der Erklärungen als Kriterien angeführt, sodass das Antwortmuster einzelner Studierender eher auf einen festen Kanon an Kriterien hinweist.

In anderen Fällen werden hingegen über die Videovignetten hinweg viele unterschiedliche Hauptkategorien zur Begründung des Globalurteils herangezogen, sodass keine systematische Anwendung von bestimmten Kategorien deutlich wird. Die Begründungen sind dabei mitunter eher vom Stimulus und dessen individuellen Auffälligkeiten geleitet. Beispielhaft für diese Form der Begründung können die Fälle BREHE95 und MATH01 betrachtet werden.

Insgesamt bedienen sich die Studierenden in den Begründungen eines großen Spektrums an Aspekten innerhalb und zwischen den Erklärungen. Die Güte der Erklärung der reduzierten Vignette aus dem Themenbereich der Genetik wird von drei Studierenden aus der Zufallsstichprobe beispielsweise folgendermaßen eingeschätzt:

„-deutlich gesprochen – Schritt für Schritt – gut visualisiert – angenehmes Sprechtempo & Lautstärke“ (ANAN85/Video G)

„Die Ursache wurde anschaulich visualisiert. Das Protein kann nicht mehr in die Membran eingebaut werden, aber was geschieht damit dann?“ (BASIO2/Video P_r)

Student:in ANAN85 nutzt einige sprecherische Aspekte (Artikulation, Sprechtempo, Lautstärke) sowie die Strukturiertheit und eine holistische Einschätzung zu den Repräsentationen. Das Beispiel von BASIO2 enthält Äußerungen mit fachlichem Bezug – genauer dem Fehlen von Inhalten, die über die formulierte Ausgangsfrage der Erklärung hinausgehen – und ebenfalls zu Repräsentationen.

Die Ergebnisse der strukturierenden qualitativen Inhaltsanalyse zeigen, dass auch Lehrkräfte die Möglichkeit zur Begründung wie Studierende eher moderat nutzen (108 von 199 möglichen Begründungsfeldern sind ausgefüllt, was 54,27 % entspricht) und dabei unterschiedliche Aspekte Aufmerksamkeit erhalten. Das visuelle Dokumentenvergleichsdiagramm (Tab. 44) verdeutlicht die Varianz der Aspekte, die zur Begründung des Globalurteils herangezogen werden für eine Zufallsstichprobe von $N = 16$.

Tabelle 44: Tabellarisch dargestelltes visuelles Dokumentenvergleichsdiagramm für eine Zufallsstichprobe von N = 16 Lehrkräften pro Videovignette

Lehrkraft	IDER28	MAGE03	CHMI75	LGÜ03	SOFRT0	INH11	MARE61	FRCH02	INMI07	JUAX07	MADI06	MAWI08	ROAN05	ROWE01	RUF32	THWO09
G _i	■		■	■	■		■	■	■				■		■	
G _k	■		■		■		■	■	■				■	■	■	■
Ö _i	■		■		■		■		■	■	■		■	■	■	■
Ö _k	■		■	■	■		■	■		■	■	■				
P _i	■		■		■		■	■	■	■						
P _k	■		■		■		■	■	■	■			■	■	■	■

Bem.: Farbzuordnungen der Kategorien: **Fach**, **Adressatenorientierung**, **Strukturiertheit**, **Repräsentationen**, **Verständliche Sprache**, **Sprech- und Körperausdruck**, Globale Aussagen, Lehrperson, Sonstiges, Merkbarkeit und Lernbarkeit; IDER28, MAGE03 usw.: Pseudonym einer Lehrkraft.

Die folgenden drei Äußerungen stammen aus den offenen Textfeldern zur komplexen Erklärung aus dem Themenbereich der Genetik und vermitteln einen Eindruck, wie die verschiedenen Aspekte von den Teilnehmer:innen aus der Zufallsstichprobe genutzt werden:

„gute Artikulation und Sprache - gutes Sprechtempo - evtl mehr Augenkontakt“
(ROAN05/Video G_k)

„Visualisierung wichtiger Stichworte. Nur Wörter, keine Bilder“ (INMI07/G_k)

„- anschaulich - einprägsam und übersichtlich - ohne große Besonderheiten“
(ROWE01/Video G_k)

Während ROAN05 auf sprecherische sowie sprachliche Aspekte der Lehrkraft achtet, bewertet INMI07 anhand der eingesetzten Visualisierungen (Aspekt ‚Repräsentationen‘) und ROWE01 anhand kriterienübergreifender Aspekte (‚Globale Aspekte‘ und ‚Merkbarkeit und Lernbarkeit‘).

Die Bewertungen von INMI07 zeigen wie bei Studierenden, dass es auch hier Personen gibt, die die sechs Erklärungen überwiegend anhand von wenigen Aspekten bewerten – auch die Bewertungen von LGÜ03 spiegeln das wider. Demgegenüber nutzen IDER28 und CHMI75 je Videovignette unterschiedliche Aspekte, um die Güte der Erklärung zu bewerten.

Didaktiker:innen nutzen die Möglichkeit zur Begründung ähnlich häufig wie Schüler:innen: Es liegen 141 von 179 möglichen Begründungen vor, was einer Nutzung des offenen Textfelds von 78,77 % entspricht. Das visuelle Dokumentenvergleichsdiagramm (Tab. 45) verdeutlicht die zur Begründung genutzten Aspekte.

Tabelle 45: Tabellarisch dargestelltes visuelles Dokumentenvergleichsdiagramm für eine Zufallsstichprobe von N = 16 Didaktiker:innen pro Videovignette

Didaktiker:in	MADI07	ROXX37	ELER15	BEHE12	EVWO09	KAGÜ08	ANTH74	MAWE02	INMI35	HEMA24	STEH54	ANRA48	ANPA29	GAA71	MAED08	FRGE13
Gi	■	■	■	■	■	■	■	■	■				■	■	■	■
Gk		■	■		■	■	■	■	■		■	■	■	■	■	■
Öi		■	■	■	■	■	■	■	■		■		■	■	■	■
Ök	■	■	■	■	■	■	■	■	■		■		■	■	■	■
Pi		■	■	■	■	■	■	■	■	■	■		■	■	■	■
Pk		■	■	■	■	■	■	■	■	■	■	■	■	■	■	■

Bem.: Farbzuordnungen der Kategorien: Fach, Adressatenorientierung, Strukturiertheit, Repräsentationen, Verständliche Sprache, Sprech- und Körperausdruck, Globale Aussagen, Lehrperson, Sonstiges, Merkbarkeit und Lernbarkeit; MADI07, RO37 usw.: Pseudonym eines Didaktikers bzw. einer Didaktikerin.

Deutlich treten auch hier Schwerpunkte in der Betrachtung und Bewertung der Videovignetten hervor: Während ANTH74 vor allem fachliche Aspekte der Erklärung zur Begründung des Globalurteils heranzieht, sind die Äußerungen von ROXX37 deutlich von Äußerungen zu den genutzten Repräsentationen geprägt. ELER15 nutzt in den Begründungen hingegen eine Vielzahl von Aspekten (Fach, Adressatenorientierung, Strukturiertheit, Repräsentationen, verständliche Sprache). Die Kriterien treten dabei jedoch ebenfalls über mehrere Videovignetten hinweg auf.

Vergleicht man die angesprochenen Hauptkategorien zwischen den Personen innerhalb einer Vignette, so lassen sich auch hier Unterschiede feststellen. Die folgenden Äußerungen veranschaulichen die individuellen Perspektiven von Teilnehmer:innen aus der Zufallsstichprobe:

„Hinweis auf Begrenztheit des aktuellen Wissensstandes und auf Notwendigkeit weiterer Forschung ist gut - allerdings wird, wie im letzten Video ein Zweck unterstellt bzw. explizit von einem Zweck des Immunsystems gesprochen; das Immunsystem soll etwas tun; besser von der biologischen Funktion für den Organismus sprechen(?) bzw. auf Unterscheidung proximaler und ultimer Ursachen hinweisen“ (ANTH74/Video G_k)

„Einbezug von Tafelbild mit Applikationen - sukzessives Entwickeln des Tafelbildes - klare Aussprache mit entsprechender Modulation“ (ANPA29/Video G_k)

„guter Ausdruck - kurze und prägnante Sätze“ (STHE54/Video G_k)

ANTH74 rekurriert bei der Begründung auf mehrere Kriterien, die alle der fachspezifischen Facette zugeordnet werden können. Demgegenüber spricht ANPA29 den Einsatz der Repräsentationen sowie den Sprech- und Körperausdruck der erklärenden Person an. Die sprachliche Gestaltung wird von STEH54 als Anhaltspunkt für die Notenvergabe angegeben.

6.2.1.5 Überblick und Einordnung der Ergebnisse zur statusgruppenspezifischen Nutzung von Kriterien zur Bewertung instruktionaler Erklärungen

Zentrale Befunde zu Statusgruppenspezifika in Bezug auf die Aspekte guten Erklärens

Alle vier befragten Statusgruppen beziehen sich bei ihren Begründungen zur Notenvergabe auf Kriterien, die den Aspekten Adressatenorientierung, Strukturiertheit, Repräsentationen, Sprech- und Körperausdruck, verständliche Sprache, fachliche Aspekte sowie Aussagen zur Lehrperson und Aussagen zur Merkbarkeit und Lernbarkeit des Erklärten zugeordnet werden können. Auch globale Aspekte der Güte einer instruktionalen Erklärung werden von allen Statusgruppen zur Bewertung genutzt. Die Häufigkeiten, mit denen die Aspekte angesprochen werden, differieren jedoch deutlich.

Weiterhin beziehen alle vier Statusgruppen die Lehrperson und deren Merkmale mit in die Bewertung der Erklärung ein. Schüler:innen (0,97 %) und Didaktiker:innen (1,23 %) nutzen Merkmale der erklärenden Person dabei häufiger als Studierende (0,26 %) und Lehrkräfte (0,40 %). Da die Merkmale der erklärenden Person nicht unmittelbar Teil der didaktischen Gestaltung einer instruktionalen Erklärung sind, wurde eine Begründung anhand derer durch Didaktiker:innen nicht erwartet. Ähnliche Ergebnisse finden sich bei Gautschi (2011), der in seinen Daten zu gutem Geschichtsunterricht ebenso Äußerungen von Didaktiker:innen findet, die sich auf die Persönlichkeitsmerkmale der Lehrkräfte beziehen. Im Rahmen des Persönlichkeitsparadigmas werden Zusammenhänge zwischen Lehrkräftepersönlichkeitsmerkmalen und Unterrichtsqualität untersucht, die Zusammenhänge sind in den durchgeführten Untersuchungen überwiegend schwach (Krauss & Bruckmaier, 2014).

Schüler:innen beziehen sich am häufigsten auf die Repräsentationen (28,37 %), auf globale Aspekte (27,15 %) sowie auf die Adressatenorientierung (20,35 %). Studierende äußern sich bei ihren Begründungen am häufigsten zu den eingesetzten Repräsentationen (24,04 %) sowie der Adressatenorientierung (18,41 %). Auch die fachlichen Aspekte (13,55 %) sowie die globalen Aspekte (13,04 %) werden häufig angeführt. Die Begründungen von Lehrkräften sind auf der Ebene der Aspekte betrachtet denen der Studierenden ähnlich: Am häufigsten nutzen sie Kriterien, die den Repräsentationen (23,81 %), der Adressatenorientierung (19,05 %) und den fachlichen Aspekten (15,87 %) zuzuordnen sind. Das Muster der Begründungen von Didaktiker:innen weicht davon in Bezug auf die Häufigkeiten ab. Didaktiker:innen verwenden am häufigsten Kriterien zu den Repräsentationen (32,19 %), zu den fachlichen Aspekten (24,32 %) und deutlich seltener zur Adressatenorientierung (11,30 %).

Die Vorannahme, dass Schüler:innen Erklärungen eher global und weniger differenziert bewerten, wohingegen die Bewertungen der Gruppe der Lehrenden (Studierende, Lehrkräfte, Didaktiker:innen) weniger häufig global und deutlich differenzierter ausfällt, kann auf Ebene der Aspekte zunächst bekräftigt werden. Innerhalb der Kategorie zu globalen Aussagen ist auffällig, dass die Häufigkeit der Nennung beginnend bei Schüler:innen (27,14 %) über Studierende (13,04 %) und Lehrkräfte (11,11 %) bis zu Didaktiker:innen (9,09 %) abnimmt.

Zentrale Befunde zu Statusgruppenspezifika in Bezug auf die Kriterien guten Erklärens

Schüler:innen

Schüler:innen bewerten auf der Ebene der einzelnen Kriterien nicht durchgehend global, sondern nutzen differenzierte Kriterien zur Begründung der Bewertung: Zu den Aspekten finden sich in den globalen Subkategorien (z. B. Strukturiertheit global, Repräsentationen global etc.) im Vergleich zu den weiteren Statusgruppen nicht auffällig viele Äußerungen. Lediglich bei den Repräsentationen äußern sich Schüler:innen häufiger global als die Statusgruppen, die der Metaperspektive zuzuordnen sind (Studierende, Lehrkräfte, Didaktiker:innen). Insgesamt bewerten Schüler:innen die Erklärungen entlang eigener Präferenzen, die jedoch auch das Lernen mit der instruktionalen Erklärung und die Verständlichkeit der Erklärung in den Blick nehmen. Sie nutzen weniger Kriterien als die weiteren Statusgruppen: Sie ziehen für die Begründung der Notenvergabe 48 unterscheidbare Kriterien heran, wobei 17 davon nur selten angeführt werden. 12 weitere Kriterien, die in den Begründungen der anderen Statusgruppen zu finden sind, kommen bei Schüler:innen nicht vor. Ihre Bewertungen können daher nicht pauschal als undifferenziert beschrieben werden, denn der Großteil der nicht oder selten genannten Kriterien bezieht sich auf den Sprech- und Körperausdruck sowie auf den fachlichen Aspekt. Gleichwohl nutzen sie am wenigsten differenzierbare Kriterien.

In diesem Zusammenhang kann auf die von Clausen (2002) getroffene Beschreibung von **Schülerurteilen Bezug genommen werden**. Er formuliert: „Für eine neutrale wissenschaftlich-differenzierende Betrachtung des Unterrichtsgeschehens sind Schülerurteile aufgrund der Tendenz zur **affektiven geprägten generalisierenden Beurteilung weniger geeignet**“ (Clausen, 2002, S. 188). In den hier vorliegenden Daten können die Äußerungen bzw. die Kriterien der Schüler:innen nicht grundlegend als generalisierend beschrieben werden. Auch Schüler:innen nennen eine Vielzahl von Kriterien, die sie zur Begründung der Notenvergabe heranziehen. Bewertungsgrundlage sind nachvollziehbarer Weise keine professionellen Wissensbestände, wie dies für die Gruppen der Metaperspektive angenommen werden kann, sodass Kriterien wie die fachliche Angemessenheit der Repräsentation, die Umsetzung der Phänomenorientierung oder die Berücksichtigung von (Fehl-)Vorstellungen nicht angelegt werden können. Die Äußerungen der Schüler:innen sind sachlicher Natur. Sie drücken die eigenen Präferenzen aus, was sich durch die Rolle als Adressat:innen natürlicherweise ergibt. Schüler:innen nehmen dabei auch das erfolgreiche Lernen mit der Erklärung bei der Bewertung im Blick, sodass die eingesetzten Repräsentationen hinsichtlich ihrer Verständlichkeit bewertet sowie der logische Aufbau der Erklärung oder das Erklärtempo als Kriterien herangezogen werden. Clausen (2002) räumt den Urteilen der Schüler:innen an anderer Stelle eine große Bedeutung ein, da die wahrgenommene affektive Qualität des Unterrichts mit der Leistung und dem Interesse der Lernenden einen engen Zusammenhang aufweist. Insgesamt sind die Bewertungen von Schüler:innen differenziert und tragen zudem aufgrund der individuellen Perspektive als Adressat:innen zum besseren Verständnis guten Erklärens bei, der über eine theoretisch begründete Gleichberechtigung der Perspektiven (Berliner, 2005; Denzin, 1970; Flick, 2011) hinausgeht. Inwiefern die Bewertungen im standardisierten Zugang mit geschlossenen Items ebenfalls als differenziert beschrieben werden können, wird in den folgenden Kapiteln aufgegriffen.

Didaktiker:innen

In den Bewertungen von Didaktiker:innen lassen sich tendenziell häufiger Bezüge zur Wissenschaftstheorie finden als in den übrigen Gruppen. Bezüge zu fachdidaktischen und lernpsychologischen Theorien werden nicht direkt vorgenommen – Kriterien, die in engem Zusammenhang mit Theorien und aktuellen Diskursen um guten bzw. effektiven Unterricht stehen, werden jedoch häufiger als in den anderen Statusgruppen aufgegriffen. Damit einhergehend zeigt sich bei den Repräsentationen ein häufiger Bezug auf Merkmale, die als lernrelevant einstufung werden können. Die Perspektive ist durch eine starke Fokussierung biologischer bzw. biologiedidaktischer Kriterien gekennzeichnet. Insgesamt werden 53 Kriterien angesprochen, davon 9 jedoch nur selten. 7 weitere Kriterien, die sich jedoch bei den anderen Statusgruppen finden lassen, werden nicht als Bewertungsgrundlage einbezogen. Damit einhergehend kann in den meisten Aspekten in der Tendenz eine breitere Bewertungsgrundlage gefunden werden. Der Sprech- und Körperausdruck wird demgegenüber von Didaktiker:innen insgesamt kaum in die Bewertung von Erklärungen einbezogen, sodass über die Differenziertheit keine Aussage getroffen werden kann. Besonders auffällig ist darüber hinaus, dass die Merkbarkeit und Lernbarkeit wie auch die interessante und motivierende Gestaltung nicht mit in die Bewertungen einbezogen werden.

Dass die interessante und motivierende Gestaltung nicht mit in die Bewertungen einbezogen wird, könnte darin begründet liegen, dass hierfür eher Erfahrungswissen notwendig ist, um die Sicht der Schüler:innen zu kennen. Bei Kriterien wie **‚Berücksichtigung von (Fehl-)Vorstellungen‘** oder **‚Verständlichkeit der Visualisierung‘** ist Erfahrungswissen hingegen weniger notwendig, da zu deren Bewertung auch theoretische bzw. empirische Bezüge genutzt werden können. Die Kategorien **‚Passung des Schwierigkeitsgrads‘**, **‚Übererklärung‘** und **‚Fokussierung auf Wesentliches‘** könnten auch deshalb weniger in den Fokus für Bewertungen durch Didaktiker:innen rücken. Darüber hinaus werden die Merkbarkeit und Lernbarkeit nicht als Maßstab zur Bewertung herangezogen. Das Fehlen von Bezügen darauf ist insofern bedeutsam, als das Kriterium in Verbindung zur Outputorientierung (Dittmer, 2010; Grotzer, 2012) steht und Didaktiker:innen durch die Nichterwähnung ein Stück weit von der schulischen Lehr-Lern-Kultur zurücktreten. Wissenschaftstheoretische Bezüge werden von Didaktiker:innen nicht sehr häufig, aber dennoch häufiger als von den anderen Statusgruppen zur Begründung der Notenvergabe angeführt. Weiterhin sind in den Begründungen von Didaktiker:innen nicht vermehrt direkte Theoriebezüge – in dem Sinne, dass beispielsweise die ‚Cognitive-Load-Theory‘ als solche benannt wird – zu erkennen. Jedoch sind Kriterien, die in engem Zusammenhang mit fachdidaktischen Theorien und aktuellen Diskursen um guten bzw. effektiven Unterricht stehen, häufiger Teil der Begründungen. Beispielsweise wird in der Fachdidaktik darüber diskutiert, dass (Fehl-)Vorstellungen von Schüler:innen zu berücksichtigende Ausgangspunkte für fachliches Lernen sind (Hammann & Asshof, 2014; Kattmann, 2015), und in der Diskussion um Lernen mit Visualisierungen und externen Repräsentationen über Gestaltungsvorschläge für verstehensförderliche Visualisierungen auf Basis der ‚Cognitive-Load-Theory‘ diskutiert (Bay et al., 2016; Mayer & Moreno, 2003). Insbesondere innerhalb des Aspekts der Repräsentationen nehmen Didaktiker:innen mitunter damit einhergehend vermehrt Bezug auf lernrelevante Merkmale. Darüber hinaus wird die Fachsprache in den Naturwissenschaftsdidaktiken kritisch in den Blick genommen und als Lerngegenstand reflektiert (Leisen, 2010; Rincke, 2010; Ziepprecht et al., 2017), sodass Didaktiker:innen theoretisch für die in den Erklärungen eingesetzte Sprache eher sensibilisiert sind bzw. das professionelle Wissen für eine Unterscheidung der Sprachebenen und Verständlichkeitscharakteristika besitzen. Der Umgang mit Fachworten wird damit einhergehend von Didaktiker:innen am häufigsten als Kriterium einbezogen.

Diese Tendenzen charakterisieren die Perspektive der Didaktiker:innen am ehesten als eine genuin biologiedidaktische Perspektive.

Lehrkräfte

Lehrkräfte beziehen sich bei der Begründung der Notenvergabe auf insgesamt 50 Kriterien, wobei 10 der Kriterien nur selten vorkommen. 10 weitere Kriterien werden von Lehrkräften nicht genutzt, lassen sich jedoch bei anderen Statusgruppen finden. Die Differenziertheit der Kriterien ist damit ungefähr mit derjenigen der Didaktiker:innen vergleichbar. Die Bewertung der Adressatenorientierung ist insbesondere durch die Kriterien des aktiven Einbezugs der Rezipient:innen sowie die Kriterien zur didaktischen Reduktion geprägt. Der Sprech- und Körperausdruck wird auch von Lehrkräften wie von Didaktiker:innen nicht nur selten, sondern auch mit wenigen Kriterien in die Bewertung einbezogen. Die Verständlichkeit der Sprache wird etwas häufiger herangezogen – in etwa so, wie von Studierenden und Didaktiker:innen –, wobei das Kriterium des präzisen Ausdrucks die Bewertungen des Aspekts durch Lehrkräfte deutlich prägt. Die Bewertung der fachlichen Angemessenheit der Erklärung ist geprägt durch häufige Nennungen der Korrektheit und Vollständigkeit, wohingegen die Auswahl des Erklärgegenstandes durch die Kriterien der Prinzipienorientierung und Phänomenorientierung in den Hintergrund tritt. Die Repräsentationen werden etwa gleich häufig anhand von lernrelevanten Merkmalen und an ihrem Vorhandensein bewertet.

Auch Lehrkräfte nutzen damit vielseitige Kriterien für die Bewertung von Erklärungen. Ihre Perspektive lässt sich vor allem in Abgrenzung zur Perspektive der Didaktiker:innen als vermehrt unterrichtsbezogen, auf die konkrete Vermittlung fokussiert und weniger fachdidaktisch charakterisieren. Die unterrichtsbezogene Perspektive wird beispielsweise durch die Stellen markiert, an denen sich Lehrkräfte bei der Bewertung der Adressatenorientierung maßgeblich an dem für die Unterrichtsvorbereitung etablierten Konzept der didaktischen Reduktion orientieren oder den aktiven Einbezug von Schüler:innen fordern. Bei der Forderung nach einem aktiven Einbezug der Lernenden steht weniger die kognitive Aktivierung (Helmke, 2007), sondern die Einbindung der Schüler:innen durch Aktivitäten (Brüning & Saum, 2019) im Fokus. Gleichzeitig beziehen Lehrkräfte zum Aspekt der Repräsentationen insbesondere neben lernrelevanten Merkmalen fast ähnlich häufig das Vorhandensein der externen Repräsentationen als Begründung für die Notenvergabe heran. Damit rekurren Sie weniger häufig implizit auf lerntheoretische Fundierungen, die über Merkmale der Verständlichkeit von Visualisierungen Aussagen treffen (Bay et al., 2016; Mayer & Moreno, 2003), sondern nehmen im Gegensatz zu Didaktiker:innen vermehrt allgemein Bezug auf den Einsatz von bildlichen Visualisierungen während der Erklärung. Betrachtet man die Kriterien zur Verständlichkeit der Sprache, wird von Lehrkräften überwiegend die Präzision des Ausdrucks in die Bewertungen eingebracht. Das Kriterium finden sich nicht prominent im naturwissenschaftsdidaktischen Diskurs zum Einsatz von Fachsprache im Unterricht (Leisen, 2010; Rincke, 2010). In Hinblick auf den zu reduzierenden Erklärgegenstand beziehen sich Lehrkräfte auf den Fachgegenstand, indem sie die fachliche Korrektheit und die Vollständigkeit des Erklärgegenstandes bewerten. In Bezug auf die Bewertung der Darstellung des Fachgegenstand ist dabei interessant, dass Lehrkräfte sich damit eher an den unmittelbaren Fakten zum Gegenstand orientieren und weniger an die Auswahl der Erklärgegenstände an sich denken. Hinzu kommt, dass Lehrkräfte sich im Gegensatz zu Didaktiker:innen auf die Merkbarkeit und Lernbarkeit beziehen und damit instruktionalen Erklärungen eine unmittelbare Funktion für den Unterricht zuweisen. Der Blick

von Lehrkräften kann auch dadurch als auf den Unterricht und die konkrete Vermittlung fokussiert beschrieben werden.

Studierende

Studierende bewerten die Erklärungen insgesamt anhand von 53 Kriterien, wovon 12 nur selten im Material auftauchen. 7 weitere Kriterien, die von anderen Statusgruppen genannt werden, werden von Studierenden nicht angeführt. Die Breite der genutzten Kriterien ist damit in etwa mit den zwei weiteren Statusgruppen der Metaperspektive vergleichbar. Das Bewertungsmuster der Studierenden ist durch wechselnde Ähnlichkeiten mit den Mustern der weiteren drei Statusgruppen gekennzeichnet, was für den Noviz:innenstatus der Studierenden in Bezug auf professionelle Wissensbestände spricht (Blömeke et al., 2015; Seidel et al., 2010; Sherin, 2001). Die Perspektive kann damit nicht eindeutig als eher die der Adressat:innen, als unterrichtsbezogen, wie von Lehrkräften oder als vorrangig biologiedidaktisch, wie von Didaktiker:innen beschrieben werden.

Studierende bewerten die Umsetzung der Adressatenorientierung in der Erklärung tendenziell mit Kriterien, die eher für die Bewertungen der Statusgruppe der Schüler:innen charakteristisch sind (z. B. Erklärtempo, Interessante und motivierende Gestaltung). Deutlich seltener beziehen sie sich auf Kriterien, für die Professionswissen notwendig sind, etwa die didaktische Reduktion, die von Lehrkräften häufig herangezogen wird, sowie auf solche zur Berücksichtigung von (Fehl-)Vorstellungen, die von Didaktiker:innen verhältnismäßig oft als Kriterium genannt werden. Den Aspekt der Repräsentationen bewerten Studierende hingegen eher entlang der Kriterien, die Lehrkräfte und Didaktiker:innen an die instruktionalen Erklärungen stellen: Während Didaktiker:innen die Verständlichkeit deutlicher häufiger als den bloßen Einsatz von Visualisierungen bewerten, beziehen sich Lehrkräfte häufiger auf das Vorhandensein von bildlichen Visualisierungen. Bei Studierenden ist das Verhältnis ausgeglichen. Die fachliche Angemessenheit der Visualisierungen wird vorrangig von Didaktiker:innen mit einbezogen, Studierende bewerten ebenso entlang des Kriteriums. Der Sprech- und Körperausdruck kann als Alleinstellungsmerkmal für die Gruppe der Studierenden gelten – sie nutzen ihn weitaus am häufigsten. Zur Bewertung dessen ziehen sie im Vergleich zu den weiteren Statusgruppen die meisten Kriterien heran, wobei mit deutlicher Mehrheit die Variation der Stimme als Kriterium für die Bewertung verwendet wird. Auf die Sprache rekurrieren Studierende etwa ähnlich häufig wie die weiteren Lehrenden. Das Bewertungsmuster ähnelt in diesem Aspekt hinsichtlich der Breite eher dem von Didaktiker:innen. Jedoch ist das am häufigsten genutzte Kriterium dasjenige des Satzbaus – bei Didaktiker:innen der Umgang mit Fachworten und bei Lehrkräften die Präzision des Ausdrucks. Die fachlichen Kriterien der Studierenden verteilen sich auf die drei Facetten (fachliche Korrektheit, Phänomen- bzw. Prinzipienorientierung, Vollständigkeit) in etwa gleich – das unterscheidet Studierende von Lehrkräften und Didaktiker:innen (Lehrkräfte nutzen häufiger die Korrektheit sowie die Vollständigkeit, Didaktiker:innen häufiger die Phänomen- bzw. Prinzipienorientierung). Die Merkbarkeit und Lernbarkeit des Gesagten verwenden Studierende ähnlich häufig wie Lehrkräfte. Insgesamt ist die Perspektive der Studierenden nicht als greifbares, eigenes Muster beschreibbar, sondern durch verschiedene Ähnlichkeiten mit den weiteren Statusgruppen in den einzelnen Aspekten am ehesten zu charakterisieren. Der Wechsel zwischen den Perspektiven ist dabei nicht unplausibel: Studierende sind zeitlich nah an der eigenen Schulzeit und damit an der Rolle der Adressat:innen bzw. im Studium regelmäßig als Adressat:innen angesprochen. Beispielsweise in Praktika erleben sie sich als unterrichtende Lehrperson und lernen in Begleitveranstaltungen die

Vorbereitung von Unterricht. Im Rahmen des Studiums erwerben sie unter anderem durch Didaktiker:innen professionelles Wissen (Blömeke et al., 2015; Seidel et al., 2010; Sherin, 2001).

Grundlegend kann festgehalten werden, dass sich die Statusgruppen hinsichtlich der selbst herangezogenen Kriterien zur Bewertung einer instruktionalen Erklärung unterscheiden, sodass verschiedene Schwerpunkte in den Perspektiven bzw. bei der Bewertung von Erklärungen erkennbar werden. Der Einbezug der Perspektiven für ein vertieftes Verständnis guter instruktionaler Erklärungen ist damit als sinnvoll zu erachten.

Aus methodischer Sicht wird eine nach Statusgruppen getrennte Analyse im Folgenden sinnvoll. Weiterhin wird die Frage aufgeworfen, ob auch in der Einschätzung der Güte durch vorgegebene Kriterien ein Unterschied erkennbar wird (siehe Kap. 6.2.2).

6.2.2 Ergebnisse zu Bewertungsunterschieden zwischen verschiedenen Statusgruppen in Hinblick auf vorgegebene Kriterien

Das Erkenntnisinteresse der Forschungsfrage besteht primär auf statusgruppenspezifischen Mustern der Einschätzung, die die Statusgruppen bzw. deren Perspektive auf die Güte der Erklärungen näher charakterisieren, und nicht auf der Einschätzung der Erklärungen im Sinne eines Erkenntnisgewinns über die Videovignetten. Dabei ist unter anderem die Frage nach der Einigkeit bzw. Uneinigkeit in der Bewertung von Interesse, da dies auch darauf hindeutet, wie gut die Statusgruppen der Metaperspektive die Wahrnehmung aus der Perspektive der Rezipient:innen einschätzen können, was wiederum für die Gestaltung adressatenorientierter Erklärungen große Bedeutung hat.

Die Voraussetzungsprüfungen für die durchgeführten einfaktoriellen Varianzanalysen werden in Anhang E und F tabellarisch zusammengestellt. Zur Prüfung der Normalverteilung wird für die größere Stichprobe der Schüler:innen der Kolmogorov-Smirnov-Test herangezogen und für die kleineren Stichproben der Lehrenden der Shapiro-Wilk-Test eingesetzt (Bühner & Ziegler, 2017; siehe Anhang E). Die einfaktorielle ANOVA ist relativ robust gegenüber Verletzungen der Normalverteilung, sodass auch bei Voraussetzungsverletzungen die Varianzanalyse durchgeführt wird (Bortz & Schuster, 2010). Eine Überschätzung der Effekte wird bei der Interpretation berücksichtigt. Die Voraussetzung der Varianzhomogenität wird mit einem Levene-Test geprüft (siehe Anhang F). Bei Verletzung der Voraussetzung wird der Welch-Test berichtet (nur bei Video Ö_k für den fachspezifischen Aspekt nicht gegeben).

6.2.2.1 Ergebnisse zur Bewertungsunterschieden zwischen Schüler:innen, Studierenden, Lehrkräften und Didaktiker:innen in Hinblick auf vorgegebene Kriterien

Die Analysen zur unterschiedlichen Einschätzung instruktionaler Erklärungen durch verschiedene Statusgruppen in Hinblick auf vorgegebene Kriterien werden im Folgenden für das Globalurteil und jeweils getrennt nach Aspekten (Adressatenorientierung, Strukturiertheit, sprachliche Verständlichkeit, Sprech- und Körperausdruck, Wahrnehmung des dargestellten Komplexitätsgrades) vorgenommen.

Dazu werden Tabellen mit den arithmetischen Mitteln sowie den dazugehörigen Standardabweichungen der vier Statusgruppen pro Videovignette zum jeweiligen Aspekt abgebildet und analysiert. Mittels der Standardabweichungen kann zum einen abgeschätzt werden, wie einig bzw. uneinig sich die Personen innerhalb der jeweiligen Statusgruppen in Bezug auf die

Einschätzung der Güte des betrachteten Aspekts zu einer Videovignette sind. Zum anderen werden durch sehr hohe sowie sehr niedrige Werte Hinweise auf besonders strittige bzw. unstrittige Momente der Güte einer instruktionalen Erklärung gegeben. Durch Betrachtung der besten sowie der schlechtesten Mittelwerte innerhalb der jeweiligen Aspekte und den Vergleich ihrer Verteilung auf die sechs Erklärungen können Hinweise auf die Einigkeit bzw. Uneinigkeit zwischen den Statusgruppen in Bezug auf die besonders gute sowie besonders schlechte Umsetzung des Aspekts in den Erklärungen gefunden werden. Der Vergleich erlaubt auf deskriptiver Ebene Aussagen darüber, ob sich beispielsweise alle oder ein Teil der Statusgruppen einig sind, dass in einer Erklärung die Realisierung der Strukturierung besonders gut bzw. besonders schlecht gelungen ist.

Im unteren Abschnitt der Tabelle sind weiterhin die entsprechenden Werte der einfaktoriellen Varianzanalyse mit dem Faktor ‚Statusgruppenzugehörigkeit‘ (post-hoc-Tests mit Korrektur nach Gabriel¹⁶⁵) angegeben, die die signifikanten Unterschiede zwischen den Einschätzungen der Statusgruppen zu einer Erklärung anzeigen. Die Ergebnisse erlauben damit einen Vergleich der Einschätzungen zu einer Erklärung und Aussagen darüber, ob sich die Einschätzungen einzelner Statusgruppen in Bezug auf die Umsetzung des jeweiligen Aspekts signifikant voneinander unterscheiden. Damit werden folglich die Einschätzungen zwischen den Statusgruppen getrennt nach Videovignetten ins Verhältnis gesetzt.

Globalurteil

Die nachstehende Tabelle 46 zeigt die Ergebnisse in Bezug auf die Einschätzungen des Globalurteils.

Tabelle 46: Nach Statusgruppen und Videovignetten differenzierte arithmetische Mittel und Standardabweichungen der Globalurteile und Ergebnisse der einfaktoriellen ANOVA mit dem Faktor Statusgruppenzugehörigkeit

Video	Video G _i M (SD)	Video G _k M (SD)	Video Ö _i M (SD)	Video Ö _k M (SD)	Video P _i M (SD)	Video P _k M (SD)
Statusgruppe						
Schüler:innen (1) (N _{max} = 122)	2,01 (0,76)	2,30 (0,74)	1,84 (0,81)	2,00 (0,85)	2,04 (0,72)	2,15 (0,79)
Studierende (2) (N _{max} = 51)	1,88 (0,81)	2,22 (0,86)	1,87 (0,89)	2,14 (0,75)	2,09 (0,81)	1,88 (0,68)
Lehrkräfte (3) (N _{max} = 34)	1,98 (0,58)	2,38 (0,98)	2,20 (0,77)	2,64 (0,95)	2,28 (0,75)	2,06 (0,88)
Didaktiker:innen (4) (N _{max} = 30)	2,01 (0,88)	1,88 (0,77)	1,99 (0,70)	2,33 (1,06)	2,06 (0,83)	1,90 (0,76)
<i>df</i> _{res}	232	231	228	231	229	224
<i>F</i> (3, <i>df</i> _{res})	0,35	2,61	1,77	5,20	0,95	1,75
η^2	0,01	0,03	0,02	0,06	0,01	0,02
<i>p</i>	0,79	0,05	0,15	≤ 0,01	0,42	0,16
post hoc (Gabriel)	(1) vs. (3)*					

Bem.: *M*: Mittelwert; *SD*: Standardabweichung; *N*_{max}: aufgrund fehlender Werte bei einzelnen Videos den Berechnungen pro Statusgruppe maximal zugrunde liegende Personenanzahl; *df*_{res}: Freiheitsgrade der Residuen; *F*: F-Wert der Teststatistik; η^2 : Eta-Quadrat; *p*: p-Wert; *: *p* ≤ 0,05; **: *p* ≤ 0,01; (1): Schüler:innen; (2): Studierende; (3): Lehrkräfte; (4): Didaktiker:innen.

¹⁶⁵ Um den Fehler 1. Art für die pro Videovignette insgesamt sechs durchgeführten Paarvergleiche zu kontrollieren und gleichzeitig die unterschiedlichen Gruppengrößen zu berücksichtigen, wurde für die Post-hoc-Tests die Korrektur nach Gabriel gewählt (Field, 2018).

Evident wird, dass die Uneinigkeit bzw. Einigkeit innerhalb der Statusgruppen gemessen an der Standardabweichung jeweils im Durchschnitt bei eineinhalb bzw. fast zwei Notenstufen liegt. Zwei Werte weisen Auffälligkeiten auf: Lehrkräfte sind sich bei der Erklärung G_i im Verhältnis zu den weiteren Vignetten sehr einig in ihrer Einschätzung ($SD = 0,58$), wohingegen Didaktiker:innen sich in Bezug auf die Einschätzung der Vignetten \ddot{O}_k am wenigsten einig sind ($SD = 1,06$). Insgesamt variieren die Standardabweichungen bei Schüler:innen ($range_{SD} = [0,72;0,85]$) am geringsten, bei Studierenden etwas mehr ($range_{SD} = [0,68;0,89]$). Bei Didaktiker:innen und Lehrkräften ist die Spannweite der Standardabweichungen aufgrund der beiden genannten Extremwerte höher (Didaktiker:innen: $range_{SD} = [0,70;1,06]$; Lehrkräfte: $range_{SD} = [0,58;0,98]$).

Der Vergleich der Einschätzungen verdeutlicht weiterhin, dass sich die Gruppen bezüglich der besten Erklärung und der schlechtesten Erklärung nicht einig sind. Die Einschätzungen zur schlechtesten Vignette lassen sich übersichtlich gruppieren: Schüler:innen und Studierende bewerten die Erklärung G_k am schlechtesten, wohingegen sich Lehrkräfte und Didaktiker:innen einig sind, dass die Erklärung \ddot{O}_k vergleichsweise schlecht ist.¹⁶⁶

Die besten Einschätzungen verteilen sich hingegen weniger systematisch über die Videovignetten: Während die Erklärung \ddot{O}_i von Schüler:innen durchschnittlich die besten Globalurteile erhält, vergeben Studierende dieser Vignette zwar ebenfalls durchschnittlich die besten Globalurteile, jedoch bestehen nur marginale Unterschiede zu den Vignetten G_i und P_k . Lehrkräfte bewerten das Video G_i mit deutlichem Abstand am besten, Didaktiker:innen hingegen die Vignetten G_k und P_k ebenfalls mit marginalen Unterschieden. Die besten Einschätzungen verteilen sich demnach zum einen zwischen den Statusgruppen wesentlich unsystematischer als die schlechten Einschätzungen und zum anderen erhalten mehr Erklärungen im Mittel ähnlich gute Bewertungen.

Die Tabelle 46 zeigt außerdem, dass sich mittels einfaktorieller Varianzanalysen in einem Fall ein signifikanter Unterschied zwischen Globalurteilen in Abhängigkeit von der Statusgruppenzugehörigkeit feststellen lässt. Er besteht bei der komplexen Erklärung zur Ökologie (\ddot{O}_k ; $F(3,229) = 5,20$; $p \leq 0,01$; $\eta^2 = 0,03$). 3,00 % der Streuung der Werte der Globalurteile um den Gesamtmittelwert können durch die Statusgruppenzugehörigkeit erklärt werden, was einem kleinen Effekt entspricht (Cohen, 1988). Die post-hoc-Tests zeigen, dass Schüler:innen die Erklärung signifikant besser bewerten als Lehrkräfte ($M_{Sus} = 2,00$; $SD_{Sus} = 0,85$; $M_{Leh} = 2,64$; $SD_{Leh} = 0,95$; $p \leq 0,01$; $d = 0,73$), wobei die Effektstärke dieses Unterschiedes als mittel einzustufen ist (Cohen, 1988).

Adressatenorientierung

Tabelle 47 zeigt analog zum vorherigen Abschnitt die Ergebnisse der Einschätzungen zur Güte des Aspekts ‚Adressatenorientierung‘.

Betrachtet man die Standardabweichungen, fallen diese für den Aspekt der Adressatenorientierung in keiner der Statusgruppen auffällig hoch aus.¹⁶⁷ Es zeichnet sich eine leichte Tendenz ab, dass Didaktiker:innen eine höhere Uneinigkeit in Bezug auf die Einschätzung der Adressatenorientierung aufweisen ($range_{SD} = [0,62;0,87]$) als Schüler:innen ($range_{SD} = [0,57;0,79]$), Studierende

¹⁶⁶ Angemerkt sei hier, dass die schlechtesten Bewertungen über alle Vignetten hinweg in allen Fällen auf komplexere Erklärvarianten entfallen. Ein Vergleich der Bewertungen wird in Kapitel 6.2.4 aufgegriffen.

¹⁶⁷ Auch hier sind die Ergebnisse vergleichbar mit denjenigen in der Gesamtstichprobe, sodass die Einigkeit der Bewertungen innerhalb der Gruppen scheinbar nicht größer als in der Gesamtstichprobe ist.

($range_{SD} = [0,49;0,80]$) und Lehrkräfte ($range_{SD} = [0,62;0,87]$). Weiterhin fällt auf, dass die Uneinigkeiten für die beiden Vignetten der Ökologie im Vergleich zu den weiteren Erklärungen für die Statusgruppen der Schüler:innen, Studierenden und Lehrkräfte tendenziell etwas geringer ausfallen als für die weiteren Vignetten in Bezug auf die Bewertungen innerhalb der Statusgruppen. Für Didaktiker:innen gilt das nicht.

Tabelle 47: Nach Statusgruppen und Videovignetten differenzierte arithmetische Mittel und Standardabweichungen des Aspekts ‚Adressatenorientierung‘ und Ergebnisse der einfaktoriellem ANOVA mit dem Faktor Statusgruppenzugehörigkeit

Video	Video G _i M (SD)	Video G _k M (SD)	Video Ö _i M (SD)	Video Ö _k M (SD)	Video P _i M (SD)	Video P _k M (SD)
Statusgruppe						
Schüler:innen (1) (N _{max} = 122)	2,93 (0,78)	2,89 (0,70)	2,71 (0,64)	2,71 (0,57)	2,67 (0,75)	2,78 (0,79)
Studierende (2) (N _{max} = 44)	2,58 (0,63)	2,77 (0,74)	2,49 (0,60)	2,54 (0,49)	2,59 (0,77)	2,52 (0,80)
Lehrkräfte (3) (N _{max} = 29)	2,82 (0,66)	3,14 (0,67)	2,69 (0,44)	2,76 (0,55)	2,73 (0,84)	2,69 (0,74)
Didaktiker:innen (4) (N _{max} = 24)	2,88 (0,82)	2,67 (0,64)	2,48 (0,63)	2,77 (0,87)	2,39 (0,66)	2,51 (0,62)
<i>df_{res}</i>	214	212	210	211	210	209
<i>F</i> (3, <i>df_{res}</i>)	2,39	2,36	1,95	1,28	1,08	1,65
η^2	0,03	0,03	0,03	0,02	0,02	0,02
<i>p</i>	0,07	0,07	0,12	0,28	0,36	0,18
post hoc (Gabriel)						

Bem.: M: Mittelwert; SD: Standardabweichung; N_{max}: aufgrund fehlender Werte bei einzelnen Videos den Berechnungen pro Statusgruppe maximal zugrunde liegende Personenanzahl; *df_{res}*: Freiheitsgrade der Residuen; *F*: F-Wert der Teststatistik; η^2 : Eta-Quadrat; *p*: p-Wert; *: $p \leq 0,05$; **: $p \leq 0,01$; (1): Schüler:innen; (2): Studierende; (3): Lehrkräfte; (4): Didaktiker:innen.

Für die besten und schlechtesten Einschätzungen der Adressatenorientierung ergibt sich eine andere Systematik als für die Globalurteile. Die Erklärung mit dem schlechtesten konstruktsspezifischen Mittelwert ist für Studierende ($M = 2,77$; $SD = 0,74$) und Lehrkräfte ($M = 3,14$; $SD = 0,67$) die komplexe Erklärung der Genetik (G_k). Für Schüler:innen ($M = 2,93$; $SD = 0,78$) und Didaktiker:innen ($M = 2,88$; $SD = 0,82$) entfallen die höchsten Mittelwerte und damit die schlechtesten Einschätzungen zur Güte der Adressatenorientierung auf die lineare Erklärung der Genetik (G_i). Bei Schüler:innen ist jedoch die Differenz zum zweitschlechtesten Mittelwert (Videovignette G_k; $M = 2,89$; $SD = 0,70$) nur gering.

Die besten konstruktsspezifischen Mittelwerte entfallen bei Studierenden ($M = 2,49$; $SD = 0,60$) und Lehrkräften ($M = 2,69$; $SD = 0,44$) auf die lineare Vignette der Ökologie (Ö_i), wobei sich für die komplexe Vignette der Physiologie (P_k) der gleiche Mittelwert der Bewertungen von Lehrkräften ($M = 2,69$; $SD = 0,74$) ergibt. Bei Schüler:innen ($M = 2,67$; $SD = 0,75$) und Didaktiker:innen ($M = 2,39$; $SD = 0,66$) entfallen sie auf die lineare Vignette der Physiologie (P_i). Auch hier ist die durchschnittlich zweitbeste Einschätzung der Schüler:innen mit marginaler Differenz zur besten (Videovignette Ö_i; $M = 2,71$; $SD = 0,64$) bei derjenigen Erklärung zu finden, die von Studierenden und Lehrkräften am besten bewertet wird.

Damit ergeben sich auf deskriptiver Ebene interessante Unterschiede hinsichtlich des Bewertungsverhaltens der Statusgruppen: Auf der einen Seite bewerten Studierende sowie Lehrkräfte die Adressatenorientierung – gemessen an der Verteilung der schlechtesten und besten

Mittelwerte – ähnlich. Auf der anderen Seite schätzen Schüler:innen und Didaktiker:innen dieselben Erklärungen als beste bzw. schlechteste ein. Die einfaktoriellen Varianzanalysen zeigen jedoch keine signifikanten Unterschiede. Damit schätzen die Statusgruppen die Adressatenorientierung der einzelnen Erklärungen ähnlich ein. In Bezug auf die Gestaltung von instruktionalen Erklärungen ist das ein positives Zeichen: Lehrende und Lernende sind sich tendenziell einig, ob eine Erklärung adressatenorientiert gestaltet ist, bzw. Lehrende scheinen die Passung für Adressat:innen angemessen einordnen zu können.¹⁶⁸

Strukturiertheit

In der folgenden Tabelle (Tab. 48) sind die entsprechenden Werte für den Aspekt der Strukturiertheit abgebildet.

Tabelle 48: Nach Statusgruppen und Videovignetten differenzierte arithmetische Mittel und Standardabweichungen des Aspekts ‚Strukturiertheit‘ und Ergebnisse der einfaktoriellen ANOVA mit dem Faktor Statusgruppenzugehörigkeit

Video	Video G _l M (SD)	Video G _k M (SD)	Video Ö _l M (SD)	Video Ö _k M (SD)	Video P _l M (SD)	Video P _k M (SD)
Statusgruppe						
Schüler:innen (1) (N _{max} = 122)	2,39 (0,79)	2,58 (0,75)	2,35 (0,69)	2,51 (0,81)	2,53 (0,75)	2,61 (0,82)
Studierende (2) (N _{max} = 44)	1,91 (0,72)	2,24 (0,82)	2,01 (0,65)	2,28 (0,76)	2,19 (0,67)	2,21 (0,71)
Lehrkräfte (3) (N _{max} = 29)	1,92 (0,67)	2,40 (0,78)	1,71 (0,61)	2,37 (1,07)	2,12 (0,83)	1,96 (0,75)
Didaktiker:innen (4) (N _{max} = 24)	1,79 (0,74)	1,98 (0,76)	1,52 (0,45)	2,22 (0,96)	2,05 (0,82)	1,82 (0,63)
<i>df_{res}</i>	214	212	210	211	210	209
<i>F</i> (3, <i>df_{res}</i>)	8,10	5,10	15,52	1,34	4,94	11,17
η^2	0,10	0,07	0,18	0,02	0,07	0,14
<i>p</i>	<i>p</i> ≤ 0,01	<i>p</i> ≤ 0,01	<i>p</i> ≤ 0,01	0,26	<i>p</i> ≤ 0,01	<i>p</i> ≤ 0,01
post hoc (Gabriel)	(1) vs. (2)** (1) vs. (3)** (1) vs. (4)**	(1) vs. (4) **	(1) vs. (2)* (1) vs. (3)** (1) vs. (4)** (2) vs. (4)*		(1) vs. (3)* (1) vs. (4)*	(1) vs. (2)* (1) vs. (3)** (1) vs. (4)**

Bem.: *M*: Mittelwert; *SD*: Standardabweichung; *N_{max}*: aufgrund fehlender Werte bei einzelnen Videos den Berechnungen pro Statusgruppe maximal zugrunde liegende Personenanzahl; *df_{res}*: Freiheitsgrade der Residuen; *F*: F-Wert der Teststatistik; η^2 : Eta-Quadrat; *p*: p-Wert; *: *p* ≤ 0,05; **: *p* ≤ 0,01; (1): Schüler:innen; (2): Studierende; (3): Lehrkräfte; (4): Didaktiker:innen.

Die Verteilungsgrößen in Tabelle 48 ergeben folgendes Bild für die Einschätzung der Strukturiertheit in Abhängigkeit von der Statusgruppe: Die Standardabweichungen weisen eine ähnliche Spannweite wie diejenigen des Aspekts der Adressatenorientierung und des Globalurteils auf (Schüler:innen: $range_{SD} = [0,69;0,82]$; Studierende $range_{SD} = [0,65;0,82]$; Lehrkräfte: $range_{SD} = [0,61;1,07]$; Didaktiker:innen: $range_{SD} = [0,45;0,96]$). Für Vignette Ö_l sind über alle Statusgruppen hinweg tendenziell eher niedrigere Werte zu verzeichnen, sodass die Einigkeit innerhalb der Statusgruppen zur Einschätzung der Erklärung relativ hoch ist. Das Ergebnis ist plausibel, da die Erklärung zur Entstehung von Populationsschwankungen bei Schneeschuhhasen

¹⁶⁸ Zu beachten ist, dass aufgrund des Studiendesings für alle Teilnehmenden fremder Unterricht bewertet wird. Die Adressatenorientierung wird von den Lehrenden also am durchschnittlichen Lernenden in der 10. Klasse gemessen.

abgesehen von der Einleitung und dem Abschluss der Erklärung einen sehr gleichförmigen Aufbau aufweist: Die Gliederung des Erklärprozesses besteht aus vier inhaltlich ähnlichen Einzelerklärungen – den vier Zyklusphasen der Populationsschwankungen, die jeweils zunächst am konkreten Beispiel mit Schneeschuhhasen und Luchsen über deren Räuber-Beute-Beziehung erklärt und anschließend in einer allgemeinen Regel (z. B. „Je mehr Beutetiere, desto mehr Räuber“) festgehalten werden. Der Zyklus kann von den Rezipient:innen vermutlich gegen Ende der Erklärung selbstständig weitergedacht werden, sodass die von den Rezipient:innen erwartete Gliederung mit derjenigen durch die Lehrkraft abgeglichen werden kann und bei Übereinstimmung tendenziell als ‚gut‘ bewertet wird. Die Visualisierung des Zyklus an der Tafel in zeitlicher Synchronität mit den Erklärschritten trägt dabei auch zur Nachvollziehbarkeit der Strukturierung und damit zum Erkennen des roten Fadens bei (siehe auch Kap. 5.2.2.5).

Vergleichsweise hoch sind hingegen die Standardabweichungen für die Vignette \ddot{O}_k für die Statusgruppen der Lehrkräfte ($SD = 1,07$) und Didaktiker:innen ($SD = 0,96$). Auch dieses Ergebnis kann anhand der Inhalte der Erklärung in einen Kontext gesetzt werden: Komplexe Phänomene und die Linearität einer mündlichen Darstellung sind schwer zu vereinen, sodass die Strukturierung für die erklärende Person herausfordernd ist.

Bezüglich der Einigkeit zur Erklärung mit der besten Strukturiertheit lässt sich ein heterogenes Bild festhalten: Von Schüler:innen erhalten die Erklärung P_k ($M = 2,61$; $SD = 0,82$) und mit nur wenig Differenz dazu die Erklärung G_k ($M = 2,58$; $SD = 0,75$) die durchschnittlich schlechtesten Einschätzungen. Studierende schätzen die Erklärung Vignette \ddot{O}_k ($M = 2,28$, $SD = 0,76$) am schlechtesten und wie Schüler:innen die Vignette G_k ($M = 2,24$; $SD = 0,82$) mit wenig Differenz als zweitschlechte ein. Bei Lehrkräften sind ebenfalls beide Erklärungen unter den schlechtesten: G_k mit $M = 2,40$ und $SD = 0,78$ sowie \ddot{O}_k mit $M = 2,37$ und $SD = 1,07$. Didaktiker:innen befinden wie Studierende die Strukturiertheit der Erklärungen \ddot{O}_k ($M = 2,22$; $SD = 0,96$) als vergleichsweise schlechteste. Zusammengefasst schätzen Schüler:innen, Studierende und Lehrkräfte die Strukturiertheit der Erklärung G_k als schlechter ein, Studierende, Lehrkräfte und Didaktiker:innen die Strukturiertheit der Erklärung \ddot{O}_k . Schüler:innen schätzen darüber hinaus die Erklärung P_k als schlecht strukturiert ein.

Die besten Mittelwerte zur Strukturiertheit verteilen sich etwas systematischer über die sechs Erklärungen: Die Erklärung \ddot{O}_l wird von Schüler:innen ($M = 2,35$; $SD = 0,69$), Lehrkräften ($M = 1,71$; $SD = 0,61$) und Didaktiker:innen ($M = 1,52$; $SD = 0,45$) als die am besten strukturierte eingeschätzt. Erklärt werden kann dies wie zu den geringen Standardabweichungen dieser Vignette mit dem gleichförmigen, stringenten Aufbau der Erklärung. Studierende bewerten die Erklärung G_l als am besten strukturiert ($M = 1,91$; $SD = 0,72$). Diese Variante (G_l) wird von Schüler:innen, Lehrkräften und Didaktiker:innen am zweitbesten eingeschätzt. Analog dazu wird die von Schüler:innen, Lehrkräften und Didaktiker:innen am strukturiertesten wahrgenommene Erklärung (\ddot{O}_l) von Studierenden am zweitbesten eingeschätzt. Alle Statusgruppen bewerten somit die linearen Varianten aus den Themengebieten der Ökologie oder der Genetik am besten oder zweitbesten hinsichtlich der Strukturiertheit.

Auffällig ist insgesamt, dass die schlechtesten Mittelwerte zur Strukturiertheit stets bei den komplexen Varianten der Erklärungen und entgegengesetzt dazu die besten Mittelwerte zur Strukturiertheit jeweils bei den linearen Varianten zu finden sind. Weitere Analysen dazu finden sich in Kapitel 6.2.4.

Auffällig ist ebenfalls, dass die Differenzen zwischen den Statusgruppen bei Vignette \ddot{O}_k am geringsten ausfallen, sodass sich die Gruppen hier bezüglich der Einschätzung der Güte der Strukturiertheit der Erklärung die größte Einigkeit zu haben scheinen. Diese Verteilung schlägt sich in den Ergebnissen der einfaktoriellen Varianzanalysen nieder.

Bei einfaktoriellen Varianzanalysen weist lediglich die komplexe Erklärung zur Ökologie keine signifikanten Unterschiede zwischen den Einschätzungen der Statusgruppen auf. Bei Vignette G_i können insgesamt 10 % der Varianz der konstruktsspezifischen Mittelwerte auf die Gruppenzugehörigkeit zurückgeführt werden (G_i ; $F(3,214) = 8,10$; $p \leq 0,01$; $\eta^2 = 0,10$), was einem mittleren Effekt entspricht (Cohen, 1988). Post-hoc-Tests zeigen, dass Unterschiede zwischen den Einschätzungen der Schüler:innen und der Studierenden ($M_{Sus} = 2,39$; $SD_{Sus} = 0,79$; $M_{Stu} = 1,91$; $SD_{Stu} = 0,72$; $p \leq 0,01$; $d = -0,62$), der Schüler:innen und der Lehrkräfte ($M_{Leh} = 1,92$; $SD_{Leh} = 0,67$; $p \leq 0,01$; $d = -0,61$) sowie der Schüler:innen und der Didaktiker:innen ($M_{Did} = 1,79$; $SD_{Did} = 0,74$; $p \leq 0,01$; $d = -0,77$) bestehen. Die ersten beiden Effekte sind nach den Vorschlägen von Cohen (1988) als mittel, der letzte als groß einzustufen.

Für Vignette G_k besteht ebenfalls ein signifikanter Unterschied (G_k ; $F(3,212) = 5,10$; $p \leq 0,01$; $\eta^2 = 0,07$). Dabei können 7 % der Varianz der konstruktsspezifischen Mittelwerte durch die Gruppenzugehörigkeit erklärt werden, sodass der Effekt als mittel einzustufen ist (Cohen, 1988). Die anschließenden post-hoc-Tests zeigen, dass der gefundene Unterschied zwischen Schüler:innen und Didaktiker:innen besteht ($M_{Sus} = 2,58$; $SD_{Sus} = 0,75$; $M_{Did} = 1,98$; $SD_{Did} = 0,76$; $p \leq 0,01$; $d = 0,80$), wobei die Effektstärke als groß einzuordnen ist (Cohen, 1988).

Die Ergebnisse der einfaktoriellen Varianzanalyse für die Vignetten \ddot{O}_i zeigen ebenfalls einen signifikanten Unterschied mit einer großen Effektstärke (Cohen, 1988) (\ddot{O}_i ; $F(3,212) = 15,52$; $p \leq 0,01$; $\eta^2 = 0,18$). Die Unterschiede liegen wie bei der linearen Erklärung zur Genetik zwischen Schüler:innen und Studierenden ($M_{Sus} = 2,35$; $SD_{Sus} = 0,69$; $M_{Stu} = 2,01$; $SD_{Stu} = 0,65$; $p = 0,02$; $d = 0,52$), Schüler:innen und Lehrkräften ($M_{Leh} = 1,71$; $SD_{Leh} = 0,61$; $p \leq 0,01$; $d = -0,95$) sowie Schüler:innen und Didaktiker:innen ($M_{Did} = 1,52$; $SD_{Did} = 0,45$; $p \leq 0,01$; $d = -1,26$). Zusätzlich unterscheiden sich hier Studierende und Didaktiker:innen signifikant bezüglich ihrer Einschätzungen ($M_{Stu} = 2,01$; $SD_{Stu} = 0,65$; $M_{Did} = 1,52$; $SD_{Did} = 0,45$; $p = 0,02$; $d = -0,83$). Die angegebenen Effektstärken beziffern den Unterschied zwischen Schüler:innen und Studierenden als mittel, die weiteren Unterschiede – jeweils zwischen Schüler:innen und Lehrkräften bzw. Didaktiker:innen und zwischen Studierenden und Didaktiker:innen – als groß.

Ein weiterer signifikanter Unterschied, der zu 7 % durch die Gruppenzugehörigkeit erklärt werden kann, wird für Vignette P_i evident (P_i ; $F(3,210) = 4,94$; $p \leq 0,01$; $\eta^2 = 0,07$). Signifikante Gruppenunterschiede bestehen mit mittleren Effekten zwischen Schüler:innen und Lehrkräften ($M_{Sus} = 2,53$; $SD_{Sus} = 0,75$; $M_{Leh} = 2,12$; $SD_{Leh} = 0,83$; $p = 0,04$; $d = -0,54$) sowie zwischen Schüler:innen und Didaktiker:innen ($M_{Sus} = 2,53$; $SD_{Sus} = 0,75$; $M_{Did} = 2,05$; $SD_{Did} = 0,82$; $p = 0,02$; $d = -0,63$).

Die Erklärung P_k weist vergleichbare Unterschiede bezüglich der Einschätzung der Strukturiertheit wie die Vignetten G_i und \ddot{O}_i auf: Es besteht ein signifikanter Unterschied zwischen den Statusgruppen (P_k ; $F(3,209) = 11,17$; $p \leq 0,01$; $\eta^2 = 0,14$), wobei 14 % der Varianz der Mittelwerte – was einem großen Effekt entspricht – durch den Faktor der Statusgruppe erklärt werden können. Signifikante Unterschiede bestehen zwischen Schüler:innen und Studierenden ($M_{Sus} = 2,61$; $SD_{Sus} = 0,82$; $M_{Stu} = 2,21$; $SD_{Stu} = 0,71$; $p = 0,02$; $d = -0,51$), Schüler:innen und Lehrkräften ($M_{Leh} = 1,96$;

$SD_{Leh} = 0,75$; $p \leq 0,01$; $d = -0,81$) sowie Schüler:innen und Didaktiker:innen ($M_{Did} = 1,82$; $SD_{Did} = 0,63$; $p \leq 0,01$; $d = -1,00$). Auch hier sind die Effektstärken der gefundenen Unterschiede zwischen Schüler:innen und Studierenden geringer (mittlerer Effekt) als zwischen Schüler:innen und Lehrkräften bzw. Didaktiker:innen (jeweils große Effekte).

Betrachtet man die Systematik der gefundenen signifikanten Unterschiede in den durchgeführten post-hoc-Tests genauer, bestehen insgesamt in zwölf Fällen überzufällige Differenzen, wobei elf dieser zwölf Fälle zwischen Schüler:innen und Statusgruppen der Metaperspektive zu finden sind (Ausnahme bei Vignette Ö). Dabei bewerten Schüler:innen die Strukturiertheit stets signifikant schlechter als die weiteren Gruppen. Eine Tendenz der Schüler:innen zur systematisch durchgehenden Vergabe schlechter Bewertungen bei allen Aspekten kann nicht festgestellt werden, sodass es sich wahrscheinlich nicht um eine allgemeine und vom jeweiligen Aspekt unabhängige Moderation der Einschätzung handelt.¹⁶⁹ Im Überblick ergibt sich ein weiteres interessantes Muster: Die Effekte der Unterschiede zwischen Schüler:innen und Studierenden sind bis auf eine Ausnahme (G) entweder nicht signifikant (G_k , \ddot{O}_k , P) oder im Gegensatz zu den weiteren Effekten von mittlerer Effektstärke (\ddot{O}_i , P_k). Damit schätzen Studierende und Schüler:innen die Güte der Strukturiertheit ähnlicher als Schüler:innen und die weiteren Statusgruppen ein.

Mitunter könnten Gründe für die Differenzen im Unterschied des Fachwissens zwischen den Schüler:innen und den Gruppen der Metaperspektive liegen: Die Bewertung, ob die Erklärung einem roten Faden folgt (Item: *St_rot*), sie Schritt für Schritt aufgebaut wird (Item: *St_bau*), sowie die Frage, ob zu viele Einzelheiten in der Erklärung aufgegriffen werden (Item: *St_ein*), könnten durch Fachwissen über den Gegenstand moderiert werden. Ein höheres Fachwissen erleichtert die selbstständige Herstellung von Zusammenhängen, die den Eindruck des roten Fadens verstärken und verändert den Blick auf das, was als Einzelheit aufgefasst wird. Wenn dies zutrifft, würde das bedeuten, dass die Gruppe der Metaperspektive in diesem Fall nicht die Perspektive der Adressat:innen übernehmen kann.

Sprech- und Körperausdruck

In der nachstehenden Tabelle (Tab. 49) finden sich die entsprechenden Werte für die Einschätzungen zum Sprech- und Körperausdruck der erklärenden Person.

Die Standardabweichungen sind innerhalb der sowie zwischen den Statusgruppen recht ähnlich und tendenziell etwas niedriger als beim Globalurteil sowie bei den Aspekten der Adressatenorientierung und Strukturiertheit (Schüler:innen: $range_{SD} = [0,58;0,83]$; Studierende: $range_{SD} = [0,58;0,76]$; Lehrkräfte: $range_{SD} = [0,53;0,90]$; Didaktiker:innen $range_{SD} = [0,57;0,79]$).

Der Spannweite der Mittelwerte innerhalb einer Statusgruppe über die sechs Vignetten hinweg weist eine ähnliche Dimension auf wie in den weiteren Aspekten. Dies ist jedoch insofern bemerkenswert, als die weiteren Aspekte stärker mit der Gestaltung der Erklärung zusammenhängen und damit eine höhere Varianz erwartet werden könnte, als dies beim Sprech- und Körperausdruck der erklärenden Lehrkraft der Fall ist. Konkret: Es wäre zu erwarten, dass die Adressatenorientierung über die sechs Videovignetten heterogener eingeschätzt wird, da jede der sechs Erklärung individuelle Umsetzungen aufweist, die bewusst durch Planung und Gestaltung der

¹⁶⁹ Lediglich der fachspezifische Aspekt wird von Schüler:innen durchgehend schlechter bewertet als von den anderen Statusgruppen.

erklärenden Person erfolgen (z. B. in Bezug auf die Anpassung des Schwierigkeitsgrads oder des Informationsgehalts). Die Umsetzung des Sprech- und Körperausdrucks könnte im Vergleich stärker als personenabhängige Variable gesehen werden und da alle Erklärungen von einer Person eingesprochen wurden, wäre infolgedessen eine geringere Varianz erwartbar (Allhoff & Allhoff, 2010; Babad, 2009). Die Einschätzungen können durch das kontext- und gegenstandsabhängige Sprechen interpretiert werden: Je nach (Erklär-)Gegenstand, zu dem der Sprecher oder die Sprecherin eine bestimmte Haltung einnimmt, und situativen Umständen variiert der Sprech- und Körperausdruck (Geißner, 1988; Neuber, 2016).

Tabelle 49: Nach Statusgruppen und Videovignetten differenzierte arithmetische Mittel und Standardabweichungen des Aspekts ‚Sprech- und Körperausdruck‘ und Ergebnisse der einfaktoriellen ANOVA mit dem Faktor Statusgruppenzugehörigkeit

Video	Video G _i M (SD)	Video G _k M (SD)	Video Ö _i M (SD)	Video Ö _k M (SD)	Video P _i M (SD)	Video P _k M (SD)
Statusgruppe						
Schüler:innen (1) (N _{max} = 70)	2,33 (0,58)	2,54 (0,62)	2,28 (0,69)	2,54 (0,83)	2,14 (0,62)	2,33 (0,63)
Studierende (2) (N _{max} = 27)	1,89 (0,67)	2,27 (0,58)	2,00 (0,62)	2,28 (0,76)	1,73 (0,69)	1,97 (0,64)
Lehrkräfte (3) (N _{max} = 19)	1,95 (0,53)	2,24 (0,90)	2,10 (0,65)	2,18 (0,84)	1,81 (0,63)	2,10 (0,80)
Didaktiker:innen (4) (N _{max} = 19)	1,97 (0,74)	1,86 (0,79)	1,67 (0,61)	1,89 (0,57)	1,90 (0,71)	1,75 (0,69)
<i>df_{res}</i>	88	126	126	88	87	124
<i>F</i> (3, <i>df_{res}</i>)	3,24	5,33	4,65	2,27	2,35	4,46
η^2	0,10	0,11	0,10	0,07	0,08	0,10
<i>p</i>	0,03	<i>p</i> ≤ 0,01	<i>p</i> ≤ 0,01	0,09	0,08	<i>p</i> ≤ 0,01
post hoc (Gabriel)	(1) vs. (2)*	(1) vs. (4)**	(1) vs. (4)**			(1) vs. (4)**

Bem.: M: Mittelwert; SD: Standardabweichung; N_{max}: aufgrund fehlender Werte bei einzelnen Videos den Berechnungen pro Statusgruppe maximal zugrunde liegende Personenanzahl; *df_{res}*: Freiheitsgrade der Residuen; *F*: F-Wert der Teststatistik; η^2 : Eta-Quadrat; *p*: p-Wert; *: *p* ≤ 0,05; **: *p* ≤ 0,01; (1): Schüler:innen; (2): Studierende; (3): Lehrkräfte; (4): Didaktiker:innen.

Betrachtet man die Mittelwerte des Aspekts pro Erklärung, kann festgehalten werden, dass die schlechtesten mittleren Einschätzungen der Schüler:innen auf die Vignetten G_k (M = 2,54; SD = 0,62) sowie Ö_k (M = 2,54; SD = 0,83) entfallen, ebenso ist dies bei Studierenden (Ö_k; M = 2,28; SD = 0,76/G_k; M = 2,27; SD = 0,68). Lehrkräfte bewerten im Mittel auch die Erklärung G_k am schlechtesten (M = 2,24; SD = 0,90). Während die drei genannten Statusgruppen ähnlich bewerten, entfällt der schlechteste Mittelwert der Didaktiker:innen auf die Erklärung G_i (M = 1,97; SD = 0,74).

Für die durchschnittlich besten Mittelwerte pro Statusgruppen ergibt sich ein ähnliches Bild: Während die Güte des Sprech- und Körperausdrucks der erklärenden Person in Vignette P_i von Schüler:innen (M = 2,14; SD = 0,62), Studierenden (M = 1,73; SD = 0,69) und Lehrkräften (M = 1,81; SD = 0,63) im Mittel am besten eingeschätzt wird, sehen Didaktiker:innen (M = 1,67; SD = 0,61) bei Vignette Ö_i im Durchschnitt den Sprech- und Körperausdruck am besten umgesetzt.

Damit sind sich die Statusgruppen mit größerer Nähe zur Schulpraxis in Bezug auf die schlechteste und die beste Umsetzung des Sprech- und Körperausdrucks in den Erklärungen einiger. Insgesamt bewerten Schüler:innen im Mittel den Sprech- und Körperausdruck durchwegs schlechter als die drei weiteren Statusgruppen. Didaktiker:innen bewerten den Sprech- und Körperausdruck hingegen

durchgehend am besten, was sich in den Ergebnisse der einfaktoriellen ANOVA niederschlägt: Tabelle 49 zeigt, dass sich mittels einfaktorieller Varianzanalysen in vier Fällen signifikante Unterschiede zwischen den Einschätzungen des Sprech- und Körperausdrucks in Abhängigkeit von der Statusgruppenzugehörigkeit feststellen lassen, wobei die meisten signifikanten Befunde zwischen Schüler:innen und Didaktiker:innen auftreten. Nur der Unterschied bei Vignette G_i ist laut post-hoc-Test auf Unterschiede zwischen Schüler:innen und Studierenden zurückzuführen: 10 % der Varianz der konstruktsspezifischen Mittelwerte – was einer mittleren Effektstärke entspricht – lassen sich auf die Statusgruppenzugehörigkeit zurückführen (G_i; $F(3,88) = 3,24$; $p = 0,03$; $\eta^2 = 0,10$). Post-hoc-Tests zeigen, dass der Unterschied zwischen Schüler:innen und Studierenden besteht ($M_{Sus} = 2,33$; $SD_{Sus} = 0,58$; $M_{Stu} = 1,89$; $SD_{Stu} = 0,67$; $p = 0,03$; $d = -0,73$) und dieser nach Cohen (1988) einem mittleren Effekt entspricht.

Vignette G_k weist in der einfaktoriellen Varianzanalyse mit einer ähnlichen Effektstärke ($\eta^2 = 0,11$; mittlerer Effekt nach Cohen (1988)) ebenfalls einen signifikanten Unterschied auf (G_k; $F(3,126) = 5,33$; $p \leq 0,01$; $\eta^2 = 0,11$), der jedoch zwischen Schüler:innen und Didaktiker:innen ($M_{Sus} = 2,54$; $SD_{Sus} = 0,62$; $M_{Did} = 1,86$; $SD_{Did} = 0,79$; $p \leq 0,01$; $d = -1,03$) besteht. Der Effekt ist im Gegensatz zu demjenigen zwischen Schüler:innen und Studierenden in Vignette G_i mit $d = -1,03$ hoch (großer Effekt nach Cohen (1988)).

Der Unterschied bei Vignette Ö_i ist ähnlich einzuordnen: $F(3,126) = 4,65$; $p \leq 0,01$; $\eta^2 = 0,10$ mit einem signifikanten Unterschied zwischen Schüler:innen und Didaktiker:innen in vergleichbarem Maße ($M_{Sus} = 2,28$; $SD_{Sus} = 0,69$; $M_{Did} = 1,67$; $SD_{Did} = 0,61$; $p \leq 0,01$; $d = -0,91$).

Für die Erklärungen Ö_k sowie P_i bestehen keine Unterschiede in den konstruktsspezifischen Mittelwerten, deren Differenzen gegen den Zufall abgesichert auf die Statusgruppenzugehörigkeit zurückgeführt werden können.

Die Einschätzungen der Videovignette P_k weisen wieder ein ähnliches Muster wie diejenigen der Erklärungen G_k und Ö_i auf: Es besteht ein signifikanter Unterschied bei der Einschätzung des Sprech- und Körperausdrucks in Abhängigkeit von der Statusgruppe (P_k; $F(3,124) = 4,46$; $p \leq 0,01$; $\eta^2 = 0,10$). Der Effekt ist mit 10 % als mittel einzustufen (Cohen, 1988). Die post-hoc-Tests zeigen signifikante Unterschiede zwischen den Bewertungen der Schüler:innen und der Didaktiker:innen ($M_{Sus} = 2,33$; $SD_{Sus} = 0,63$; $M_{Did} = 1,75$; $SD_{Did} = 0,69$; $p \leq 0,01$; $d = -0,90$).

Insgesamt bestehen somit hauptsächlich signifikante Unterschiede in der Einschätzung der Güte des Sprech- und Körperausdrucks zwischen Schüler:innen und Didaktiker:innen (bei 3 von 6 Vignetten). Diese weisen stets eine identische Ausprägung bezüglich der Tendenz der Einschätzung auf: Schüler:innen bewerten den Sprech- und Körperausdruck schlechter als Didaktiker:innen, wobei der Effekt der Statusgruppenzugehörigkeit jeweils als groß einzuordnen ist.

Verständliche Sprache

Die folgende Tabelle (Tab. 50) zeigt die deskriptiven Einschätzungen in Bezug auf die Verständlichkeit der Sprache je Statusgruppe und Erklärung inklusive der Ergebnisse der einfaktoriellen Varianzanalyse mit dem Faktor der Statusgruppenzugehörigkeit.

Tabelle 50: Nach Statusgruppen und Videovignetten differenzierte arithmetische Mittel und Standardabweichungen des Aspekts ‚Verständliche Sprache‘ und Ergebnisse der einfaktoriellen ANOVA mit dem Faktor Statusgruppenzugehörigkeit

Video	Video G _i M (SD)	Video G _k M (SD)	Video Ö _i M (SD)	Video Ö _k M (SD)	Video P _i M (SD)	Video P _k M (SD)
Statusgruppe						
Schüler:innen (1) (N _{max} = 122)	2,65 (0,81)	2,79 (0,78)	2,33 (0,83)	2,47 (0,78)	2,74 (0,79)	2,72 (0,84)
Studierende (2) (N _{max} = 44)	2,78 (0,78)	2,91 (0,89)	2,26 (0,81)	2,62 (0,80)	2,85 (0,82)	2,70 (0,82)
Lehrkräfte (3) (N _{max} = 29)	2,53 (0,87)	2,64 (0,63)	2,01 (0,72)	2,30 (0,71)	2,59 (0,92)	2,49 (0,86)
Didaktiker:innen (4) (N _{max} = 24)	2,72 (0,85)	2,72 (0,77)	1,90 (0,60)	2,55 (0,93)	2,65 (0,93)	2,33 (0,79)
<i>df_{res}</i>	214	212	210	211	210	209
<i>F</i> (3, <i>df_{res}</i>)	0,62	0,75	2,77	1,03	0,64	1,78
η^2	0,01	0,01	0,04	0,01	0,01	0,03
<i>p</i>	0,60	0,52	0,04	0,38	0,59	0,15
post hoc (Gabriel)						

Bem.: M: Mittelwert; SD: Standardabweichung; N_{max}: aufgrund fehlender Werte bei einzelnen Videos den Berechnungen pro Statusgruppe maximal zugrunde liegende Personenanzahl; *df_{res}*: Freiheitsgrade der Residuen; *F*: F-Wert der Teststatistik; η^2 : Eta-Quadrat; *p*: p-Wert; *: $p \leq 0,05$; **: $p \leq 0,01$; (1): Schüler:innen; (2): Studierende; (3): Lehrkräfte; (4): Didaktiker:innen.

Die Standardabweichungen unterscheiden sich von den bisher analysierten Aspekten nicht auffällig. Lediglich die Spannweite der Standardabweichungen von Schüler:innen umfasst einen überraschend geringen Bereich ($range_{SD} = [0,78;0,84]$), sodass bei keiner der Erklärungen größere und geringere Einigkeit bzw. Uneinigkeit in Bezug auf die Einschätzung durch Schüler:innen festgestellt werden kann. Die Spannweiten der Standardabweichungen von Studierenden ($range_{SD} = [0,78;0,89]$), Lehrkräften ($range_{SD} = [0,63; 0,92]$) und Didaktiker:innen ($range_{SD} = [0,60;0,93]$) sind mit denen der anderen Aspekte vergleichbar.

Betrachtet man die angegebenen Mittelwerte und vergleicht diese zunächst pro Erklärung zwischen den Gruppen, lässt sich feststellen, dass kaum ausgeprägte Differenzen vorhanden sind – wie dies beispielsweise in den Aspekten der Strukturiertheit oder des Sprech- und Körperausdrucks zu verzeichnen war. Das heißt, die Statusgruppen schätzen die Güte der Erklärung bezüglich der abgebildeten Facetten der sprachlichen Verständlichkeit sehr ähnlich ein. Die übereinstimmende Bewertung ist insbesondere in Hinblick auf die Bedeutung der Sprache als zentrales Kommunikationsmittel sowie die Schwierigkeiten der Fachsprache für Schüler:innen relevant. Eine ähnliche Einschätzung der Verständlichkeit der Sprache deutet darauf hin, dass die Statusgruppen der Metaperspektive – also diejenigen, die Unterricht gestalten – die sprachliche Verständlichkeit in Bezug auf die hier erhobenen Parameter für Schüler:innen adäquat einschätzen können.

Die Einigkeit in den Einschätzungen der Erklärungen, die sich für die Verteilung der Mittelwerte angedeutet hat, verfestigt sich weiter, wenn die besten und schlechtesten konstruktsspezifischen Mittelwerte analysiert werden. Die Erklärung, die von allen Statusgruppen am besten bezüglich der sprachlichen Verständlichkeit beurteilt wurde – bei den Gruppen der Metaperspektive sogar mit einer beachtlichen Differenz zum nächstbesten Mittelwert –, ist die lineare Erklärung der Ökologie (Ö_i). Ein Grund dafür könnte auf Ebene des Textes die Wiederholung im Aufbau sein, da jede Schwankung der Räuber- und Beutepopulation mit einem stets gleich aufgebauten Text erfolgt. Auf Ebene der Sätze ist ebenso ein repetitiver Satzbau zu erkennen, da stets ähnliche Satzkonstruktionen

für das Herstellen der Zusammenhänge zwischen der Größe der Räuber- und Beutepopulation verwendet werden. Auf Ebene der Worte ist damit einhergehend die Varianz verwendeter Worte geringer, sodass die Verständlichkeit dadurch besonders hoch eingeschätzt werden könnte.

Die tendenziell am schlechtesten eingeschätzte Erklärung ist ebenfalls ausnahmslos die komplexe Erklärung der Genetik (G_k). Didaktiker:innen bewerten die reduzierte Erklärung der Genetik jedoch im Mittel ebenso schlecht ($M_{Did} = 2,72$).

Aufgrund der kaum vorhandenen Differenzen zwischen den Statusgruppen, die sich auf deskriptiver Ebene zeigen, ergeben sich in der einfaktoriellen Varianzanalyse in fünf von sechs Fällen keine signifikanten Zusammenhänge zwischen der Einschätzung und der Zugehörigkeit zu einer der Statusgruppen. Für die Erklärung \ddot{O}_l ergibt sich der Fall, dass die Statusgruppenzugehörigkeit mit der Vergabe des Globalurteils inferenzstatistisch signifikant ($p = 0,04$) assoziiert ist, die post-hoc-Tests jedoch nicht aufklären können, welche Statusgruppen sich inferenzstatistisch signifikant voneinander unterscheiden lassen. Die in der Einschätzung gefundene Varianz kann daher nicht eindeutig auf die Gruppenzugehörigkeit zurückgeführt bzw. durch diese aufgeklärt werden.

Fachspezifischer Aspekt

Die letzte Tabelle (Tab. 51) zeigt die Einschätzungen in Bezug auf den fachspezifischen Aspekt pro Erklärung und getrennt nach Statusgruppen.¹⁷⁰

Tabelle 51: Nach Statusgruppen und Videovignetten differenzierte arithmetische Mittel und Standardabweichungen des fachspezifischen Aspekts und Ergebnisse der einfaktoriellen ANOVA mit dem Faktor Statusgruppenzugehörigkeit

Video	Video G_l <i>M (SD)</i>	Video G_k <i>M (SD)</i>	Video \ddot{O}_l <i>M (SD)</i>	Video \ddot{O}_k <i>M (SD)</i>	Video P_l <i>M (SD)</i>	Video P_k <i>M (SD)</i>
Statusgruppe						
Schüler:innen (1) ($N_{max} = 122$)	3,20 (0,91)	3,30 (0,93)	3,10 (0,84)	3,18 (0,92)	3,26 (0,99)	3,30 (0,98)
Studierende (2) ($N_{max} = 44$)	2,61 (0,74)	2,96 (0,82)	2,54 (0,74)	2,75 (0,63)	2,68 (0,73)	2,79 (0,77)
Lehrkräfte (3) ($N_{max} = 29$)	2,93 (0,68)	3,44 (0,78)	2,95 (0,70)	3,08 (0,65)	2,99 (0,82)	3,11 (0,75)
Didaktiker:innen (4) ($N_{max} = 24$)	3,14 (0,78)	3,03 (0,74)	2,77 (0,81)	3,10 (0,80)	3,25 (0,85)	3,12 (0,74)
df_{res}	215	212	210	211	210	209
$F(3, df_{res})$	5,72	2,58	5,46	4,09	4,64	3,55
η^2	0,07	0,04	0,07	0,04	0,06	0,05
p	$\leq 0,01$	0,06	$\leq 0,01$	0,01	$\leq 0,01$	0,02
post hoc (Gabriel)	(1) vs. (2)**		(1) vs. (2)**	(1) vs. (2)*	(1) vs. (2)**	(1) vs. (2)**

Bem.: *M*: Mittelwert; *SD*: Standardabweichung; N_{max} : aufgrund fehlender Werte bei einzelnen Videos den Berechnungen pro Statusgruppe maximal zugrunde liegende Personenanzahl; df_{res} : Freiheitsgrade der Residuen; *F*: F-Wert der Teststatistik; η^2 : Eta-Quadrat; *p*: p-Wert; *: $p \leq 0,05$; **: $p \leq 0,01$; (1): Schüler:innen; (2): Studierende; (3): Lehrkräfte; (4): Didaktiker:innen.

Die Standardabweichungen des fachspezifischen Aspekts sind ähnlich ausgeprägt wie in den weiteren Aspekten. Auffällig ist jedoch, dass die Werte für Schüler:innen tendenziell höher ausfallen.

¹⁷⁰ Weitere Analysen finden sich hinsichtlich der fachspezifischen Kriterien in Kapitel 6.2.4 in Hinblick auf die vierte Forschungsfrage zu den Unterschieden in der Einschätzung der linearen und komplexen Erklärungen.

Die Spannweiten für die Standardabweichungen sind für keine Statusgruppe auffällig, sodass über alle Erklärungen die durchschnittlichen Abweichungen fast zwei Notenstufen betragen.

Beim Vergleich der berechneten Mittelwerte des Aspekts wird sichtbar, dass Schüler:innen durchschnittlich auch die schlechtesten Bewertungen (Ausnahme bildet die Erklärung G_k , die von Lehrkräften die schlechtesten Bewertungen bekommt) vergeben. Innerhalb der Gruppe der Studierenden sind über alle sechs Videovignetten hingegen die niedrigsten Werte zu finden. Sie schätzen den fachspezifischen Aspekt demnach durchwegs am besten ein.

Die Erklärung, die für Schüler:innen ($M = 3,10$; $SD = 0,84$), Studierende ($M = 2,54$; $SD = 0,74$) und Didaktiker:innen ($M = 2,77$; $SD = 0,81$) hinsichtlich des fachspezifischen Aspekts am besten abschneidet, ist die lineare Variante aus dem Themenbereich der Ökologie (\ddot{O}). Für Lehrkräfte ist dies mit marginalem Unterschied die zweitbeste Erklärung ($M = 2,95$; $SD = 0,70$) – die reduzierte Vignette zum Themenbereich der Genetik schneidet bei Lehrkräften im Mittel etwas besser ab ($M = 2,93$; $SD = 0,68$). Damit sind sich alle Statusgruppen in Hinblick auf die beste Umsetzung des fachspezifischen Aspekts einig.

Die im Mittel am schlechtesten bewertete Erklärungen sind für Schüler:innen die Vignetten G_k ($M = 3,30$; $SD = 0,93$) und P_k ($M = 3,30$; $SD = 0,98$) sowie mit marginalem Unterschied auch P_i ($M = 2,26$; $SD = 0,99$). Studierende ($M = 2,96$; $SD = 0,82$) und Lehrkräfte ($M = 3,44$; $SD = 0,78$) bewerten im Durchschnitt ebenfalls Vignette G_k am schlechtesten. Didaktiker:innen schätzen hingegen Vignette P_i am schlechtesten ein ($M = 3,25$; $SD = 0,85$). Damit sind sich die Perspektiven mit angenommenem stärkerem Praxisbezug einig bezüglich der schlechtesten Umsetzung des fachspezifischen Aspekts und unterscheiden sich von der Einschätzung der Didaktiker:innen.

Die Ergebnisse der einfaktoriellen Varianzanalysen in der letzten Zeile der Tabelle 51 zeigen, dass in fünf von sechs Vignetten signifikante Unterschiede zu finden sind, die durch die Statusgruppenzugehörigkeit erklärt werden können. Die Effektstärken der signifikanten Zusammenhänge mit dem getesteten Faktor der Statusgruppenzugehörigkeit sind mit $\eta^2 = 0,07$ für die Vignetten G_i ($F(3,215) = 5,72$; $p \leq 0,01$) und \ddot{O}_i ($F(3,210) = 5,46$; $p \leq 0,01$) sowie mit $\eta^2 = 0,06$ für die Vignette P_i ($F(3,210) = 4,64$; $p \leq 0,01$) als mittel einzustufen (Cohen, 1988), die Effektstärken für die Unterschiede innerhalb der Vignetten \ddot{O}_k ($F(3,211) = 4,09$; $p = 0,01$) sowie P_k ($F(3,209) = 3,55$; $p = 0,02$) mit $\eta^2 = 0,04$ bzw. mit $\eta^2 = 0,05$ als klein (Cohen, 1988).¹⁷¹ Die post-hoc-Tests zeigen aufgrund der vorher beschriebenen deskriptiven Unterschiede zwischen Schüler:innen und Studierenden erwartungsgemäß zwischen diesen beiden Gruppen signifikante Bewertungsunterschiede: auf einem Signifikanzniveau von $p \leq 0,01$ für die Vignetten G_i ($M_{Sus} = 3,20$; $SD_{Sus} = 0,91$; $M_{Stu} = 2,61$; $SD_{Stu} = 0,74$; $d = -0,68$), \ddot{O}_i ($M_{Sus} = 3,10$; $SD_{Sus} = 0,84$; $M_{Stu} = 2,54$; $SD_{Stu} = 0,74$; $d = -0,69$), P_i ($M_{Sus} = 3,26$; $SD_{Sus} = 0,99$; $M_{Stu} = 2,68$; $SD_{Stu} = 0,73$; $d = -0,62$) sowie P_k ($M_{Sus} = 3,30$; $SD_{Sus} = 0,98$; $M_{Stu} = 2,79$; $SD_{Stu} = 0,77$; $d = -0,55$). Die Varianzhomogenität als Voraussetzung der Analyse ist für Vignette \ddot{O}_k nicht gegeben (Levene Tests mit $p = 0,04$). Für dieses Video wurde deshalb der Welch-Test verwendet (Field, 2018). Das Ergebnis zeigt auf einem Signifikanzniveau von $p = 0,02$ ebenfalls einen Unterschied zwischen Schüler:innen und Studierenden ($M_{Sus} = 3,18$; $SD_{Sus} = 0,92$; $M_{Stu} = 2,75$; $SD_{Stu} = 0,63$; $d = -0,50$). Nach den Vorschlägen von Cohen (1988) sind die Effektstärken alle als mittel einzuordnen.

¹⁷¹ Zu bemerken ist, dass die größeren Effektstärken bei den reduzierten Erklärungen zu finden sind. Die kleineren Effektstärken kommen zustande, da die Studierenden – ebenso wie Lehrkräfte – die komplexeren Vignetten stets schlechter als die reduzierten Vignetten bewerten. Daher sind die Mittelwertunterschiede auf deskriptiver Ebene kleiner, sodass folgend auch die Effekte kleiner werden.

Insgesamt schätzen damit Schüler:innen in fünf von sechs Fällen die Erklärungen signifikant schlechter hinsichtlich des fachspezifischen Aspekts ein als Studierende, auf die für den Aspekt insgesamt bei allen Vignetten mit die niedrigsten Mittelwerte und damit die besten Bewertungen entfallen.

6.2.2.2 Überblick und Einordnung der Ergebnisse zu Bewertungsunterschieden zwischen verschiedenen Statusgruppen in Hinblick auf vorgegebene Kriterien

Zu den Bewertungsunterschieden der Statusgruppen in Hinblick auf vorgegebene Kriterien lassen sich folgende Punkte festhalten:

Die Einigkeit bzw. Uneinigkeit innerhalb der Statusgruppen gemessen an der Standardabweichung beträgt über alle Aspekte und Videovignetten hinweg etwa eineinhalb bis zwei Notenstufen. Somit zeigt die Einigkeit bzw. Uneinigkeit in der Bewertung der Aspekte keine auffälligen Muster. Nur in wenigen Fällen beträgt die Standardabweichung weniger als eine Notenstufe (Studierende für die Adressatenorientierung in \ddot{O}_k $SD = 0,49$; Lehrkräfte für die Adressatenorientierung in \ddot{O}_l $SD = 0,44$; Didaktiker:innen für die Strukturiertheit in \ddot{O}_l $SD = 0,45$) und ebenfalls selten mehr als zwei Notenstufen (Didaktiker:innen im Globalurteil in \ddot{O}_k $SD = 1,06$, Lehrkräfte für die Strukturiertheit in \ddot{O}_k $SD = 1,07$).

Die nachstehende Tabelle (Tab. 52) stellt die Verteilung der besten und schlechtesten Bewertungen für das Globalurteil sowie die fünf weiteren Aspekte pro Videovignette und Statusgruppe zusammen, die in Kapitel 6.2.2.1 bereits beschrieben wurden.

Tabelle 52: Verteilung der besten und schlechtesten Bewertungen für das Globalurteil und die fünf Aspekte guten Erklärens pro Statusgruppe und Videovignette

Videovignette	G_l		G_k		\ddot{O}_l		\ddot{O}_k		P_l		P_k	
	+	-	+	-	+	-	+	-	+	-	+	-
Schüler:innen		ad		GLO (ad) se sw fa	GLO (ad) st sw fa		(ad)		se	se	(fa)	st fa
Studierende	(GLO) st			GLO ad (st) (se) sw fa	GLO ad fa			st se	se			(GLO)
Lehrkräfte	GLO			ad st se sw fa	ad st sw (fa)			GLO (st)	se			ad
Didaktiker:innen		ad	GLO		st se sw fa			GLO st	ad			(GLO) fa

Bem.: GLO: Globalurteil; ad: Adressatenorientierung; st: Strukturiertheit; se: Sprech- und Körperausdruck; sw: verständliche Sprache; fa: fachspezifischer Aspekt; (ad), (st) etc: zweitbeste bzw. zeitschlechteste Bewertung innerhalb des Aspekts mit nur marginalem Unterschied zur besten bzw. schlechtesten Bewertung.

Folgende Aussagen können anhand der Analyse zur Verteilung der besten bzw. schlechtesten Einschätzungen getroffen werden:

Schüler:innen und Studierende sind sich gemessen am Globalurteil einig, welche Erklärung am besten ist. Lehrkräfte und Didaktiker:innen bewerten jeweils unterschiedliche Erklärungen global beste. Bezüglich der schlechtesten Erklärung gemessen am Globalurteil sind sich ebenfalls Schüler:innen und Studierende einig. Lehrkräfte und Didaktiker:innen schätzen die gleiche Erklärung global als schlechteste ein. Im Vergleich der globalen Bewertungen zwischen den Videovignetten kann auf der hier fokussierten deskriptiven Ebene die Vorannahme, dass Schüler:innen und Studierende die Güte instruktionaler Erklärungen ähnlicher einschätzen als Schüler:innen und Lehrkräfte bzw. Schüler:innen und Didaktiker:innen, zunächst bestärkt werden.

Über alle Aspekte hinweg sind ergibt sich kein einheitliches Bewertungsmuster der Statusgruppen, in welchen Erklärungen die Aspekte am besten bzw. am schlechtesten umgesetzt sind. Betrachtet man die Aspekte getrennt, so ergibt sich folgendes Bild: Bezüglich der Adressatenorientierung sind sich auf der einen Seite Schüler:innen und Didaktiker und auf der anderen Seite Studierende und Lehrkräfte in Bezug auf die beste sowie auf die schlechteste Realisierung der Adressatenorientierung einig. Die Strukturiertheit erhält die besten Bewertungen von den Statusgruppen der Schüler:innen, Lehrkräften und Didaktiker:innen bei der gleichen Erklärung, sodass sich diese Statusgruppen einiger zu sein scheinen. Die schlechteste Strukturiertheit wird hingegen von Studierenden und Didaktiker:innen beim selben Video gesehen, von Schüler:innen und Lehrkräften bei jeweils anderen Vignetten. Beim Sprech- und Körperausdruck sind sich Schüler:innen, Studierende und Lehrkräfte eher einig, welche Vignette die beste Umsetzung enthält. Die schlechteste Umsetzung des Sprech- und Körperausdrucks in einer der sechs Erklärungen wird hingegen von Schüler:innen, Studierenden und Didaktiker:innen einheitlich eingeschätzt. Die Verständlichkeit der Sprache weist das einheitlichste Bewertungsmuster auf: Alle Statusgruppen sind sich sowohl hinsichtlich der besten als auch der schlechtesten Realisation einig. Ebenso verhält es sich mit der besten Umsetzung des fachlichen Aspekts im Vergleich der sechs Erklärungen. Bei der schlechtesten Variante weisen nur noch Schüler:innen, Studierende und Lehrkräfte Einigkeit auf.

Insgesamt verteilen sich so die Bewertungsunterschiede in unterschiedlicher Art und Weise je nach betrachtetem Aspekt, sodass die Vermutungen zu den Hintergründen dieser Bewertungsunterschiede nicht weiter systematisiert bzw. verallgemeinert werden können. Die Vorannahme, dass sich die Bewertungen von Schüler:innen und Lehrkräften bzw. Didaktiker:innen unähnlicher sind, kann mit dem Fokus auf die besten und schlechtesten Bewertungen auf Ebene der Konstrukte nicht bestätigt werden.

Weiterhin fällt auf, dass bei Schüler:innen und Studierenden diejenigen Erklärungen, die das beste bzw. schlechteste Globalurteil erhalten, auch mehr Aspekte die besten bzw. schlechtesten Mittelwerte aufweisen. Für Lehrkräfte und Didaktiker:innen gilt das nicht. Inwieweit die erfassten Aspekte das Globalurteil tatsächlich vorhersagen können bzw. welche Relevanz die Aspekte für die verschiedenen Statusgruppen haben, wird in Kapitel 6.2.3 mit hierarchischen linearen Regressionsmodellen analysiert, sodass die deskriptiven Befunde überprüft bzw. verallgemeinert werden können.

Die folgende Tabelle (Tab. 53) stellt die gefundenen signifikanten Bewertungsunterschiede pro Aspekt und Videovignette bzw. die Ergebnisse der entsprechenden post-hoc-Tests, die die Rückführung der signifikanten Varianzunterschiede auf spezifische Statusgruppen erlauben, dar.

Aus Gründen der Übersichtlichkeit sind die Werte der einfaktoriellen ANOVA (Faktor: Statusgruppe) und der post-hoc-Tests nicht mehr angegeben.

Tabelle 53: Übersicht zu den signifikanten Bewertungsunterschieden getrennt nach Aspekten und Videovignetten

Videovignette	G _i	G _k	Ö _i	Ö _k	P _i	P _k
Aspekt						
Globalurteil				(1) vs. (3)*		
Adressatenorientierung						
Strukturiertheit	(1) vs. (2)** (1) vs. (3)** (1) vs. (4)**	(1) vs. (4)**	(1) vs. (2)* (1) vs. (3)** (1) vs. (4)** (2) vs. (4)*		(1) vs. (3)** (1) vs. (4)**	(1) vs. (2)* (1) vs. (3)** (1) vs. (4)**
Sprech- und Körperausdruck	(1) vs. (2)*	(1) vs. (4)**	(1) vs. (4)**			(1) vs. (4)**
Verständliche Sprache						
Fachspezifischer Aspekt	(1) vs. (2)**		(1) vs. (2)**	(1) vs. (2)**	(1) vs. (2)**	(1) vs. (2)**

Bem.: (1): Schüler:innen; (2): Studierende; (3): Lehrkräfte; (4): Didaktiker:innen; *: $p \leq 0,05$; **: $p \leq 0,01$.

Signifikante Unterschiede bezüglich der Bewertung finden sich vor allem zwischen Schüler:innen und den verschiedenen Gruppen der Metaperspektive. Nur in einem Fall unterscheiden sich die Einschätzungen von Studierenden und Didaktiker:innen, also zwischen Gruppen der Metaperspektive, signifikant. Besonders häufig sind hingegen signifikante Unterschiede zwischen Schüler:innen und Didaktiker:innen sowie zwischen Schüler:innen und Studierenden zu verzeichnen.

In den Fällen der signifikanten Bewertungsunterschiede vergeben Schüler:innen fast ausschließlich die schlechteren Bewertungen – eine Ausnahme bildet der Unterschied im Globalurteil.

Die signifikanten Bewertungsunterschiede bestehen nicht bei allen Aspekten gleichermaßen: Die Strukturiertheit einer Erklärung wird insbesondere von Schüler:innen und den Gruppen der Metaperspektive unterschiedlich eingeschätzt. Der Sprech- und Körperausdruck wird von Schüler:innen und Didaktiker:innen auffällig häufig unterschiedlich bewertet, der fachspezifische Aspekt von Schüler:innen und Studierenden. Hohe Einigkeit zwischen allen befragten Statusgruppen besteht bei der Bewertung der Adressatenorientierung sowie der Verständlichkeit der Sprache.

Damit einhergehend kann die Vorannahme, dass die größten Differenzen zwischen Schüler:innen und Lehrkräften bzw. Schüler:innen und Didaktiker:innen liegen, auf inferenzstatistischer Ebene so nicht bestätigt werden. Lediglich für den Aspekt des Sprech- und Körpersudrucks zeigt sich in drei Videos die angenommene Differenz zwischen Schüler:innen, die durchgehend schlechter bewerten, und Didaktiker:innen, nicht jedoch für Schüler:innen und Lehrkräfte. Für die Strukturiertheit liegen zwar auch Differenzen zwischen Schüler:innen und Lehrkräften bzw. Didaktiker:innen vor, jedoch bestehen auch signifikante Unterschiede zwischen Schüler:innen und Studierenden in der selben Ausprägung (Schüler:innen bewerten die Strukturiertheit schlechter als Studierende).

Bei der Betrachtung der Unterschiede zwischen Schüler:innen und Studierenden bezüglich des fachspezifischen Aspekts müssen die Bewertungstendenzen mit einbezogen werden: Studierende beurteilen diesen Aspekt als einzige Statusgruppe durchgehend besser als Schüler:innen, Lehrkräfte und Didaktiker:innen. Post-hoc-Tests ergeben jedoch lediglich für die Bewertungen von

Schüler:innen und Studierenden signifikante Unterschiede. Formuliert werden muss demnach, dass Studierende die Wahrnehmung des dargestellten Komplexitätsgrade über alle sechs Erklärungen hinweg sehr positiv bewerten.

Die fehlenden Unterschiede in der Bewertung der Adressatenorientierung sind für die Unterrichtsgestaltung ein positives Zeichen, da (angehende) Lehrkräfte und Ausbilder:innen von Lehrkräften ihre Schülerschaft hinsichtlich deren Wissen und Fähigkeiten angemessen einschätzen.

6.2.3 Ergebnisse zu maßgeblichen Einflussfaktoren auf die globale Bewertung instruktionaler Erklärungen

Für die Beantwortung der Frage zu maßgeblichen Einflussfaktoren auf die Güte instruktionaler Erklärungen wurden hierarchische lineare Regressionsmodelle genutzt. Dabei wird der Zusammenhang zwischen Globalurteil (als abhängige Variable) und den Bewertungen der erfassten Aspekte inklusive der Persönlichkeitswirkung der erklärenden Lehrkraft (als unabhängige Variablen; Prädiktoren) allgemein beschrieben. Durch die hierarchische Modellierung wird berücksichtigt, dass die vorliegenden Daten (also die Bewertungen der Aspekte) nicht unabhängig sind: Die Bewertungen pro Person sind sich ähnlicher als die Bewertungen verschiedener Personen und die Merkmale der Videovignetten beeinflussen die Bewertung, sodass sich diese Bewertungen innerhalb der Vignetten ähnlicher sind als zwischen den Vignetten. Damit geht ein Teil der Varianz systematisch auf Personen und Vignetten zurück. Die Daten sind deshalb nach Personen (intraindividueller Faktor) und Videovignetten (interindividueller Faktor) strukturiert. Die Intraklassenkorrelation (intra-class-correlation (= ICC)) dient dabei als Maß für den Einfluss dieser zufälligen Effekte (Field, 2018).

Vor den Regressionsmodellen (statusgruppenübergreifend in Kap. 6.2.3.1 und getrennt nach Statusgruppen in 6.2.3.2.) sind jeweils die Korrelationen zwischen den erfassten Variablen tabellarisch dargestellt, um die Stärken der Zusammenhänge zwischen den unabhängigen Variablen (Aspekte guten Erklärens) sowie jeweils den unabhängigen Variablen und dem Globalurteil zu zeigen.

6.2.3.1 Ergebnisse zu statusgruppenübergreifend maßgeblichen Einflussfaktoren auf die globale Bewertung instruktionaler Erklärungen

Die nachstehende Tabelle (Tab. 54) zeigt zunächst die Interskalenkorrelationen und die Zusammenhänge mit dem Globalurteil über alle Erklärungen und alle Statusgruppen hinweg.

Dabei wird deutlich, dass es sich ausschließlich um positive Zusammenhänge handelt – sowohl zwischen den Aspekten als auch zwischen den Aspekten und dem Globalurteil.

Die bivariate Korrelation der Aspekte der Persönlichkeitswirkung und des Sprech- und Körperausdrucks ist auffällig hoch. Dieses Ergebnis ist nachvollziehbar, da die Wirkung der Person – hier erfasst über die empfundene Sympathie, die wahrgenommene Begeisterung und Natürlichkeit der erklärenden Person – anhand äußerlich wahrnehmbarer Merkmale getroffen wird, wenn die Person nicht näher bekannt ist. Dazu zählen stimmlich-sprecherische (paraverbale) und den Körperausdruck betreffende (extraverbale) Merkmale, die auch bei der Facette des Sprech- und Körperausdrucks zugrunde liegen. Auffällig hoch ist weiterhin die Korrelation der fachspezifischen Facette sowie der Adressatenorientierung. Es kann vermutet werden, dass eine inhaltliche Nähe der

Aspekte besteht: So kann beispielsweise die Anzahl der Einflussfaktoren, die zur Erklärung des Phänomens herangezogen wird, Teil der didaktischen Reduktion sein, die die Passung zum Wissen der Adressat:innen auf inhaltlicher Ebene sicherstellt. Auch das Item zur Faszination durch das aufgezeigte Phänomen hat einen Überschneidungsbereich zur Orientierung an Adressat:innen und deren Interessen. Trotz vereinzelter Korrelationskoeffizienten, die als starke Zusammenhänge eingeordnet werden können (Cohen, 1988), lassen die Werte darauf schließen, dass die Aspekte in ausreichendem Maß unterschiedliche Varianz erfassen, sodass sie in das Regressionsmodell eingehen können.

Tabelle 54: Interskalenkorrelationen und deren Zusammenhänge mit dem Globalurteil über alle Videos und alle Statusgruppen (Pearson-Produkt-Moment-Korrelation)

Aspekt	Adressatenorientierung	Strukturiertheit	Sprech- und Körperausdruck	Verständliche Sprache	Fachspezifische Facette	Persönlichkeitswirkung
Adressatenorientierung						
Strukturiertheit	0,56**					
Sprech- u. Körperausdruck	0,31**	0,41**				
Verständliche Sprache	0,50**	0,56**	0,32**			
Fachspezifische Facette	0,60**	0,50**	0,38**	0,31**		
Persönlichkeitswirkung	0,29**	0,28**	0,73**	0,21**	0,36**	
Globalurteil	0,39**	0,35**	0,28**	0,27**	0,32**	0,26**

Bem.: *: $p \leq 0,05$; **: $p \leq 0,01$.

Die Stärke der Zusammenhänge der Aspekte und des Globalurteils variiert von $r = 0,26$ (Persönlichkeitswirkung) bis $r = 0,39$ (Adressatenorientierung). Jede der erfassten Variablen besitzt daher in unterschiedlichem Maße Voraussagekraft für die Gesamtbewertung einer instruktionalen Erklärung. Da alle Facetten signifikante Zusammenhänge mit dem Globalurteil aufweisen, werden sie dementsprechend alle in das hierarchische lineare Regressionsmodell aufgenommen.

Die Ergebnisse des hierarchischen linearen Regressionsmodells sind Tabelle 55 zu entnehmen.¹⁷²

Da alle Aspekte sowie das Globalurteil auf derselben sechsstufigen Ratingskala eingeschätzt wurden, ist der unstandardisierte Regressionskoeffizient (b) in der originalen Metrik interpretierbar. An einem Beispiel verdeutlicht das: Wird von den Teilnehmenden der Aspekt der Adressatenorientierung auf der Ratingskala um eine Stufe besser bewertet, so verbessert sich das Globalurteil um den Wert 0,30. Grundlegend zeigt das Modell, dass 18,69 % der erfassten Varianz auf Personen und Videovignetten zurückgehen. Der relativ hohe Anteil der damit erfassten Varianz zeigt, dass die Implementierung der Ebene bzw. die hierarchische Modellierung notwendig ist.

¹⁷² Die Anzahl der Beobachtungen ergibt sich durch die Anzahl der Teilnehmenden, die zu sechs Videovignetten Einschätzungen abgegeben haben, wobei nur für maximal die Hälfte der Videovignetten vollständige Datensätze vorliegen, da der Sprech- und Körperausdruck zugunsten einer kürzeren Bearbeitungszeit (siehe Kap. 5.2.1) für drei Ausschnitte aus Vignetten eingeschätzt wurde. Daher reduziert sich die maximal einzubeziehende Datenmenge für die linearen hierarchischen Regressionsmodelle.

Tabelle 55: Hierarchische lineare Regressionsmodelle mit der abhängigen Variable Globalurteil und unter Berücksichtigung der nach Teilnehmenden und den sechs Videovignetten geordneten Datenstruktur

$N_{Beob.} = 587$		
ICC	%	18,69
Konstante	b	0,60**
Adressatenorientierung	b	0,30**
Strukturiertheit	b	0,15**
Sprech- und Körperausdruck	b	0,06
Verständliche Sprache	b	-0,04
Fachspezifische Facette	b	0,03
Persönlichkeitswirkung	b	0,08
R^2 (marginal konditional)		0,20 0,31

Bem.: $N_{Beob.}$: Anzahl der Beobachtungen; ICC: Intraklassenkorrelation; b : unstandardisierter Regressionskoeffizient; R^2 : Determinationskoeffizient; *: $p \leq 0,05$; **: $p \leq 0,01$.

Insgesamt können durch die in das Modell implementierten Faktoren (Adressatenorientierung, Strukturiertheit, Sprech- und Körperausdruck, verständliche Sprache, Wahrnehmung des dargestellten Komplexitätsgrads, Persönlichkeitswirkung) 20 % der festen Effekte ($R^2_{marginal} = 0,20$) aufgeklärt werden. Feste und zufällige Effekte zusammen ergeben 31 % ($R^2_{konditional} = 0,31$). Die Varianzaufklärung ist als mittlerer Effekt zu beschreiben (Cohen, 1988) und damit das Modell zufriedenstellend, um die Güte einer Erklärung zu beschreiben bzw. vorherzusagen.

Aus den Ergebnissen wird ersichtlich, dass statusgruppenübergreifend die Adressatenorientierung und die Strukturiertheit der instruktionalen Erklärung signifikante Prädiktoren für die Einschätzung der globalen Güte der Erklärung sind, wobei die Adressatenorientierung mit $b = 0,30$ doppelt so viel Einfluss auf das Globalurteil einer Erklärung nimmt als die Strukturiertheit mit $b = 0,15$.

Auch bei der Analyse der selbst herangezogenen Kriterien (offenes Textfeld) zeigen sich die Adressatenorientierung und die Strukturiertheit als zwei häufige Bezugspunkte für die Bewertung instruktionaler Erklärungen.¹⁷³ Während bei den selbst herangezogenen Kriterien auch häufig auf die fachlichen Aspekte rekurriert wird, zeigt sich der implementierte Aspekt ‚Wahrnehmung der dargestellten Komplexität‘ im linearen hierarchischen Regressionsmodell nicht als signifikanter Prädiktor. Mit der statusgruppenspezifischen Analyse der selbstgewählten Kriterien konnte gezeigt werden, dass insbesondere Studierende, Lehrkräfte und Didaktiker:innen auf fachliche Aspekte eingehen, sodass vermutet werden kann, dass der fachliche Aspekt erst bei der statusgruppengetrennten Modellierung als Prädiktor auftritt (siehe Kap. 6.2.3.2).

Im Gegensatz dazu können die Aspekte des Sprech- und Körperausdrucks, der verständlichen Sprache sowie Indikatoren zur Wahrnehmung des dargestellten Komplexitätsgrades und der Kontrollvariable der Persönlichkeitswirkung die Güte einer Erklärung nicht signifikant vorhersagen.¹⁷⁴ Kontraintuitiv ist dabei der negative Regressionskoeffizient für den Prädiktor der

¹⁷³ In dieser Zusammenfassung wurden die Aspekte, die im standardisierten Zugang aufgenommen wurden, ausgelassen. Globale Äußerungen und der Bezug auf den Einsatz von Repräsentationen sind im Textmaterial ebenfalls häufig.

¹⁷⁴ Nimmt man die Persönlichkeitswirkung aus dem Modell, so ändert sich die aufgeklärte Varianz insgesamt kaum ($R^2_{marginal} = 0,19$; $R^2_{konditional} = 0,31$), jedoch wird der Prädiktor des Sprech- und Körperausdrucks nun signifikant ($p \leq 0,01$; $b = 0,13$). Der hohe Anteil geteilter Varianz zeigte sich bereits in den hohen Korrelationen der beiden Aspekte (siehe Tab. 50). Das Ergebnis ist ein starkes empirisches Argument für die Implementation der Persönlichkeitswirkung als Kontrollvariable, insbesondere, um den Einfluss des Sprech- und Körperausdrucks nicht zu überschätzen. Zur Vollständigkeit die unstandardisierten Regressionskoeffizienten der weiteren Faktoren des Modells ohne Persönlichkeitswirkung: Konstante ($p \leq 0,01$; $b = 0,32$), Adressatenorientierung ($p \leq 0,01$; $b = 0,30$), Strukturiertheit ($p \leq 0,01$; $b = 0,14$), Verständliche Sprache ($p = 0,36$; $b = -0,04$), Fachspezifische Facette ($p = 0,43$; $b = 0,04$).

verständlichen Sprache ($b = -0,04$). Dieser deutet an, dass eine schlechtere Bewertung der sprachlichen Verständlichkeit zu einer besseren globalen Bewertung der instruktionalen Erklärung führt. Das Regressionsgewicht ist jedoch marginal, sodass das Globalurteil um eine Notenstufe besser ausfallen würde, wenn die sprachliche Verständlichkeit um 25 Stufen schlechter wäre. Die Bewertung der sprachlichen Verständlichkeit, wie sie hier operationalisiert wurde, hat somit über alle Statusgruppen hinweg keinen Einfluss auf die Voraussage der Güte der instruktionalen Erklärung.

6.2.3.2 Ergebnisse zu maßgeblichen Einflussfaktoren auf die globale Bewertung instruktionaler Erklärungen für Lernende und Lehrende

Die folgende Tabelle (Tab. 56) zeigt wie im vorherigen Abschnitt zunächst die Interskalenkorrelationen, nun jedoch nach Gruppen getrennt. Da sich zwischen den Gruppen der Metaperspektive (Studierende, Lehrkräfte, Didaktiker:innen) in den einfaktoriellem ANOVAS keine systematischen, signifikanten Unterschiede in den Bewertungen der Vignetten gezeigt haben, werden die Gruppen nicht mehr einzeln betrachtet, sondern zur SLD-Gruppe (=Studierende, Lehrkräfte, Didaktiker:innen) zusammengefasst.

Tabelle 56: Interskalenkorrelationen und deren Zusammenhänge mit dem Globalurteil über alle Videos (Pearson Produkt-Moment-Korrelation) nach SuS- und SLD-Gruppe getrennt

Aspekt	Adressatenorientierung	Strukturiertheit	Sprech- und Körperausdruck	Verständliche Sprache	Fachspezifische Facette	Persönlichkeitswirkung	Globalurteil
Adressatenorientierung		0,54**	0,29**	0,45**	0,53**	0,31**	0,31**
Strukturiertheit	0,58**		0,30**	0,53**	0,49**	0,29**	0,31**
Sprech- u. Körperausdruck	0,30**	0,43**		0,32**	0,28**	0,69**	0,33**
Verständliche Sprache	0,55**	0,62**	0,31**		0,24**	0,25**	0,25**
Fachspezifische Facette	0,70**	0,45**	0,44**	0,42**		0,26**	0,21**
Persönlichkeitswirkung	0,26**	0,22**	0,77**	0,17**	0,47**		0,28**
Globalurteil	0,48**	0,42**	0,26**	0,29**	0,49**	0,25**	

Bem.: Korrelationswerte oberhalb der Diagonale für Schüler:innen, unterhalb der Diagonale für Studierende, Lehrkräfte und Didaktiker:innen (SLD); *: $p \leq 0,05$; **: $p \leq 0,01$.

Auch bei der nach Gruppen getrennten Betrachtung der bivariaten Korrelation ergeben sich ausschließlich positive, signifikante Zusammenhänge zwischen den Aspekten sowie zwischen den Aspekten und der globalen Bewertung.

Für Korrelationen, die bei statusgruppenübergreifender Betrachtung aufgefallen sind (Sprech- und Körperausdruck und Persönlichkeitswirkung; Adressatenorientierung und fachspezifische Facette (= Wahrnehmung des dargestellten Komplexitätsgrades)), ergibt sich hier ein differenziertes Bild: Der hohe Korrelationskoeffizient, der sich bei statusgruppenübergreifender Analyse zwischen den Aspekten der Persönlichkeitswirkung und dem Sprech- und Körperausdruck zeigte, fällt bei der Gruppe der Metaperspektive sogar etwas höher aus ($r = 0,77$) als bei der Gruppe der Schüler:innen ($r = 0,69$).

Des Weiteren ist der Zusammenhang zwischen den Aspekten der Adressatenorientierung und der Wahrnehmung des dargestellten Komplexitätsgrades für die Gruppe der Metaperspektive deutlich höher ($r = 0,70$) als für die Gruppe der Schüler:innen ($r = 0,53$). Beide Aspekte enthalten für die Gruppe der Metaperspektive mehr Items als für die Gruppe der Schüler:innen: Schüler:innen werden in Hinblick auf die Realisierung der Adressatenorientierung nicht nach der Angemessenheit der didaktischen Reduktion gefragt (siehe Kap. 5.2.4.1). Die fachspezifische Facette enthält für die Gruppe der Metaperspektive zusätzlich das Item „Die Erklärung zeigt, wie es in der Natur wirklich ist“, da nur aus der Metaperspektive heraus das entsprechende Fachwissen für eine Einschätzung verfügbar ist. Auch das Item „Der Inhalt der Erklärung hilft Schüler:innen, andere Beispiele zu verstehen“ wird nur den Personen der Metaperspektive präsentiert, da hierfür fachdidaktisches Wissen notwendig ist, über das Schüler:innen nicht verfügen. Während das Item zur didaktischen Reduktion in inhaltlicher Nähe zur fachspezifischen Facette steht, da hier auch die Angemessenheit der Anzahl der Faktoren abgefragt wird, ist eine Verbindung der zwei Items aus der fachspezifischen Facette zur Adressatenorientierung nicht ersichtlich.

Für die fachspezifische Facette kann insgesamt formuliert werden, dass die Zusammenhänge dieser mit den anderen abgefragten Aspekten guten Erklärens in der Gruppe der Metaperspektive überwiegend stärker ausfallen als in der Gruppe der Schüler:innen. Eine Ausnahme bildet die Korrelation von Strukturiertheit und fachspezifischer Facette – hier besteht ein stärkerer Zusammenhang bei Schüler:innen. Insbesondere der starke Zusammenhang zwischen der fachspezifischen Facette und der Persönlichkeitswirkung ist auffällig. Aufgrund der rezipierten Literatur oder anderweitigen Vorannahmen kann hierzu keine Vermutung geäußert werden.

Vergleicht man die Zusammenhänge der Aspekte mit dem Globalurteil zwischen der Gruppe der Schüler:innen und der Metaperspektive, so fällt auf, dass die Korrelationskoeffizienten der Metaperspektive für die Facetten Adressatenorientierung, Strukturiertheit, verständliche Sprache sowie die fachspezifische Facette höher ausfallen als bei Schüler:innen. Begründet werden kann dies zum Teil durch eine höhere Itemanzahl für die Gruppe der Metaperspektive (Adressatenorientierung, fachspezifische Facette).

Die Ergebnisse des hierarchischen linearen Regressionsmodells sind Tabelle 57 einzusehen.

Tabelle 57: Hierarchische lineare Regressionsmodelle mit der abhängigen Variable Globalurteil und unter Berücksichtigung der nach Teilnehmenden und den sechs Videovignetten geordneten Datenstruktur

		Schüler:innen ($N_{Beob.} = 302$)	SLD-Gruppe ($N_{Beob.} = 285$)
ICC	%	14,04	28,50
Konstante	b	0,69**	0,38
Adressatenorientierung	b	0,23**	0,25**
Strukturiertheit	b	0,22**	0,17*
Sprech- und Körperausdruck	b	0,18*	-0,02
Verständliche Sprache	b	-0,09	-0,02
Fachspezifische Facette	b	-0,05	0,21**
Persönlichkeitswirkung	b	0,06	0,06
R^2 (marginal konditional)		0,16 0,25	0,29 0,44

Bem.: $N_{Beob.}$ Anzahl der Beobachtungen; ICC: Intraklassenkorrelation; b : unstandardisierter Regressionskoeffizient; R^2 : Determinationskoeffizient; *: $p \leq 0,05$; **: $p \leq 0,01$.

Betrachtet man die Intraklassenkorrelation (ICC), zeigt das Modell der Schüler:innen ($ICC_{Sus} = 14,04\%$) im Vergleich zum Modell der Lehrenden ($ICC_{SLD} = 28,50\%$) geringere Werte. Der ICC gibt insgesamt an, welcher Anteil der Varianz auf Personen und Videovignetten zurückgeführt werden kann. Die jeweiligen ICC zeigen auch, dass insbesondere für die Gruppe der Metaperspektive die Berücksichtigung der hierarchischen Struktur im Modell sinnvoll ist. Aus dem Korrelationsmaß kann abgeleitet werden, dass Lehrende stringenter bzw. stabiler bewerten als Schüler:innen.

Insgesamt können durch die in die Modelle implementierten Faktoren (Adressatenorientierung, Strukturiertheit, Sprech- und Körperausdruck, verständliche Sprache, Wahrnehmung des dargestellten Komplexitätsgrads, Persönlichkeitswirkung) bei Schüler:innen 16 % der festen Effekte ($R^2_{\text{marginal}} = 0,16$) und 25 % der festen und zufälligen Effekte ($R^2_{\text{konditional}} = 0,25$) erklärt werden. Die Varianzaufklärung entspricht daher einem mittleren Effekt (Cohen, 1988), sodass für die Statusgruppe der Schüler:innen das Modell zur Beschreibung der Güte einer instruktionalen Erklärung als zufriedenstellend beschreiben werden kann. Für Lehrende fallen beide Werte deutlich höher aus: Der Anteil für feste Effekte liegt bei 29 % ($R^2_{\text{marginal}} = 0,29$) und der Anteil für feste und zufällige Effekte bei 44 % ($R^2_{\text{konditional}} = 0,44$). Die Varianzaufklärung für die Stichprobe der Lehrenden entspricht daher einem großen Effekt nach Cohen (1988).

Die Varianzaufklärung entspricht daher einem großen Effekt (Cohen, 1988). Betrachtet man das Verhältnis der beiden $R^2_{\text{konditional}}$, fällt auf, dass bei Schüler:innen ein geringerer Anteil der Varianz auf Personen und Videovignetten (also die zufälligen Effekte) zurückzuführen ist als bei Lehrenden. Damit kann mithilfe des Modells auch gezeigt werden, was sich in den deskriptiven Befunden (siehe Kap. 6.2.2) andeutet: Schüler:innen bewerten über alle Videovignetten hinweg ähnlich und damit weniger fallspezifisch als Lehrende. In Hinblick auf die beiden R^2_{marginal} zeigt sich, dass die Varianzaufklärung, die systematisch auf die Faktoren zurückzuführen ist, bei Lehrenden höher ausfällt als bei Schüler:innen. Dabei muss beachtet werden, dass bei Schüler:innen zwei Items weniger enthalten waren (Item: *Fa_hil*; *Ad_red*), sodass grundsätzlich etwas weniger Varianz erfasst werden kann. Insgesamt zeigt sich jedoch bereits in den niedrigeren Korrelationen der Aspekte mit dem Globalurteil (siehe Tab. 56), dass Schüler:innen weniger stringent bewerten.

Interessante Ergebnisse zeigen sich bei der statusgruppenspezifischen Betrachtung der Regressionskoeffizienten: Für Schüler:innen sind die Adressatenorientierung, die Strukturiertheit sowie der Sprech- und Körperausdruck signifikante Einflussfaktoren auf die Güte der Erklärung. Dabei sind die Adressatenorientierung mit $b = 0,23$ und die Strukturiertheit mit $b = 0,22$ etwa gleich starke Prädiktoren, der Sprech- und Körperausdruck mit $b = 0,18$ hat im Vergleich dazu etwas weniger Einfluss auf die Bewertung instruktionaler Erklärungen. Dass sich bei Schüler:innen die Persönlichkeitswirkung der erklärenden Person nicht als signifikanter Einflussfaktor zeigt, kann auch an dem hohen Anteil der geteilten Varianz des Aspekts mit dem Sprech- und Körperausdruck liegen (siehe Tab. 56).

Für Lehrende sind ebenfalls die Adressatenorientierung und die Strukturiertheit maßgebliche Einflussfaktoren, jedoch nicht der Sprech- und Körperausdruck. Zusätzlich stellt die fachspezifische Facette (die Wahrnehmung des dargestellten Komplexitätsgrades) einen starken Einflussfaktor auf die Einschätzung der Güte einer instruktionalen Erklärung dar. Das Ergebnis ist insofern überraschend, als sich die Aspekte der Adressatenorientierung und der fachspezifische Aspekt zur Wirkung der dargestellten Komplexität einen hohen Anteil an Varianz teilen (siehe Tab. 56; $r = 0,70$), sodass erwartbar gewesen wäre, dass einer der beiden Aspekte im Modell als signifikanter Einflussfaktor evident wird. Auch die Stärke der Zusammenhänge variiert: Die Adressatenorientierung ist mit $b = 0,25$ mit vergleichbarer Größe zu derjenigen der Schüler:innen

der stärkste Einflussfaktor, die fachliche Facette mit $b = 0,21$ der zweitgrößte und die Strukturiertheit mit $b = 0,17$ der geringste signifikante Einflussfaktor, der bei Schüler:innen mit $b = 0,22$ höher ausfällt.

Dass für Schüler:innen die Wahrnehmung der dargestellten Komplexität kein relevanter Prädiktor für die Güte der Erklärung ist, ist insbesondere in Hinblick auf die in Kapitel 3.4 herausgearbeiteten Herausforderungen beim Lernen bemerkenswert. Eine genauere Betrachtung der Unterschiede zwischen den Einschätzungen von Lernenden und Lehrenden in Abhängigkeit des dargestellten Komplexitätsgrades folgt in Kapitel 6.2.3.

Bereits bei der Analyse der selbst gewählten Kriterien aus dem offenen Textfeld zur Begründung der Notenvergabe (siehe Kap. 6.2.1.3 bzw. in der Zusammenfassung in Kap. 6.2.1.5) hat sich gezeigt, dass Studierende, Lehrkräfte und Didaktiker:innen bei der Bewertung der instruktionalen Erklärungen häufiger auf fachliche Aspekte zurückgreifen als Schüler:innen. Die Adressatenorientierung war für alle Statusgruppen ein häufiger Bezugspunkt für die Bewertung. Der Einfluss des Sprech- und Körperausdrucks für Schüler:innen war auf Grundlage der Äußerungen im offenen Textfeld nicht so abzuschätzen.

6.2.3.3 Überblick und Einordnung der Ergebnisse zu maßgeblichen Einflussfaktoren auf die globale Bewertung instruktionaler Erklärungen

Anhand der Regressionsmodelle kann gezeigt werden, dass durch die Faktoren (Adressatenorientierung, Strukturiertheit, Sprech- und Körperausdruck, verständliche Sprache, Wahrnehmung des dargestellten Komplexitätsgrads, Persönlichkeitswirkung) ein zufriedenstellender Anteil der Varianz aufgeklärt werden kann. Über alle Statusgruppen hinweg können 20 % der Varianz durch die Faktoren vorhergesagt werden, was einem mittleren bis großen Effekt entspricht. Bei nach Statusgruppen getrennten Modellen zeigt sich, dass der Anteil der Varianzaufklärung mit 16 % für Schüler:innen deutlich geringer ausfällt als für Lehrende mit 29 %.

Über alle Statusgruppen hinweg zeigen sich die Adressatenorientierung und die Strukturiertheit als maßgebliche Einflussfaktoren auf die Güte instruktionaler Erklärungen. Damit kann die Vorannahme bestätigt werden, dass die Adressatenorientierung ein wesentlicher Einflussfaktor ist.

Trennt man nach Statusgruppen der Lernenden, den eigentlichen Adressat:innen von Erklärungen, und der Lehrenden (Studierende, Lehrkräfte und Didaktiker:innen), zeigen sich Unterschiede. Für Schüler:innen sind (nach absteigendem Einfluss geordnet) die Adressatenorientierung, die Strukturiertheit sowie der Sprech- und Körperausdruck maßgebliche Einflussfaktoren. Die Vorannahmen können daher insofern bestätigt werden, als die Adressatenorientierung und der Sprech- und Körperausdruck wesentlich zur Güte einer Erklärung beitragen. Die Persönlichkeitswirkung zeigt sich im Modell hingegen nicht als relevanter Prädiktor, wobei ein hoher Anteil an geteilter Varianz mit dem Sprech- und Körperausdruck dazu beitragen könnte.

Für Lehrende sind hingegen – ebenfalls nach absteigendem Regressionsgewichten geordnet – die Adressatenorientierung, die Wahrnehmung des dargestellten Komplexitätsgrades sowie die Strukturiertheit einer Erklärung maßgebliche Prädiktoren. Damit können die Vorannahmen insofern bestätigt werden, als für Lernende und für Lehrende die Adressatenorientierung wesentlich zur Güte einer Erklärung beiträgt. Die Vorannahmen für die Gruppe der Lehrenden können damit bestätigt werden, wobei die Strukturiertheit nicht als relevanter Prädiktor angenommen wurde.

Die Ergebnisse verdeutlichen, was bereits die Auswertungen der Begründungen für die Notenvergabe (offenes Textfeld) gezeigt haben (siehe Kap. 6.2.1.3. bz. 6.2.1.5): Die befragten Statusgruppen legen unterschiedliche Kriterien an instruktionale Erklärungen an. Die unterschiedlichen Maßstäbe sind dabei nicht gegeneinander auszuspielen bzw. als besser und schlechter zu bewerten – sie sind Hinweis darauf, dass die Sicht der Schüler:innen bei Forschungsansätzen berücksichtigt werden sollte, wenn es um die Einschätzung der Güte aus einer normativen Sicht geht (Baumert, 2002; Berliner, 2005; Gautschi, 2011; Hattie, 2015; Kleickmann et al., 2019). Nur Schüler:innen können dazu verhelfen, zu verstehen, wie Lernen, Unterricht, Erklärungen usw. aus ihrer Perspektive auf sie wirken. Die Wahrnehmung des Lernangebots ist als Filter vor den Lernaktivitäten ein relevanter Faktor für Lernerfolg (Helmke, 2007).

6.2.4 Ergebnisse zu den Bewertungs- und Begründungsunterschieden zwischen linearen und komplexen instruktionalen Erklärungen in Abhängigkeit von der Statusgruppe

Die vorangegangenen Kapitel fokussierten darauf, welche Kriterien von verschiedenen Statusgruppen an instruktionale Erklärungen gestellt werden. Die Videovignetten wurden dabei allgemein als Stimuli betrachtet. Bisher unbeachtet blieb damit die experimentelle Manipulation der Videovignetten: Drei Erklärungen stellen einen linearen Ursache-Wirkungs-Zusammenhang dar, bei dem eine Ursache zu einer Wirkung führt. Drei weitere Erklärungen zeigen hingegen ein komplexeres Wirkungsgefüge auf und die erklärende Person verweist explizit auf Nichtwissen und unsicheres Wissen (zur Konzeptualisierung und Operationalisierung siehe Kap. 5.2.2.2). Der Unterschied rückt in diesem Kapitel in den Fokus. Da sich in den vorangegangenen Analysen relevante statusgruppenspezifische Unterschiede gezeigt haben, wird auf die Darstellung von statusgruppenübergreifenden Analysen verzichtet.

Dazu werden zunächst die statusgruppenspezifischen Unterschiede in der Bewertung komplexer und linearer Erklärungen anhand des Globalurteils analysiert (Kap. 6.2.4.1) sowie anschließend zur Übersicht zusammengefasst und eingeordnet (Kap. 6.2.4.2). Um weitere Anhaltspunkte zur Wahrnehmung der Güte von komplexen und linearen Erklärungen zu erhalten, werden danach die Ergebnisse der evaluativen qualitativen Inhaltsanalyse in Hinblick auf Unterschiede zwischen linearen und komplexen Erklärungen ausgewertet (Kap. 6.2.4.2) und abschließend zusammengefasst (Kap. 6.2.4.3).

6.2.4.1 Ergebnisse zu den Bewertungsunterschieden zwischen linearen und komplexen instruktionalen Erklärungen in Abhängigkeit von der Statusgruppe

Für die Analysen zu unterschiedlichen Einschätzungen instruktionaler Erklärungen in Abhängigkeit des dargestellten Komplexitätsgrades durch verschiedene Statusgruppen wird das Globalurteil herangezogen.¹⁷⁵ In den durchgeführten Varianzanalysen mit gemischtem Design (mixed ANOVA) wird der Komplexitätsgrad als Messwiederholungsfaktor modelliert: Zwei Erklärungen im selben Themenbereich (Genetik, Ökologie, Physiologie) greifen ähnliche Phänomene auf, die einmal einen linearen Ursache-Wirkungs-Zusammenhang und einmal ein komplexes Ursache-Wirkungs-Gefüge aufzeigen. Die Statusgruppen gehen als vierstufiger Zwischensubjektfaktor (Schüler:innen, Studierende, Lehrkräfte, Didaktiker:innen) ein. Das Globalurteil wird als abhängige Variable, der Komplexitätsgrad sowie die Statusgruppe der Teilnehmenden werden als unabhängige Variablen im

¹⁷⁵ In den vorangegangenen Analysen wurde evident, dass das Globalurteil hoch mit den fünf Aspekten korreliert, sodass die Analyse zu Bewertungsunterschieden anhand des Globalurteils durchgeführt werden kann.

Modell berücksichtigt. Neben dem Haupteffekt des Komplexitätsgrads ist auch der Interaktionseffekt von besonderem Interesse.

Für die Prüfung der Voraussetzung der Normalverteilung wird – wie für die einfaktoriellem ANOVAs – für Schüler:innen der Kolmogorov-Smirnov-Test und für die Gruppen der Lehrenden der Shapiro-Wilk-Test durchgeführt (Bühner & Ziegler, 2017). Die Ergebnisse, die bereits in Kapitel 6.2.2 als Grundlage dienen, sind tabellarisch in Anhang E einzusehen. Die mixed ANOVA ist gegenüber Verletzungen bei der vorliegenden Stichprobengröße insgesamt als robust einzuschätzen (Bortz & Schuster, 2010). Für die Prüfung der Voraussetzung der Gleichheit der Varianzen wurde der Levene-Test eingesetzt. Die Ergebnisse sind ebenfalls im Anhang (Anhang F) einzusehen (die Varianzgleichheit ist für alle Vignetten gegeben).

Die nachstehende Tabelle (Tab. 58) enthält im oberen Teil die deskriptiven Statistiken des Globalurteils (arithmetische Mittel, Standardabweichungen) und im unteren Teil die Ergebnisse der Varianzanalysen mit gemischtem Design.

Tabelle 58: Nach Statusgruppen und Erklärthema sowie Komplexitätsgrad differenzierte mittlere Bewertungen (inkl. Standardabweichungen) und Varianzanalyse mit gemischtem Design

Themenbereich Komplexitätsgrad Statusgruppe	Genetik		Ökologie			Physiologie			
	linear M (SD)	komplex M (SD)	linear M (SD)	komplex M (SD)	linear M (SD)	komplex M (SD)			
Schüler:innen ($N_{max} = 122$)	2,01 (0,76)	2,30 (0,74)	1,84 (0,81)	2,00 (0,85)	2,04 (0,72)	2,15 (0,79)			
Studierende ($N_{max} = 51$)	1,88 (0,81)	2,22 (0,86)	1,87 (0,89)	2,14 (0,75)	2,09 (0,81)	1,88 (0,68)			
Lehrkräfte ($N_{max} = 34$)	1,98 (0,58)	2,38 (0,98)	2,20 (0,77)	2,64 (0,95)	2,28 (0,75)	2,06 (0,88)			
Didaktiker:innen ($N_{max} = 31$)	2,01 (0,87)	1,88 (0,77)	1,99 (0,69)	2,33 (1,05)	2,06 (0,82)	1,90 (0,75)			
ANOVA	df	F	η^2	df	F	η^2	df	F	η^2
Komplexitätsgrad	1	8,51**	0,04	1	11,96**	0,05	1	2,47	0,01
Statusgruppe	3	1,39	0,02	3	5,25**	0,07	3	0,72	0,01
Kompl.grad x Statusgr.	3	2,71*	0,03	3	0,74	0,01	3	2,08	0,03
Residuen	229			226			222		

Bem.: N_{max} : aufgrund fehlender Werte bei einzelnen Videos den Berechnungen pro Statusgruppe maximal zugrunde liegende Personenanzahl; M: Mittelwert; SD: Standardabweichung; ANOVA: Varianzanalyse mit gemischtem Design; df: Freiheitsgrade; F: F-Wert; η^2 : partielles Eta-Quadrat; *: $p \leq 0,05$; **: $p \leq 0,01$.

Themenbereich Genetik

Auf deskriptiver Ebene lässt sich anhand der Mittelwerte feststellen, dass sich Schüler:innen, Studierende und Lehrkräfte einig sind, dass die komplexe Erklärvariante schlechter als die lineare Variante ist. Nur Didaktiker:innen bewerten die komplexe Variante besser als die lineare Erklärvariante. Inferenzstatistisch lässt sich bei gegebener Varianzhomogenität (Levene-Tests mit $p = 0,06$ für Video G₁ und $p = 0,13$ für Video G₂) ein signifikanter Interaktionseffekt feststellen. Das bedeutet, dass der dargestellte Komplexitätsgrad in Abhängigkeit von der Statusgruppe beurteilt wird. Anhand der angegebenen Mittelwerte wird bereits deutlich, dass es sich dabei um die Bewertung der Didaktiker:innen handelt, die die lineare Erklärvariante im Gegensatz zu allen

weiteren Statusgruppen deutlich schlechter als das komplexe Pendant bewerten. Der folgende Profilplot kann den numerisch gezeigten Interaktionseffekt noch einmal illustrieren.¹⁷⁶

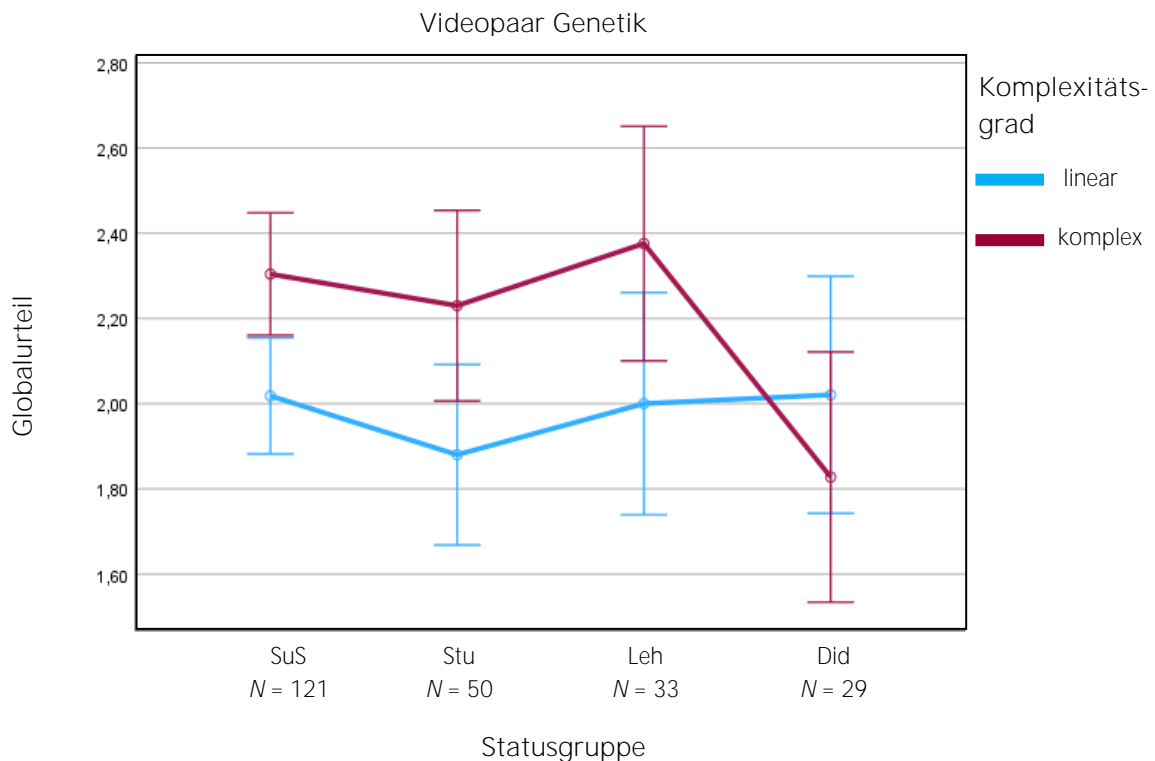


Abbildung 25: Profilplot zur deskriptiven Darstellung des Interaktionseffekts (Abhängigkeit des Faktors ‚Komplexitätsgrad‘ vom Faktor ‚Statusgruppe‘) für den Themenbereich Genetik mit dem Faktor ‚Komplexitätsgrad‘ als trace-Faktor und dem Faktor ‚Statusgruppe‘ auf der x-Achse inklusive des Fehlerbalkens für das 95-%-Konfidenzintervall

Im Themenbereich Genetik besteht weiterhin ein signifikanter Haupteffekt für den Faktor Komplexitätsgrad. Der Haupteffekt bedeutet, dass sich der Gesamtmittelwert der Globalurteile über alle Statusgruppen zur linearen Erklärvariante signifikant vom Gesamtmittelwert der Globalurteile über alle Statusgruppen zur komplexen Erklärvariante unterscheidet. Mit $\eta^2 = 0,04$ ist der Effekt als Maß für die aufgeklärte Varianz als annähernd mittel einzustufen (Cohen, 1988). Aufgrund des gefundenen Interaktionseffekts kann jedoch nicht formuliert werden, dass die komplexe Erklärvariante allgemein signifikant schlechter beurteilt wird.

Um den signifikanten Effekt weiter zu beschreiben, werden Paarvergleiche (t-Tests für verbundene Stichproben) durchgeführt. Die Ergebnisse sind in Tabelle 59 abgebildet.

Tabelle 59: Effektstärken zu den Mittelwertsunterschieden (t-Tests) im Themenbereich Genetik zwischen linearen und komplexen Erklärungen getrennt nach Statusgruppen

Statusgruppe	t	df	p	d
Schüler:innen	-3,62	120	≤ 0,01	0,87
Studierende	-2,77	49	≤ 0,01	0,89
Lehrkräfte	-2,00	32	≤ 0,05	1,08
Didaktiker:innen	1,07	28	n.s.	0,97

Bem.: t: T-Wert; df: Freiheitsgrade; p: p-Wert; d: Effektstärke nach Cohen.

¹⁷⁶ Ein weiterer Profilplot mit dem Faktor ‚Globalurteil‘ auf der Kategorienachse (x-Achse) bestätigt, dass es sich um einen disordinalen Interaktionseffekt handelt.

Für Schüler:innen zeigt sich ein statistisch signifikanter Bewertungsunterschied ($t = -3,62$; $p \leq 0,01$; $N = 121$). Dabei wird die lineare Erklärung ($M = 2,02$; $SD = 0,75$) von den teilnehmenden Schüler:innen signifikant besser als die komplexe Erklärung ($M = 2,30$; $SD = 0,74$) bewertet. Die Effektstärke nach Cohen (1988) liegt bei $d = 0,87$ und entspricht damit einem starken Effekt.

Bei Studierenden zeigt sich ein aufgrund der deskriptiven Ergebnisse erwartbarer signifikanter Bewertungsunterschied ($t = -2,77$; $p \leq 0,01$; $N = 50$). Auch sie bewerten die lineare Vignette ($M = 1,88$; $SD = 0,81$) signifikant besser als die komplexe Vignette ($M = 2,23$; $SD = 0,86$). Die Effektstärke ist mit $d = 0,89$ ebenfalls als groß einzuordnen (Cohen, 1988).

Auch Lehrkräfte bewerten die lineare und komplexe Erklärvarianten signifikant unterschiedlich ($t = -2,00$; $p \leq 0,05$; $N = 33$), wobei ebenfalls die komplexe Variante schlechter bewertet wird ($M = 2,38$; $SD = 0,98$) als die lineare Variante ($M = 2,00$; $SD = 0,57$). Die Effektstärke beträgt $d = 1,08$ und ist damit als groß einzuordnen (Cohen, 1988).

Der deskriptive Unterschied in der Bewertung der Didaktiker:innen, der gegenläufig zu den Unterschieden der drei weiteren Statusgruppen ist, lässt sich inferenzstatistisch nicht gegenüber einem zufälligen Effekt absichern ($t = 1,07$; $p = 0,29$; $N = 29$).

Themenbereich Ökologie

Auf Ebene der deskriptiven Statistiken ergibt sich anhand der Mittelwerte ein eindeutiges Bild: Alle Statusgruppen bewerten die komplexe Erklärvariante schlechter als die lineare Erklärvariante. Im Unterschied zum Themenbereich der Genetik bewerten also auch Didaktiker:innen wie die weiteren drei Statusgruppen die lineare Erklärung besser. Der Profilplot (Abb. 26) zeigt die Ergebnisse noch einmal grafisch.

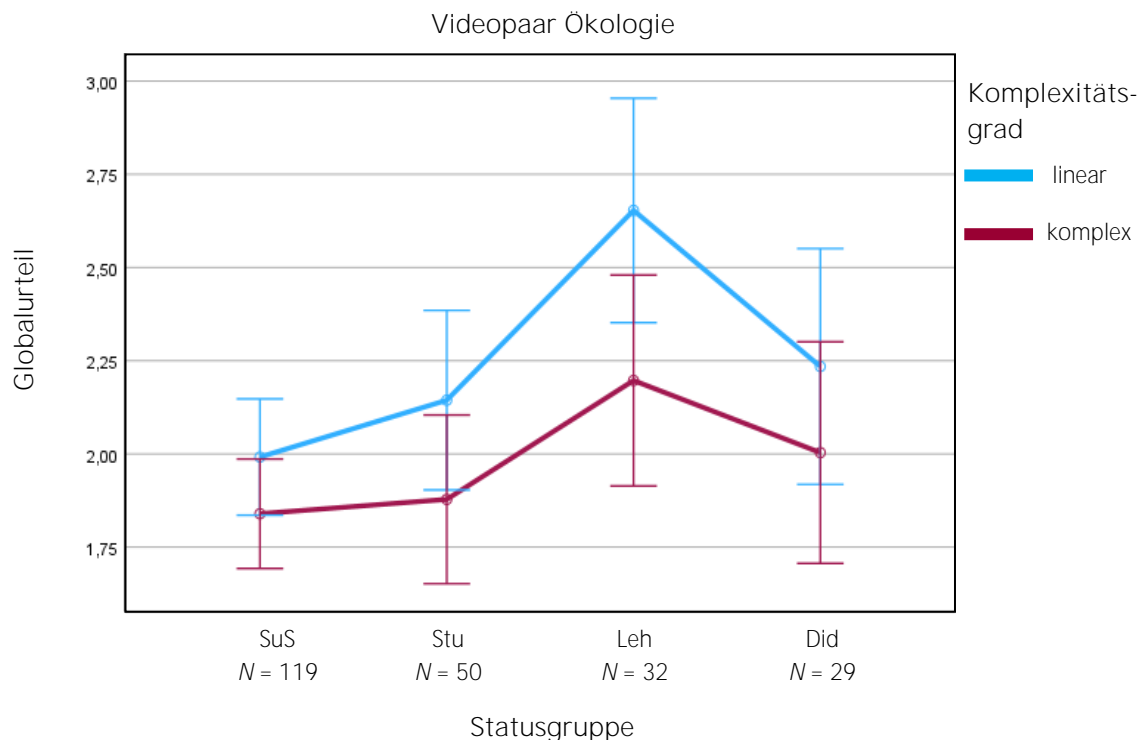


Abbildung 26: Profilplot zur deskriptiven Darstellung des Interaktionseffekts (Abhängigkeit des Faktors ‚Komplexitätsgrad‘ vom Faktor ‚Statusgruppe‘) für den Themenbereich Ökologie mit dem Faktor ‚Komplexitätsgrad‘ als trace-Faktor und dem Faktor ‚Statusgruppe‘ auf der x-Achse inklusive des Fehlerbalkens für das 95%-Konfidenzintervall

Zur Prüfung der Voraussetzungen für die Analyse wurden Levene-Tests durchgeführt, die zeigen, dass Varianzhomogenität gegeben ist ($p = 0,26$ für Video Ö_i und $p = 0,47$ für Video Ö_k). Da kein signifikanter Interaktionseffekt besteht, können die Haupteffekte interpretiert werden. Der Haupteffekt des Zwischensubjektfaktors ‚Statusgruppe‘ ist in Hinblick auf die Fragestellung nicht sinnvoll zu interpretieren. Er gibt an, ob sich der Mittelwert über beide Erklärvarianten einer Statusgruppe von den Mittelwerten über beide Erklärvarianten der anderen Statusgruppen signifikant unterscheidet. Damit können folglich keine Aussagen über die statusgruppenabhängigen Bewertungsunterschiede zwischen den Erklärvarianten getroffen werden.

Der Haupteffekt des Messwiederholungsfaktors ‚Komplexitätsgrad‘ ist ebenfalls signifikant und mit $\eta^2 = 0,05$ als mittlerer Effekt einzustufen (Cohen, 1988). Der Unterschied in der Darstellung (komplex vs. linear) erklärt demnach über alle Statusgruppen hinweg 5 % der Varianz, die in den Daten gefunden wird. Aufgrund der Mittelwerte wird deutlich, dass die komplexen Erklärvarianten über alle Statusgruppen hinweg schlechter bewertet werden als die linearen Pendanten. Um den Effekt detaillierter beschreiben zu können, werden anschließend nach Statusgruppen getrennte Paarvergleiche (t -Test für verbundene Stichproben) präsentiert.

Table 60: Effektstärken zu den Mittelwertsunterschieden (t -Tests) im Themenbereich Ökologie zwischen linearen und komplexen Erklärungen getrennt nach Statusgruppen

Statusgruppe	t	df	p	d
Schüler:innen	-1,60	118	n.s.	0,87
Studierende	-2,14	49	$\leq 0,05$	0,88
Lehrkräfte	-2,02	31	$\leq 0,05$	1,28
Didaktiker:innen	-1,20	28	n.s.	0,97

Bem.: t : T -Wert; df : Freiheitsgrade; p : p -Wert; d : Effektstärke nach Cohen.

Dabei wird der Bewertungsunterschied von Schüler:innen nicht signifikant ($t = -1,60$; $p = 0,11$; $N = 119$).

Für die Statusgruppe der Studierenden lässt sich hingegen ein signifikanter Bewertungsunterschied finden ($t = -2,14$; $p \leq 0,05$; $N = 50$). Dabei wird das komplexe Video ($M = 2,14$; $SD = 0,75$) signifikant schlechter bewertet als das lineare Video ($M = 1,88$; $SD = 0,89$). Die ermittelte Effektstärke beträgt $d = 0,88$, sodass es sich um einen großen Effekt handelt (Cohen, 1988).

Bei Lehrkräften zeigt sich ebenfalls ein derartig gerichteter, signifikanter Bewertungsunterschied ($t = -2,02$; $p \leq 0,05$; $N = 32$): Das komplexe Video wird schlechter bewertet ($M = 2,20$; $SD = 0,76$) als das lineare Video ($M = 2,65$; $SD = 0,98$). Cohens d beträgt 1,28, was einem starken Effekt entspricht (Cohen, 1988).

Wie bei Schüler:innen kann bei Didaktiker:innen der Bewertungsunterschied nicht gegenüber dem Zufall abgesichert werden ($t = -1,20$; $p = 0,24$; $N = 29$).

Damit ergibt sich auf deskriptiver Ebene ein einheitliches Bewertungsmuster, innerhalb dessen die komplexen Erklärungen von allen Statusgruppen schlechter bewertet werden, was aber auf inferenzstatistischer Ebene nur für die Statusgruppen der Studierenden und der Lehrkräfte abgesichert werden kann.

Themenbereich Physiologie

Bereits auf der Ebene der deskriptiven Statistik ergeben sich für dieses Vignettenpaar interessante Auffälligkeiten: Die im Mittel besseren Bewertungen von Studierenden, Lehrkräften und Didaktiker:innen sammeln sich bei den komplexen Erklärungen. Nur Schüler:innen bewerten auch in diesem Themenbereich die komplexe Erklärung schlechter als die lineare und bewerten daher mit gegenläufiger Tendenz zu allen weiteren Statusgruppen. Der folgende Profilplot kann die Bewertungsunterschiede noch einmal grafisch verdeutlichen.

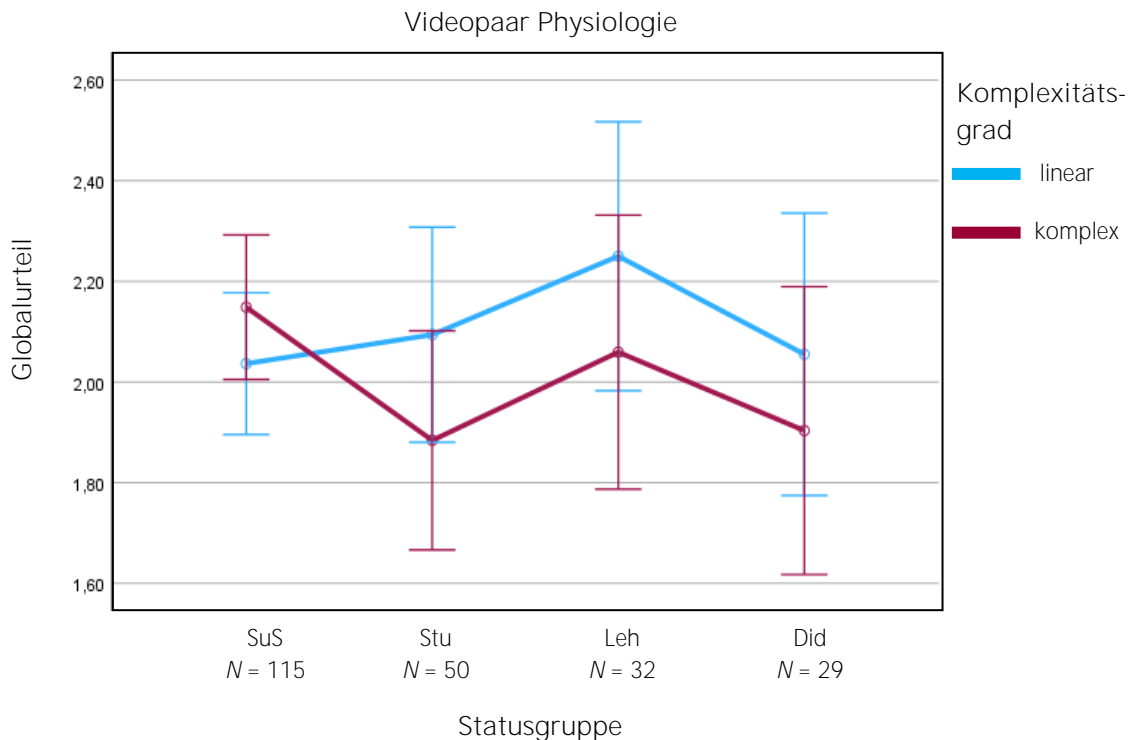


Abbildung 27: Profilplot zur deskriptiven Darstellung des Interaktionseffekts (Abhängigkeit des Faktors ‚Komplexitätsgrad‘ vom Faktor ‚Statusgruppe‘) für den Themenbereich Physiologie mit dem Faktor ‚Komplexitätsgrad‘ als trace-Faktor und dem Faktor ‚Statusgruppe‘ auf der x-Achse inklusive des Fehlerbalkens für das 95%-Konfidenzintervall

Insgesamt ergeben sich für den Themenbereich der Physiologie keine signifikanten Unterschiede hinsichtlich der Haupteffekte. Darüber hinaus lässt sich auch kein signifikanter Interaktionseffekt feststellen. Zur Vollständigkeit ist untenstehend auch für den Themenbereich der Physiologie die Tabelle (Tab. 61) zu den Effektstärken der Mittelwertsunterschiede pro Statusgruppe abgebildet.

Tabelle 61: Effektstärken zu den Mittelwertsunterschieden (t-Tests) im Themenbereich Physiologie zwischen linearen und komplexen Erklärungen getrennt nach Statusgruppen

Statusgruppe	t	df	p	d
Schüler:innen	-1,34	114	n.s.	0,87
Studierende	1,58	49	n.s.	0,88
Lehrkräfte	1,08	31	n.s.	1,00
Didaktiker:innen	1,04	28	n.s.	0,79

Bem.: t: T-Wert; df: Freiheitsgrade; p: p-Wert; d: Effektstärke nach Cohen.

6.2.4.2 Überblick und Einordnung der Ergebnisse zu Unterschieden der globalen Bewertung von linearen und komplexen instruktionalen Erklärungen in Abhängigkeit von der Statusgruppe

Die Vorannahme (siehe Kap. 5.1), dass Erklärungen mit komplexen Ursache-Wirkungs-Beziehungen und Thematisierung von Nichtwissen sowie unsicherem Wissen durchgehend schlechter bewertet werden als lineare Erklärungen, in denen nur eine Ursache angeführt wird und keine Wissenslücken aufgegriffen werden, bestätigt sich anhand der Daten nicht gänzlich, sondern muss differenziert betrachtet werden.

Die Daten zeigen drei unterscheidbare Fälle von Bewertungen der Videopaare:

- a) Die komplexe Erklärung wird signifikant schlechter bewertet als die lineare Erklärung (insgesamt fünfmal),
- b) die komplexe Erklärung wird tendenziell schlechter bewertet als die lineare Erklärung (insgesamt dreimal)
- c) oder die lineare Erklärung wird tendenziell schlechter bewertet als die komplexe Erklärung (insgesamt viermal).

Auffällig ist, dass die lineare Variante nie signifikant schlechter bewertet wird als die komplexe Variante.

Die unterscheidbaren Fälle verteilen sich dabei nicht gleich über die befragten Statusgruppen:

- a) Schüler:innen bewerten im Themenbereich der Genetik die komplexe Variante signifikant schlechter, in den zwei weiteren Themenbereichen zur Ökologie und zur Physiologie bewerten sie die komplexen Erklärungen tendenziell schlechter.
- b) Studierende und Lehrkräfte bewerten die komplexen Erklärungen der Themenbereiche Genetik und Ökologie ebenfalls schlechter als die linearen Erklärvarianten – alle vier Unterscheide sind signifikant. Im Gegensatz zu Schüler:innen bewerten Studierende und Lehrkräfte die komplexe Erklärung im Themenbereich der Physiologie tendenziell besser.
- c) Die Bewertungen der Videopaare von Didaktiker:innen weisen keine signifikanten Unterschiede auf: Sie bewerten die komplexe Erklärvariante aus dem Themenbereich der Ökologie tendenziell schlechter als die linearen Erklärvarianten sowie die komplexen Erklärungen aus den Themenbereichen Genetik und Physiologie tendenziell besser als die linearen Pendanten.

Die Vorannahme, dass komplexe Erklärungen stets schlechter bewertet werden, trifft damit in ihrer Tendenz nur für Lernende zu. Insbesondere die Gruppe der Didaktiker:innen bewertet hingegen kaum entlang dieser Vorannahme.

Die Vorannahme begründet sich in der Vermutung, dass komplexe Ursache-Wirkungs-Zusammenhänge Lernende und Lehrende vor Herausforderungen stellen. Im Theorieteil der Arbeit (Kap. 3.4) wurde dazu herausgearbeitet, dass insbesondere Noviz:innen dazu neigen, Faktoren eines Systems zu reduzieren, in linearen Zusammenhängen zu denken und von einer zentralen Ursache ausgehend eine schrittweise Abfolge zu rekonstruieren (Jacobson, 2000; Resnick, 1995). Weiterhin sind indirekte Zusammenhänge deutlich schwieriger zu verstehen als direkte Zusammenhänge zwischen Faktoren (Mambrey et al., 2020). Darüber hinaus gehen Noviz:innen eher von einer vollständigen Vorhersagbarkeit der Effekte aus und denken kaum in probabilistischen Zusammenhängen (Jacobson, 2000). Eine schlechtere Bewertung könnte auch darin begründet sein,

dass der ‚cognitive load‘ durch die höhere Anzahl involvierter Faktoren höher ist (Grotzer et al., 2017). Vermutet werden kann auch, dass durch die Erwähnung von Nichtwissen und unsicherem Wissen die Eignung für eine Prüfung des Stoffes abnimmt und deshalb die Erklärungen, die Wissenslücken aufgreifen bzw. explizit thematisieren, nicht zur Kultur des Schulfaches passen (Dittmer, 2010).

Ordnet man die oben dargestellten Befunde nicht nach Statusgruppen, sondern nach Themenbereichen, wird deutlich, dass sich die unterscheidbaren Fälle nicht gleich über die drei Themenbereiche verteilen:

- a) Der Themenbereich der Ökologie scheint am eindeutigsten zu sein, denn hier bewerten alle Statusgruppen die komplexe Variante schlechter als die lineare Variante – Studierende und Lehrkräfte signifikant, Schüler:innen und Didaktiker:innen tendenziell.
- b) Die komplexen Videos aus dem Themenbereich der Genetik evozieren zwar bei Schüler:innen, Studierenden und Lehrkräften signifikant schlechtere globale Bewertungen, für Didaktiker:innen stellt die lineare Variante jedoch die bessere dar.
- c) Die komplexe Erklärvariante im Themenbereich der Physiologie wird ausschließlich von Schüler:innen tendenziell schlechter bewertet. Alle weiteren drei Statusgruppen bewerten das lineare Pendant tendenziell besser.

In dieser Anordnung der Ergebnisse wird deutlich, dass die implementierten Merkmale von Komplexität nicht zwangsläufig bei allen Teilnehmenden zu schlechten Bewertungen führen.

Anhaltspunkte für die Beschreibung der gefundenen Unterschiede bzw. für die Suche nach Gründen und damit für die Herausforderungen, komplexe Phänomene in instruktionalen Erklärungen gut zu vermitteln, können die Begründungen zur Notenvergabe im offenen Textfeld des Fragebogens liefern (siehe folgendes Kap. 6.2.4.3).

6.2.4.3 Ergebnisse zu den Begründungsunterschieden zwischen linearen und komplexen instruktionalen Erklärungen in Abhängigkeit von der Statusgruppe

Für ein detaillierteres Bild zu Unterschieden in der Bewertung von linearen und komplexen Erklärungen werden die Ergebnisse der evaluativen qualitativen Inhaltsanalyse herangezogen. Die Ergebnisse des offenen Textfeldes können Ansätze für Gründe aufzeigen, warum die Bewertungen zwischen linearen und komplexen Erklärungen differieren. Da die Analyse und Diskussion aller Kategorien weder übersichtlich noch zielführend ist, werden anhand von Vorannahmen und Auffälligkeiten des Materials Kategorien ausgewählt. Die Auffälligkeiten ergaben sich zum einen bereits während des Codierungsprozesses und sind nicht ausschließlich anhand von Häufigkeiten, sondern auch an besonders prägnanten Formulierungen und solchen, die in besonderer Weise den theoretischen Hintergrund aus der Perspektive der Teilnehmenden beleuchten, festzumachen. Zum anderen werden Auffälligkeiten berichtet, die durch die Darstellung der Daten in unterschiedlichen Formaten in Erscheinung traten (Gruppenvergleiche (Kreuztabellen), Code-Matrix-Browsern). Insgesamt sind nicht nur die vorgefundenen Äußerungen zur Begründung, sondern auch fehlende Referenzen auf bestimmte Kriterien von Interesse. Die Zusammenfassung und Einordnung der Befunde befinden sich im darauffolgenden Kapitel (Kap. 6.2.4.4)

Schüler:innen

Schüler:innen bewerten die komplexen Varianten stets schlechter als die linearen Varianten, wobei der Unterschied im Themenbereich der Genetik signifikant ist (siehe Kap. 6.2.3.1). Mit Blick auf die Begründungen für die Notenvergabe ergeben sich bei diesen zwei Vignetten (G_k und G_l) insbesondere in den globalen Äußerungen auffällige Unterschiede. So wird deutlich, dass die schlechter bewertete komplexe Erklärung (G_k) häufiger als weniger verständlich, klar, übersichtlich und anschaulich beschrieben wird. Die lineare Variante (G_l), die im Globalurteil besser abschneidet, wird hingegen als verständlicher, klarer, nachvollziehbarer und anschaulicher beschrieben. In Bezug auf die eingesetzten Repräsentationen scheinen diese bei der komplexen Erklärung für einige Schüler:innen etwas weniger verständlich zu sein, was folgende Beispiele illustrieren können:

„und ich finde auch dass diese übersicht mit den ganzen Pfeilen hier sehr unpassend und extrem unübersichtlich war“ (TAST6/ G_k)

„man kann die Pfeile benennen und noch Ffaktoren oder so auf die Seite schreiben, weil ich finde es war etwas unübersichtlich dargestellt“ (SAST39a/ G_k)

Der Vergleich zwischen gesunden und an MS erkrankten Personen, der bei der linearen Erklärung aufgegriffen wird, wird häufig mit positiver Valenz in die Begründungen mit aufgenommen – eine ähnliche Vorgehensweise gibt es bei der komplexen Erklärung der Genetik nicht. Schüler:innen schreiben beispielsweise:

„durch Vergleich der Unterschied sehr gut verständlich“ (MAHE 11/ G_l)

„gut mit Vergleich bei einem gesunden Menschen (CHRO28/ G_l)

Die für diese beiden Videovignetten gefundenen Unterschiede können nicht auf alle komplexen bzw. alle linearen Erklärungen übertragen werden. Sie sind lediglich für die Vignetten zur Genetik charakteristisch.

Allgemein tritt beim Vergleich der Begründungen zur Bewertung linearer und komplexer Erklärungen Folgendes hervor:

Die interessante und motivierende Gestaltung wird von den Schüler:innen in etwa ähnlich häufig für die komplexe Erklärung wie für das jeweilige lineare Pendant genannt – sie scheint also nicht systematisch zwischen den zwei Varianten zu variieren. So beschreibt ANM27 die Erklärung G_l als *„ein bisschen trocken erklärt“* und ANMA29 mit den Worten *„aber es war nicht sooo einfach, ständig zuzuhören“*. Ähnliche Äußerungen finden sich bei Video G_k : *„Langweilig vorgetragen, aber falls man es schafft aufzupassen recht informativ“ (KOPA5)* und *„Erklärung im allgemeinen allerdings etwas trocken“ (BAFE23)*. Bezüge zur interessanten und motivierenden Gestaltung mit positiver Valenz finden sich ebenso im Material: entweder allgemein wie bei AMAT7 mit *„war interessant“ (P)* oder mit Begründung wie bei CHER8 mit *„sodass man selber vergleichen konnte, was wiederum zum nachdenken angeregt hat“ (G)*.

Unerwartet ist, dass Schüler:innen kaum Bezug auf das implementierte Nichtwissen und unsichere Wissen in den komplexen Erklärvarianten nehmen. Es wäre anzunehmen gewesen, dass dieses Merkmal der Vignetten häufiger in den Äußerungen herangezogen wird, da es kein prominenter Teil der Unterrichtskultur ist und daher in den Fokus für die Bewertung rücken könnte. Aufmerksamkeit muss auch darauf gerichtet werden, was Schüler:innen nicht zur Begründung der Notenvergabe anführen. In Kapitel 6.2.1.3 wurde bereits allgemein aufgegriffen, dass auf die

Kategorie zum Gefühl der Unsicherheit bzw. Sicherheit, das durch die Erklärung hervorgerufen werden kann, und auf die Kategorie zur Berücksichtigung von Nichtwissen bei Schüler:innen keine Codierungen entfallen. Damit scheint die Integration von Nichtwissen und unsicherem Wissen in instruktionalen Erklärungen nicht global auffallend störend für Schüler:innen zu sein. Eine Kategorie mit Bezug zu Wissenslücken ist diejenige der ‚Vagheit‘: Die wenigen Ausnahmen in dieser Kategorie können illustrieren, dass die komplexen Erklärungen von Schüler:innen als vager wahrgenommen werden als die linearen Pendanten. Zum Video Ö_k äußert sich JOCH4 beispielsweise mit „*kein klares Ergebnis*“. Hingegen führen beispielsweise ANST94 und CHMA9 als Argument für die Notenvergabe bei linearen Varianten die Abwesenheit von Vagheit an: So greift ANST94 auf, dass es ein klares Ergebnis gibt („*klares Ergebnis*“; P_i), und CHMA9 führt aus, dass es bei der linearen Variante aus der Ökologie Merksätze gibt („*super fand ich die Merksätze*“; Ö_i), mit denen die Erklärung wahrscheinlich eindeutig und abgeschlossen wirkt. Die Äußerungen deuten an, dass Schüler:innen von Erklärungen auch eine Konkretheit in Form von Ergebnissen erwarten. Offen bleibt durch die Kürze der Antworten, ob das ‚klare Ergebnis‘ eine Eigenschaft des Erklärgegenstandes sein muss oder durch die Tätigkeiten der Lehrkraft hergestellt werden kann (z. B. indem noch ausführlicher darauf eingegangen wird, dass bei komplexen Wirkungsgefügen oft kein klares Ergebnis existiert oder eine sichere Aussage getroffen werden kann bzw. das Ergebnis das Zusammentragen und Abwägen von Faktoren ist).

Insbesondere im Themenbereich der Ökologie fällt auf, dass Schüler:innen darauf Bezug nehmen, dass die Zusammenhänge bei der linearen Erklärung für sie deutlich werden. CHRO28 schreibt *dazu* „*Das Je ... desto war gut nach zu vollziehen*“ und TAGE15 meint „*du Zusammenhänge waren klar*“. Die lineare Erklärung aus der Ökologie ist dabei besonders, da sie eine hohe Repetition aufweist – die Räuber-Beute-Beziehung bzw. die damit einhergehenden Populationsgrößen werden in vier Teilabschnitten durch die Abhängigkeit der Räuber von den Beutetieren erklärt. Bei den komplexen Erklärvarianten wird darauf weniger häufig Bezug genommen – ANAN23 beschreibt die Erklärungen G_k und P_k sogar als zusammenhangslos („*aber relativ zusammenhangslos*“). Dabei muss einbezogen werden, dass die Beispiele für lineare Erklärungen stets nur einen Zusammenhang thematisieren, der vertieft dargestellt wird, während komplexe Erklärungen mehrere Zusammenhänge aufgreifen, sodass sie nicht vergleichbar vertieft und detailliert dargestellt werden – aus zeitlichen Gründen, aber auch aufgrund von bestehendem Nichtwissen über die Mechanismen und Faktoren.

Überraschend sind vor dem Hintergrund der Äußerungen zur Vagheit und zur Deutlichkeit der Zusammenhänge die Äußerungen von Schüler:innen zur Merkbarkeit und Lernbarkeit des Gesagten. Während aufgrund der Schwierigkeit, multikausale Ursache-Wirkungs-Gefüge in eine erzählbare Ordnung zu bringen, eher davon auszugehen wäre, dass komplexe Erklärungen deshalb auch schlechter merkbar und lernbar erscheinen, äußern sich Schüler:innen konträr dazu – die Merkbarkeit und Lernbarkeit werden häufiger negativ bei linearen Erklärungen und häufiger positiv bei komplexen Erklärungen erwähnt. So beschreibt CHJO26 die lineare Erklärung zur Entstehung von Mukoviszidose (G_i) mit „*nicht besonders gut zu behalten*“ und GAMA44 erwähnt als Begründung für die Notenvergabe „*es bleibt nicht alles hängen*“ (G_i). Die komplexe Erklärung zur Ökologie (Ö_k) – also die Entstehung von Populationsschwankungen bei Erdkröten – wird hingegen ausschließlich als gut merkbar beschrieben, beispielsweise von IRAL10 mit „*Durch das Beispiel gut einpräglich*“. In Verbindung mit den Äußerungen zur Vagheit könnte ein Grund dafür sein, dass Schüler:innen die linearen Erklärungen als solche mit einem klaren, in Tests abprüfbar Ergebnis wahrnehmen. Der Eindruck, sich das Gesagte merken zu müssen, könnte daraufhin höher sein. Mitunter werden im Gegensatz dazu die Konzepte und Fakten in den komplexen Erklärungen nicht

als solche wahrgenommen – die Möglichkeit einer Verstehensillusion (Wittwer & Renkl, 2008) besteht dabei eher als bei linearen Erklärungen.

Eine Auffälligkeit in Bezug auf die Häufigkeit der Nennung bei komplexen Erklärungen im Vergleich zu linearen Erklärungen ergibt sich in Hinblick auf die Verständlichkeit der Visualisierung. Sie wird insgesamt bei komplexen Erklärungen häufiger in den Begründungen angeführt (21 Äußerungen bei linearen Erklärungen, 35 Äußerungen bei komplexen Erklärungen). Die Verständlichkeit der Visualisierung wird dabei in den Begründungen der komplexen Erklärungen häufiger mit negativer Valenz angegeben (26 Äußerungen zu schlechter Verständlichkeit bei komplexen Erklärungen, 9 Äußerungen bei linearen Erklärungen). Über alle drei komplexen Erklärungen hinweg finden sich negative Äußerungen, die auf die Übersichtlichkeit des Tafelbildes und fehlende Informationen zum Beispiel in Bezug auf die Verbindungen zwischen den Faktoren durch Pfeile gerichtet sind:

„zu wenig Informationen an der Tafel“ (BAKA23/P_k)

„In dem Fall haben die Pfeile nicht gereicht. Was sie gesagt hat war zwar gut und verständlich wenn man aber den Faden verloren hat kann man es sich auch nicht mehr durch das Tafelbild erklären“ (CHER8/P_k)

„jedoch hätte mehr an der Tafel stehen können“ (BAKA23/G_k)

„und ich finde auch dass diese übersicht mit den ganzen Pfeilen hier sehr unpassend und extrem unübersichtlich war“ (TAST6/G_k)

Die Äußerungen geben Hinweise, dass die Visualisierungen bei komplexen Erklärungen mehr in den Aufmerksamkeitsfokus rücken und mitunter wichtiger für das Verständnis sind. Gleichzeitig scheint durch die hohe Anzahl an Faktoren und Zusammenhängen die Übersichtlichkeit zur Herausforderung sowie eine Beschriftung zur Nachvollziehbarkeit relevanter zu werden.

Mit der Codierung der Valenzen wird unabhängig vom Grad der dargestellten Komplexität in den Erklärungen etwas Weiteres deutlich: Schüler:innen schätzen die Passung der Erklärung für sie als Adressat:innen heterogen ein. Die Erklärung Ö₁ wird beispielsweise hinsichtlich des Schwierigkeitsgrades von einigen Schüler:innen als passend beschrieben: *„war entspannt und man konnte es sich denken“* (ULBE67) oder *„sehr einfach erklärt“* (KACL28). Andere Schüler:innen begründen ihr Globalurteil für dasselbe Video mit der fehlenden Passung des Schwierigkeitsgrades: *„Das Thema ist eher Grundschulstandard“* (IRHA21) oder *„Das Niveau, auf dem erklärt wurde, war für eine gymnasiale Anstalt seeeehr einfach gehalten ...“* (NAWO9). Diese Heterogenität lässt sich auch in der Kategorie zum Erklärtempo wiederfinden: Während ANMA15 die Erklärung G_k mit *„wurde langsam erklärt, sodass man Zeit hatte, es zu verstehen“* beschreibt, äußert sich CHKA89 gegenteilig mit *„Sehr viele Informationen auf einmal“* und EVJO38 mit *„teilweise etwas zu langsam, dadurch verliert man etwas das Interesse zuzuhören“*. Herausforderungen für Lehrkräfte liegen damit insbesondere auch darin, die heterogenen Voraussetzungen, Wahrnehmungen und Erwartungen an instruktionale Erklärungen zu antizipieren, zu berücksichtigen oder gegebenenfalls durch die weiterführenden Aktivitäten aufzufangen.

Studierende

Studierende bewerten die komplexen Erklärungen aus den Themenbereichen der Genetik und Ökologie signifikant schlechter. Die komplexe Erklärung aus dem Themenbereich der Physiologie wird tendenziell etwas besser bewertet.

Auffällige Unterschiede für das Videopaar aus dem Bereich der Physiologie, die die tendenziell positive Bewertung der komplexen Erklärung begründen könnten, sind im fachlichen Aspekt zu finden. Aus Sicht der Studierenden ist beispielsweise die Vollständigkeit für die lineare Erklärung nicht durchgehend gegeben. Dabei beziehen sich die Studierenden jedoch auf unterschiedliche Punkte wie die fehlende Ausführung zur besseren Aufnahme von Vitamin A („*nicht wirklich erklärt nicht klar, warum auf einmal Vitamin A durch die Dünndarmschleimhaut aufgenommen werden kann*“ (GNAI04/P)) sowie die zu oberflächliche Ausführung chemischer Grundlagen zur Verdauung („*Bei der Erklärung zur Löslichkeit und Aufnahme von Vit. A nicht genug chemisches Verständnis von Verdauung vermittelt. (Vergleich von ohne Fette <=> in Emulgierung mit Fett könnte genauer sein)*“ (INST04/P)).

Auch die Korrektheit der Erklärung wird im Gegensatz zu den weiteren Vignetten für die lineare Erklärung der Physiologie von Studierenden zum Teil negativ bewertet. So JUMA08 „*Aussage: unser Darm ist ein wässrige Millieu - würde ich anders formulieren*“ und NGAI04 „*Funktion der Mizelle falsch erklärt*“. Wie im Fall der Vollständigkeit beziehen sich die Studierenden jeweils auf unterschiedliche Punkte innerhalb der Erklärung. Die Äußerungen deuten damit nicht an, dass eine Vignette einen gravierenden fachlichen Fehler aufweist, der zum Ausschluss der Vignette führen müsste.

Hinzu kommt, dass Studierende die komplexe Erklärung der Physiologie in Hinblick auf die realisierte Phänomenorientierung und die Darstellung der Komplexität des Phänomens deutlich besser als die lineare Variante bewerten (in der linearen Variante wird darauf nicht Bezug genommen). Sie formulieren das beispielsweise wie folgt:

„darauf hingewiesen, dass meist viele Faktoren eine Rolle spielen“ (JUMA08/P_k)

„+es wird dieses mal das Zusammenspiel der verschiedenen Faktoren angesprochen“
(SUJO15/P_k)

„Es wird sehr deutlich wie komplex das System ist“ (BRHE95/P_k)

Zuletzt ergibt sich eine Auffälligkeit in Bezug auf die Deutlichkeit der Zusammenhänge – im Gegensatz zu Schüler:innen bewerten Studierende diese bei der komplexen Erklärung der Physiologie besser, wohingegen bei der linearen Erklärung nicht darauf Bezug genommen wird (z. B. „*Abhängigkeit des Folsäurehaushalts von verschiedenen Faktoren wurde gut verdeutlicht*“ (HEJO14/P_k)).

Die in der linearen Variante schlechter bewertete Vollständigkeit und fachliche Korrektheit sowie die in der komplexen Variante besser bewertete Phänomenorientierung und Deutlichkeit der Zusammenhänge können als Hinweise für die Bewertungsdifferenz zwischen der linearen und der komplexen Erklärung im Themenbereich der Physiologie herangezogen werden. Auch in den Regressionsanalysen zeigt sich der fachliche Aspekt für die Gruppe der Metaperspektive als relevanter Prädiktor für das Globalurteil. Die aufgegriffenen Kategorien gehen über die Operationalisierung des fachlichen Aspekts im standardisierten Zugang hinaus.

In Hinblick auf die Unterschiede zwischen den zwei weiteren Videopaaren, bei denen die komplexen Erklärungen schlechter als die linearen Erklärungen bewertet werden, fällt auf, dass Studierende die Verweise auf Nichtwissen mit positiver Valenz in die Begründungen aufnehmen. Auch für Studierende scheinen jedoch die Verweise auf Nichtwissen und unsicheres Wissen nicht prominent in den Fokus zu rücken, da sich lediglich zwei Studierende über alle drei komplexen Vignetten hinweg dazu äußern:

„Es wird deutlich, dass es in der Wissenschaft noch viel zu erforschen gibt und man noch längst nicht alles weiß“ (BREH95/G_k)

„Die verschiedenen bisher bekannten Gründe und Ursachen genannt“ (JUO24/G_k)

Beide Begründungen finden sich zur komplexen Erklärung aus dem Themenbereich der Genetik (Entstehung von Multipler Sklerose) und sind als eher beschreibend zu charakterisieren: Sie sind positiv konnotiert, greifen jedoch vorwiegend das auf, was die erklärende Person sagt. Es ist kaum ein normatives Moment zu erkennen – sie fügen beispielsweise nicht an, dass es ein wichtiges Lernziel für ein adäquates Verständnis des Phänomens oder biowissenschaftlicher Forschung ist, was bei Biologiedidaktiker:innen (siehe Abschnitt zu Didaktiker:innen) deutlicher wird.

Wie Schüler:innen greifen Studierende die Vagheit komplexer Erklärungen auf, jedoch finden sich die wenigen Äußerungen ausschließlich zur komplexen Vignette aus dem Themenbereich der Genetik. Die Studierenden beziehen sich darauf, dass die Erklärung nicht konkret die Entstehung von Multipler Sklerose darstellt. So schreiben beispielsweise SUJO15 *„Die Erklärung sagt nur aus, dass Faktoren zu multipler Sklerose führt (das kann man sich denken)“* und ANAN85 *„erklärt nicht wirklich die Entstehung der Krankheit sondern mehr die auslösenden Faktoren“*. Die beiden exemplarischen Äußerungen werfen die Frage auf, warum die auslösenden Faktoren nicht als Erklärung für die Entstehung der Multiplen Sklerose verstanden werden.¹⁷⁷ Wenn die Äußerung von SUJO15 so verstanden wird, man könne sich denken, dass mehrere Faktoren Auslöser für Multiple Sklerose sind, verweist sie auf die Bedeutung für ein Verständnis, dass Multikausalität eine Herausforderung für Schüler:innen ist und gleichzeitig einen wichtiger Teil zum Verständnis der belebten Welt beiträgt (Jacobson, 2000; Jacobson & Wilensky, 2006; Resnick, 1995).

Die Kategorie mit Bezug zur Multikausalität bzw. Komplexität biologischer Phänomene ist die mit ‚Phänomenorientierung‘ benannte Kategorie. Diese wird von Studierenden bei komplexen Erklärungen häufiger adressiert. ELUL12 hebt beispielsweise positiv hervor: *„Problematik der komplexen Krankheitsentstehung von MS wird klar“* (G_k). Insbesondere bei der komplexen Variante im Themenbereich der Physiologie wird die Phänomenorientierung von den Studierenden mit positiver Valenz in den Begründungen angeführt. Interessant ist, dass Äußerungen mit negativer Valenz bei der komplexen Erklärung im Themenbereich der Genetik auftreten – nicht jedoch bei der linearen Variante im Themenbereich der Genetik. Die zwei folgenden Beispiele finden sich im Material:

„Es fehlt ob sich die Faktoren gegenseitig beeinflussen“ (SUJO15/G_k)

„komplexes Thema sehr vereinfacht dargestellt“ (ANAN85/G_k)

¹⁷⁷ Die unmittelbare Ursache für Multiple Sklerose – die Schädigung von Nervenzellen – wird in der Erklärung zu Beginn benannt. Es wird weiterhin aufgegriffen, dass die Schädigung durch das eigene Immunsystem zustande kommt. Anschließend wird zu denjenigen Faktoren übergeleitet, die zur Dysfunktion des Immunsystems führen.

So scheint die Komplexität biologischer Phänomene bei linearen Erklärungen nicht mitgedacht zu werden – zumindest wird sie nicht als Kriterium für die Notenvergabe genannt. Das Fehlen von Elementen der Komplexität (SUJO15) oder die Einschränkung (ANAN85) werden hingegen in einzelnen Fällen erkannt und zur Begründung herangezogen. Vermutet werden könnte, dass die Phänomenorientierung eher bei komplexen Erklärungen und die Prinzipienorientierung eher bei linearen Erklärungen zur Begründung herangezogen wird. Eine derartige Tendenz findet sich jedoch nicht im Material.

Als weiterer Unterschied zwischen komplexen und linearen Erklärungen ließ sich bei Schüler:innen die Verständlichkeit der Visualisierungen anführen. Studierende bewerten die Verständlichkeit der Repräsentationen in den linearen Erklärungen in etwa gleich wie in den komplexen Varianten. Das Kriterium wird dabei hauptsächlich mit negativer Valenz angeführt – Schüler:innen beziehen es hingegen häufiger mit positiver Valenz in ihre Bewertungen ein. Folgende Beispiele geben Einblick in die Einschätzungen und Anhaltspunkte der Studierenden:

„nachvollziehbares, nacheinander aufgebautes Tafelbild (gilt auch für vorheriges Video)“
(SUNO37/Ö_k)

„wäre schön gewesen, wenn sie die Abbildungen an der Tafel noch beschriftet hätte“
(REWE05/P_k)

„Zu Beginn wird ein Teil der DNA angehängt --> würde darunterschreiben, dass es ein Gen symbolisiert, denn wenn man kurz nicht aufpasst wird durch das Bild nicht klar, dass es sich um ein Gen handelt“ (MAAL02/G_i)

Auch Studierende nehmen in den Kritikpunkten Bezug auf fehlende Beschriftungen von Teilen der Visualisierung. Dabei wird häufig – wie bei MAAL02 – auf die Funktion des Tafelbildes hingewiesen, dass die Adressat:innen bei Unaufmerksamkeit wieder in die Erklärung einsteigen können.

Ein Bewertungsunterschied zwischen linearen und komplexen Erklärungen findet sich hingegen in der Kategorie zum Erklärtempo. Dieses wird von Studierenden bei linearen Erklärungen besser bewertet als bei komplexen Erklärungen. Die negativen Bewertungen von Studierenden zeigen dabei interessanterweise die gleiche Heterogenität wie diejenigen von Schüler:innen:

„lange Dauer für wenig Information“ (URFR07/G_k)

„etwas zu schnell“ (JUMA08/G_k)

„nicht unbedingt für die Klassenstufe so geeignet, zu langsam“ (KAMA21/P_k)

„etwas schnell“ (INMA18/P_k)

Das Erklärtempo wird negativ bewertet, wobei die Beispiele zeigen, dass es von einigen Studierenden als zu hoch und von anderen Studierenden als zu niedrig eingeschätzt wird.

Lehrkräfte

Lehrkräfte bewerten wie Studierende die komplexen Erklärungen aus den Themenbereichen der Genetik und der Ökologie signifikant schlechter. Die komplexe Erklärung aus dem Themenbereich der Physiologie erhält – ebenfalls wie bei Studierenden – tendenziell ein besseres Globalurteil.

Auffällige Unterschiede für das Videopaar aus dem Bereich der Physiologie ergeben sich lediglich in zwei Kategorien. So wird die fachliche Vollständigkeit von drei Lehrkräften bei der linearen Erklärung negativ angeführt, wobei sich alle Äußerungen auf die bessere Aufnahme von Vitamin A in den Körper durch die Bindung in Mizellen beziehen:

„Ich würde mich als Schüler fragen, warum Fett besser in die Darmwand gelangt. Das habe ich aufgrund der Erklärung nicht verstanden“ (INWO07/P)

„Zusammenhang Fettlöslichkeit und Aufnahme in Zelle nicht klar genug“ (ASBR09/P)

Die komplexe Erklärung aus dem Themenbereich der Physiologie enthält hingegen keine negativen, aber auch keine positiven Begründungen zur fachlichen Vollständigkeit.

Eine Auffälligkeit im Unterschied zwischen den Videos der Physiologie ergibt sich weiterhin für die allgemein gehaltene Kategorie zum Einsatz bildlicher Visualisierungen beim Aspekt der Repräsentationen. Während die lineare Erklärung eine Äußerung mit eher negativer Valenz („*langsam frage ich mich, ob jedesmal Applikationen nötig sind*“ (MAED32/P)) und eine mit positiver Valenz („*Das Tafelbild war wieder sehr anschaulich*“ (INWO07/P)) aufweist, erhält die komplexe Variante ausschließlich Äußerungen mit positiver Valenz (z. B. „*gut nur durch Visualisierung*“ (FRCH02/P_k) und „*gute Übersicht an der Tafel*“ (RUF32/P_k)).

Neben den beschriebenen Tendenzen in den Kategorien zum Einsatz bildlicher Visualisierungen sowie zur fachlichen Vollständigkeit, die sich von den weiteren zwei Videopaaren unterscheiden, lässt sich feststellen, dass in der Kategorie zur Vagheit der Erklärung keine Äußerungen für das Videopaar der Physiologie enthalten sind. Die Vagheit ist jedoch insgesamt ein von den Lehrkräften herangezogener Unterschied in den Videopaaren der Ökologie und der Genetik. In der Gesamtschau lassen die Begründungen von Lehrkräften aufgrund der geringeren Anzahl an Äußerungen im Vergleich zu den Statusgruppen der Schüler:innen und Studierenden keine weiteren Ansätze für Gründe zur tendenziell besseren Bewertung der komplexen Erklärung der Physiologie zu.

Bei der Betrachtung der Unterschiede zwischen komplexen und linearen Erklärungen aus den Themenbereichen der Genetik und der Ökologie ergeben sich folgende Auffälligkeiten:

Die komplexen Erklärungen werden von Lehrkräften in Hinblick auf die Verweise auf Nichtwissen und die entstehende Vagheit negativ bewertet. Die folgenden Äußerungen von Lehrkräften können den Einbezug der Vagheit bei der Bewertung komplexer Erklärungen verdeutlichen:

„Aber inhaltlich wenig ergiebig außer: Immunsystem greift Nervenzellen an. All die Ursachen wirken recht willkürlich ohne jeden Mechanismus“ (MAED33/G_k)

„es liegen genaue Zahlen vor konkrete Situation für die Stadt Merklingen herausarbeiten“ (ANMA14/Ö_k)

Die Vagheit scheint sich dabei für Lehrkräfte nicht in einem unpräzisen sprachlichen Ausdruck zu manifestieren: Im Material konnten überwiegend Äußerungen mit positiver Valenz in Bezug auf die

Präzision des Ausdrucks gefunden werden und das sowohl bei linearen als auch bei komplexen Erklärungen mit gleicher Häufigkeit. Die Lehrkräfte beschreiben den sprachlichen Ausdruck dabei am häufigsten mit dem Adjektiv ‚klar‘ und etwas seltener mit ‚deutlich‘.

Bezüge zum Erwähnen von nicht Nichtwissen in den Erklärungen finden sich selten im Material. MARE61 beschreibt die Erklärung (P_k) allgemein mit „zu viel Blackbox“. MAED22 geht etwas ausführlicher auf die Wirkung von Nichtwissen ein: „Sinnvolles Lernziel könnte nur sein: Wir kennen viele wichtige Zusammenhänge noch nicht, vielleicht will jemand von euch so etwas erforschen“ (MAED33/ G_k). Während die Äußerung von MARE61 eher andeutet, dass Eindeutigkeit in der Erklärung fehlt, steht bei MAED33 das Nichtwissen im scheinbaren Gegensatz zu einer Erklärung mit fachlichen Lernzielen. Die zu Beginn der Erklärungen aufgeworfenen Fragestellungen können durch das vorhandene Nichtwissen nicht beantwortet werden, sodass die Übereinstimmung des Erklärziels (z. B. Wie entsteht Multiple Sklerose?) insofern nicht mit dem Erklärinhalt überstimmt, als am Ende der Erklärung auch eine vollständige Erklärung steht. Die Abwesenheit von Nichtwissen wird hingegen nicht in die Bewertungen mit einbezogen. So finden sich weder positiv noch negativ codierte Äußerungen bei den linearen Erklärungen.¹⁷⁸

Zuletzt fällt im Material auf, dass Lehrkräfte nur selten auf die Phänomenorientierung innerhalb des fachlichen Aspekts eingehen (die geringe Häufigkeit der Bezugnahme ist insbesondere dann auffällig, wenn die hohen Häufigkeiten der Biologiedidaktiker:innen als Vergleich herangezogen werden; siehe nächster Abschnitt). Unerwartet ist dabei vor allem, dass die Phänomenorientierung bei den komplexen Erklärungen aus den Themenbereichen der Ökologie und der Physiologie sowohl positiv als auch negativ geäußert werden:

„Zusammenspiel Faktoren wird deutlich“ (ASBR09/ $Ö_k$)

„viele Aspekte mit einbezogen“ (RUF32/ P_k)

„anschaulich, aber etwas zu komplex“ (ROAN05/ P_i)

Während sich ASBR09 und RUF32 positiv konnotiert zum phänomenorientierten Aufzeigen eines multikausalen Ursache-Wirkungs-Gefüges äußern, ist ROAN05 der Meinung, dass das Ursache-Wirkungs-Gefüge eine zu hohe Komplexität aufweist. Bei Lehrkräften findet sich ausschließlich der Bezug zur Phänomenorientierung, nicht jedoch ein solcher zur Prinzipienorientierung.

Didaktiker:innen

Im standardisierten Zugang zeigt sich bei Biologiedidaktiker:innen keine signifikant unterschiedliche Bewertung der komplexen Erklärungen gegenüber den linearen Erklärungen gemessen am Globalurteil. Im Themenbereich der Ökologie erhielt das komplexe Video tendenziell schlechtere Bewertungen, in den Themenbereichen der Genetik und der Physiologie bekamen hingegen die linearen Varianten tendenziell bessere Bewertungen.

¹⁷⁸ Interessant ist vor dem Hintergrund der Äußerungen zu Nichtwissen, dass Lehrkräfte die Merkbarkeit und Lernbarkeit nur bei komplexen Erklärungen als Begründung für die Notenvergabe anführen. Die linearen Erklärungen evozieren hingegen keine Äußerungen dazu.

In Bezug auf die tendenziell bessere Bewertung der linearen Erklärvariante im Themenbereich der Ökologie lassen sich in den Begründungen für die Vergabe des Globalurteils aus dem offenen Textfeld folgende Hinweise finden:

Die Bezugnahme auf das Beispiel als Kriterium zum Aspekt der Repräsentationen scheint im Falle der komplexen Erklärung etwas schlechter gelungen zu sein. Dabei wird auf die zu geringe Einbindung des Diagramms, das die konkreten Zahlen der Erdkröten in Merklingen zeigt, verwiesen und die Erläuterung von Einflussfaktoren auf die Anzahl von Erdkröten als zu gering bewertet. Beispielsweise äußern sich zwei Teilnehmende dazu folgendermaßen:

„Das Diagramm hätte bei der Erklärung der einzelnen Einflussfaktoren wieder mit eingebunden werden können; z. B. Hypothesen über die Anzahl der stehenden Gewässer oder der Fressfeinde für bestimmte Abschnitte des Graphen aufstellen“
(ANTH74/Ö_k)

„liegen zum Diagramm Daten vor, was in den einzelnen Jahren passiert ist (also z. B. Bau neuer Straßen oder sehr trockene Sommer), um eine Verknüpfung zwischen möglichen Umweltveränderungen und der Anzahl der Kröten abzulesen? Wenn ja, wäre eine Verknüpfung hilfreich gewesen“ (BEUL05/Ö_k)

Gleichzeitig scheint aus Sicht der Didaktiker:innen die Verbindung der eingesetzten Repräsentationen¹⁷⁹ im Fall der linearen Erklärung besser als im komplexen Pendant realisiert worden zu sein. Die zwei folgenden Beispiele können die Bezugspunkte der Didaktiker:innen verdeutlichen:

„Dopplung der Erklärung auf mehreren Kanälen (Liniendiagramm sowie Fussdiagramm) sowie immer wieder Rückbezug zwischen beiden Darstellungen“
(MOBE36/Ö)

„gut: Repräsentationsformen systematisch aufeinander bezogen (Übersetzung von Diagramm mit linearer Zeitachse in zyklisches Diagramm)“ (EVWO09/Ö)

Auffällig ist, dass die Phänomenorientierung als Kriterium zum fachlichen Aspekt für das lineare Pendant häufiger mit negativer Valenz in die Begründungen aufgenommen wird. Diese Bewertung läuft entgegen der globalen Bewertungstendenz. Beispiele für eine negative Einschätzung der Phänomenorientierung finden sich im Folgenden:

„weniger gut: monokausale Erklärung an sehr einfachem Modell; Komplexität natürlicher Räuber-Beute-Zusammenhänge und der mathematischen Modellierung wird wenig deutlich“ (ASBR09/Ö)

„Flux of Nature fehlt Festigung von Balance of Nature“ (MAED08/Ö)

„Schön wäre noch zu erwähnen, dass die Luchse natürlich nicht nur Hasen fressen (ingeschränktes Modell)“ (PAWI03/Ö)

Positive Äußerungen von Biologiedidaktiker:innen zur Phänomenorientierung im komplexen Video sind exemplarisch folgende:

¹⁷⁹ Diagramm zur Entwicklung der Individuenzahlen der Schneeschuhhasen, Tafelbild zum allgemeinen Zusammenhang zwischen Anzahl der Räuber und Beutetiere in einem Gebiet.

„Hinweis auf Problem der Erklärung eines Phänomens wenn viele Einflussfaktoren beteiligt sind, ist wichtig --> fördert eine realistische Sichtweise“ (GEHA01/Ö_k)

„Multikausalität wird so gut deutlich“ (FRGE13/Ö_k)

Obwohl sich für die Gruppe der Lehrenden die fachspezifische Facette im standardisierten Zugang als großer Einflussfaktor auf die Güte einer Erklärung erwies (siehe Kap. 6.2.3.2), tritt hier die weniger gut gelungene Phänomenorientierung nicht gleichzeitig mit einer schlechteren Globalbewertung auf. Grund dafür könnte sein, dass das gewählte Beispiel für die lineare Erklärung als solches mit Modellcharakter etabliert ist. Der Unterschied in den zwei Kategorien zum Aspekt der Repräsentationen (‘Ausreichende Bezugnahme auf das Beispiel’ und ‘Verbindung der eingesetzten Repräsentationen’) fällt im Material auf und könnte zur schlechteren Bewertung der komplexen Variante im Themenbereich der Ökologie beitragen.

In Hinblick auf die tendenziell schlechtere Bewertung der linearen Erklärvariante in den zwei weiteren Themenbereichen (Genetik und Physiologie) lassen sich in den Begründungen folgende Auffälligkeiten feststellen:

Die befragten Biologiedidaktiker:innen äußern sich zur Verständlichkeit der Visualisierungen bei den linearen Erklärungen deutlich häufiger negativ. Dabei wird oft aufgegriffen, dass Beschriftungen von Elementen in der Visualisierung fehlen (z. B. *„keine klare Beschriftung der Teile“* (INMI35/G_i)) oder die Leserichtung bei der Visualisierung nicht durchgehend beachtet wurde (z. B. *„Leserichtungswechsel (oben: links, rechts; unten: Kreis)“* (BEHE29/Ö_i)). Die Bewertungstendenzen stehen den Bewertungen der Schüler:innen in diesem Kriterium entgegen. Sie äußerten sich vorrangig zu den komplexen Varianten negativ zu Verständlichkeit der Visualisierung, beziehen sich jedoch ebenfalls auf die aus ihrer Sicht mangelnde Beschriftung von Teilen.

Zum fachlichen Aspekt ergeben sich bei Biologiedidaktiker:innen die meisten und größten Differenzen in den geäußerten Begründungen. So bewerten Biologiedidaktiker:innen die Phänomenorientierung nicht nur bei der komplexen Variante der Ökologie besser, sondern auch den zwei weiteren komplexen Varianten. Dabei äußert sich beispielsweise EVWO09 zur Erklärung G_k: *„gut: Multikausalität statt OGOD (one gene, one disease), auch unter Einbezug des Spannungsfelds Gen-Umwelt-Interaktionen“*. In Bezug auf die häufig vorkommende, fachlich nicht angemessene Vorstellung (one gene, one disease) bezieht die Person die Multikausalität positiv in die Bewertung der Erklärung mit ein. Für das Video Ö_k nimmt GEHA01 ebenfalls positiv auf das Kriterium Bezug und hebt dabei die Förderung einer *„realistischen Sichtweise“* hervor: *„Hinweis auf Problem der Erklärung eines Phänomens wenn viele Einflussfaktoren beteiligt sind, ist wichtig --> fördert eine realistische Sichtweise“*. Die Äußerungen kontrastieren damit Alltagsvorstellungen, die linear sein können (z. B. ‘one gene, one disease’) mit einer phänomenorientierten, multikausalen, realistischen Vorstellung über natürliche Phänomene. Auch bei Video P_k äußert sich GEHA01 positiv und schätzt dabei die Eignung für die Schüler:innen als gelungen ein bzw. ist sich scheinbar der Problematik komplexer Phänomene bewusst, da eine mögliche Überforderung der Schüler:innen nicht gesehen wird: *„Herausheben der möglichen Schwachstellen --> Darstellung / Andeuten der Komplexität ohne Überforderung hervorzuheben“*.

Negativ wird die Phänomentorientierung lediglich auffällig oft bei der linearen Erklärung der Ökologie aufgegriffen – die weiteren Vignetten zeigen keine Häufung negativer Nennungen. Die Prinzipienorientierung wird auch von Didaktiker:innen im Wesentlichen nicht als Kriterium für komplexe oder lineare Erklärungen herangezogen.

Deutlich tritt auch ein Bewertungsunterschied zum Kriterium zur Berücksichtigung von Nichtwissen hervor. Folgende Beispiele geben zunächst einen Überblick zu den Begründungen:

„gut: Unsicherheit bei Erklärungen im naturwissenschaftlichen Erkenntnisgewinnungsprozess wird hervorgehoben, insbesondere unter Einbezug aktueller Forschung“ (EVWO09/G_k)

„offene, naturwissenschaftlich interessante Fragestellungen werden aufgeworfen“ (GEHA01/G_k)

„gut: Unsicherheit bei Erklärungen einbezogen (statt positivistischer richtig/falsch-Dichotomie“ (EVWO09/Ö_k)

Während Lehrkräfte in ihren Begründungen das Nichtwissen negativ aufgreifen und dabei zum Beispiel auf das Fehlen eines sinnvollen Lernziels oder auf eine zu große ‚Blackbox‘ ohne eigentliche Ursache verweisen, ordnen Biologiedidaktiker:innen das aufgezeigte Nichtwissen anders ein: Insbesondere in den ersten zwei Beispielen wird deutlich, dass ein Bezug zu ‚Nature of Science‘ besteht bzw. Nichtwissen und Unsicherheit als Teil der naturwissenschaftlichen Forschung und deren Erkenntnisse ein normaler Zustand sind. Mit dem Bezug zu ‚Nature of Science‘ wird Nichtwissen und unsicheres Wissen auch als wissenschaftspropädeutischer Lerngegenstand wahrgenommen.

Auch Biologiedidaktiker:innen beziehen sich mit negativer Valenz vereinzelt auf die Vagheit der komplexen Erklärungen. MAPA89 beschreibt den Eindruck der Erklärung P_k beispielsweise mit *„Die Erklärung ist noch nicht genügend greifbar“* und MAWE02 die Erklärung Ö_k mit *„aber für Kinder eventuell zu diffus da nur potentielle Faktoren genannt werden, nicht aber tatsächliche aufgetretene Faktoren oder Situationen“*. DONO32 geht bei der Erklärung G_k etwas genauer darauf ein, warum sie einen zu wenig konkreten Eindruck macht: *„Aber: Irgendwie war das zu viel versprochen, weil letztlich eben nicht erklärt wurde wie MS entsteht. Das heißt die (inhaltliche) Güte der Erklärung lässt zu wünschen übrig: Irgendwie multifaktoriell / weiß man nicht genau“*. Aufgrund der Äußerungen ist auch bei Didaktiker:innen davon auszugehen, dass der Eindruck von Vagheit aus verschiedenen Gründen entsteht. Bei der komplexen Erklärung der Ökologie wird von MAWE02 wie beispielsweise auch von der Lehrkraft ANMA14 bemängelt, dass der konkrete Bezug auf die Situation der Entwicklung der Erdkrötenpopulation in der Stadt Merklingen, der zur Einführung der Problemstellung genutzt wird, zu wenig ist. DONO32 geht hingegen in ähnlicher Weise wie der/die Studierende ANAN85 und die Lehrkraft MAED33 auf die fehlenden, konkreten Mechanismen ein, die jedoch überwiegend nicht bekannt und somit in gesamtwissenschaftliches Nichtwissen einzuordnen sind (siehe Kap. 5.2.2.5).

Konsequenzen werden im folgenden Kapitel nach einer zusammenfassenden Beschreibung der Perspektiven erörtert.

6.2.4.4 Überblick und Einordnung der Ergebnisse zu Begründungsunterschieden zwischen linearen und komplexen instruktionalen Erklärungen in Abhängigkeit von der Statusgruppe

Komplexe Erklärungen scheinen bei Adressat:innen zwar keine Unsicherheit hervorzurufen, jedoch werden sie teilweise als vage wahrgenommen und dabei wird der Wunsch nach einem Ergebnis angedeutet. Zusätzlich werden die Zusammenhänge in kurzen instruktionalen Erklärungen für

Schüler:innen zum Teil nicht deutlich genug. Lineare Erklärungen sind hingegen durch die Abwesenheit von Vagheit geprägt. Die Fülle an wahrgenommenen konkreten Ergebnissen kann mitunter Ursache dafür sein, dass Schüler:innen die Merkbarkeit und Lernbarkeit von linearen Erklärungen als schwieriger einschätzen. Ein weiterer Aspekt, der bei komplexen Erklärungen zur Herausforderung werden kann, sind die Visualisierungen. In den Daten deutet sich an, dass die übersichtliche Darstellung schwieriger ist, was auch Mambrey et al. (2020) aufzeigen. Gleichzeitig gewinnt damit mitunter die Verständlichkeit der Visualisierung an Bedeutung. Die abgebildeten Tendenzen deuten zusammengenommen zum einen didaktische Herausforderungen für Lehrkräfte an – zum Beispiel durch die Verständlichkeit der Visualisierungen und die deutliche Herausarbeitung von Zusammenhängen. Zum anderen geben sie auch vereinzelt Hinweise darauf, dass Erklärungen, die die Komplexität biologischer Phänomene aufzeigen und Nichtwissen sowie unsicheres Wissen thematisieren, eher weniger den Erwartungen entsprechen, die Schüler:innen an Erklärungen im Biologieunterricht stellen. In Kapitel 6.2.1.5 wurde dieser Befund bereits in Beziehung zu Erwartungen gegenüber Erklärungen im Unterricht bzw. der Lehr-Lern-Kultur gestellt (Dittmer, 2010; Grotzer, 2012). So äußern Schüler:innen das Bedürfnis nach konkreten Ergebnissen und bewerten die Merkbarkeit und Lernbarkeit von Erklärungen.

Bei Studierenden konzentrieren sich die auffälligen Bewertungsunterschiede hauptsächlich auf Kriterien, die dem fachspezifischen Aspekt zugeordnet werden können. So geht die schlechtere Bewertung der linearen Erklärung aus dem Themenbereich Physiologie mit schlechterer Bewertung der fachlichen Vollständigkeit und der fachlichen Korrektheit einher, die bessere Bewertung des komplexen Pendants hingegen mit einer besser wahrgenommenen Deutlichkeit der Zusammenhänge und der Phänomenorientierung. Bei den zwei Videopaaren, bei denen die komplexen Erklärungen im standardisierten Zugang schlechter bewertet werden (Genetik und Ökologie), lassen sich in den Begründungen kaum Auffälligkeiten finden, die Unterschiede zwischen den Varianten markieren und Erklärungsansätze für die schlechtere Bewertung der komplexen Varianten liefern. Während die Phänomenorientierung und das Aufzeigen von multifaktoriellen Wirkungsgefügen in den komplexen Varianten eher positiv bewertet werden, werden die komplexen Erklärungen zum Teil auch als zu vage beschrieben. Die explizite Erwähnung von Wissenslücken wird von Studierenden positiv konnotiert in die Begründungen aufgenommen. Durch die Ergebnisse wäre sogar eine positive Bewertung der komplexen Varianten denkbar. Über alle Vignetten hinweg wird von einigen Studierenden das Erklärtempo in den komplexen Varianten schlechter, in den linearen Varianten hingegen besser bewertet.

Über alle Vignetten hinweg ergibt sich bei Lehrkräften kein einheitliches Muster, welche Kriterien bei den Erklärvarianten systematisch Anlass für eine gute bzw. schlechte globale Bewertung sind. Grund dafür sind auch die wenigen Äußerungen, die auf jede Vignette entfallen, und mitunter fallspezifische Bewertungen für jede Vignette. Für die bessere Bewertung der komplexen Variante im Themenbereich der Physiologie geben die vereinzelt Äußerungen zur mangelnden Vollständigkeit und der besser bewertete Einsatz von Repräsentationen Hinweise. Darüber hinaus wird die komplexe Erklärung nicht als vage beschrieben, was bei den anderen beiden Videopaaren der Fall ist. Bei den zwei weiteren Videopaaren, bei denen die komplexen Erklärvarianten schlechter bewertet werden, tritt in den Äußerungen hervor, dass für Lehrkräfte eine unangemessene Vagheit entsteht. Auch die explizite Thematisierung von Nichtwissen scheint für diese Statusgruppe nicht angemessen zu sein. Die negative Bewertung des Nichtwissens ist ein Alleinstellungsmerkmal der Lehrkräfte – Studierende und vor allem Didaktiker:innen erwähnen den Einbezug positiv. Die Phänomenorientierung, die bei Studierenden noch positiv Erwähnung findet, wird von Lehrkräften kaum in die Begründungen einbezogen und zusätzlich heterogen, also auch negativ, bewertet.

Für Didaktiker:innen scheinen Gründe für die schlechtere Bewertung der komplexen Variante in der Ökologie zu sein, dass das Beispiel nicht ausführlich genug in der Erklärung adressiert wird. Im Gegensatz dazu wird im linearen Pendant, das besser bewertet wird, die Verbindung der Repräsentationen auffällig häufig positiv zur Begründung der Notenvergabe herangezogen. Insgesamt ist bei Didaktiker:innen die Phänomenorientierung ein häufiger Bezugspunkt für die Bewertung. Innerhalb des Themenbereichs der Ökologie ist die Phänomenorientierung jedoch für die lineare Variante ausnahmsweise häufiger mit negativer Valenz in den Begründungen zu finden, sodass sich hier eine gegenläufige Tendenz zum Globalurteil zeigt. Mitunter ist dies durch die maximale Reduktion eines Ökosystems auf zwei Tierarten, die in einem Räuber-Beute-Beziehungen stehen, bedingt, wobei es sich jedoch um ein modellartiges Beispiel handelt. Bei den weiteren zwei weiteren Themenbereichen (Physiologie und Genetik) wird die Phänomenorientierung entsprechend dem Globalurteil bei den komplexen Erklärungen mit positiver Valenz in den Begründungen aufgenommen. Eine Besonderheit ergibt sich für die Bewertung des Nichtwissens: Es wird von Didaktiker:innen deutlich positiv wahrgenommen (was sich bei Studierenden auch andeutete, von Lehrkräften jedoch gegenläufig bewertet wird). Dabei wird an einigen Stellen ein Bezug zur ‚Nature of Science‘ in den Begründungen deutlich.

Fasst man die Befunde der Metaperspektive weiter zusammen, so zeigen sich Unterschiede in der Bewertung der Phänomenorientierung sowie der Bewertung von Nichtwissen und unsicherem Wissen in instruktionalen Erklärungen. Während durch die Ergebnisse der inhaltlich strukturierenden qualitativen Inhaltsanalyse bereits deutlich wurde, dass Lehrkräfte kaum auf die Phänomenorientierung eingehen (siehe Kap. 6.2.1.3 bzw. 6.2.1.5), zeigt sich anhand der evaluativen qualitativen Inhaltsanalyse eine Tendenz, dass bei Studierenden und Didaktiker:innen die Phänomenorientierung, also das Aufzeigen der Komplexität, im Zusammenhang mit positiven Bewertungen von Erklärungen steht.

Darüber hinaus wird von Lehrkräften das Erwähnen von Nichtwissen und unsicherem Wissen negativ bewertet – bei Studierenden finden sich vereinzelt positive Bewertungen zu diesem Punkt, bei Didaktiker:innen deutlicher positive Bewertungen. Vor allem im Kontrast der Antworten von Lehrkräften und Didaktiker:innen deutet sich an, dass Lehrkräfte bei der Bewertung die unmittelbaren Lernziele und mitunter das Abprüfen von Lerninhalten als Bezugspunkt sehen. Didaktiker:innen verweisen bei ihren Begründungen auch auf das Lernen über die Biologie und betten Nichtwissen und unsicheres Wissen damit den Nature-of-Science-Diskurs ein. Die Ergebnisse vertiefen die Differenzierung der Perspektiven, die in Kapitel 6.2.1.5 anhand der angesprochenen Kriterien erarbeitet wurde: Während Didaktiker:innen am ehesten mit einer genuin biologiedidaktischen Perspektive auf Erklärungen blicken, scheinen Lehrkräfte eher auf die Realisierung des Unterrichts und die konkrete Vermittlung zu achten. Studierende zeigten über alle Kriterien hinweg wechselnde Ähnlichkeiten mit Lehrkräften und Didaktiker:innen – in Falle der Phänomenorientierung und des Nichtwissens in instruktionalen Erklärungen besteht die Ähnlichkeit eher zu Didaktiker:innen.

Über alle Statusgruppen hinweg wird mit negativer Valenz darauf Bezug genommen, dass in den komplexen Erklärungen in unangemessener Weise Vagheit enthalten ist. Während Schüler:innen darauf verweisen, dass ein konkretes Ergebnis fehlt, begründen Lehrende die Vagheit meist inhaltlich, wobei sich die Begründungen über die Statusgruppen der Lehrenden hinweg ähneln, jedoch zwischen den Vignetten unterscheiden: So beziehen sich Begründungen zur komplexen Erklärung der Ökologie darauf, dass allgemein auf mögliche Faktoren eingegangen wird, die die Anzahl der Erdkröten moderieren, jedoch nicht auf die konkreten Zahlen in der Stadt Merklingen,

die als Einstiegsbeispiel diene, bezogen werden. Aus didaktischer Sicht könnte deshalb darüber nachgedacht werden, wie datenbasiert die Situation der Stadt Merklingen – vielleicht auch im weiteren Unterrichtsverlauf – zumindest in nähernder Abschätzung besser aufgenommen werden kann. Die Erklärungen würden dann verstärkt auf konkrete **Zahlen, aber dennoch ‚nur‘** auf möglichen Faktoren eingehen, was gleichermaßen mit vermehrten Abwägungen zur besten Erklärung für ein Sinken oder Ansteigen der Anzahl von Erdkröten einhergeht (probabilistische Erklärungen; ‚how-possibly‘-Erklärungen, Potochnik, 2013).

Die Vagheit in der komplexen Erklärung zur Genetik wird anders begründet: So beziehen sich die Lehrenden darauf, dass die zu Beginn der Erklärungen aufgeworfenen Fragestellung – bedingt durch das vorhandene gesamtwissenschaftliche Nichtwissen – nicht beantwortet wird. Die didaktische Konsequenz aus dieser von einigen Teilnehmenden wahrgenommenen Diskrepanz müsste demnach ein Überdenken der Fragestellung oder eine weitere Verdeutlichung des Nichtwissens und der Multikausalität innerhalb der Erklärung sein, wodurch die Fragestellung nicht eindeutig und abgeschlossen beantwortet werden kann. Gleichzeitig sind die Äußerungen Ausdruck dessen, dass die Frage nach der Entstehung Fakten nach sich zieht – Wissenslücken und unbekannte Faktoren werden weniger erwartet.

7 Zusammenfassung und Diskussion

In diesem Kapitel werden die zentralen Befunde noch einmal zusammengestellt, wobei hier insbesondere auch auf eine Zusammenführung der Ergebnisse zur quer liegenden Daten- bzw. Analysestruktur der Statusgruppen und des Komplexitätsgrades geachtet wird (Kap. 7.1). Die Befunde werden dabei in die Diskussion um instruktionale Erklärungen, wie sie in den Kapiteln 1 bis 4 entfaltet wurde, eingeordnet. Anschließend werden Limitationen erörtert (Kap. 7.2). Vor diesem Hintergrund werden zuletzt die Bedeutung bzw. Implikationen der Befunde für die schulische Praxis, die Lehrkräftebildung sowie die Forschung zu diesem Themenfeld abgeleitet (Kap. 7.3).

7.1 Zusammenfassung und Einordnung zentraler Befunde

Zielsetzung der vorliegenden Untersuchung war, instruktionale Erklärungen im Biologieunterricht zu konzeptualisieren und Kriterien guten Erklärens unter Berücksichtigung verschiedener Perspektiven – derjenigen von Schüler:innen, Studierenden, Lehrkräften und Didaktiker:innen – empirisch zu bestimmen. Folgende Befunde können in Hinblick auf die Ziele der Forschungsarbeit festgehalten werden.

Begriffsbestimmung zum instruktionalen Erklären im Biologieunterricht

Mit dem Ziel der fachspezifischen Begriffsbestimmung instruktionaler Erklärungen wurde einer Forderung nachgekommen, Unterricht bzw. Merkmale guten, effektiven oder qualitätvollen Unterrichts nicht ausschließlich fachübergreifend und allgemein zu denken, sondern fachspezifische bzw. fachdidaktische Besonderheiten in Überlegungen zu berücksichtigen (Baumert & Kunter, 2006; Helmke, 2007; Praetorius & Gräsel, 2021; Schilcher et al., 2021). Da biologiedidaktische

Überlegungen dazu fehlen, wurde neben einer sprachwissenschaftlichen sowie einer lerntheoretischen, allgemeindidaktischen Perspektive auf Erklären auch eine fachliche, wissenschaftstheoretische Perspektive in die Konzeptualisierung einbezogen. Letztere bezieht sich nicht auf den Vermittlungsprozess von Wissen, wie es beim unterrichtlichen Erklären der Fall ist, sondern auf Erklärungen der wissenschaftlichen Disziplin und deren Formen, Eigenschaften und Gültigkeit.

Die Konzeptualisierung aus den drei genannten Perspektiven nimmt zusammenfassend auf folgende Punkte Bezug (siehe Kap. 2.4):

Instruktionale Erklärungen im Biologieunterricht sind vielgestaltige Sprachhandlungen von Lehrkräften, die an Lernende gerichtet sind und die Initiierung einer Wissensvermittlung, eines Verstehensprozesses oder eines Fähigkeitserwerbs zum Ziel haben, wobei sich die Erklärgegenstände stets der direkten Wahrnehmung entziehen. Bei der Erklärung biologischer Phänomene orientieren sich Lehrkräfte an wissenschaftlichen Erklärungsformen der Biologie und modellieren damit implizit, welche Erklärungen aus Sicht des Faches angemessen sind.

Neben den fachübergreifenden Formen instruktionaler Erklärungen (Begriffserklärungen, Handlungserklärungen, Erklärungen zum Zweck des Erklärinhalts) sind im Fach Biologie kausale (komplexe und lineare kausale Erklärungen, Wahrscheinlichkeitsabhängige kausale Erklärungen), mechanistische, funktionale und naturhistorische Erklärungen fachlich angemessen bzw. notwendig, um die fachlichen Phänomene erklären zu können.

Als Ergebnis der Arbeit kann daher auch formuliert werden, dass sich der Einbezug der wissenschaftstheoretischen Perspektive zur Erarbeitung von Fachspezifika als gewinnbringend gezeigt hat: Durch den Diskurs über die Merkmale zu erklärender Phänomene, fachspezifischer Erkenntnisgewinnung und daraus resultierender Formen der Erklärungen in der Biologie konnten die Formen angemessener Erklärungen im Biologieunterricht sowie die Facette der fachlichen Korrektheit fachspezifisch differenziert werden. Damit können wissenschaftstheoretische Überlegungen nicht nur zur Ableitung von Inhalten für einen wissenschaftspropädeutischen Unterricht dienlich (Abrams et al., 2001; Braaten & Windschitl, 2011; Potochnik, 2013; Zabel, 2009), sondern auch für die fachspezifische Ausrichtung von fachdidaktischer Forschung von Nutzen sein.

Aspekte und Kriterien guter instruktionaler Erklärungen im Biologieunterricht

Das zweite Ziel der Untersuchung bezog sich auf die Bestimmung von Kriterien guten Erklärens im Biologieunterricht. Dabei wurden verschiedene Perspektiven einbezogen, um die normativen Standards der beteiligten Gruppen zu berücksichtigen (Berliner, 2005) und ein tieferes Verständnis für den Gegenstand zu ermöglichen (Denzin, 1970; Flick, 2011), wobei die Perspektiven gleichberechtigt aufeinander bezogen werden (Gläser & Laudel, 2010). Für die empirische Identifikation von Kriterien guten Erklärens und die Bestimmung ihrer Relevanz wurde ein Instrument entwickelt, das die Erfassung aus verschiedenen Perspektiven erlaubt. Neben geschlossenen Items, mit denen Bewertungsunterschiede und maßgebliche Faktoren auf die Güte einer Erklärung analysiert werden können, erlaubt ein offenes Item, in dem die Teilnehmenden die Begründung zur globalen Bewertung eingeben können, Kriterien guten Erklärens zu erfassen.

Die Ergebnisse dieser Begründungen zeigen, dass der Einsatz von Repräsentationen, die Adressatenorientierung, die Strukturiertheit, der Sprech- und Körperausdruck und die sprachliche Verständlichkeit Aspekte guten Erklärens sind. Die Ergebnisse decken sich mit den in der Literatur

diskutierten Aspekten (z. B. Geelan, 2012; Odora, 2014; Schopf & Zwischenbrugger, 2015; Sevia & Gonsalves, 2008; Treagust & Harrison, 2000; Wagner & Wörn, 2011; Wragg, 1993). Auch die Merkbarkeit und Lernbarkeit des Gesagten spielen für die Bewertungen eine Rolle. Der Aspekt wird in der Literatur nicht direkt als Aspekt guten Erklärens geführt, sondern spiegelt implizit die Funktion von instruktionalen Erklärungen – bzw. von Vermittlung allgemein – im Schulunterricht wider. Die Lehrperson als solche tritt in den Begründungen in wenigen Fällen auf. Die Häufigkeiten, mit denen auf die genannten Aspekte Bezug genommen wird, variieren je nach Statusgruppe (siehe Abschnitt ‚Statusgruppe‘), wobei die Nennung vermutlich von mehreren Faktoren moderiert wird (z. B. Anlage des Studiendesigns, Kürze der Erklärungen, fehlendes Wissen).

Von 60 Kriterien, die durch die Literatursynopse differenziert wurden, nutzen die Teilnehmenden 40, wobei 12 weitere selten herangezogen werden (aus Gründen der Übersichtlichkeit werden die Kriterien an dieser Stelle nicht mehr aufgelistet – eine Übersicht findet sich in Kap. 6.2.1.1). Über die literaturbasiert erarbeiteten Kriterien hinaus können im Material sechs weitere gefunden werden, die sich den Aspekten zuordnen lassen: (Aktiver) Einbezug der Rezipient:innen (Adressatenorientierung), Übereinstimmung des Erklärziels mit dem Erklärinhalt (Strukturiertheit), ästhetische Gestaltung der realistischen Bilder (Repräsentationen), ausreichende Bezugnahme auf das Beispiel (Repräsentationen), Sprechflüssigkeit (Sprech- und Körperausdruck) und Zusammenhänge zwischen den Faktoren innerhalb der Erklärung werden deutlich (fachspezifischer Aspekt). Hinzu kommen globale Aspekte: Die instruktionale Erklärung sollte verständlich, anschaulich, übersichtlich, nachvollziehbar, klar, sachlich und informativ sein. Außerdem werden die Lehrperson und ihre Merkmale (z. B. Humor, Autorität, Kompetenz, Stil) sowie die Merkbarkeit und Lernbarkeit mit in die Bewertung einbezogen.

Die Forderung nach einem aktiven Einbezug der Rezipient:innen muss dabei reflektiert werden und ist vermutlich auch gegenstandsabhängig – die oberflächliche Sichtbarkeit von Aktivitäten seitens der Rezipient:innen (z. B. durch Redebeteiligungen einzelner Schüler:innen) ist dabei kein Anzeichen für einen Wissenserwerb (Brüning & Saum, 2019).

Einige in der Literatur aufgegriffenen Kriterien lassen sich in den Begründungen der Teilnehmenden nicht finden. Unter den fehlenden Kriterien ist unter anderem auffällig, dass der Lebensweltbezug nicht als Kriterium herangezogen wird. Vor allem in praxisnahen Quellen wird dieser angeführt (z. B. Pauli, 2015; Schopf & Zwischenbrugger, 2015; Wellenreuther, 2013).

Konzeptualisierung und Operationalisierung guter instruktionaler Erklärungen im Biologieunterricht

Bereits auf Grundlage der Literaturrecherche fällt auf, dass Kriterien guter instruktionaler Erklärungen auf verschiedenen Abstraktionsebenen angesiedelt sind. Das zeigt beispielsweise die Gegenüberstellung des Kriteriums des Lebensweltbezugs, der an der Auswahl des Themas bzw. der Problemstellung oder der Wahl der Beispiele festgemacht werden kann, auf der einen Seite und des Kriteriums der Berücksichtigung der Kapazität des Arbeitszeitgedächtnisses, das sich als latentes Merkmal und komplexes theoretisches Konstrukt darstellt und daher nicht direkt innerhalb der Erklärung beobachtbar ist, auf der anderen Seite. Weiterhin sind die Kriterien auf unterschiedliche Grade subjektiv. So kann das Vorhandensein einer Zusammenfassung eher dichotom mit ‚Ja‘ oder ‚Nein‘ beantwortet werden und ist weniger subjektiv als die Fokussierung auf das Wesentliche, dessen Güte eher auf einem Kontinuum bewertet wird und subjektiver zu sein scheint, da die

Bewertung eher von den Voraussetzungen der bewertenden Person abhängig ist.¹⁸⁰ Das macht die **Konzeptualisierung und Operationalisierung des Konstrukts ‚Gutes instruktionales Erklären‘ und der Aspekte** insgesamt zur Herausforderung und äußert sich auch in einer Vielzahl an Vorschlägen (u.a. Findeisen, 2017; Kulgemeyer & Tomczyszyn, 2015; Schopf & Zwischenbrugger, 2015; Sevia & Gonsalves, 2008; Wagner & Wörn, 2011).

Empirische Argumente für die hier realisierte Konzeptualisierung finden sich an verschiedenen Stellen in den erhobenen Daten: Die hohe Varianzaufklärung in den hierarchischen linearen Regressionsmodellen spricht für die angemessene Auswahl der Aspekte bzw. der Kriterien und damit auch für die Konstruktvalidität. Weiterhin ist bis auf die Verständlichkeit der Sprache, die jedoch aufgrund der geringen Itemanzahl und im Mittel niedrigen Itemtrennschärfen als Konstrukt zu überarbeiten ist, jeder der fünf Aspekte für mindestens eine der Statusgruppen ein maßgeblicher Einflussfaktor auf die Güte einer Erklärung (siehe Kap. 6.2.3), was ebenfalls ein empirisches Argument für die Auswahl der Konstruktvalidität ist. Die im standardisierten Zugang gewählten Aspekte und Kriterien lassen sich zudem überwiegend – wenn auch zum Teil selten – in den von den Teilnehmenden selbst herangezogenen Kriterien im offenen Textfeld finden, was für eine gute Konstruktvalidität spricht. Nur die Facette „Berücksichtigung der Auswirkung des Gesagten“ im fachspezifischen Aspekt wird nicht herangezogen. Die im Zuge der Studie durchgeführte Augenscheinvalidierung¹⁸¹ legt darüber hinaus nahe, dass die vorgenommene Konzeptualisierung der fünf Aspekte sowie der Persönlichkeitswirkung durch die gewählten Kriterien trennscharf ist.

Damit liegt erstmals ein Vorschlag für eine validierte Konzeptualisierung guter instruktionaler Erklärungen vor, die Spezifika des Biologieunterrichts bzw. biologischer Phänomene berücksichtigt.

Maßgebliche Faktoren

Ein weiteres Ziel der Studie war die Bestimmung von maßgeblichen Einflussfaktoren auf die Güte einer Erklärung. Die Ergebnisse der Studie liefern erstmals empirische Ansatzpunkte bzw. Argumente unter anderem für die Auswahl von Fokussen für weitere Forschungsansätze zur Effektivität von Erklärungen sowie die Aus- und Weiterbildung von Biologielehramtsstudierenden und Biologielehrkräften (siehe Kap. 7.2).

Über alle Statusgruppen hinweg zeigten sich die Adressatenorientierung sowie die Strukturiertheit als zentrale Aspekte, die zur Güte einer Erklärung beitragen. Damit lässt sich empirisch abbilden, was in der Literatur häufig aufgrund theoretischer Überlegungen angeführt wird: Die Adressatenorientierung ist zentrales Merkmal instruktionaler Erklärungen (Hargie, 2013; Schilcher

¹⁸⁰ Schopf und Zwischenbrugger (2015) unterscheiden zwischen Elementen einer guten Erklärung und Merkmalen einer guten Erklärung: „Während Elemente Bausteine darstellen, die eine verständliche Erklärung beinhalten sollte, sind Merkmale Attribute, die die einzelnen Elemente bzw. die gesamte Erklärung erfüllen sollte/n“ (Schopf & Zwischenbrugger, 2015, S. 12). Diese Idee könnte auch für die Unterscheidung in Hinblick auf die oben beschriebenen Unterschiede der Kriterien hilfreich sein. Während jedoch die Begriffe ‚Element‘ und ‚Merkmal‘ gut auf das oben angeführte Beispiel passen könnten, deutet sich in den Beispielen, die die beiden Autorinnen zuordnen, ein Unterschied an, der eine direkte Übertragung nicht möglich macht: Als ‚Elemente‘ werden sowohl das Vorhandensein von Visualisierungen als auch das Anknüpfen an bereits vorhandene relevante Eingangsvoraussetzungen gefasst. Letzteres wäre nach obiger Auslegung ein Kriterium mit höherer Subjektivität bei der Bewertung. Unter ‚Merkmale‘ werden sowohl die Konzentration auf das Wesentliche, jedoch auch die fachliche Richtigkeit gezählt.

¹⁸¹ Für die Validierung bzw. die Bestimmung der Trennschärfe der fächerübergreifenden Aspekte Adressatenorientierung, Strukturiertheit, sprachliche Verständlichkeit, Sprech- und Körperausdruck sowie die Wirkung der erklärenden Person wurden Lehramtsstudierende, Promovierende und Professor:innen der empirischen Bildungsforschung ($N = 98$) der Universität Regensburg befragt, die die Kriterien den fünf Aspekten zuordnen sollten (siehe auch Kap. 6.1.2 bzw. Schilcher, Krauss, Lindl & Hilbert, im Druck).

et al., 2017; Wittwer & Renkl, 2008). Die Strukturiertheit lässt sich ebenfalls prominent im Diskurs um Kriterien guter oder effektiver Erklärungen (Brown, 2006; Hargie, 2013; Wagner & Wörn, 2011; Wragg, 1993) und darüber hinaus auch in Überlegungen zu Merkmalen guten Unterrichts als fachübergreifende Kriterien finden (Helmke, 2007; Meyer, 2004). In den im deutschsprachigen Raum am häufigsten rezipierten drei Basisdimensionen von Unterrichtsqualität (kognitive Aktivierung, konstruktive Unterstützung und Klassenführung) lassen sich die zwei Aspekte nicht prominent finden. Kleickmann, Steffensky und Praetorius (2020) greifen hingegen beide Aspekte als Merkmale der kognitiven Unterstützung (**im Original „cognitive support“**, Kleickmann et al., 2020, S. 38) im Unterricht auf, die sie als vierte Basisdimension der Unterrichtsqualität beschreiben. Sie **definieren diese vierte Dimension folgendermaßen**: „In sum, cognitive support aims to reduce complexity and cognitive demands by means of structuring content and promoting clarity, so that students can master the respective tasks and successfully gain **understanding**“ (Kleickmann et al., 2020, S. 38). Insofern könnten Ergebnisse der Unterrichtsqualitätsforschung zum Teil auf instruktionale Erklärungen übertragbar sein. In Hinblick auf den maßgeblichen Einfluss der Strukturiertheit auf die Einschätzung der instruktionalen Erklärungen ist zu beachten, dass zwischen Lernenden und Lehrenden erhebliche Unterschiede in den Einschätzungen zu verzeichnen sind (siehe Kap. 6.2.2).

In der statusgruppenspezifischen Analyse zeigte sich für Schüler:innen und die Gruppe der Metaperspektive jeweils noch ein weiterer Aspekt als relevant, der die Unterschiedlichkeit der Perspektiven bzw. Einschätzungen markiert: für Schüler:innen der Sprech- und Körperausdruck und für die Gruppe der Metaperspektive der fachspezifische Aspekt zur Wahrnehmung des dargestellten Komplexitätsgrades.

Dabei überrascht die Relevanz des Sprech- und Körperausdrucks für die Bewertung instruktionaler Erklärungen durch Schüler:innen, da ihnen eher affektiv geprägte und globale Beurteilungen zugesprochen werden (Clausen, 2002). Der Sprech- und Körperausdruck, wie er hier operationalisiert wurde, ist jedoch im Gegensatz zur Wirkung der erklärenden Person eher beschreibend und detailliert auf Merkmale bezogen – wenn auch auf eher oberflächlich wahrnehmbare wie beispielsweise die Zugewandtheit der erklärenden Person.

In den FALKE-Studien anderer Fächer zeigt sich bezüglich der maßgeblichen Einflussfaktoren ein anderes Bild: Während für Erklärungen im Schulfach Musik bei Schüler:innen der Sprech- und Körperausdruck sowie die sprachliche Verständlichkeit maßgebliche Faktoren sind (Frei, Puffer & Hofmann, im Druck), erweisen sich im Fach Deutsch bei Schüler:innen der Sprech- und Körperausdruck sowie der fachspezifische Aspekt (Erklärtiefe) als maßgebliche Einflussfaktoren (Gaier, 2022). Im Fach Physik erhält hingegen keiner der Aspekte im Regressionsmodell einen signifikanten Regressionskoeffizienten (Heinze & Rincke, im Druck). Für die Gruppe der Metaperspektive zeigt sich im Fach Musik nur die Adressatenorientierung als signifikanter Prädiktor (Frei et al., im Druck). Im Fach Deutsch ergibt sich hingegen ein ähnliches Bild wie im Fach Biologie – die Adressatenorientierung und Strukturiertheit sowie der fachspezifische Aspekt zeigen sich anhand der Daten als maßgebliche Prädiktoren für die Gesamtbewertung (Gaier, 2022). Im Fach Physik sind die Adressatenorientierung und die Strukturiertheit maßgebliche Einflussfaktoren für die Gruppe der Metaperspektive (Heinze & Rincke, im Druck). Vermutet werden kann, dass die Differenzen zum Teil auf die unterschiedlichen Konzeptualisierungen und Operationalisierungen der fachspezifischen Aspekte sowie das Alter der befragten Schüler:innen zurückführbar sind. Die Unterschiede können jedoch auch durch die Inhalte der Erklärungen bedingt sein. Damit wären die unterschiedlichen Prädiktoren ein weiteres Argument, Unterrichtsforschung nicht ausschließlich aus

einer fachübergreifenden Perspektive heraus zu denken – aus der Bildungsforschung, der Pädagogik oder der Psychologie –, sondern auch aus der fachlichen bzw. fachdidaktischen Perspektive zu betrachten (Baumert & Kunter, 2006; Helmke, 2007; Praetorius & Gräsel, 2021; Schilcher et al., 2021).

Ob die maßgeblichen Einflussfaktoren auf die Güte instruktionaler Erklärungen auch lernrelevante Einflussfaktoren sind, kann anhand der erhobenen Daten nicht abgeschätzt werden. Hinweise darauf liefern die bisher wenigen empirischen Daten anderer Studien. Beispielsweise analysiert Wragg (1993, S. 135) „high-scoring explanations“ (Erklärungen von Lehrkräften, nach denen die Schüler:innen besonders hohe Werte in einem Test aufweisen) zu Merkmalen von Insekten und beschreibt diese mit folgenden Kriterien:

clearly structured, with the 'keys' or central ideas made distinct and lined to one another in a logical sequence and shape, [...] the teacher's subjects knowledge was factually correct. Language choice were sensible, neither too banal nor at too high a level of abstraction [...]. Voice was clear and well modulated, with effective movement, facial expression an animation, and there was fluency and pace. (Wragg, 1993, S. 135–136)

Die von Wragg (1993) aufgezählten Kriterien lassen sich in die Konzeptualisierungen der vorliegenden Studie zu den Aspekten der Strukturiertheit, der sprachlichen Verständlichkeit und des Sprech- und Körperausdrucks einordnen. Die Adressatenorientierung ist hingegen nicht explizit aufgeführt.

Die von Kleickmann et al. (2020) durchgeführte Untersuchung zur vierten Basisdimension von Unterrichtsqualität legt nahe, dass ein Einfluss der Adressatenorientierung und Strukturiertheit auf die Lernleistung von Schüler:innen zu erwarten ist.

Statusgruppen

Der explorative Charakter der Studie zeigt sich insbesondere auch in der Befragung der Statusgruppen. Ziel war dabei insbesondere, Erkenntnisse darüber zu gewinnen, welche Kriterien die verschiedenen Personengruppen zur Bewertung heranziehen und an welchen Stellen sich die Statusgruppen hinsichtlich der Bewertung einig bzw. uneinig sind, sowie dadurch ein umfassendes Verständnis von guten instruktionalen Erklärungen zu generieren. Die Befunde dazu können auch zur Diskussion über die perspektivenspezifische Validität, die vor allem im Rahmen der Unterrichtsqualitätsforschung geführt wird, beitragen (z. B. Clausen, 2002; Kleickmann et al., 2019).

Die Daten des standardisierten Forschungszugangs deuten darauf hin, dass es nicht vollkommen unterschiedliche Vorstellungen darüber gibt, welche instruktionalen Erklärungen gut sind. So zeigt sich im standardisierten Zugang nur ein signifikanter Unterschied in der globalen Bewertung zwischen Schüler:innen und Lehrkräften. Auf der Ebene der fünf Aspekte guten Erklärens finden sich hingegen unterschiedliche Bewertungsmaßstäbe zwischen den Statusgruppen: Schüler:innen schätzen die Strukturiertheit häufig signifikant schlechter ein als die Gruppen der Metaperspektive. Der Sprech- und Körperausdruck wird von Schüler:innen häufig signifikant schlechter als von Didaktiker:innen eingeschätzt. Die Wahrnehmung des dargestellten Komplexitätsgrades (fachspezifischer Aspekt) wird von Schüler:innen ebenfalls in einigen Fällen signifikant schlechter bewertet als von Studierenden. Keine signifikanten Unterschiede konnten für den Aspekt der Verständlichkeit der Sprache sowie für die Adressatenorientierung gefunden werden.

In den Begründungen der Notenvergabe – also im offenen Textfeld – wird jedoch deutlich, dass sich die Statusgruppen hinsichtlich der selbst-herangezogenen Kriterien unterscheiden, sodass verschiedene Schwerpunkte in den Perspektiven erkennbar werden: Schüler:innen beziehen sich am häufigsten auf den Einsatz von Repräsentationen, auf globale Aspekte sowie auf die Adressatenorientierung. Auf der Ebene der Kriterien konnte allgemein herausgearbeitet werden, dass Schüler:innen entlang eigener Präferenzen bewerten und dabei weniger Kriterien als die weiteren Statusgruppen nutzen, insbesondere zum Sprech- und Körperausdruck sowie zum fachlichen Aspekt. Der Befund von Clausen (2002), der Schüler:innenurteile zur Unterrichtsqualität als affektiv geprägt generalisierend vor allem unter globalem Einbezug ihrer Lehrkräfte beschreibt, kann damit nicht gänzlich bestätigt werden. Letzterer Effekt lässt sich mitunter auch dadurch abmildern, dass die Schüler:innen in der hier vorliegenden Untersuchung – wie auch alle weiteren Statusgruppen – fremde Unterrichtsausschnitte bewerten, also nicht direkt am Unterrichtsgeschehen beteiligt sind und damit keine Beziehung zur erklärenden Person haben. Festzustellen ist hingegen, dass Schüler:innen beispielsweise in Bezug auf Repräsentationen eher oberflächlich Wahrnehmbares in die Bewertungen einfließen lassen und sich deutlich häufiger auf **globale Aspekte (wie „verständlich“, „anschaulich“) beziehen.**

Studierende rekurren in ihren Begründungen häufig auf die Aspekte zum Einsatz von Repräsentationen und zur Adressatenorientierung sowie deutlich seltener auf den fachlichen Aspekt. In Bezug auf die Kriterien kann festgehalten werden, dass Studierende ähnlich differenziert wie Lehrkräfte und Didaktiker:innen bewerten. In ihren Bewertungen zeichnet sich dabei ihr Status als Noviz:innen ab: Es zeigen sich wechselnde Ähnlichkeiten mit den Bewertungsmustern von Adressat:innen auf der einen Seite und Lehrkräften und Didaktiker:innen auf der anderen Seite. Für Forschungsfragen zu gutem oder effektivem Unterricht scheint diese Perspektive damit am wenigsten notwendig zu sein.

Lehrkräfte sprechen wie Studierende am häufigsten Kriterien an, die den Aspekten der Repräsentationen und der Adressatenorientierung sowie dem fachlichen Aspekt zuzuordnen sind. Insgesamt sind die selbst herangezogenen Kriterien dabei vielfältig. Ihre Perspektive kann anhand der Begründungen als eher unterrichtsbezogen und auf die konkrete Vermittlung fokussiert beschrieben werden. Dabei wird von Lehrkräften häufig auf den aktiven Einbezug von Schüler:innen verwiesen – die fehlende sichtbare Aktivität bei instruktionalen Erklärungen scheint für Lehrkräfte problematisch zu sein, was auch in der Literatur zum Spannungsverhältnis zwischen instruktionaler Erklärung bzw. Instruktion allgemein und einem konstruktivistischem Lernverständnis aufgegriffen wird (Gaier, 2022; Hattie, 2015; Leisen, 2007). Dass Lehrkräfte die Merkbarkeit und Lernbarkeit des Gesagten in ihren Begründungen aufgreifen und Nichtwissen negativ anführen, legt vor allem im Kontrast zu den Didaktiker:innen den Schluss nahe, dass für Lehrkräfte instruktionalen Erklärungen eine faktenorientierte, wissensvermittelnde Funktion im Unterricht zu kommt.

Didaktiker:innen bewerten Erklärungen hingegen am häufigsten anhand von Kriterien, die den Aspekten der Repräsentationen oder dem fachlichen Aspekt zuzuordnen sind. Deutlich seltener wird die Adressatenorientierung angesprochen. Insgesamt nutzen Didaktiker:innen für die Bewertung das breiteste Spektrum an Kriterien, wobei durch die gewählten Kriterien bzw. Begründungen sichtbar wird, dass Didaktiker:innen im Gegensatz zu Lehrkräften vor allem biologische und biologiedidaktische Kriterien verwenden. Unter den fachlichen Bezügen sind auch solche, die die Wissenschaftstheorie bzw. Überlegungen zur Wissenschaftspropädeutik einbeziehen, die sich in den anderen Statusgruppen kaum finden lassen. Die Perspektive kann verallgemeinernd am ehesten als genuin biologiedidaktisch beschrieben werden. Hervorzuheben sind zwei weitere Ergebnisse: Es

finden sich keine expliziten Verweise in den Begründungen zur Merkbarkeit und Lernbarkeit des Gesagten und die Erwähnung von Nichtwissen wird meist positiv in die Bewertung der Erklärung aufgenommen. Damit kann angenommen werden, dass Didaktiker:innen zum einen eine andere Auffassung von fachlicher Korrektheit haben bzw. Nichtwissen insgesamt anders als Lehrkräfte bewerten. Zum anderen kann in Bezug auf instruktionale Erklärungen im Biologieunterricht davon ausgegangen werden, dass Didaktiker:innen sich etwas mehr von der Lehrplan-, Prüfungs- und Outputorientierung, wie sie beispielsweise von Dittmer (2010) oder Grotzer (2012) aufgegriffen wird, distanzieren können und das Lernen über biologische Phänomene breiter denken.

Aufgrund der Unterschiede ist der Einbezug der Perspektiven als sinnvoll zu erachten und entlang Denzins (1970) und Flicks (2011) Vorschlag für ein vertieftes Verständnis des Gegenstands – gute instruktionale Erklärungen im Biologieunterricht – unverzichtbar.

Komplexe und lineare instruktionale Erklärungen im Biologieunterricht

Komplexe Erklärungen sind im Unterricht vermutlich eher unüblich (Grotzer, 2012; Jacobson & Wilensky, 2006; Potochnik, 2013; siehe auch Kap. 3.5 zur Analyse der Bildungsstandards, KMK, 2020). Die Form der Erklärung mit ihren Merkmalen – multikausale Ursache-Wirkungs-Gefüge, Wechselwirkungen zwischen Faktoren, Nichtwissen und unsicheres Wissen – ist auch in der Biologie unverzichtbar (Grotzer, 2012; Kampourakis & McCain, 2020; Mainzer, 2008; Mitchell, 2008; Potochnik, 2013). Die wissenschaftstheoretisch fundierte Konzeptualisierung komplexer Erklärung bzw. deren Merkmale sind damit ein erster Schritt für die Implementierung von Lerngelegenheiten in die schulische Praxis und die Ausbildung von Lehrkräften.

Da diese Form der Erklärung Herausforderungen für Lernende und Lehrende darstellen kann, sind Befunde zur Wahrnehmung bzw. zur Bewertung von besonderem Interesse. Auf Ebene der Globalurteile bewerten vor allem Schüler:innen die komplexen Varianten schlechter, Studierende und Lehrkräfte bewerten sie überwiegend schlechter und Didaktiker:innen bewerten konträr dazu eher lineare Erklärungen schlechter. Die Begründungen zur Notenvergabe lassen erste Einblicke in die Ursachen für die Bewertungen zu: Über alle Statusgruppen hinweg wird eine entstehende Vagheit negativ bewertet. Während Schüler:innen ein konkretes Ergebnis fehlt, tritt bei der Gruppe der Metaperspektive in den Vordergrund, dass nur mögliche Faktoren aufgegriffen werden. Referenz finden diese Äußerungen unter anderem bei Jacobson (2000), der beschreibt, dass beim Denken in und über komplexe Systeme häufig von vollständiger Vorhersagbarkeit der Effekte ausgegangen („completely predictable“; Jacobson, 2000, S. 17) und nicht an Zufall und probabilistische Zusammenhänge gedacht wird. Als didaktische Herausforderung kann daher gelten, Wissenslücken aufzugreifen und dabei keine Vagheit zu erzeugen.

Die Teilnehmenden äußern sich weiterhin nicht derart in Bezug auf die Erwähnung von Wissenslücken in den Erklärungen, dass eine Verunsicherung oder sogar ein Vertrauensverlust in die Wissenschaft anzunehmen ist, wie zum Teil vermutet wird (Böschen et al., 2008; Kampourakis & McCain, 2020; Maier et al., 2018; Varwig, 2020), mitunter, da die Erklärgegenstände nicht unmittelbar in einen Prozess der Bewertung oder Entscheidungsfindung eingebunden sind. Allerdings decken sich die Befunde auch nicht mit denen von Retzbach und Maier (2015), die eine leichte Erhöhung des Interesses feststellen konnten, wenn Wissenslücken explizit gemacht werden.

Für die Gruppe der Adressat:innen zeichnet sich ab, dass die Visualisierungen bei komplexen Erklärungen vermehrt in den Fokus rücken und für ein Verständnis an Bedeutung gewinnen. Das

Ergebnis deckt sich mit den Überlegungen bzw. Befunden, dass Visualisierungen allgemein wichtig für ein Verständnis von Erklärungen sind, (z. B. Schopf & Zwischenbrugger, 2015; Sevian & Gonsalves, 2008; Wragg, 1993) und im Speziellen in Hinblick auf die Merkmale komplexer Erklärungen damit, dass Visualisierungen eine strukturierende Funktion zukommt (Hargie, 2013; Mambrey et al., 2020) und sie den ‚cognitive load‘ reduzieren können (Grotzer et al., 2017).

Während sich bei Studierenden und Didaktiker:innen Überlegungen zur ‚Nature of Science‘ widerspiegeln, vor deren Hintergrund der Einbezug von Nichtwissen und unsicherem Wissen eher positiv bewertet wird, spiegelt sich bei Schüler:innen und Lehrkräften eher wider, dass fehlende eindeutige Ergebnisse nicht zu den Erwartungen an eine gute Erklärung gehören. In den Bewertungen der komplexen Erklärungen finden sich Tendenzen, die auf den Wunsch nach Abgeschlossenheit (Webster & Kruglanski, 1994) und Eindeutigkeit (Krüger et al., 2013) hinweisen. Die Einforderung von Eindeutigkeit und Abgeschlossenheit bzw. die negative Bewertung von Nichtwissen und unsicherem Wissen steht dabei in einem Spannungsverhältnis zum Verständnis der fachlichen Korrektheit. Es stellt sich die Frage, ob Erklärgegenstände auch als fachlich korrekt dargestellt gelten können, wenn Wissenslücken nicht kommuniziert werden (siehe Kap. 4.6; siehe Kap. 7.3 für Implikationen).

In Hinblick auf gutes Erklärens scheint es so zu sein, dass neben der grundlegend Anpassung an die Adressat:innen, der gelungenen Strukturierung, dem angemessenen Einsatz des Sprech- und Körperausdrucks sowie der adäquaten fachlichen Darstellung jeweils spezifische Herausforderungen vom jeweiligen Erklärgegenstand ausgehen und damit einhergehend Gestaltungskriterien bzw. -merkmale variieren. Auch Wragg (1993) kommt zu diesem Schluss:

As is often the case, however, despite some communalities, high-scoring explanations were also different from each other, with their own unique qualities. The general conclusions above are not universal prescriptions, but rather interesting illuminations from one intensive study. (Wragg, 1993, S. 136)

Für komplexe Erklärungen zeigen sich insbesondere die Implementierung von Nichtwissen und unsicherem Wissen vor dem Hintergrund der Erwartungen an instruktionale Erklärungen, die Vermeidung von Vagheit in Erklärungen, die Wissenslücken thematisieren, und die Erstellung unterstützender und verständlicher Visualisierungen als Herausforderungen.

7.2 Reflexion des Vorgehens, Limitationen und Ausblick

Design und Forschungsmethodik

Durch das Studiendesign können Kriterien guter instruktionaler Erklärungen aus verschiedenen Perspektiven erfasst werden. Von Vorteil dürfte dabei sein, dass alle befragten Statusgruppen bzw. Personen Erklärungen bewerten, in die sie nicht unmittelbar involviert sind. Kleickmann et al. (2019) sehen die Beurteilung von Unterricht durch Lehrkräfte und Schüler:innen, die selbst am zu beurteilenden Unterricht teilnehmen bzw. diesen gestalten – wie es häufig in Studien zur Unterrichtsqualität realisiert wird –, als große Schwäche für die Validität der Ergebnisse. Diese Schwäche konnte durch das Vorgehen umgangen werden.

Ein Nachteil des Vorgehens ist die Einschränkung hinsichtlich der ökologischen Validität: Bei den **Vignetten handelt es sich um sogenannte ‚staged-videos‘** (Piwowar et al., 2018), also nachgestellte Szenen bzw. instruktionale Erklärungen, die so oder so ähnlich auch im Biologieunterricht

vorkommen können. Durch die gegebene Einschränkung scheint die Überprüfung anhand realen Unterrichts sinnvoll. Ob die hier aufgegriffene experimentelle Manipulation zur Herausarbeitung von Unterschieden zwischen komplexen und linearen Erklärungen in realen Unterrichtssettings aufgegriffen werden kann, ist fraglich. Mitunter müsste sich allgemein auf instruktionale Erklärungen beschränkt werden, da komplexe Erklärungen vermutlich nur selten Teil der Unterrichtskultur sind – zumindest lassen die Untersuchungen aus anderen Ländern (Grotzer, 2012; Jacobson & Wilensky, 2006; Potochnik, 2013) und die Analyse der Bildungsstandards vermuten (KMK, 2020).

Durch das Studiendesign ist sowohl die Abfrage selbstgewählter Kriterien durch ein offenes Textfeld als auch die Bewertung von Erklärungen durch vorgegebene Kriterien anhand geschlossener Items möglich. Die Abfrage eines Globalurteils nach dem ersten Betrachten der sechs Vignetten und die Abfrage der Bewertung anhand von vorgegebenen Kriterien nach dem zweiten Betrachten der sechs Vignetten erlaubt außerdem erstmals, die Relevanz von Kriterien für die Güte einer Erklärung zu bestimmen. Die Bearbeitungszeit des Fragebogens ist durch die zwei Messzeitpunkte (bzw. Durchgänge) jedoch relativ hoch. Bei einer Erweiterung der abgefragten Kriterien (siehe Abschnitt zu Konzeptualisierung und Operationalisierung im standardisierten Zugang in diesem Kapitel) oder Erweiterung der Fragestellung für die Begründung des Globalurteils (siehe Abschnitt zu selbst herangezogenen Kriterien bzw. Begründung der Notenvergabe in diesem Kapitel) müsste daher mitunter die Anzahl der Vignetten reduziert werden.

Durch den Verzicht auf einen Wissenstest können die Befunde der Studie ausschließlich zur Beschreibung ‚guter‘ instruktionaler Erklärungen im Sinne Berliners (2005) herangezogen werden. Eine Überprüfung, inwiefern Erklärungen, die von den verschiedenen Statusgruppen als gut bewertet werden, auch effektiv sind, steht aus und ist im Zuge weiterer Forschung zu instruktionalen Erklärungen wünschenswert.

Was für die Frage nach gutem instruktionalen Erklären weniger große Einschränkungen mit sich bringt, könnte für die Frage nach effektivem instruktionalem Erklären eine größere Herausforderung werden: Da Erklärungen der Lehrkraft nicht isoliert im Unterricht stehen, sondern eingebettet sind, also beispielsweise weitere unterrichtliche Aktivitäten folgen, die die Verarbeitung des Gesagten stützen (Hattie, 2015; Leinhardt, 1990; Webb et al., 1995; Wörn, 2014), entsteht ein komplexes Gefüge, das zur Lernwirksamkeit von Erklärungen beitragen kann. Eine Herauslösung aus dem Unterrichtskontext und eine Verdichtung der Erklärsituation, wie sie in dieser Studie realisiert wurde, führt daher mitunter zu weniger validen Ergebnissen für die Überprüfung der Effektivität der Erklärungen.

In Bezug auf Reihenfolgeeffekte, also die Bewertung der Erklärungen in Abhängigkeit von ihrer Position im Fragebogen, zeigen sich kleine bis mittlere Effekte, die die Entscheidung für eine systematische Randomisierung der Videoreihenfolge im Fragebogen bestärken. Es kann davon ausgegangen werden, dass sich die Effekte durch die Bildung von Mittelwerten über beide Versionen ausmitteln. Bei weiteren Einsätzen des Fragebogens wäre dennoch zu überlegen, eine vollständige Randomisierung der Vignetten oder weitere Versionen des Fragebogens einzusetzen, da sich in der Antwortbereitschaft für das offene Textfeld (Begründung zur Vergabe des Globalurteils) eine abnehmende Tendenz zeigt.

In Hinblick auf die angewendeten statistischen Verfahren könnte statt der ANOVA mit gemischtem Design zur Analyse von Bewertungsunterschieden zwischen linearen und komplexen Erklärungen auch ein „linear mixed regression model“ (Hilbert, Stadler, Lindl, Naumann & Bühner, 2019, S. 101)

angewendet werden. In das Regressionsmodell wird die experimentelle Manipulation der Videovignetten (komplex vs. linear) als kategoriale Variable, als **sogenannte ‚Dummy-Codierung‘** aufgenommen („dummy coding“, Hilbert et al., 2019, S. 104). Erhalten bleibt die kategoriale Variable ‚Statusgruppe‘. Das ‚linear mixed regression model‘ hat laut Hilbert et al. (2019) gegenüber der ‚mixed ANOVA‘ unter anderem die Vorteile, dass die Effekte direkt anhand der Regressionskoeffizienten interpretiert werden können (die mixed ANOVA gibt als Omnibus-Test nur an, dass es signifikante Unterschiede gibt, durch weitere Tests können diese aufgeklärt werden, was jedoch zu Einbußen hinsichtlich der statistischen Power führt), eine höhere statistische Power erreicht wird und das Verfahren robuster gegenüber Voraussetzungsverletzungen ist bzw. weniger Modellvoraussetzungen besitzt.

Statusgruppen

Der Einbezug verschiedener Statusgruppen wurde – **anhand der Definition von ‚gut‘** im Sinne Berliners (2005) und einer multiperspektivischen Annäherung an einen Forschungsgegenstand für ein vertieftes Verständnis im Sinne Denzins (1979) und Flicks (2011) – theoretisch begründet. Durch die Reflexion der Datenauswertung ergibt sich auch ein forschungsmethodisches Argument für den Einbezug der verschiedenen Statusgruppen: Die Antwortverhalten und Bewertungsmuster erzeugen kontrastierbare Fälle, die nicht nur neue Fragestellungen aufwerfen, sondern auch den Blick auf Dinge lenken, die bei der Betrachtung nur einer Statusgruppe unauffällig erschienen wären. Komparative Ansätze, bei denen mindestens zwei systematisch unterscheidbare Untersuchungsgegenstände kontrastiert werden bzw. hinsichtlich deren Gemeinsamkeiten und Unterschiede analysiert werden, sind bereits eine etablierte Methode beispielsweise in der Kommunikationswissenschaft (Esser, 2012) oder Literaturwissenschaft (Tainio & Winkler, 2014).

Selbst herangezogene Kriterien bzw. Begründung der Notenvergabe

Die Bereitschaft zur Begründung der Notenvergabe war über alle Statusgruppen vorhanden und vor allem bei Didaktiker:innen und Schüler:innen sehr hoch, sodass insgesamt ausreichend Datenmaterial entstand, um Kriterien herauszuarbeiten.

Damit konnten – wie intendiert – der Kriterienkatalog erweitert bzw. validiert sowie Unterschiede zwischen den Perspektiven herausgearbeitet werden. Darüber hinaus war es auch möglich, Ansätze für Gründe für das Bewertungsverhalten bei linearen und komplexen Erklärungen zu finden. Um der Kürze einiger Antworten entgegenzuwirken und damit die Zuordnung zu Kriterien differenzierter sowie zuverlässiger und valider durchführen zu können, wäre zu eruieren, ob die Frage nach dem **Grund der Notenvergabe** („Warum hast du diese Note vergeben?“, „Warum haben Sie diese Note vergeben?“) um einen Bezugspunkt oder ein Beispiel in der Erklärung (z. B. „An was machst du das fest?“, „An was machen Sie das fest“) erweitert werden sollte. Der zusätzliche Mehraufwand könnte jedoch zu einer geringeren Beteiligung an der optionalen Angabe führen. Für eine Vertiefung einzelner Aspekte (z. B. Sicht auf die Erwähnung von Nichtwissen von der Statusgruppe der Lehrkräfte; Rolle der Visualisierungen für Schüler:innen bei komplexen Erklärungen) könnten deswegen auch leitfadengestützte Interviews mit einer Auswahl der eingesetzten Vignetten als Stimuli in Betracht gezogen werden. Insgesamt sind die Länge und Differenziertheit der Antworten jedoch zufriedenstellend und für den intendierten Zweck überwiegend gut geeignet. Ob das offene Textfeld so auch von jüngeren Schüler:innen genutzt wird, ist fraglich.

In Hinblick auf die Begründung der Notenvergabe ergeben sich zwei weitere, miteinander verschränkte Limitationen. Zum einen muss bei Aussagen, die diese Datengrundlage verwenden, stets im Blick behalten werden, dass die Äußerungen sowohl Merkmale der Personen als auch solche der Vignetten unter den Rahmenbedingungen der Erhebung widerspiegeln: In begrenzter Zeit äußern die Teilnehmenden eine Auswahl von Kriterien. Dabei können nur Kriterien genannt werden, die auch bekannt sind, sodass bestehendes Wissen die Wahrnehmung der Umsetzung von Erklärungen verändert und die Äußerung von Kriterien beeinflusst. Die Merkmale der Vignetten moderieren dabei die Wahrscheinlichkeit, dass ein Kriterium genannt wird, da auf besonders auffällige Merkmale (störend oder exzellent gelungen) auch häufiger referiert wird. Zum anderen ergeben sich damit einhergehend Einschränkungen in Hinblick auf die Triangulation der Begründungen zur Notenvergabe mit dem Antwortverhalten im standardisierten Zugang: Nur unter den oben genannten Einschränkungen sind valide Aussagen über die Relevanz der Kriterien für die Güte einer Erklärung zu treffen. Die Methodentriangulation kann daher als wechselseitige Ergänzung realisiert werden (Flick, 2011), wobei sowohl individuelle Sichtweisen von Subjekten als auch die Struktur des Phänomens in den Blick genommen werden (Fielding & Fielding, 1986). Die verschiedenen Datenquellen (geschlossene Items und optionale Begründung im offenen Textfeld) können hingegen nicht unmittelbar zur Beantwortung derselben Forschungsfrage (z. B. im Sinne Denzins, 1970) genutzt werden. Für den erstmaligen Versuch einer biologiespezifischen Konzeptualisierung und die Forschungsfragen mit explorativem Charakter ist das offene Textfeld zur Explikation selbstgewählter Kriterien als unverzichtbar einzuschätzen und bringt einen deutlichen Mehrwert zur Beschreibung guten instruktionalen Erklärens im Biologieunterricht.

Aufgrund der geringen Anzahl der Nennung je Kategorie – insbesondere bei den Daten der evaluativen qualitativen Inhaltsanalyse – sind die gezogenen Schlussfolgerungen als Tendenzen zu verstehen, die es in weiteren Forschungsarbeiten zu vertiefen gilt.

Konzeptualisierung und Operationalisierung guten Erklärens im standardisierten Forschungszugang

Auch hinsichtlich der Konzeptualisierung und Operationalisierung des guten Erklärens im standardisierten Zugang ergibt sich eine Einschränkung: Für die Konzeptualisierung und Operationalisierung wurden die vier fachübergreifenden Aspekte (Adressatenorientierung, Strukturiertheit, Sprech- und Körperausdruck, verständliche Sprache) sowie der fachspezifische Aspekt (Wahrnehmung des dargestellten Komplexitätsgrades) gewählt. Sie lassen sich auch anhand der selbst gewählten Kriterien der Teilnehmenden rekonstruieren (siehe Kap. 6.2.1.1). Im geschätzten hierarchischen linearen Regressionsmodell können durch die Aspekte bzw. Faktoren über alle Statusgruppen hinweg 20 % der Varianz und bei Lehrenden sogar 29 % der Varianz erklärt werden (siehe Kap. 6.2.3.1 bzw. 6.2.3.2). Jedoch wurde die Nutzung externer Repräsentationen nicht mit aufgenommen, da für den Aspekt aufgrund der hohen Fachspezifität keine fachübergreifende Konzeptualisierung und Operationalisierung möglich waren – unter anderem, weil in unterschiedlichen Fächern stark unterschiedliche Repräsentationen gebraucht werden und die Formulierung konkreter Items damit erschwert ist (siehe Kap. 5.2.4). Jedoch ist davon auszugehen, dass die Nutzung von externen Repräsentationen ein Teil effektiver instruktionaler Erklärungen ist (Sevian & Gonsalves, 2008; Wragg, 1993), sie zum Aufbau interner Repräsentationen und damit zum Verstehen des Erklärgegenstandes beiträgt (Schnotz & Bannert, 2003) sowie die Strukturierung einer Erklärung unterstützen kann (Hargie, 2013). Auch in Begründungen zur Vergabe des Globalurteils wird der Aspekt der Repräsentationen – insbesondere von Schüler:innen

– häufig angesprochen (siehe Kap. 6.2.1.1 bzw. 6.2.1.3). Eine Erweiterung um den Aspekt scheint daher für eine Weiterentwicklung des Modells sinnvoll.

In Hinblick auf die Konzeptualisierung und Operationalisierung der Aspekte guten Erklärens ist insbesondere die sprachliche Verständlichkeit zu prüfen. In der ursprünglichen Operationalisierung (Kriterien auf Wort, Satz und Textebene) zeigten sich negative Trennschärfen für das Item, das auf die Verständlichkeit der Sprache auf Textebene referiert (Item: *Sw_ver*; „*Manchmal hat die Lehrerin extra gesagt, dass etwas sehr wichtig ist*“). Um alle Ebenen zu repräsentieren, müsste der Aspekt erneut operationalisiert werden.

Um weitere Erkenntnisse über die faktorielle Struktur bzw. das Messmodell ‚gutes instruktionales Erklären im Biologieunterricht‘ zu erhalten, können weiterführend auf Basis der Daten konfirmatorische Faktorenanalysen durchgeführt werden.

Komplexe und lineare instruktionale Erklärungen

Mit der experimentellen Manipulation der Videovignetten wurde ein zentraler Unterschied hinsichtlich verschiedener Formen von Erklärungen biologischer Phänomene aufgegriffen. Dabei konnte gezeigt werden, dass die Formen von verschiedenen Statusgruppen unterschiedlich wahrgenommen werden. Die Daten zeigen weiterhin eindrücklich, dass nicht ausschließlich die operationalisierten Merkmale ‚Anzahl der Faktoren‘ bzw. ‚Multikausalität‘ sowie ‚Erwähnung von Nichtwissen und unsicherem Wissen‘ ausschlaggebend für die Bewertung sind, sondern weitere Faktoren zur Bewertung der Erklärungen führen. Mithilfe der Begründungen zur Notenvergabe im offenen Textfeld konnten Hinweise auf die unterschiedlichen Bewertungen gefunden werden. Unter Einsatz von leitfadengestützten Interviews müsste diesen Ansatzpunkten weiter nachgegangen werden, um die Unterschiede zwischen komplexen und linearen Erklärungen und insbesondere den Herausforderungen bei komplexen Erklärungen im Umgang mit Nichtwissen und unsicherem Wissen differenzierter herauszuarbeiten. Dabei scheint neben den eingesetzten Repräsentationen insbesondere auch das Wissen über Nichtwissen als wissenschaftspropädeutischer Inhalt ein relevanter Ansatzpunkt für die Unterschiede in den Bewertungen zu sein. Die Ursachen für entstehende Vagheit, die negativ bewertet wird, sollten ebenfalls in den Blick genommen werden.

7.3 Fazit und Implikationen

Aus den theoretischen und empirischen Befunden der Studie lassen sich sowohl für Forschungsvorhaben als auch für die schulische Praxis sowie die Aus- und Weiterbildung von Lehrkräften Implikationen ableiten, die über den Themenbereich der guten instruktionalen Erklärungen im Biologieunterricht hinausgehen.

Für Forschungsvorhaben

Neben den bereits in Kapitel 7.2 aufgegriffenen Forschungsdesiderata ergeben sich Implikationen zur Durchführung von Forschungsvorhaben:

Da sich die maßgeblichen Faktoren der Güte instruktionaler Erklärungen zum Teil auch in der Forschung zur Unterrichtsqualität zeigen (Kleickmann, Steffensky & Paetorius, 2020), scheinen das

Zusammendenken und Rückbezüge auf jeweilige empirische Befunde sinnvoll, insbesondere dann, wenn bei der Effektivität von Erklärungen – wie von Webb et al. (1995) angeführt – auch die nachfolgenden Aktivitäten in die Bewertung mit einfließen.

Die Perspektive der Schüler:innen ist für die Unterrichtsforschung eine zentrale (z. B. Clausen, 2002; Helmke, 2012; Seidel, 2014), wobei die valide Messung von Unterrichtsqualität durch Schüler:innen mit standardisierten Instrumenten mit Problemen behaftet ist (z. B. Clausen, 2002; Fauth et al., 2020; Kleickmann et al., 2019; Kunter & Baumert, 2007). Auch in den Daten des standardisierten Zugangs dieser Studie zeigen sich wenig fallspezifische Bewertungen durch Schüler:innen, sondern eher einheitliche Bewertungen über Videovignetten hinweg. In den offenen Textfeldern sind die Begründungen jedoch abhängig vom Aspekt als differenziert und insgesamt als informativ bewerten (ein ähnlicher Befund findet sich bei Clausen (2002) und Gautschi (2011)). Vor diesem Hintergrund lohnt es sich, einen Einbezug offener Formate vermehrt in den Blick zu nehmen.

Die Berücksichtigung der Statusgruppe der Studierenden ist aufgrund der wechselnden Ähnlichkeiten mit dem Antworttendenzen von Schüler:innen und vor allem Lehrkräften und Didaktiker:innen in Hinblick auf die Fragestellung jeweils zu reflektieren.

Für die Praxis an Schulen

Für die Praxis an Schulen sind insbesondere die Ergebnisse des Vergleichs von komplexen und linearen instruktionalen Erklärungen interessant. Die Ergebnisse sprechen dafür, dass Schüler:innen ein Umgang mit Nichtwissen und unsicherem Wissen zugetraut werden kann. Von ihren Lehrkräften sollten Nichtwissen und unsicheres Wissen als ‚normale‘ Bestandteile der biologischen Phänomene bzw. des Wissens über diese Phänomene in den Unterricht implementiert werden. Der routinierte Umgang mit Nichtwissen ist eine zentrale Voraussetzung für eine Teilhabe der Schüler:innen an aktuellen Diskussionen (Dittmer & Ehras, 2024).

Für die Aus- und Weiterbildung von Lehrkräften

Für die Aus- und Weiterbildung von Lehrkräften können drei Implikationen bzw. Forderungen abgeleitet werden: Erstens sollten Erklärungen vermehrt in den Fokus der universitären Ausbildung von angehenden Lehrkräften gerückt werden. Da sich andeutet, dass maßgebliche Faktoren der Güte einer Erklärung auch in Bezug auf die Unterrichtsqualität insgesamt als maßgebliche Faktoren der kognitiven Unterstützung angenommen werden können (Kleickmann et al., 2020), bieten sich instruktionale Erklärungen als Unterrichtsminiaturen für Studierende an – damit kann die inhaltliche Begründung anhand der Relevanz der Sprachhandlung des Erklärens wie beispielsweise im Core-Practice-Ansatz (Asen-Molz et al., 2022; McDonald et al., 2013) durch eine methodisch-didaktische Begründung erweitert werden.

Zweitens bietet die Integration des Erklärens in die Curricula zur Aus- und Weiterbildung von Lehrkräften einen weiteren Vorteil: Das Erklären im Unterricht ist eng verschränkt mit den Erkenntnismethoden und Erklärungsformen der wissenschaftlichen Disziplin. Anhand instruktionaler Erklärungen im Biologieunterricht lassen sich deshalb auch wissenschaftstheoretische bzw. erkenntnistheoretische Inhalte (z. B. Methoden der Erkenntnisgewinnung, die Rolle von Gesetzen, Schwierigkeiten bei Vorhersagen etc.) exemplarisch und durch den Erklärgegenstand an konkrete Beispiele gebunden erarbeiten. Insbesondere durch den Vergleich und die Diskussion mit anderen

Fächern treten die Charakteristika anhand der Beispiele deutlich hervor und erleichtern damit mitunter den Zugang zur Wissenschaftstheorie (Asen-Molz et al., im Druck).

Drittens sollte Wissen über Nichtwissen und unsicheres Wissen erworben werden. Das Wissen bezieht sich dabei auf zwei Dimensionen: Nichtwissen und unsicheres Wissen sollten als Teile des Wissens über das Phänomen, also den Fachgegenstand, mit erworben werden. Diese Forderung adressiert die Fachwissenschaft, aber vor allem auch die Biologiedidaktik, die im Zuge der didaktischen Reduktion bzw. der didaktischen Rekonstruktion eine solche Bezugnahme auf Wissenslücken und die Berücksichtigung der Grenzen des Wissens einfordern kann. Die zweite Dimension betrifft das Wissen über die Kommunikation und die divergierenden Interpretationen von Nichtwissen und unsicherem Wissen. Für eine politische Teilhabe sollten Lehrkräfte Wissen darüber besitzen und schließlich an ihre Schüler:innen vermitteln, wie in öffentlichen Diskussionen Wissenslücken kommuniziert und zum Teil instrumentalisiert werden, wie unterschiedliche Meinungen und Positionen aufgrund von Nichtwissen zustande kommen. Die unterschiedlichen Einschätzungen von Nichtwissen und unsicherem Wissen zu erkennen und ihren Effekt in der Kommunikation abschätzen zu können, ist ein wesentlicher Teil der eigenen Meinungsbildung und damit der gesellschaftlichen Teilhabe.

Abbildungsverzeichnis

Abbildung 1: Zentrale Begriffe und unterscheidbare Ebenen im Kontext von verschiedenen Wissensstadien.....	49
Abbildung 2: Faktoren zur Gestaltung einer (guten) instruktionalen Erklärung aus Sicht der Lehrkraft	99
Abbildung 3: Aufbau und Ablauf des eingesetzten Fragebogens.....	108
Abbildung 4: Standbild aus dem Video zur Entstehung von Vitamin-A-Mangel als Beispiel für das Setting aller Aufnahmen	114
Abbildung 5: Durchlaufene Schritte der Videovignettenentwicklung in Anlehnung an Piwowar et al. (2018) und Dieker et al. (2009).....	116
Abbildung 6: Screenshot aus dem Video zur Entstehung von Mukoviszidose (Minute 2:20).....	118
Abbildung 7: Screenshot aus dem Video zur Entstehung von Multipler Sklerose (Minute 2:36) ..	120
Abbildung 8: Screenshot aus dem Video zur Entstehung von Populationsschwankungen bei Schneeschuhhasen (Minute 3:06)	123
Abbildung 9: Screenshot aus dem Video zur Entstehung von Populationsschwankungen bei Erdkröten (Minute 2:43)	126
Abbildung 10: Screenshot aus dem Video zur Entstehung eines Vitamin-A-Mangels (Minute 2:55)	129
Abbildung 11: Screenshot aus dem Video zur Entstehung eines Folsäuremangels (Minute 2:51) ..	132
Abbildung 12: Systematisch randomisierte Videoreihenfolge in den zwei Versionen des Fragebogens.....	148
Abbildung 13: Ablaufschema einer inhaltlich-strukturierenden qualitativen Inhaltsanalyse nach Kuckartz (2018, S. 100); in grüner Umrandung sind die Modifikationen in Hinblick auf das vorliegende Dantematerial dargestellt.....	161
Abbildung 14: Ergebnis der Kategorienbildung bzw. angewandtes Kategoriensystem.....	168
Abbildung 15: Grafik zur Reihenfolge der Videovignetten in den zwei Versionen des Fragebogens	169
Abbildung 16: Relative Häufigkeiten der Nennungen aller Subkategorien über alle Statusgruppen ($N_{\max} = 238$) und Videovignetten ($n = 6$) hinweg	184
Abbildung 17: Relative Häufigkeiten der Nennungen der Subkategorien aus der Hauptkategorie ‚Adressatenorientierung‘ getrennt nach Statusgruppen (Schüler:innen $N_{\max}=122$; Studierende $N_{\max}=51$; Lehrkräfte $N_{\max}=34$; Didaktiker:innen $N_{\max}= 30$) über alle Videovignetten hinweg ...	208
Abbildung 18: Relative Häufigkeiten der Nennungen der Subkategorien aus der Hauptkategorie ‚Strukturiertheit‘ getrennt nach Statusgruppe (Schüler:innen $N_{\max} = 122$; Studierende $N_{\max} = 51$; Lehrkräfte $N_{\max} = 34$; Didaktiker:innen $N_{\max} = 30$) über alle Videovignetten hinweg.....	212
Abbildung 19: Relative Häufigkeiten der Nennungen der Subkategorien aus der Hauptkategorie ‚Repräsentationen‘ getrennt nach Statusgruppe (Schüler:innen $N_{\max} = 122$; Studierende $N_{\max} = 51$; Lehrkräfte $N_{\max} = 34$; Didaktiker:innen $N_{\max} = 30$) über alle Videovignetten hinweg	214
Abbildung 20: Relative Häufigkeiten der Nennungen der Subkategorien aus der Hauptkategorie ‚Sprech- und Körperausdruck‘ getrennt nach Statusgruppe (Schüler:innen $N_{\max} = 122$; Studierende $N_{\max} = 51$; Lehrkräfte $N_{\max} = 34$; Didaktiker:innen $N_{\max} = 30$) über alle Videovignetten hinweg.....	218
Abbildung 21: Relative Häufigkeiten der Nennungen der Subkategorien aus der Hauptkategorie ‚Verständliche Sprache‘ getrennt nach Statusgruppe (Schüler:innen $N_{\max} = 122$; Studierende $N_{\max} = 51$; Lehrkräfte $N_{\max} = 34$; Didaktiker:innen $N_{\max} = 30$) über alle Videovignetten hinweg	220

Abbildung 22: Relative Häufigkeiten der Nennungen der Subkategorien aus der Hauptkategorie ,Fachlicher Aspekte' getrennt nach Statusgruppe (Schüler:innen $N_{\max} = 122$; Studierende $N_{\max} = 51$; Lehrkräfte $N_{\max} = 34$; Didaktiker:innen $N_{\max} = 30$) über alle Videovignetten hinweg	223
Abbildung 23: Relative Häufigkeiten der Nennungen der Subkategorien aus der Hauptkategorie ,Globale Aspekte' getrennt nach Statusgruppe (Schüler:innen $N_{\max} = 122$; Studierende $N_{\max} = 51$; Lehrkräfte $N_{\max} = 34$; Didaktiker:innen $N_{\max} = 30$) über alle Videovignetten hinweg.....	226
Abbildung 24: Relative Häufigkeiten der Nennungen in den Kategorien ,Lehrperson', ,Merkbarkeit und Lernbarkeit' sowie ,Sonstiges' getrennt nach Statusgruppe (Schüler:innen $N_{\max} = 122$; Studierende $N_{\max} = 51$; Lehrkräfte $N_{\max} = 34$; Didaktiker:innen $N_{\max} = 30$) über alle Videovignetten hinweg.....	227
Abbildung 25: Profilplot zur deskriptiven Darstellung des Interaktionseffekts (Abhängigkeit des Faktors ,Komplexitätsgrad' vom Faktor ,Statusgruppe') für den Themenbereich Genetik mit dem Faktor ,Komplexitätsgrad' als trace-Faktor und dem Faktor ,Statusgruppe' auf der x-Achse inklusive des Fehlerbalkens für das 95%-Konfidenzintervall.....	263
Abbildung 26: Profilplot zur deskriptiven Darstellung des Interaktionseffekts (Abhängigkeit des Faktors ,Komplexitätsgrad' vom Faktor ,Statusgruppe') für den Themenbereich Ökologie mit dem Faktor ,Komplexitätsgrad' als trace-Faktor und dem Faktor ,Statusgruppe' auf der x-Achse inklusive des Fehlerbalkens für das 95%-Konfidenzintervall.....	264
Abbildung 27: Profilplot zur deskriptiven Darstellung des Interaktionseffekts (Abhängigkeit des Faktors ,Komplexitätsgrad' vom Faktor ,Statusgruppe') für den Themenbereich Physiologie mit dem Faktor ,Komplexitätsgrad' als trace-Faktor und dem Faktor ,Statusgruppe' auf der x-Achse inklusive des Fehlerbalkens für das 95%-Konfidenzintervall.....	266

Tabellenverzeichnis

Tabelle 1: Relevante Unterschiede der drei Disziplinen Sprachwissenschaft, Didaktik und Wissenschaftstheorie der Biologie in Hinblick auf (instruktionale) Erklärungen.....	19
Tabelle 2: Übersicht zu den differenzierbaren Erklärungsformen im Unterricht und verallgemeinert exemplarische Fragestellungen zu den jeweiligen Erklärungsformen.....	39
Tabelle 3: Formen externer Repräsentationen sowie Beispiele für die jeweiligen Formen.....	78
Tabelle 4: Übersicht zu den Kriterien guter instruktionaler Erklärungen.....	96
Tabelle 5: Übersicht zu den Themen der Videovignetten.....	111
Tabelle 6: Itemformulierungen zum Aspekt ,Adressatenorientierung'.....	142
Tabelle 7: Itemformulierungen zum Aspekt ,Strukturiertheit'.....	143
Tabelle 8: Itemformulierungen für den Aspekt ,Sprachliche Verständlichkeit'.....	144
Tabelle 9: Itemformulierungen für den Aspekt ,Sprech- und Körperausdruck'.....	145
Tabelle 10: Itemformulierungen zum Aspekt der Wahrnehmung des dargestellten <i>Komplexitätsgrades</i>	147
Tabelle 11: Überblick über die Zusammensetzung der Gesamtstichprobe.....	154
Tabelle 12: Auszug aus dem Codebuch zur exemplarischen Kategoriendefinition ,Didaktische Reduktion'.....	163
Tabelle 13: Ergebnisse der Interkorrelationen der Items zur wahrgenommenen Begeisterung, Sympathie und Natürlichkeit der erklärenden Person bei Schüler:innen ($N_{\max} = 70$) und Lehrenden (SLD; $N_{\max} = 60$) sowie zugehörige Mittelwerte und Standardabweichungen.....	171
Tabelle 14: Ergebnisse des selbsteingeschätzten Vorwissens der Schüler:innen pro Video.....	172

Tabelle 15: Relative und absolute Häufigkeiten der Äußerungen innerhalb der Hauptkategorien von Studierenden mit und ohne Seminarbesuch über alle Videovignetten	173
Tabelle 16: Interne Konsistenzen der sechs Skalen für Schüler:innen – Cronbachs Alpha und mittlere part-whole korrigierte Itemtrennschärfen inkl. Standardabweichungen	175
Tabelle 17: Interne Konsistenzen der sechs Skalen für Studierende – Cronbachs Alpha und mittlere part-whole korrigierte Itemtrennschärfen inkl. Standardabweichungen	176
Tabelle 18: Interne Konsistenzen der sechs Skalen für Lehrkräfte – Cronbachs Alpha und mittlere part-whole korrigierte Itemtrennschärfen inkl. Standardabweichungen	176
Tabelle 19: Interne Konsistenzen der sechs Skalen für Didaktiker:innen – Cronbachs Alpha und mittlere part-whole korrigierte Itemtrennschärfen inkl. Standardabweichungen	177
Tabelle 20: Korrelationskoeffizienten r nach Pearson für die vier skalenbezogenen Globalurteile der Aspekte und den implementierten Items der Aspekte (ohne die aufgrund negativer Trennschärfen ausgeschlossenen Items) je Videovignette getrennt nach den Statusgruppen der Schüler:innen ($N_{\max} = 122$) und der Lehrenden (Studierende, Lehrkräfte, Fachdidaktiker:innen ($N_{\max} = 97$))	179
Tabelle 21: Überblick zur Struktur des Ergebnisteils anhand der Kapitelnummerierung, der dort adressierten Forschungsfragen und der dafür verwendeten Daten bzw. Auswertungsmethoden	182
Tabelle 22: Übersicht zum Aspekt ‚Adressatenorientierung‘ mit den aus der Literatur rekonstruierten Facetten und Kriterien, den Items des standardisierten Zugangs und den entsprechenden Subkategorien aus dem Kategoriensystem	187
Tabelle 23: Übersicht zum Aspekt ‚Strukturiertheit‘ mit den aus der Literatur rekonstruierten Facetten und Kriterien, den Items des standardisierten Zugangs und den entsprechenden Subkategorien aus dem Kategoriensystem	190
Tabelle 24: Übersicht zum Aspekt ‚Repräsentationen‘ mit den aus der Literatur rekonstruierten Facetten und Kriterien sowie den entsprechenden Subkategorien aus dem Kategoriensystem	192
Tabelle 25: Übersicht zum Aspekt ‚Sprech- und Körperausdruck‘ mit den aus der Literatur rekonstruierten Facetten und Kriterien, den Items des standardisierten Zugangs und den entsprechenden Subkategorien aus dem Kategoriensystem	196
Tabelle 26: Übersicht zum Aspekt ‚Verständliche Sprache‘ mit den aus der Literatur rekonstruierten Facetten und Kriterien, den Items des standardisierten Zugangs und den entsprechenden Subkategorien aus dem Kategoriensystem	198
Tabelle 27: Übersicht zum fachlichen Aspekt mit den aus der Literatur rekonstruierten Facetten und Kriterien, den Items des standardisierten Zugangs und den entsprechenden Subkategorien aus dem Kategoriensystem	200
Tabelle 28: Übersicht zu den statusgruppenspezifischen (Spalte 2-5) und statusgruppenübergreifenden (Spalte 6) relativen Häufigkeiten der angesprochenen Hauptkategorie bzw. der angesprochenen Aspekte guten Erklärens über alle Videovignetten hinweg.....	206
Tabelle 29: Häufigkeiten der Äußerungen, die der Hauptkategorie „Adressatenorientierung“ zugeordnet werden können getrennt nach Statusgruppe über alle Videovignetten hinweg ..	207
Tabelle 30: Absolute und relative Häufigkeiten der Äußerungen, die den verschiedenen Facetten des Aspekts der ‚Adressatenorientierung‘ zugeordnet werden können, getrennt nach Statusgruppe über alle Videovignetten hinweg	210
Tabelle 31: Häufigkeiten der Äußerungen, die der Hauptkategorie ‚Strukturiertheit‘ zugeordnet werden können, getrennt nach Statusgruppe über alle Videovignetten hinweg.....	211

Tabelle 32: Absolute und relative Häufigkeiten der Äußerungen, die den verschiedenen Facetten des Aspekts der ‚Strukturiertheit‘ zugeordnet werden können, getrennt nach Statusgruppe über alle Videovignetten hinweg.....	213
Tabelle 33: Häufigkeiten der Äußerungen, die der Hauptkategorie ‚Repräsentationen‘ zugeordnet werden können, getrennt nach Statusgruppe über alle Videovignetten hinweg.....	214
Tabelle 34: Absolute und relative Häufigkeiten der Äußerungen zum Aspekt der ‚Repräsentationen‘ getrennt nach Verweisen auf das Vorhandensein von Repräsentationen, oberflächliche Gestaltungsmerkmale und lernrelevante Gestaltungsmerkmale getrennt nach Statusgruppe über alle Videovignetten hinweg	216
Tabelle 35: Häufigkeiten der Äußerungen, die der Hauptkategorie ‚Sprech- und Körperausdruck‘ zugeordnet werden können, getrennt nach Statusgruppe über alle Videovignetten hinweg .	217
Tabelle 36: Absolute und relative Häufigkeiten der Äußerungen, die den verschiedenen Facetten des Aspekts der ‚Sprech- und Körperausdruck‘ zugeordnet werden können, getrennt nach Statusgruppe über alle Videovignetten hinweg	219
Tabelle 37: Häufigkeiten der Äußerungen, die der Hauptkategorie ‚Verständliche Sprache‘ zugeordnet werden können, getrennt nach Statusgruppe über alle Videovignetten hinweg .	220
Tabelle 38: Absolute und relative Häufigkeiten der Äußerungen, die den verschiedenen Facetten des Aspekts ‚Verständliche Sprache‘ zugeordnet werden können, getrennt nach Statusgruppe über alle Videovignetten hinweg.....	221
Tabelle 39: Häufigkeiten der Äußerungen, die der Hauptkategorie ‚Fachlicher Aspekte‘ zugeordnet werden können, getrennt nach Statusgruppe über alle Videovignetten hinweg.....	222
Tabelle 40: Absolute und relative Häufigkeiten der Äußerungen, die den verschiedenen Facetten des fachlichen Aspekts zugeordnet werden können, getrennt nach Statusgruppe über alle Videovignetten hinweg.....	224
Tabelle 41: Häufigkeiten der Äußerungen, die der Hauptkategorie ‚Globale Aspekte‘ zugeordnet werden können, getrennt nach Statusgruppe über alle Videovignetten hinweg.....	225
Tabelle 42: Tabellarisch dargestelltes visuelles Dokumentenvergleichsdiagramm für eine Zufallsstichprobe von N = 16 Schüler:innen pro Videovignette	228
Tabelle 43: Tabellarisch dargestelltes visuelles Dokumentenvergleichsdiagramm für eine Zufallsstichprobe von N = 16 Studierenden pro Videovignette.....	229
Tabelle 44: Tabellarisch dargestelltes visuelles Dokumentenvergleichsdiagramm für eine Zufallsstichprobe von N = 16 Lehrkräften pro Videovignette	231
Tabelle 45: Tabellarisch dargestelltes visuelles Dokumentenvergleichsdiagramm für eine Zufallsstichprobe von N = 16 Didaktiker:innen pro Videovignette.....	232
Tabelle 46: Nach Statusgruppen und Videovignetten differenzierte arithmetische Mittel und Standardabweichungen der Globalurteile und Ergebnisse der einfaktoriellen ANOVA mit dem Faktor Statusgruppenzugehörigkeit	239
Tabelle 47: Nach Statusgruppen und Videovignetten differenzierte arithmetische Mittel und Standardabweichungen des Aspekts ‚Adressatenorientierung‘ und Ergebnisse der einfaktoriellen ANOVA mit dem Faktor Statusgruppenzugehörigkeit	241
Tabelle 48: Nach Statusgruppen und Videovignetten differenzierte arithmetische Mittel und Standardabweichungen des Aspekts ‚Strukturiertheit‘ und Ergebnisse der einfaktoriellen ANOVA mit dem Faktor Statusgruppenzugehörigkeit.....	242
Tabelle 49: Nach Statusgruppen und Videovignetten differenzierte arithmetische Mittel und Standardabweichungen des Aspekts ‚Sprech- und Körperausdruck‘ und Ergebnisse der einfaktoriellen ANOVA mit dem Faktor Statusgruppenzugehörigkeit	246

Tabelle 50: Nach Statusgruppen und Videovignetten differenzierte arithmetische Mittel und Standardabweichungen des Aspekts ‚Verständliche Sprache‘ und Ergebnisse der einfaktoriellen ANOVA mit dem Faktor Statusgruppenzugehörigkeit	248
Tabelle 51: Nach Statusgruppen und Videovignetten differenzierte arithmetische Mittel und Standardabweichungen des fachspezifischen Aspekts und Ergebnisse der einfaktoriellen ANOVA mit dem Faktor Statusgruppenzugehörigkeit	249
Tabelle 52: Verteilung der besten und schlechtesten Bewertungen für das Globalurteil und die fünf Aspekte guten Erklärens pro Statusgruppe und Videovignette.....	251
Tabelle 53: Übersicht zu den signifikanten Bewertungsunterschieden getrennt nach Aspekten und Videovignetten	253
Tabelle 54: Interskalenkorrelationen und deren Zusammenhänge mit dem Globalurteil über alle Videos und alle Statusgruppen (Pearson-Produkt-Moment-Korrelation)	255
Tabelle 55: Hierarchische lineare Regressionsmodelle mit der abhängigen Variable Globalurteil und unter Berücksichtigung der nach Teilnehmenden und den sechs Videovignetten geordneten Datenstruktur	256
Tabelle 56: Interskalenkorrelationen und deren Zusammenhänge mit dem Globalurteil über alle Videos (Pearson Produkt-Moment-Korrelation) nach SuS- und SLD-Gruppe getrennt	257
Tabelle 57: Hierarchische lineare Regressionsmodelle mit der abhängigen Variable Globalurteil und unter Berücksichtigung der nach Teilnehmenden und den sechs Videovignetten geordneten Datenstruktur	258
Tabelle 58: Nach Statusgruppen und Erklärthema sowie Komplexitätsgrad differenzierte mittlere Bewertungen (inkl. Standardabweichungen) und Varianzanalyse mit gemischtem Design.....	262
Tabelle 59: Effektstärken zu den Mittelwertsunterschieden (t-Tests) im Themenbereich Genetik zwischen linearen und komplexen Erklärungen getrennt nach Statusgruppen	263
Tabelle 60: Effektstärken zu den Mittelwertsunterschieden (t-Tests) im Themenbereich Ökologie zwischen linearen und komplexen Erklärungen getrennt nach Statusgruppen	265
Tabelle 61: Effektstärken zu den Mittelwertsunterschieden (t-Tests) im Themenbereich Physiologie zwischen linearen und komplexen Erklärungen getrennt nach Statusgruppen	266
<i>Tabelle 62: Übersicht zu Mittelwerten und Standardabweichungen der einzelnen Items sowie Itemtrennschärfen der in der Skala berücksichtigten Items und Cronbachs Alpha als Maß für die interne Konsistenz der Skalen für die Statusgruppe der Lehrenden bzw. der Metaperspektive (Studierende, Lehrkräfte, Didaktiker:innen; N_{max} = 97)</i>	<i>357</i>

Literaturverzeichnis

- Abrams, E., Southerland, S. & Cummins, C. (2001). The how's and why's of biological change: How learners neglect physical mechanisms in their search for meaning. *International Journal of Science Education*, 23(12), 1271–1281.
- Allhoff, D.-W. & Allhoff, W. (2010). *Rhetorik & Kommunikation. Ein Lehr- und Übungsbuch* (15. aktualisierte Auflage). München: Ernst Reinhardt Verlag.
- Anacker, M. (2007). Das Erkenntnisproblem und der Wissensbegriff in der philosophischen Tradition. In R. Schützeichel (Hrsg.), *Handbuch Wissenssoziologie und Wissensforschung* (Erfahrung – Wissen – Imagination, Bd. 15, S. 353–374). Konstanz: UVK-Verlagsgesellschaft mbH.
- Apel, H. (2009). Behalten und Verstehen von Hörfunknachrichten: medientheoretische Hintergründe und empirische Belege zum Einfluss der Prosodie. In L. C. Anders & I. Bose (Hrsg.), *Aktuelle Forschungsthemen der Sprechwissenschaft 1. Sprach-, Sprech- und*

- Stimmstörungen/Sprache und Sprechen von Hörfunknachrichten* (Bd. 30, S. 89–127). Frankfurt am Main: Peter Lang.
- Asen-Molz, K., Ehras, C. & Krauss, S. (im Druck). Im interdisziplinären Tandem kriterienbasiert Erklären lehren und lernen. In A. Schilcher, S. Krauss, A. Lindl & S. Hilbert (Hrsg.), *Fachspezifische Lehrerkompetenzen im Erklären*. Weinheim: Beltz Juventa.
- Asen-Molz, K., Knott, C. & Schilcher, A. (2022). Erklären als Core Practice. Über die Förderung von Erklärkompetenz angehender Lehrkräfte. *Journal für LehrerInnenbildung*, 3(22), 30–43.
- Aubusson, P. J., Harrison, A. G. & Ritchie, S. M. (2006). Metaphor and Analogy: Serious thought in science education. In P. J. Aubusson, A. G. Harrison & S. M. Ritchie (Hrsg.), *Metaphor and Analogy in Science Education* (Bd. 30, S. 1–10). Dordrecht: Springer.
- Babad, E. (2009). Teaching and Nonverbal Behavior in the Classroom. In L. J. Saha & A. G. Dworkin (Hrsg.), *International Handbook of Research on Teachers and Teaching* (Springer international handbooks of education, Bd. 21, S. 817–828). Boston, MA: Springer US.
- Babad, E., Avni-Babad, D. & Rosenthal, R. [Robert]. (2003). Teachers' brief nonverbal behaviors in defined instructional situations can predict students' evaluations. *Journal of Educational Psychology*, 95(3), 553–562.
- Baddeley, A. D. & Graham Hitch (1974). Working Memory. In G. H. Bower (Hrsg.), *The psychology of learning and motivation* (Vol. 8, S. 47–89). New York: Academic Press.
- Bartelborth, T. (2007). *Erklären*. Berlin: Walter de Gruyter.
- Bar-Yam, Y. (1997). *Dynamics of complex systems* (Studies in nonlinearity). Reading, MA: Addison-Wesley.
- Baumert, J. (2002). Deutschland im internationalen Bildungsvergleich. In N. Killius, J. Kluge & L. Reisch (Hrsg.), *Die Zukunft der Bildung* (Edition Suhrkamp, Bd. 2289, S. 100–150). Frankfurt am Main: Suhrkamp.
- Baumert, J. & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 9(4), 469–520.
- Baumert, J. & Kunter, M. (2011). Das Kompetenzmodell von COACTIV. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (S. 29–53). Münster: Waxmann.
- Bay, A. W., Thiede, B. & Wirtz, M. A. (2016). Die Theorie der kognitiven Belastung (Cognitive Load Theory). In P. Gretsch & L. Holzäpfel (Hrsg.), *Lernen mit Visualisierungen. Erkenntnisse aus der Forschung und deren Implikationen für die Fachdidaktik* (1. Auflage, neue Ausgabe, S. 123–137). Münster: Waxmann.
- Bechtel, W. (2011). Mechanism and Biological Explanation. *Philosophy of Science*, 78(4), 533–557.
- Bechtel, W. (2013). Understanding Biological Mechanisms: Using Illustrations from Circadian Rhythm Research. In K. Kampourakis (Hrsg.), *The philosophy of biology. A companion for educators* (History, Philosophy and Theory of the Life Sciences, S. 487–510). Dordrecht: Springer.
- Becker-Mrotzek, M. (2004). *Schreibentwicklung und Textproduktion. Der Erwerb der Schreibtätigkeit am Beispiel der Bedienungsanleitung*. Radolfzell: Verlag für Gesprächsforschung.
- Begon, M., Howarth, R. W. & Townsend, C. R. (2017). *Ökologie*. Berlin Heidelberg: Springer Spektrum.
- Berliner, D. C. (2005). The Near Impossibility of Testing for Teacher Quality. *Journal of Teacher Education*, 56(3), 205–213.
- Bleck, V. (2019). *Lehrerenthusiasmus. Entwicklung, Determinanten, Wirkungen*. Wiesbaden: Springer.
- Blömeke, S., Gustafsson, J.-E. & Shavelson, R. J. (2015). Beyond dichotomies: Competence viewed as a continuum. *Zeitschrift für Psychologie*, 233(1), 3–13.

- Bögeholz, S. & Barkmann, J. (2005). Rational choice and beyond: Handlungsorientierende Kompetenz für den Umgang mit faktischer und ethischer Komplexität. In R. Klee, A. Sandmann & H. Vogt (Hrsg.), *Lehr- und Lernforschung in der Biologiedidaktik* (Band 2, S. 211–224). Innsbruck: Studien Verlag.
- Bohl, T. (2016). Schülerselbstbewertung. In I. Kunze & C. Solzbacher (Hrsg.), *Individuelle Förderung in der Sekundarstufe I und II* (S. 177–182). Baltmannsweiler: Schneider Verlag Hohengehren.
- Bohnsack, R. (2005). Standards nicht-standardisierter Forschung in den Erziehungs- und Sozialwissenschaften. *Zeitschrift für Erziehungswissenschaft*, 8(Beiheft 4), 63–81.
- Bortz, J. & Döring, N. (2006). *Forschungsmethoden und Evaluation. für Human- und Sozialwissenschaftler*. Heidelberg: Springer Medizin Verlag.
- Bortz, J. & Schuster, C. (2010). *Statistik für Human- und Sozialwissenschaftler*. Berlin, Heidelberg: Springer.
- Bosch, B. & Boeck, K. de. (2016). Searching for a cure for cystic fibrosis. A 25-year quest in a nutshell. *European Journal of Pediatrics*, 175(1), 1–8.
- Böschen, S., Kastenhofer, K., Rust, I., Soentgen, J. & Wehling, P. (2008). Entscheidungen unter Bedingungen pluraler Nichtwissenskulturen. In R. Mayntz, Neidhardt, Friedhelm: Weingart, Peter & U. Wengenroth (Hrsg.), *Wissensproduktion und Wissenstransfer: Wissen im Spannungsfeld von Wissenschaft, Politik und Öffentlichkeit* (S. 197–219). Bielefeld: transcript Verlag.
- Braaten, M. & Windschitl, M. (2011). Working toward a stronger conceptualization of scientific explanation for science education. *Science Education*, 95(4), 639–669.
- Brandstetter, M., Florian, C. & Sandmann, A. (2016). Abbildungsmerkmale, Vorwissen, Cognitive Load und die Validierung eines Instrumentes zum Verstehen von Prozessdarstellungen in der Biologie. In U. Gebhard & M. Hammann (Hrsg.), *Lehr- und Lernforschung in der Biologiedidaktik. Bildung durch Biologieunterricht* (S. 301–318). Innsbruck: Studien Verlag.
- Brechman, J. M., Lee, C.-J. & Cappella, J. N. (2009). Lost in Translation? A Comparison of Cancer-Genetics Reporting in the Press Release and its Subsequent Coverage in Lay Press. *Science Communication*, 30(4), 453–474.
- Brown, G. (2006). Explaining. In O. Hargie (Hrsg.), *The handbook of communication skills* (S. 195–228). New York: Routledge.
- Brown, G. & Atkins, M. (1986). Explaining in professional contexts. *Research Papers in Education*, 1(1), 60–86.
- Brüning, L. & Saum, T. (2019). *Direkte Instruktion. Kompetenzen wirksam vermitteln*. Essen: Neue Deutsche Schule Verlagsgesellschaft mbH.
- Bühner, M. (2011). *Einführung in die Test- und Fragebogenkonstruktion*. München: Pearson.
- Bühner, M. & Ziegler, M. (2017). *Statistik für Psychologen und Sozialwissenschaftler. Grundlagen und Umsetzung mit SPSS und R*.
- Campbell, N. A. & Reece, J. B. (2009). *Biologie*. München: Pearson Studium.
- Clausen, M. (2002). *Unterrichtsqualität: Eine Frage der Perspektive?* (Pädagogische Psychologie und Entwicklungspsychologie, Bd. 29). Münster: Waxmann.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. New York: Erlbaum.
- Collins, M. L. (1978). Effects of Enthusiasm Training on Preservice Elementary Teachers. *Research in Teacher Education*, 29(1), 53–57.
- Da Costa, D. S., Hygino, J., Ferreira, T.B., Kasahara, T. M., Barros, P. O., Monteiro, C., Oliveira, A. et al. (2016). Vitamin D modulates different IL-17-secreting T cell subsets in multiple sclerosis patients. *Journal of Neuroimmunology*, (299), 8–18.
- Dagher, Z. & Cossman, G. (1992). Verbal explanations given by science teachers. Their nature and implications. *Journal of Research in Science Teaching*, 29(4), 361–374.

- Dendrou, C. A., Fugger, L. & Friese, M. A. (2015). Immunopathology of multiple sclerosis. *Nature Reviews. Immunology*, 15(9), 545–558.
- Denzin, N. K. (1970). *The Research Act. A Theoretical Introduction to Sociological Methods*. New York: McGraw-Hill.
- Dieker, L. A., Lane, H. B., Allsopp, D. H., O'Brien, C., Butler, T. W., Kyger, M. et al. (2009). Evaluating Video Models of Evidence-Based Instructional Practices to Enhance Teacher Learning. *Teacher Education and Special Education: The Journal of the Teacher Education Division of the Council for Exceptional Children*, 32(2), 180–196.
- Dittmer, A. (2010). *Nachdenken über Biologie. Über den Bildungswert der Wissenschaftsphilosophie in der akademischen Biologielehrerbildung*. Wiesbaden: Springer Fachmedien.
- Dittmer, A. & Ehras, C. (2024). Nature of Bioscience. Komplexität und Ethik als Merkmale eines biologiedidaktischen NOS-Ansatzes. In B. Reinisch, D. Krüger & D. Mahler (Hrsg.), *Biologiedidaktische Nature of Science-Forschung: Zukunftsweisende Praxis*. Springer.
- Dittmer, A. & Langlet, J. (im Druck). Kultur der Naturwissenschaften. In H. Gropengießer, U. Harms & U. Kattmann (Hrsg.), *Fachdidaktik Biologie. Begründet von D. Eschenhagen, U. Kattmann, D. Rodi*. Hallbergmoos: Aulis-Verlag.
- Dittmer, A. & Zabel, J. (2019). Das Wesen der Biologie verstehen: Impulse für den wissenschaftspropädeutischen Biologieunterricht. In J. Groß, M. Hammann, P. Schmiemann & J. Zabel (Hrsg.), *Biologiedidaktische Forschung: Erträge für die Praxis* (S. 93–110). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Dunlosky, J. & Lipko, A. R. (2007). Metacomprehension. A Brief History and How to Improve its Accuracy. *Current Directions in Psychological Science*, 16(4), 228–232.
- Ehlich, K. (2009). Erklären verstehen – Erklären und Verstehen. In R. Vogt (Hrsg.), *Erklären – Gesprächsanalytische und fachdidaktische Perspektiven* (Linguistik, Bd. 52, S. 11–24). Tübingen: Stauffenburg-Verlag.
- Ehras, C., Asen-Molz, K., Frei, M., Schilcher, A. & Krauss, S. (2021). Erklären lernen – Ein Seminarkonzept zur Förderung von Erklärkompetenz durch Videografie als Reflexionsanlass. In E. Matthes, S. T. Siegel & T. Heiland (Hrsg.), *Lehrvideos – das Bildungsmedium der Zukunft? Erziehungswissenschaftliche und fachdidaktische Perspektiven* (S. 203–212). Bad Heilbrunn: Julius Klinkhardt.
- Elgin, M. (2006). There may be strict empirical laws in biology, after all. *Biology and Philosophy*, 21(1), 119–134.
- Endruweit, G. (2014). Subjekt, soziales. In G. Endruweit, G. Trommsdorff & N. Burzan (Hrsg.), *Wörterbuch der Soziologie* (S. 521–522). Konstanz: UVK Verlagsgesellschaft mbH.
- Erickson, F. & Shultz, J. (1982). The counselor as gatekeeper: Social interactions in interviews. *Anthropology & Education Quarterly*, 14(4), 297–300.
- Ernst, P. (2011). *Germanistische Sprachwissenschaft: eine Einführung in die synchrone Sprachwissenschaft des Deutschen*. Wien: Facultas Verlag.
- Esser, F. (2012). Fortschritte und Herausforderungen der komparativen Kommunikationswissenschaft. In B. Stark, M. Magin, O. Jandura & M. Maurer (Hrsg.), *Methodische Herausforderungen komparativer Forschungsansätze* (S. 18–45). Köln: Halem Verlag.
- Farez, M. F. & Correale, J. (2011). Immunizations and risk of multiple sclerosis: systematic review and meta-analysis. *Journal of Neurology*, 258(7), 1197–1206.

- Fauth, B., Göllner, R., Lenke, G., Praetorius, A.-K. & Wagner, W. (2020). Who sees what? Conceptual considerations on the measurement of teaching quality from different perspectives. *Zeitschrift Für Pädagogik, Beiheft Nr. 66*, 138–155.
- Feldman, K. A. (2007). Identifying Exemplary Teachers and Teaching: Evidence from Student Ratings. In R. P. Perry & J. C. Smart (Hrsg.), *The scholarship of teaching and learning in higher education – an evidence-based perspective* (S. 93–143). Dordrecht: Springer.
- Feynman, R. P. (1994). *Six Easy Pieces. Essentials of Physics Explained by Its Most Brilliant Teacher* (Helix Book). Reading, MA: Addison-Wesley.
- Field, A. (2018). *Discovering statistics using IBM SPSS statistics*. London: SAGE.
- Fielding, N. & Fielding, J. (1986). *Linking Data* (Qualitative research methods, Bd. 4). Beverly Hills, Californien: SAGE.
- Findeisen, S. (2017). *Fachdidaktische Kompetenzen angehender Lehrpersonen. Eine Untersuchung zum Erklären im Rechnungswesen*. Wiesbaden: Springer.
- Flick, U. (2011). *Triangulation. Eine Einführung* (Qualitative Sozialforschung, Bd. 12). Wiesbaden: Springer Fachmedien.
- Frei, M., Asen-Molz, K., Hilbert, S., Schilcher, A. & Krauss, S. (2020). Die Wirksamkeit von Erklärvideos im Rahmen der Methode Flipped Classroom. In *Bildung.Schule.Digitalisierung* (S. 278–283). Münster: Waxmann.
- Frei, M., Puffer, G. & Hofmann, B. (im Druck). Allgemeine Musiklehre – „gut“ erklärt. In A. Schilcher, S. Krauss, A. Lindl & S. Hilbert (Hrsg.), *Fachspezifische Lehrerkompetenzen im Erklären*. Weinheim: Beltz Juventa.
- Friedrich, M. (2017). *Textverständlichkeit und ihre Messung. Entwicklung und Erprobung eines Fragebogens zur Textverständlichkeit*. Münster: Waxmann.
- Friesen, M., Kuntze, S. & Vogel, M. (2018). Videos, Texte oder Comics? Die Rolle des Vignettenformats bei der Erhebung fachdidaktischer Analysekompetenz zum Umgang mit Darstellungen im Mathematikunterricht. In J. Rutsch, M. Rehm, M. Vogel, M. Seidenfuß & T. Dörfler (Hrsg.), *Effektive Kompetenzdiagnose in der Lehrerbildung. Professionalisierungsprozesse angehender Lehrkräfte untersuchen* (S. 153–177). Wiesbaden: Springer.
- Gaehtgens, P. (1994). Das Kreislaufsystem. In R. Klinke & S. Silbernagel (Hrsg.), *Lehrbuch der Physiologie* (S. 135–182). Stuttgart: Georg Thieme Verlag.
- Gaier, L. (2022). *Gut erklärt? FALKE-D: Eine empirische Studie zu Erklärungen im Deutschunterricht aus unterschiedlichen Beobachterperspektiven*. Regensburg: Universitätsbibliothek Regensburg.
- Gautschi, P. (2011). *Guter Geschichtsunterricht* (Forum Historisches Lernen). Schwalbach/Ts.: Wochenschau Verlag.
- Geelan, D. (2012). Teacher Explanations. In B. J. Fraser (Hrsg.), *Second international handbook of science education* (Springer international handbooks of education, Bd. 24, S. 987–999). Dordrecht: Springer.
- Geelan, D. (2013). Teacher Explanation of Physics Concepts. A Video Study. *Research in Science Education*, 43(5), 1751–1762.
- Gehlhaar, K.-H. (2008). Lebende Organismen. In H. Gropengießer & U. Kattmann (Hrsg.), *Fachdidaktik Biologie* (S. 298–311). Köln: Aulis.
- Geißner, H. (1988). *Sprechwissenschaft. Theorie der mündlichen Kommunikation* (2., durchgesehene Auflage). Frankfurt am Main: Scriptor-Verlag.
- Gettier, E. L. (1963). Is Justified True Belief Knowledge? *Analysis*, 23(6), 121–123.
- Gigerenzer, G. (2019). **Rationales Entscheiden unter Ungewissheit \neq Rationales Entscheiden unter Risiko**. In B. Fleischer, R. Lauterbach & K. Pawlik (Hrsg.), *Rationale Entscheidungen unter*

- Unsicherheit* (Abhandlungen der Akademie der Wissenschaften in Hamburg, Bd. 8, S. 1–14). Berlin: Walter de Gruyter.
- Gläser, J. & Laudel, G. (2010). *Experteninterviews und qualitative Inhaltsanalyse. Als Instrumente rekonstruierender Untersuchungen*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Glynn, S. M. (1994). Teaching Science With Analogy: A Strategy for Teachers and Textbook Authors. *Reading Research Report*, (15), 1–25.
- Gogolin, I. & Michel, U. (2010). Kooperation und Vernetzung. Eine Dimension "Durchgängiger Sprachbildung". *Diskurs Kindheits- und Jugendforschung*, (4), 273–284.
- Goodwin, C. (1994). Professional Vision. *American Anthropologist*, 96(3), 606–633.
- Gordon, R. A., Druckman, D., Rozelle, R. M. & James, C. B. (2006). Non-verbal behavior as communication: Approaches, issues and research. In O. Hargie (Hrsg.), *The handbook of communication skills* (S. 73–120). New York: Routledge.
- Grotzer, T. A. (2012). *Learning causality in a complex world. Understandings of consequence*. Lanham, Maryland: Rowman & Littlefield Education.
- Grotzer, T. A., Solis, S. L., Tutwiler, M. S. & Cuzzolino, M. P. (2017). A study of students' reasoning about probabilistic causality. Implications for understanding complex systems and for instructional design. *Instructional Science*, 45(1), 25–52. <https://doi.org/10.1007/s11251-016-9389-6>
- Guenther, L. (2017). *Evidenz und Medien. Journalistische Wahrnehmung und Darstellung wissenschaftlicher Unsicherheit*. Wiesbaden: VS Verlag.
- Guenther, L. & Ruhrmann, G. (2016). Scientific Evidence and Mass Media: Investigating the Journalistic Intention to Represent Scientific Uncertainty. *Public Understanding of Science*, 25(8), 927–943.
- Halbfas, H. (2012). *Religiöse Sprachlehre. Theorie und Praxis*. Ostfildern: Patmos-Verlag.
- Hammann, M. & Asshof, R. (2014). *Schülervorstellungen im Biologieunterricht. Ursachen für Lernschwierigkeiten*. Seelze: Klett Kallmeyer.
- Hargie, O. (2013). *Die Kunst der Kommunikation. Forschung, Theorie, Praxis* (K. Beifuss, Übers.). Bern: Huber.
- Harren, I. (2009). Formen von Begriffsarbeit. wie im Unterrichtsgespräch Inhalte und Fachtermini verknüpft werden. In R. Vogt (Hrsg.), *Erklären – Gesprächsanalytische und fachdidaktische Perspektiven* (Linguistik, Bd. 52, S. 151–168). Tübingen: Stauffenburg-Verlag.
- Harren, I. (2015). *Fachliche Inhalte sprachlich ausdrücken lernen*. Verfügbar unter: <http://verlag-gespraechsforschung.de/2015/pdf/unterricht.pdf>
- Harris, M. & Rosenthal, R. [R.] (2005). Effects of Teacher Nonverbal Behavior on Student Outcomes. In R. E. Riggio & R. S. Feldman (Hrsg.), *Applications of nonverbal communication* (S. 157–192). Mahwah, NJ: Lawrence Erlbaum Associates.
- Harroud, A., Mitchell, R., Morris, J., Forgetta, V., Smith, G. D., Baranzini, S. et al. (2021). Childhood obesity and multiple sclerosis susceptibility: A Mendelian randomization study. *Canadian Journal of Neurological Sciences*, 27(14), 2150-2158.
- Hattie, J. (2015). *Lernen sichtbar machen*. Überarbeitete deutschsprachige Ausgabe von "Visible Learning" besorgt von Wolfgang Beywl und Klaus Zierer. Baltmannsweiler: Schneider Verlag Hohengehren GmbH.
- Heering, P. & Kremer, K. (2018). Nature of Science. In D. Krüger, I. Parchmann & H. Schecker (Hrsg.), *Theorien in der naturwissenschaftsdidaktischen Forschung* (S. 105–120). Berlin: Springer Spektrum.

- Heinze, J. & Rincke, K. (im Druck). Physik erklären. Zur Sprachlichkeit instruktionaler Erklärungen im Physikunterricht. In A. Schilcher, S. Krauss, A. Lindl & S. Hilbert (Hrsg.), *Fachspezifische Lehrerkompetenzen im Erklären*. Weinheim: Beltz Juventa.
- Helmke, A. (2007). *Unterrichtsqualität. Erfassen, Bewerten, Verbessern*. Seelze: Kallmeyer.
- Helmke, A. (2012). *Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts* (Schule weiterentwickeln, Unterricht verbessern Orientierungsband). Seelze: Kallmeyer.
- Hempel, C. & Oppenheim, P. (1948). Studies in the Logic of Explanation. *Philosophy of Science*, 14(2), 135–174.
- Hilbert, S., Stadler, M., Lindl, A., Naumann, F. & Bühner, M. (2019). Analyzing longitudinal intervention studies with linear mixed models. *Testing, Psychometrics, Methodology in Applied Psychology (TPM)*, 26(1), 101–119.
- Hohenstein, C. (2006). *Erklärendes Handeln im wissenschaftlichen Vortrag. Ein Vergleich des Deutschen mit dem Japanischen* (Studien Deutsch, Bd. 36). Dissertation an der Universität Hamburg 2003. München: Iudicium.
- Hopf, C. (2016). *Schriften zu Methodologie und Methoden qualitativer Sozialforschung* (1. Auflage). Wiesbaden: Springer Fachmedien.
- Jacobson, M. J. (2000). Problem Solving About Complex Systems. Differences Between Experts and Novices. In *Fourth International Conference of the Learning Sciences* (S. 14–21). Mahawah: Erlbaum.
- Jacobson, M. J. & Wilensky, U. (2006). Complex Systems in Education: Scientific and Educational Importance and Implications for the Learning Sciences. *The Journal of the Learning Sciences*, 15(1), 11–34.
- Janich, N. & Rhein, L. (2018). Einleitung. In N. Janich & L. Rhein (Hrsg.), *Unsicherheit als Herausforderung für die Wissenschaft. Reflexionen aus Natur-, Sozial- und Geisteswissenschaften* (Wissen-Kompetenz-Text, Band 13, S. 7–14). Berlin: Peter Lang.
- Jänig, W. & Vaupel, P. (2019). Allgemeine Aspekte des Gastrointestinaltrakts. In R. Bradnes, F. Lang & R. F. Schmidt (Hrsg.), *Physiologie des Menschen. mit Pathophysiologie* (S. 471–485). Berlin: Springer.
- Kampourakis, K. (Hrsg.). (2013). *The philosophy of biology. A companion for educators* (History, Philosophy and Theory of the Life Sciences). Dordrecht: Springer.
- Kampourakis, K. & McCain, K. (2020). *Uncertainty. How it makes science advance*. New York: Oxford University Press.
- Kampourakis, K. & Zogza, V. (2009). Preliminary Evolutionary Explanations: A Basic Framework for Conceptual Change and Explanatory Coherence in Evolution. *Science and Education*, 18, 1313–1340.
- Kattmann, U. (1995). Konzeption eines naturgeschichtlichen Biologieunterrichts: Wie Evolution Sinn macht. *Zeitschrift für Didaktik der Naturwissenschaften*, 1, 29–42.
- Kattmann, U. (2005). Lernen mit anthropomorphen Vorstellungen? – Ergebnisse von Untersuchungen zur Didaktischen Rekonstruktion in der Biologie. *Zeitschrift für Didaktik der Naturwissenschaften*, 11, 165–174.
- Kattmann, U. (2008a). Biologie als Wissenschaft im Unterrichtsfach. In H. Gropengießer & U. Kattmann (Hrsg.), *Fachdidaktik Biologie* (S. 39–63). Köln: Aulis.
- Kattmann, U. (2008b). Vielfalt und Funktionen von Unterrichtsmedien. In H. Gropengießer & U. Kattmann (Hrsg.), *Fachdidaktik Biologie* (292–297). Köln: Aulis.
- Kattmann, U. (2015). *Schüler besser verstehen. Alltagsvorstellungen im Biologieunterricht*. Hallbergmoos: Aulis.

- Keller, M. M., Hoy, A. W., Goetz, T. & Frenzel, A. C. (2016). Teacher Enthusiasm. Reviewing and Redefining a Complex Construct. *Educational Psychology Review*, 28(4), 743–769.
- Kennedy, J. (1996). Classroom explanatory discourse: A look at how teachers explain things to their students. *Language Awareness*, 5(1), 26–39.
- Kiel, E. (1999). *Erklären als didaktisches Handeln*. Würzburg: Ergon-Verlag.
- Kintsch, W. & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85(5), 363–394.
- Kitcher, P. (1981). Explanatory Unification. *Philosophy of Science*, 48(4), 507–531.
- Kleickmann, T., Praetorius, A.-K. & Riecke-Baulecke, T. (2019). Beurteilung von Unterrichtsqualität. In F. Zimmermann, J. Möller & T. Riecke-Baulecke (Hrsg.), *Basiswissen Lehrerbildung. Schulische Diagnostik und Leistungsbeurteilung* (S. 207–223). Hannover: Klett Kallmeyer.
- Kleickmann, T., Steffensky, M. & Praetorius, A.-K. (2020). Quality of teaching in science education. More than Three Basic Dimensions? *Zeitschrift für Pädagogik, Beiheft 66*, 37–55.
- Klein, J. (2009). Erklären-WAS, Erklären-WIE, Erklären-WARUM. Typologie und Komplexität zentraler Akte der Welterschließung. In R. Vogt (Hrsg.), *Erklären – Gesprächsanalytische und fachdidaktische Perspektiven* (Linguistik, Bd. 52, S. 25–36). Tübingen: Stauffenburg-Verlag.
- Kleinknecht, M. & Schneider, J. (2013). What do teachers think and feel when analyzing videos of themselves and other teachers teaching? *Teaching and Teacher Education*, 33, 13–23.
- KMK. (2020). *Bildungsstandards im Fach Biologie für die Allgemeine Hochschulreife*. Hürth: Wolters Kluwer.
- Koch, P. & Oesterreicher, W. (1985). Sprache der Nähe – Sprache der Distanz. Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgeschichte. In O. Deutschmann, H. Flasche, B. König, M. Kruse, W. Pabst & W.-D. Stempel (Hrsg.), *Romanistisches Jahrbuch* (Bd. 36, S. 15–43). Berlin, New York: Walter de Gruyter.
- Koller, H.-C. (2007). Bildung als Entstehung neuen Wissens? Zur Genese des Neuen in transformatorischen Bildungsprozessen. In H.-R. Müller & W. Stravoravdis (Hrsg.), *Bildung im Horizont der Wissensgesellschaft* (S. 49–66). Wiesbaden: Springer VS.
- Kort, K. (2021). *Babypuder-Skandal: Oberstes US-Gericht bestätigt Milliardenstrafe für Johnson & Johnson*. Handelsblatt. Verfügbar unter: <https://www.handelsblatt.com/technik/medizin/asbestverseuchung-babypuder-skandal-oberstes-us-gericht-bestaetigt-milliardenstrafe-fuer-johnson-und-johnson/27246470.html>
- Kotthoff, H. (2009). Erklärende Aktivitätstypen in Alltags- und Unterrichtskontexten. In J. Sprechels (Hrsg.), *Erklären im Kontext. Neue Perspektiven aus der Gesprächs- und Unterrichtsforschung* (S. 120–146). Baltmannsweiler: Schneider Verlag Hohengehren.
- Krammer, K. & Reusser, K. (2005). Unterrichtsvideos als Medium der Aus- und Weiterbildung von Lehrpersonen. *Beiträge zur Lehrerinnen- und Lehrerbildung*, 23(1), 35–50.
- Krapp, A. & Weidenmann, B. (2006). *Pädagogische Psychologie. Ein Lehrbuch*.
- Krauss, S. & Bruckmaier, G. (2014). Das Experten-Paradigma in der Forschung zum Lehrerberuf. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (S. 217–237). Münster: Waxmann.
- Krauss, S., Lindl, A., Schilcher, A., Fricke, M., Göhring, A., Hofmann, B. et al. (Hrsg.). (2017). *Falko: Fachspezifische Lehrkompetenzen. Konzeption von Professionswissenstest in den Fächern Deutsch, Englisch, Latein, Physik, Musik, Evangelische Religion und Pädagogik*. Münster: Waxmann.
- Krohs, U. (2004). *Eine Theorie biologischer Theorien. Status und Gehalt von Funktionsaussagen und informationstheoretischen Modellen*. Berlin: Springer.

- Krohs, U. (2007). Der Funktionsbegriff in der Biologie. In A. Bartels & M. Stöckler (Hrsg.), *Wissenschaftstheorie. Ein Studienbuch* (S. 287–306). Paderborn: Mentis.
- Krüger, Ruhrig, J. & Höttecke, D. (2013). Lehrerperspektiven auf unsichere Evidenz II: Ergebnisse einer Gruppendiskussionsstudie. In S. Bernholt (Hrsg.), *Inquiry-based Learning – Forschendes Lernen. Gesellschaft für Didaktik der Chemie und Physik, Jahrestagung in Hannover 2012* (Bd. 33, S. 728–730). IPN Kiel.
- Kuckartz, U. (2018). *Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung*. Weinheim/Basel: Beltz Juventa.
- Kuhlmann, M. (2007). Theorien komplexer Systeme. Nicht-fundamental und doch unverzichtbar? In A. Bartels & M. Stöckler (Hrsg.), *Wissenschaftstheorie. Ein Studienbuch* (S. 307–356). Paderborn: Mentis.
- Kuhn, D. (2007). Is direct instruction an answer to the right question? *Educational Psychologist*, 42(2), 109–113.
- Kulgemeyer, C. (2018). Wie gut erklären Erklärvideos? Ein Bewertungs-Leitfaden. *Computer + Unterricht*, 109, 8–11.
- Kulgemeyer, C. & Schecker, H. (2013). Students Explaining Science – Assessment of Science Communication Competence. *Research in Science Education*, 43(6), 2235–2256.
- Kulgemeyer, C. & Tomczyszyn, E. (2015). Physik erklären. Messung der Erklärens-fähigkeit angehender Physiklehrkräfte in einer simulierten Unterrichtssituation. *Zeitschrift für Didaktik der Naturwissenschaften*, 21, 111–126.
- Kunter, M. & Baumert, J. (2007). Who is the expert? Construct and criteria validity of student and teacher ratings of instruction. *Learning Environments Research*, 9(3), 231–251.
- Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S. & Neubrand, M. (Hrsg.). (2011). *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV*. Münster: Waxmann.
- Kunter, M., Klusmann, U., Baumert, J., Richter, D., Voss, T. & Hachfeld, A. (2013). Professional competence of teachers: Effects on instructional quality and student development. *Journal of Educational Psychology*, 105(3), 805–820.
- Lägel-Gunga, E., Schilcher, A., Kranich, W. & Gegner, C. (im Druck). Einfluss des Sprech- und Körperausdrucks von Lehrkräften auf die wahrgenommene Qualität von schulischen Erklärungen. In A. Schilcher, S. Krauss, A. Lindl & S. Hilbert (Hrsg.), *Fachspezifische Lehrerkompetenzen im Erklären*. Weinheim: Beltz Juventa.
- Lakoff, G. & Johnson, M. (2011). *Leben in Metaphern. Konstruktion und Gebrauch von Sprachbildern* (Systemische Horizonte). Heidelberg: Auer.
- Lamnek, S. (2010). *Qualitative Sozialforschung*. Weinheim: Beltz.
- Lange, M. (2013). What Would Natural Laws in the Life Sciences Be? In K. Kampourakis (Hrsg.), *The philosophy of biology. A companion for educators* (History, Philosophy and Theory of the Life Sciences, S. 67–85). Dordrecht: Springer.
- Langlet, J. (2001). Wissenschaft – entdecken & begreifen. *Unterricht Biologie*, 25(268), 4–12.
- Langlet, J. (2004). Wie leben wir mit Metaphern im Biologieunterricht? In H. Gropengießer, A. Janßen-Bartels & E. Sander (Hrsg.), *Lehren fürs Leben. Didaktische Rekonstruktion in der Biologie ; Ulrich Kattmann zur Verabschiedung aus dem Dienst der Carl-von-Ossietzky-Universität Oldenburg* (S. 51–59). Köln: Aulis-Verl. Deubner.
- Langlet, J. (2008). Wissenschaftspropädeutik. In H. Gropengießer & U. Kattmann (Hrsg.), *Fachdidaktik Biologie* (S. 66–83). Köln: Aulis.

- Larreamendy-Joerns, J. & Muñoz, T. (2010). Learning, Identity, and Instructional Explanations. In M. K. Stein & L. Kucan (Hrsg.), *Instructional Explanations in the Disciplines* (S. 23–40). Boston, MA: Springer US.
- Lederman, N. G. (2007). Nature of science: Past, present and future. In S. K. Abell & N. G. Lederman (Hrsg.), *Handbook of Research on Science Education* (S. 831–880). Mahwah, NJ: Lawrence Erlbaum Associates.
- Leinhardt, G. (1987). Development of an Expert Explanation: An Analysis of a Sequence of Substraction Lessons. *Cognition and Instruction*, 4(4), 225–282.
- Leinhardt, G. (1990). *Towards understanding instructional Explanation*.
- Leinhardt, G. (2001). Instructional Explanations: A Commonplace for Teaching and Location for Contrast. In V. Richardson (Hrsg.), *Handbook of research on teaching* (S. 333–357). Washington, DC: American Educational Research Association.
- Leisen, J. (2007). Das Erklären im Unterricht. *Der Mathematische und naturwissenschaftliche Unterricht (MNU)*, 60(8), 459–462.
- Leisen, J. (2010). *Handbuch Sprachförderung im Fach. Sprachsensibler Fachunterricht in der Praxis*. Bonn: Varus.
- Leisen, J. (2013). Trägst du noch vor oder erklärst du schon? Der Lehrer als Erzähler oder als Erklärer. *Unterricht Physik*, 24(135/136), 26–32.
- Lennox, J. G. & Kampourakis, K. (2013). Biological Teleology: The Need for History. In K. Kampourakis (Hrsg.), *The philosophy of biology. A companion for educators* (History, Philosophy and Theory of the Life Sciences, S. 421–454). Dordrecht: Springer.
- Luhmann, N. (1997). *Die Gesellschaft der Gesellschaft* (Bd. 2). Frankfurt am Main: Suhrkamp.
- Maier, M., Guenther, L., Ruhrmann, G., Berkela, B. & Milde, J. (2018). Kommunikation ungesicherter wissenschaftlicher Evidenz – Herausforderungen für Wissenschaftler, Journalisten und Publikum. In N. Janich & L. Rhein (Hrsg.), *Unsicherheit als Herausforderung für die Wissenschaft. Reflexionen aus Natur-, Sozial- und Geisteswissenschaften* (Wissen-Kompetenz-Text, Band 13, S. 93–111). Berlin: Peter Lang.
- Mainzer, K. (1997). *Gehirn, Computer, Komplexität*. Berlin: Springer.
- Mainzer, K. (2008). *Komplexität* (UTB, Bd. 3012). Paderborn: Wilhelm Fink Verlag GmbH & Co.
- Mambrey, S., Timm, J., Landskron, J. J. & Schmiemann, P. (2020). The impact of system specifics on systems thinking. *Journal of Research in Science Teaching*, 57(10), 1632–1651.
- Mayer, R. E. & Moreno, R. (2003). Nine Ways to Reduce Cognitive Load in Multimedia Learning. *Educational Psychologist*, 38(1), 43–52.
- Mayr, E. (2002). Die Autonomie der Biologie. *Naturwissenschaftliche Rundschau*, 55(1), 23–29.
- Mayring, P. (2010). *Qualitative Inhaltsanalyse. Grundlagen und Techniken* (Beltz Pädagogik). Weinheim: Beltz.
- McComas, W. F. & Olson, J. K. (1998). The nature of science in international science education standards documents. In W. F. McComas (Hrsg.), *The Nature of Science in Science Education* (S. 41–52). Dordrecht: Kluwer Academic Publishers.
- McDonald, M., Kazemi, E. & Kavanagh, S. S. (2013). Core Practices and Pedagogies of Teacher Education. *Journal of Teacher Education*, 64(5), 378–386.
- McKinney, W. C., Larkins, Guy A., Kazelskis, Richard, Ford, M. J., Allen, J. A. & Davis, J. C. (1983). Some Effects of Teacher Enthusiasm on Student Achievement in Fourth Grade Social Studies. *The Journal of Educational Research*, 76(4), 249–253.
- McLaughlin, P. (2020). Funktionen. In G. Toepfer & U. Krohs (Hrsg.), *Philosophie der Biologie. Eine Einführung* (S. 19–35). Frankfurt am Main: Suhrkamp.

- Medienpädagogischer Forschungsverbund Südwest. (2020). *JIMplus 2020. Lernen und Freizeit in der Corona-Krise*. Verfügbar unter: https://www.mpfs.de/fileadmin/files/Studien/JIM/JIMplus_2020/JIMplus_2020_Corona.pdf
- Mehren, M., Mehren, R., Ohl, U. & Resenberger. (2015). Die doppelte Komplexität geographischer Themen. Eine lohnenswerte Herausforderung für Schüler und Lehrer. *Geographie und Schule*, 37(216), 4–11.
- Melief, Koper, Endert, Moller, Hamann, Utidehaag et al. (2015). Glucocorticoid receptor haplotypes conferring increased sensitivity (Bcll and N363S) are associated with faster progression of multiple sclerosis. *Journal of Neuroimmunology*, (299), 84–89.
- Menold, N. & Bogner, K. (2015). *Gestaltung von Ratingskalen in Fragebögen* (GESIS Survey Guidelines). Mannheim: GESIS – Leibniz-Institut für Sozialwissenschaften.
- Merzyn, W. (2015). *Guter Physikunterricht. Die Sicht von Schülern, Lehrern und Wissenschaftlern*. PhyDid B – Didaktik Der Physik – Beiträge Zur DPG-Frühjahrstagung. Verfügbar unter: <https://ojs.dpg-physik.de/index.php/phydid-b/article/view/589>
- Mettler, G. (2018). *Amphibienschutz*. Verfügbar unter: <https://naturschutz-merkligen.de/aktivitaeten/amphibienschutz>
- Meyer, H. (2004). *Was ist guter Unterricht?* Berlin: Cornelsen Scriptor.
- Meyer, L., Seidel, T. & Prenzel, M. (2006). Wenn Lernsituationen zu Leistungssituationen werden: Untersuchung zur Fehlerkultur in einer Videostudie. *Schweizerische Zeitschrift für Bildungswissenschaften*, 28(1), 21–41.
- Mitchell, S. D. (2008). *Komplexitäten. Warum wir erst anfangen, die Welt zu verstehen*. Frankfurt am Main: Suhrkamp.
- Mittelstrass, J. (2007). Geozentrisch, geozentrisches Weltsystem. In J. Ritter, K. Gründer & G. Garbiel (Hrsg.), *Historisches Wörterbuch der Philosophie online*. Basel: Schwabe AG.
- Morek, M. (2012). *Kinder Erklären. Interaktionen in Familie und Unterricht im Vergleich* (Linguistik, Bd. 60). Tübingen: Stauffenburg-Verl.
- Morek, M. & Heller, V. (2012). Bildungssprache. Kommunikative, epistemische, soziale und interaktive Aspekte ihres Gebrauchs. *Zeitschrift für angewandte Linguistik*, 57(1), 67–101.
- Morton, V. & Watson, D. R. (2001). The impact of impaired vocal quality on children's ability to process spoken language. *Log Phon Vocol*, 26(1), 17–25.
- Neuber, B. (2002). *Prosodische Formen in Funktion. Leistungen der Suprasegmentalia für das Verstehen, Behalten und die Bedeutungs(re)konstruktion* (Hallesche Schriften zur Sprechwissenschaft und Phonetik, Bd. 7). Frankfurt am Main: Peter Lang GmbH.
- Neuber, B. (2016). Paraverbale und nonverbale Anteile der rhetorischen Kommunikation. In I. Bose, U. Hirschfeld, B. Neuber & E. Stock (Hrsg.), *Einführung in die Sprechwissenschaft. Phonetik, Rhetorik, Sprechkunst* (Narr-Studienbücher, S. 134–139). Tübingen: Narr Francke Attempto.
- Nitz, S., Nerdel, C. & Prechtel, H. (2012). Entwicklung eines Erhebungsinstruments zur Erfassung der Verwendung von Fachsprache im Biologieunterricht. *Zeitschrift für Didaktik der Naturwissenschaften*, 18, 117–139.
- Nückles, M., Wittwer, J. & Renkl, A. (2005). Information About a Layperson's Knowledge Supports Experts in Giving Effective and Efficient Online Advice to Laypersons. *Journal of Experimental Psychology*, 11(4), 219–236.
- Odora, R. J. (2014). Using Explanation as a Teaching Method. How prepared are high school technology teachers in free state province, South Africa? *Journal of Social Sciences*, 38(1), 71–81.
- Ogborn, J. (2011). *Explaining science in the classroom*. Buckingham: Open University Press.

- Ohl, U. (2018). Herausforderungen und Wege eines systematischen Umgangs mit komplexen Themen in der schulischen Nachhaltigkeitsbildung. In T. Pyhel (Hrsg.), *Zwischen Ohnmacht und Zuversicht? Vom Umgang mit Komplexität in der Nachhaltigkeitskommunikation* (DBU-Umweltkommunikation, Bd. 10, S. 131–146). München: Oekom Verlag.
- Ohlhoff, D. (2002). *Das freundliche Selbst und der angreifende Feind. Politische Metaphern und Körperkonzepte in der Wissensvermittlung der Biologie*. metaphorik.de. Verfügbar unter: <https://www.metaphorik.de/de/journal/03/das-freundliche-selbst-und-der-angreifende-feind-politische-metaphern-und-koerperkonzepte-der.html>
- Osborne, J. F., Ratcliffe, M., Collins, S., Millar, R. & Duschl, R. (2001). *What should we teach about science? A Delphi study*. Evidence-based in Science Education (EPSE) Research Network.
- Pabst-Weinschenk, M. (2011). Hörverstehen und Sprechdenken. In M. Pabst-Weinschenk (Hrsg.), *Grundlagen der Sprechwissenschaft und der Sprecherziehung*. (S. 58–82). München: Reinhardt UTB.
- Paivio, A. (1986). *Mental Representations*. New York: Oxford University Press.
- Pauli, C. (2015). Einen Sachverhalt erklären. *Pädagogik*, 67(3), 44–47.
- Pietrzik, K., Golly, I. & Loew, D. (2008). *Handbuch Vitamine*. Stuttgart: Gustav Fischer.
- Piwovar, V., Barth, V. L., Ophardt, D. & Thiel, F. (2018). Evidence-based scripted videos on handling student misbehavior: the development and evaluation of video cases for teacher education. *Professional Development in Education*, 44(3), 369–384.
- Potochnik, A. (2013). Biological Explanation. In K. Kampourakis (Hrsg.), *The philosophy of biology. A companion for educators* (History, Philosophy and Theory of the Life Sciences, S. 49–65). Dordrecht: Springer.
- Praetorius, A.-K. & Gräsel, C. (2021). Noch immer auf der Suche nach dem heiligen Gral: Wie generisch oder fachspezifisch sind Dimensionen der Unterrichtsqualität? *Unterrichtswissenschaft*, 49(2), 167–188.
- Prüfer, P. & Rexroth, M. (2005). *Kognitive Interviews* (ZUMA How-to-Reihe 15). Mannheim: Zentrum für Umfragen, Methoden und Analysen.
- Rädiker, S. & Kuckartz, U. (2019). *Analyse qualitativer Daten mit MAXQDA. Text, Audio und Video* (Lehrbuch). Wiesbaden: Springer VS.
- Rahmstorf, S. & Schellnhuber, H.-J. (2019). *Der Klimawandel. Diagnose, Prognose, Therapie* (Beck'sche Reihe C.-H.-Beck-Wissen, Bd. 2366). München: C. H. Beck.
- Raithel, V. & Wrede, B. (2003). Warum haben es „Schnellsprecher“ schwer? *Forschung an der Universität Bielefeld*, 25, 35–41.
- Ravetz, J. R. (1987). Usable Knowledge, Usable Ignorance. Incomplete Science with Policy Implications. *Knowledge*, 9(1), 87–116.
- Rehm, M. & Bölsterli, K. (2014). Entwicklung von Unterrichtsvignetten. In D. Krüger (Hrsg.), *Methoden in der naturwissenschaftsdidaktischen Forschung* (Lehrbuch, S. 213–225). Berlin: Springer Spektrum.
- Renkl, A. (2001). Explorative Analysen zur effektiven Nutzung von instruktionalen Erklärungen beim Lernen aus Lösungsbeispielen. *Unterrichtswissenschaft*, 29, 41–63.
- Renkl, A. (2004). Vom trägen zum aktiven Wissen. Träges Wissen auf pädagogisch-psychologischer Sicht. *Schulmagazin 5-10*, (4), 5–8.
- Renkl, A., Wittwer, J., Große, C., Hauser, S., Hilbert, T., Nückles, M. et al. (2007). Instruktionale Erklärungen beim Erwerb kognitiver Fertigkeiten: Sechs Thesen zu einer oft vergeblichen Bemühung. In I. Hosenfeld & F.-W. Schrader (Hrsg.), *Schulische Leistung – Grundlagen, Bedingungen, Perspektiven* (S. 203–223). Münster: Waxmann.
- Resnick, M. (1995). Beyond the centralized mindset. *Journal of the Learning Sciences*, 5(1), 1–22.

- Retzbach, A. & Maier, M. (2015). Communicating Scientific Uncertainty: Media Effects on Public Engagement With Science. *Communication Research*, 42(3), 429–456.
- Rhein, L. (2018). Thematisierung von Nichtwissen und Unsicherheiten in wissenschaftlichen Diskussionen. In N. Janich & L. Rhein (Hrsg.), *Unsicherheit als Herausforderung für die Wissenschaft. Reflexionen aus Natur-, Sozial- und Geisteswissenschaften* (Wissen-Kompetenz-Text, Band 13, S. 71–92). Berlin: Peter Lang.
- Richter, K. & Rost, J.-M. (2014). *Komplexe Systeme*. Frankfurt am Main: Fischer.
- Rincke, K. (2010). Alltagssprache, Fachsprache und ihre besonderen Bedeutungen für das Lernen. *Zeitschrift für Didaktik der Naturwissenschaften*, 16, 235–260.
- Salmon, W. (1984). *Scientific Explanation and the Causal Structure of the World*. Princeton: Princeton University Press.
- Savant, A. P. & McColley, S. A. (2017). Cystic fibrosis year in review 2016. *Pediatric Pulmonology*, (52), 1092–1102.
- Schilcher, A., Krauss, S., Kirchhoff, P., Lindl, A., Hilbert, S., Asen-Molz, K. et al. (2021). FALKE: Experiences From Transdisciplinary Educational Research by Fourteen Disciplines. *Frontiers in Education*, 5(579982), 1–12.
- Schilcher, A., Krauss, S., Lindl, A. & Hilbert, S. (Hrsg.) (im Druck). *Fachspezifische Lehrerkompetenzen im Erklären*. Weinheim: Beltz Juventa.
- Schilcher, A., Krauss, S., Rincke, K. & Hilbert, S. (2017). Ein Ausblick – Aus FALKE wird FALKE. Fachspezifische Lehrerkompetenz im Erklären. In S. Krauss, A. Lindl, A. Schilcher, M. Fricke, A. Göhring, B. Hofmann et al. (Hrsg.), *Falko: Fachspezifische Lehrerkompetenzen. Konzeption von Professionswissenstest in den Fächern Deutsch, Englisch, Latein, Physik, Musik, Evangelische Religion und Pädagogik*. Münster: Waxmann.
- Schlosser, G. & Weingarten, M. (Hrsg.). (2002). *Formen der Erklärung in der Biologie* (Studien zur Theorie in der Biologie, Bd. 2). Berlin: VWB – Verlag für Wissenschaft und Bildung.
- Schmelzing, S. (2010). *Das fachdidaktische Wissen von Biologielehrkräften: Konzeptionalisierung, Diagnostik, Struktur und Entwicklung im Rahmen der Biologielehrerbildung*. Berlin: Logos Verlag.
- Schmiemann, P. (2011). Fachsprache in biologischen Testaufgaben. *Zeitschrift für Didaktik der Naturwissenschaften*, 17, 115–136.
- Schmölzer-Eibinger, S. (2013). Sprache als Medium des Lernens im Fach. In M. Becker-Mrotzek, K. Schramm, E. Thürmann & H. J. Vollmer (Hrsg.), *Sprache im Fach. Sprachlichkeit und fachliches Lernen* (Fachdidaktische Forschungen, Band 3, S. 25–40). Münster: Waxmann.
- Schnotz, W. (1994). *Aufbau von Wissensstrukturen. Untersuchungen zur Kohärenzbildung beim Wissenserwerb mit Texten* (Fortschritte der psychologischen Forschung, Bd. 20). Weinheim: Beltz Psychologie Verlags Union.
- Schnotz, W. (2001). Wissenserwerb mit Multimedia. *Unterrichtswissenschaft*, 29(4), 292–318.
- Schnotz, W. (2003). Bild- und Sprachverarbeitung aus psychologischer Sicht. In K. Sachs-Hombach (Hrsg.), *Was ist Bildkompetenz? Studien zur Bildwissenschaft* (Bildwissenschaft, Bd. 10, S. 25–42). Wiesbaden: Deutscher Universitäts-Verlag.
- Schnotz, W. (2005). An integrated model of text and picture comprehension. In R. E. Mayer (Hrsg.), *The Cambridge handbook of multimedia learning* (S. 49–69). Cambridge: Cambridge University Press.
- Schnotz, W. & Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning and Instruction*, 13(2), 141–156.
- Schopf, C. & Zwischenbrugger, A. (2015). Verständliche Erklärungen im Wirtschaftsunterricht. Eine Heuristik basierend auf dem Verständnis der Fachdidaktiker/innen des Wiener Lehrstuhls für Wirtschaftspädagogik. *Zeitschrift für ökonomische Bildung*, (3), 1–31.

- Schreier, M. (2014). Varianten qualitativer Inhaltsanalyse: Ein Wegweiser im Dickicht der Begrifflichkeiten. *Forum: Qualitative Sozialforschung*, 15(1), Artikel 18, 1–27.
- Schurz, G. (2004). Erklären und Verstehen: Tradition, Transformation und Aktualität einer klassischen Kontroverse. In F. Jaeger & J. Straub (Hrsg.), *Handbuch der Kulturwissenschaften* (Bd. 2, S. 156–174). Stuttgart: J. B. Metzler Verlag.
- Schwering, M. (2020). *Jürgen Habermas über Corona: „So viel Wissen über unser Nichtwissen gab es noch nie“*. Frankfurter Rundschau. Verfügbar unter: <https://www.fr.de/kultur/gesellschaft/juergen-habermas-coronavirus-krise-covid19-interview-13642491.html>
- Seidel, T. (2014). Angebots-Nutzungs-Modelle in der Unterrichtspsychologie. Integration von Struktur- und Prozessparadigma. *Zeitschrift für Pädagogik*, 60(6), 850–866.
- Seidel, T., Blomberg, G. & Stürmer, K. (2010). „Observer“ – Validierung eines videobasierten Instruments zur Erfassung der professionellen Wahrnehmung von Unterricht. Projekt OBSERVE. *Zeitschrift für Pädagogik*, (56), 296–306.
- Seidel, T. & Prenzel, M. (2007). Wie Lehrpersonen Unterricht wahrnehmen und einschätzen. Erfassung pädagogisch-psychologischer Kompetenzen bei Lehrpersonen mit Hilfe von Videosequenzen. *Zeitschrift für Erziehungswissenschaft*, (Sonderheft 8), 201–214.
- Seidel, T. & Thiel, F. (2017). Standards und Trends der videobasierten Lehr-Lernforschung. *Zeitschrift für Erziehungswissenschaft*, 20(1), 1–21.
- Seifried, J. & Wuttke, E. (2017). Der Einsatz von Videovignetten in der wirtschaftspädagogischen Forschung: Messung und Förderung von fachwissenschaftlichen und fachdidaktischen Kompetenzen angehender Lehrpersonen. In C. Gräsel & K. Trempler (Hrsg.), *Entwicklung von Professionalität pädagogischen Personals. Interdisziplinäre Betrachtungen, Befunde und Perspektiven* (S. 303–322). Wiesbaden: Springer.
- Sevian, H. & Gonsalves, L. (2008). Analysing how Scientists Explain their Research. A rubric for measuring the effectiveness of scientific explanations. *International Journal of Science Education*, 30(11), 1441–1467.
- Sherin, M. G. (2001). Developing a professional vision of classroom events: Teaching elementary school mathematics. In T. Wood, B. S. Nelson & J. Warfield (Hrsg.), *Beyond classical pedagogy: Teaching elementary school mathematics* (S. 75–93). Hillsdale: Erlbaum.
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4–14.
- Sick, U. (2014). *Poltern. Theoretische Grundlagen, Diagnostik, Therapie* (Forum Logopädie). Stuttgart: Thieme.
- Spreckels, J. (2009). Erklären im Kontext: Neue Perspektiven. In J. Spreckels (Hrsg.), *Erklären im Kontext. Neue Perspektiven aus der Gesprächs- und Unterrichtsforschung* (S. 1–11). Baltmannsweiler: Schneider Verlag Hohengehren.
- Stadler, F. (2015). *Der Wiener Kreis. Ursprung, Entwicklung und Wirkung des Logischen Empirismus im Kontext* (Studien zum Wiener Kreis, Bd. 20). Wien: Springer International Publishing.
- Stamann, C., Jansse, M. & Schreier, M. (2016). Qualitative Inhaltsanalyse – Versuch einer Begriffsbestimmung und Systematisierung. *Forum: Qualitative Sozialforschung*, 17(3), Artikel 16, 1–15.
- Steinke, I. (2007). Qualitätssicherung in der qualitativen Forschung. In U. Kuckartz (Hrsg.), *Qualitative Datenanalyse: computergestützt. Methodische Hintergründe und Beispiele aus der Forschungspraxis* (S. 176–189). Wiesbaden: VS Verlag.

- Steinke, I. (2012). Gütekriterien qualitativer Forschung. In U. Flick, E. von Kardoff & I. Steinke (Hrsg.), *Qualitative Forschung. Ein Handbuch* (Rowohlts Enzyklopädie, S. 319–331). Reinbek bei Hamburg: Rowohlt Taschenbuch Verlag.
- Stiensmeier-Pelster, J. & Schöne, C. (2008). Fähigkeitsselbstkonzept. In W. Schneider & M. Hasselhorn (Hrsg.), *Handbuch der Pädagogischen Psychologie* (Bd. 10, S. 62–73). Göttingen: Hogrefe.
- Stigler, J. W., Gonzales, P., Kawanaka, T., Knoll, S. & Serrano, A. (1999). *The TIMSS Videotape Classroom Study. Methods and Findings from an Exploratory Research Project on Eighth-Grade Mathematics Instruction in Germany, Japan, and the United States*. Washington, DC: National Center for Education Statistics.
- Strevens, M. (2004). The Causal and Unification Approaches to Explanation Unified – Causally. *NOUS*, 38(1), 154–176.
- Strube, P. (1986). Explanation and explaining in science education. *Research in Science Education*, 16(1), 135–140.
- Stukenbrock, A. (2009). Erklären – Zeigen – Demonstrieren. In J. Spreckels (Hrsg.), *Erklären im Kontext. Neue Perspektiven aus der Gesprächs- und Unterrichtsforschung* (S. 160–177). Baltmannsweiler: Schneider Verlag Hohengehren.
- Sweller, J. (2011). Cognitive Load Theory. In J. P. Mestre & B. H. Ross (Hrsg.), *Psychology of Learning and Motivation* (S. 37–76). New York: Academic Press.
- Taddicken, M., Reif, A. & Hoppe, I. (2018). Wissen, Nichtwissen, Unwissen, Unsicherheit: Zur Operationalisierung und Auswertung von Wissensitems am Beispiel des Klimawissens. In N. Janich & L. Rhein (Hrsg.), *Unsicherheit als Herausforderung für die Wissenschaft. Reflexionen aus Natur-, Sozial- und Geisteswissenschaften* (Wissen-Kompetenz-Text, Band 13, S. 113–142). Berlin: Peter Lang.
- Tainio, L. & Winkler, I. (2014). The construction of ideal reader in German and Finnish textbooks for literacy education. *L1-Educational Studies in Language and Literature*, 14, 1–25.
- Thim-Mabrey, C. & Lindl, A. (im Druck). Erklärqualität aus sprachwissenschaftlicher Sicht. Die Wahrnehmung der sprachlichen Verständlichkeit von Erklärvideos aus elf verschiedenen Fächern. In A. Schilcher, S. Krauss, A. Lindl & S. Hilbert (Hrsg.), *Fachspezifische Lehrerkompetenzen im Erklären*. Weinheim: Beltz Juventa.
- Thompson, J., Braaten, M. & Windschitl, M. (2009). *Learning Progressions as vision tools for advancing teachers' pedagogical performance*. Paper presented at the Learning Progressions in Science (LeaPS) Conference. Iowa City.
- Toepfer, G. (2004). *Zweckbegriff und Organismus. Über die teleologische Beurteilung biologischer Systeme*. Würzburg: Königshausen und Neumann.
- Tourangeau, R., Rips, L. & Rasinski, K. (2000). *The psychology of survey response*. Cambridge: Cambridge University Press.
- Trautmann, M. & Wischer, B. (2016). Lehren. Überlegungen zu einem nur wenig benutzten Begriff. In A. Feindt, W. Herget, M. Trautmann, B. Wischer & K. Zierer (Hrsg.), *Lehren* (Friedrich-Jahresheft, Bd. 34, S. 4–7). Seelze: Friedrich Verlag GmbH.
- Treagust, D. F. & Harrison, A. G. (2000). In search of explanatory frameworks. An analysis of Richard Feynman's lecture 'Atoms in motion'. *International Journal of Science Education*, 22(11), 1157–1170.
- Trommler, F., Gresch, H. & Hammann, M. (2017). Students' reasons for preferring teleological explanations. *International Journal of Science Education*, 40(2), 159–187.
- Van Es, E. A. & Sherin, M. G. (2002). Learning to Notice: Scaffolding New Teachers' Interpretations of Classroom Interactions. *Journal of Technology and Teacher Education*, 10(4), 571–596.

- Van Fraassen, B. C. (1980). *The scientific image* (Clarendon library of logic and philosophy). Oxford, New York: Clarendon Press; Oxford University Press.
- Van Vliet, L. M., Hillen, M. A., van der Wall, E., Plum, N. & Bensing, J. M. (2013). How to create and administer scripted video-vignettes in an experimental study on disclosure of a palliative breast cancer diagnosis. *Patient Education and Counseling*, 91(1), 56–64.
- Varwig, C. (2020). Kommunizieren oder verscheigen – Wie geht man mit wissenschaftlicher Unsicherheit um? In J. Schnurr & A. Mäder (Hrsg.), *Wissenschaft und Gesellschaft: Ein vertrauensvoller Dialog. Positionen und Perspektiven der Wissenschaftskommunikation heute* (S. 209–214). Berlin: Springer.
- Vogt, R. (2009). Die Organisation von Erklärprozessen im Unterricht. In R. Vogt (Hrsg.), *Erklären – Gesprächsanalytische und fachdidaktische Perspektiven* (Linguistik, Bd. 52, S. 203–225). Tübingen: Stauffenburg-Verlag.
- Voigt-Zimmermann, S. (2011). Zum Einfluss gestörter Lehrerstimmen auf den Verstehensprozess von Schülern. In I. Bose & B. Neuber (Hrsg.), *Interpersonelle Kommunikation. Analyse und Optimierung* (Hallesche Schriften zur Sprechwissenschaft und Phonetik, Bd. 39, S. 269–275). Frankfurt am Main: Peter Lang.
- Wagner, A. & Wörn, C. (2011). *Erklären lernen – Mathematik verstehen. Ein Praxisbuch mit Lernangeboten*. Seelze: Kallmeyer.
- Webb, N. M., Troper, J. D. & Fall, R. (1995). Constructive Activity and Learning in Collaborative Small Groups. *Journal of Educational Psychology*, 87(3), 406–423.
- Webster, D. M. & Kruglanski, A. W. (1994). Individual differences in need for cognitive closure. *Journal of Personality and Social Psychology*, 67(6), 1049–1062.
- Wehling, P. (2006). *Im Schatten des Wissens? Perspektiven der Soziologie des Nichtwissens. Theorie und Methode*. Konstanz: UVK Verlagsgesellschaft mbH.
- Wehling, P. (2008). Wissen und seine Schattenseite: Die wachsende Bedeutung des Nichtwissens in (vermeintlichen) Wissensgesellschaften. In T. Brüsemeister & K.-D. Eubel (Hrsg.), *Evaluation, Wissen und Nichtwissen* (S. 17–34). Wiesbaden: VS Verlag für Sozialwissenschaften/GWV Fachverlage GmbH Wiesbaden.
- Wehling, P. (2009). Nichtwissen – Bestimmungen, Abgrenzungen, Bewertungen. *Erwägen Wissen Ethik*, 20(1), 95–106.
- Wehling, P. (2012). Nichtwissenskulturen und Nichtwissenskommunikation in den Wissenschaften. In N. Janich, A. Nordmann & L. Schebeck (Hrsg.), *Nichtwissenskommunikation in den Wissenschaften. Interdisziplinäre Zugänge* (Wissen-Kompetenz-Text, Bd. 1, S. 73–92). Frankfurt am Main: Peter Lang.
- Weidenmann, B. (1994). Informierende Bilder. In B. Weidenmann (Hrsg.), *Wissenserwerb mit Bildern. Instruktive Bilder in Printmedien, Film/Video und Computerprogrammen* (S. 9–58). Bern: Huber.
- Weinhuber, M., Lachner, A., Leuders, T. & Nückles, M. (2019). Mathematics is practice or argumentation: Mindset priming impacts principle- and procedure-orientation of teachers' explanations. *Journal of Experimental Psychology*, 25(4), 618–646.
- Wellenreuther, M. (2013). *Lehren und Lernen – aber wie? Empirisch-experimentelle Forschungen zum Lehren und Lernen im Unterricht* (Grundlagen der Schulpädagogik, Bd. 50). Baltmannsweiler: Schneider Verlag Hohengehren.
- Witt, P. L., Wheelless, L. R. & Allen, M. (2004). A meta-analytical review of the relationship between teacher immediacy and student learning. *Communication Monographs*, 71(2), 184–207.

- Wittwer, J. & Renkl, A. (2008). Why Instructional Explanations Often Do Not Work. A Framework for Understanding the Effectiveness of Instructional Explanations. *Educational Psychologist*, 43(1), 49–64.
- Wörn, C. (2014). *Unterrichtliche Erklärsituationen. Eine empirische Studie zum Lehrerhandeln und zur Kommunikation in Mathematikunterricht der Sekundarstufe I* (Didaktik in Forschung und Praxis, Bd. 74). Hamburg: Dr. Kovac.
- Wouters, A. (2013). Explanation in Biology. In W. Dubitzky, O. Wolkenhauer, K.-H. Cho & H. Yokota (Hrsg.), *Encyclopedia of systems biology* (Bd. 2, S. 706–708). New York, NY [u. a.]: Springer Reference.
- Wragg, E. C. (1993). *Primary Teaching Skills*. London.
- Zabel, J. (2009). *Biologie verstehen: die Rolle der Narration beim Verstehen der Evolutionstheorie* (Didaktisches Zentrum der Carl von Ossietzky Universität Oldenburg, Bd. 24). Oldenburg: Universität Oldenburg.
- Ziepprecht, K., Jäger, D. S. & Schwanewedel, J. (2017). Charakteristika von Repräsentationen im Biologieunterricht aus Sicht von Schülerinnen und Schülern der Sekundarstufe II. *Biologie lehren und lernen – Zeitschrift für Didaktik der Biologie*, 21, 23–41.

Anhang

A) Transkripte der sechs Videovignetten

Die Transkripte sowie die nachstehende Vorbemerkung zum Transkript sind von Prof. Dr. Christiane Thim-Mabrey verfasst und wurden den Projektmitarbeitenden zur Verfügung gestellt. Sie dienen als Grundlage für die ausführliche Auswertung der Vignetten aus sprachwissenschaftlicher Sicht (Thim-Mabrey & Lindl, im Druck).

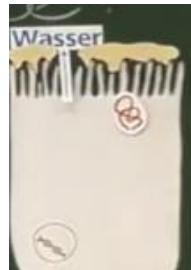
Vorbemerkung zur Weise der Verschriftlichung: Für den gesprochenen Text wird die sonst orthographische Groß-/Kleinschreibung nicht verwendet. Markiert werden sprecherische Merkmale, die im Deutschen inhaltsstrukturierende und bedeutungsstabilisierende Funktionen erfüllen können: Betonungen, Vokaldehnungen, umgangssprachliche und dialektale Wortverschleifungen, Pausen, syntaktische und inhaltliche Einheiten (zu den verwendeten Markierungsmitteln: siehe unterhalb der Transkription). Außerdem werden Visualisierungen notiert, da sie ebenfalls solche Funktionen erfüllen. Die Visualisierungsmittel in der rechten Spalte werden, soweit möglich, graphisch parallel angeordnet, um die zeitliche Synchronizität abzubilden.

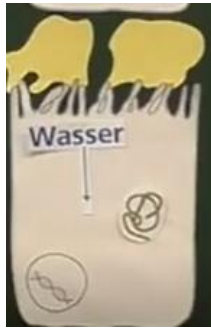
Markierungen: Bedeutung der Markierungsmittel	
Zeilenumbruch	Am Ende einer syntaktischen Einheit
Leerzeile	Nach einer inhaltlichen Zäsur
akZENT	Betonung (Primär- und Hauptakzent)
ak!ZENT!	besonders starke Betonung
(.)	Mikropause
(-), (--), (---)	kurze, mittlere, längere Pause
:, ::, :::	Dehnung eines Vokals (je nach Länge)
ne=das könn=wir	Verschleifungen innerhalb von Einheiten
...//	vollständiger Satzabbruch
.../	Neuansatz für einen Reformulierungsversuch
[...]	unverständliche Passage

Gj: Wie entsteht Mukoviszidose?

Die Vignette befindet sich auf der beigelegten CD unter dem Dateinamen „Genetik_linear“.

Nr.	Gesprochener Text	Art der Visualisierung: Tafelbild mit Wortkarten
01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63	<p>was der TEXT jetzt allerdings noch NICHT erklärt hat, is, (.) was die URsache der mukoviszidose ist (.) oder, waRUM der schleim auf einmal ZÄher is, und das schauen wir uns jetzt noch geMEINSam an. (--)</p> <p>ja, mukoviszidose is eine geNETische krankheit. das HEI:ßt, wir haben einen (-) abschnitt auf unserer dnA, also ein GEN, (.) das dafür verANTwortlich ist, dass (.) MENschen an mukoviszidOse erkranken. (.) und um das zu verSTehen, müssen wir uns nochmal ALLgemein dran erinnern, wie des mit den GENen funktioniert. (-)</p> <p>das GE:N auf unserer dnA: wird also ABgelesen (.) und die informaTIONen AUF diese:m GEN werden dafür geNUTZT, ein proteIN herzustellen (.) zu synthetiSIEren sagt man. (.) und bei geSUNden patienten .../ bei geSUNden menschen (.) wird (.) DIEses protein IN eine SCHLEIMhautzelle eingebaut. die sieht ungefähr !SO! aus. (.) also das protein wird (.) !EIN!gebaut und sorgt dann INdirekt DA:für, dass AUS der zelle (.) ausreichend WASSer gegeben wird und DA:durch (-) der SCHLEI:M eben Relativ FLÜSSig bleibt und durch diese (.) FLIMmerhärchen hier oben WIE am fließband nach ab.../ Abtransportiert werden kann und dann ABgehustet werden kann. (-)</p> <p>bei einer person, die mukoviszidOse hat, (.) SIEHT dieses gen (.) etwas ANders aus. es is verÄNDert. (-) dieses GE:N wird aber DENnoch ABgelesen, ganz normal, (-) UND es entsteht !AUCH! ein proteIN. (.) ein protein wird synthetiSIERT, !A:!ber das sieht anders AUS, WEIL (-) das gen andere informaTIONen liefert. (.) und DIEses veränderte gen (-) - hier analog die SCHLEIMhautzelle - KANN DO:RT eben nicht mehr EINGebaut werden. (.) es übernimmt also nich mehr seine EIgentliche funkTION, weil es eben anders AUSSieht. DAS HEIßt; (.) das WASSer bleibt IN der ZELLE (.) und daraus FOLgend (-) ist der SCHLEIM eben</p>	<p>L schreibt die Überschrift <i>Wie entsteht Mukoviszidose?</i> an Tafel. L heftet das Bild eines DNA-Strangs links an die Tafel.</p> <p>L zeichnet einen Pfeil vom Gen aus nach rechts zu einem weiteren Gen, klebt ein gelbes Oval auf das zweite Gen und klebt ein Bild eines Proteins auf das zweite Gen.</p> <p>L deutet auf das zweite Gen.</p> <p>L zeichnet einen weiteren Pfeil nach rechts, klebt das Schaubild daneben und klebt das Bild eines Proteins und das Wort <i>Wasser</i> über das Schaubild.</p> <p>L heftet das Bild eines zweiten DNA-Strangs links an die Tafel. Zeichnet einen Pfeil vom Gen aus nach rechts zu einem weiteren Gen. L klebt ein gelbes Oval auf das zweite Gen.</p>



<p>64 65 66 67 68 69 70 71 72 73 74 75 76 77</p>	<p>sehr sehr zäh und kann NICHT abtransportiert werden. (.)</p> <p>ja, und solche URsachen zu FINDen, wie hier in DEM fall (.) ein GEN, is für medIZIner und forschER natürlich toTA:L spannend, weil !DA:!durch neue beHANDlungsmethoden und medikaMENTE (.) auf den WE:G gebracht werden. (.) und welche beHANDlungsmethoden es GIBT, des schauen wir uns [...]</p>	<p>L klebt ein Bild eines Proteins auf das zweite Gen und zeichnet einen Pfeil nach rechts.</p> <p>L erstellt dort ein Schaubild der veränderten Zelle.</p> 
--	---	---

G_k: Wie entsteht Multiple Sklerose?


Die Vignette befindet sich auf der beigelegten CD unter dem Dateinamen „Genetik_komplex“.

Nr.	Gesprochener Text	Art der Visualisierung: Tafelbild (mit Wortkarten)
<p>01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37</p>	<p>was der text noch NICHT (.) erklärt hat, is, WIE multiple sklerose EIgentlich entSTEHT, und das wollen wir jetzt geMEINSam noch machen. (---)</p> <p>ja, die multiple skleROse is eine sogenannte autoimM:Ukrankheit und das beDEUTet, dass unser (-) immUNsystem (.) des körpers, des EI:gentlich (.) viren und bakterien UNSchädlich machen soll, DIE zum beispiel (.) über ne WUNde in unseren KÖRper gelangen, auf EINmal EIgene KÖRperzellen (.) angreift, un im falle der multiplen skleROse sind !DAS! eben NERvenzellen. (--)</p> <p>aber dass unser immUNsystem auf EINmal nich=mehr SO funktioniert, wie es SOLL, hat ja !AUCH! wieder eine ursache (-) und !DA:! FORschen (.) mediziner und ÄRZte sehr intenSI:V daran. (-) sie HAM zum beispiel schon einige GE:ne identifiZIERT, die dazu betragen. (--) also nicht nur EIN gen, sondern tatsächlich MEHrere. (-)</p> <p>und um das zu verSTEHen, wie GENE zu=ner KRANKheit beitragen können, müssen wir uns dran erINnern, (.) dass GENE - also (-) ein ABSchnitt auf unserer dnA - informaTIONen enthält, (.) die dann ABgelesen werden</p>	<p>Tafelanschrift: <i>Wie entsteht Multiple Sklerose?</i></p> <p>L heftet einen Zettel mit <i>Immunsystem</i> und einem Blitz an die Tafel.</p> <p>L bringt einen Zettel mit <i>Nervenzellen</i> rechts daneben an und zeichnet einen Pfeil von <i>Immunsystem</i> zu <i>Nervenzellen</i>. L deutet auf <i>Immunsystem</i>.</p> <p>L bringt drei Zettel mit <i>Gen</i> und einem DNA-Strang links neben <i>Immunsystem</i> an und zeichnet einen unterbrochenen Pfeil von den Genen zu <i>Immunsystem</i>.</p> <p>L deutet auf die DNA-Stränge.</p>

<p>38 und (.) aus diesen informaTIONen, (.) 39 ähm, proteine synthetisiert 40 also zuSAMmengebaut werden. 41 und diese proteIne, 42 die aus diesen GENen entSTEHen (.) 43 !SIND! eben am imMUNsystem mit beteiligt. 44 (.) 45 wenn die verÄNDert sind, 46 funktioniert des immunsystem ANders. (-) 47 des sin aber nicht die EINzigen faktoren, 48 die man im blick hat, SONdern (--) 49 50 man hat AUCh schon rausgefunden, 51 dass (.) patienten mit multiplen skleROse 52 sich von geSUNden DARin unterscheiden, (.) 53 dass sie (.) FRÜher bereits Andere 54 krankheiten hatten, (-) 55 dass ihr vitamin-D(.)-spiegel im Körper 56 also wie viel vitamin D im körper is, (.) 57 ANders is 58 U:ND auch ds geWICHT scheint eine rolle zu 59 spielen. (--) 60 aber (.) WIE diese fakTOren geNAU dazu 61 FÜHren, 62 dass das immunsystem nun (.) auf einmal die 63 eigenen NERvenzellen (.) angreift, 64 das hat man (.) noch gar=nich rausgefunden, 65 das WEIß man nicht. 66 !UND! man GEHT auch davon AUS, 67 dass es noch ganz viele WEITere faktoren 68 gibt, 69 die man einfach noch NICHT entDECKT hat. 70 (--) 71 72 ja, für mediZIner un ärzte is=es natürlich 73 !UN!glaublich spannend, 74 diese fakTOren aufzudecken, 75 denn nur SO kann man neue 76 behANDlungsmethoden und medikaMENTe auf den 77 weg bringen. 78 aber im FALL der multiplen skleROse is des 79 gar=nich so EINFach, 80 !WEIL! es eben so viele (.) !UN!bekannte 81 noch gibt 82 also (.) welche faktoren !NOCH! dazukommen 83 !O:!der wie das jetzt geNAU zusammenhängt. 84 (-) 85 ähm, welche behANDlungsmethoden es schon 86 gibt,...//</p>		<p>L deutet auf die drei Zettel links (<i>Gen</i>, <i>DNA-Stränge</i>) und dann auf <i>Immunsystem</i>.</p> <p>L zeichnet einen unterbrochenen Pfeil vom freien Raum unter den Genen zu <i>Immunsystem</i>.</p> <p>L heftet einen Zettel mit <i>Infektionen</i> an die freie Stelle. L bringt unter <i>Infektionen</i> einen Zettel mit <i>Vitamin D</i> an und zeichnet einen unterbrochenen Pfeil von dort zu <i>Immunsystem</i>. L bringt einen Zettel mit <i>Starkes Übergewicht</i> unter <i>Vitamin D</i> an und zeichnet einen unterbrochenen Pfeil von dort zu <i>Immunsystem</i>. L bringt neben <i>Starkes Übergewicht</i> einen Zettel mit einem Fragezeichen an und zeichnet einen unterbrochenen Pfeil von dort zu <i>Immunsystem</i>. L deutet auf das Fragezeichen.</p> <p>L deutet auf die Überschrift.</p> <p>L deutet auf das Fragezeichen.</p> <p>L deutet auf die einzelnen Faktoren.</p>
--	--	---

Ö: Waum gibt es Schwankungen bei der Anzahl der Schneeschuhhasen?

Die Vignette befindet sich auf der beigelegten CD unter dem Dateinamen „Ökologie_linear“.

Nr.	Gesprochener Text	Art der Visualisierung: Tafelbild (mit Wortkarten); Schaubild, welches permanent oberhalb der Tafel sichtbar ist
01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52	<p>un was vielleicht jetz=n BISSchen verBLÜFFT oder (.) erSTAUNlich is; dass gar nicht mal GLEICH viele HASen dort VORKommen, SONdern (.) wir tatsächlich SEHR sehr starke SCHWANKungen haben. wenn ihr KUCKT (.) hier (-) ACHTzehnhundertVIERunsechzig; da ham wir ne ZIEMlichen (.) PEAK, (.) da sind sehr VIELe HASen (.) in diesem gebiet geSICHTet worden. ich hab hier=n BILD dabei (.) für die VIElen hasen, die man gefUNden hat. und nur !WE:nige jahre später, hier, ACHTzehnhundertSIEBzig, gibt es nur noch SEHR, sehr !WE:nige hasen. (.) ja:; (.) und was man sich jetz FRAGen kann, is; (.) wie KOMMT es EIgentlich, dass (-) es zuNÄCHST sehr sehr VIEle hasen gibt, (-) und nur !WE:nige jahre SPäter es nur noch sehr sehr WEnige sind (.) UND (.) genauso UMgekehrt; (.) warum STEIGT eigentlich die ANzahl dann auch WIEder? und genau !DIE! ursache wollen wir uns=n bisschen geNAUer ansehen, diesen GRUND, warum es diese SCHWANKungen gibt. und !DA:!für müssen wir uns den Lebensraum der SCHNEESchuhhasen bisschen geNAUer ansehen (.) und wenn man das !TUT!, dann (.) SIEHT man, dass es dort LUCHse gibt. und LUCHse gehören zu der familie der KATzen und ernähren sich RÄÜberisch, das heißt von FLEISCH (.) und die !HAUPT!nahrungsquelle sind tatsÄCHlich diese SCHNEESchuhhasen. wenn ihr euch jetz VORstellt, was es für einen RÄÜber beDEUTet, wenn er SEHR sehr viele beutetiere VORfindet; (-) naJA (.) er kann sich WUNderbar ernÄHren, das heißt; (.) !WE:nige tiere sterben am HUNgertod, er kann seine NACHkommen sehr sehr gut</p>	 <p>L zeigt auf das Schaubild.</p> <p>L heftet ein Bild mit sechs Hasen an die Tafel.</p> <p>L heftet ein Bild mit zwei Hasen an die Tafel.</p> <p>L zeichnet einen Pfeil von dem Bild mit den vielen zu dem Bild mit den wenigen Hasen. L zeichnet einen Pfeil in die umgekehrte Richtung.</p> <p>L bringt ein Bild mit sechs Luchsen in der Mitte des Pfeils an, der von den vielen zu den wenigen Hasen führt.</p>

<p>53 DURCHbringen, 54 das heißt, (.) 55 er verMEHRT sich sehr stark. (-) 56 57 also die ANzahl der luchse STEIGT. 58 also kann man SA:gen; 59 je mehr BEUte in einem geBIET sind, 60 desto mehr RÄUÜber (.) entSTEHEen dort auch 61 oder kommen DURCH. 62 ja, (.) 63 jetz ham wir hier SEHR sehr viele LUCHse, 64 (.) 65 die !NOCH! mehr HA:sen fressen, 66 ihre BEUte, (.) 67 des HEIßt, 68 die !AN!zahl der hasen wird wieder 69 !AB!nehmen, 70 weil SEHR viele SCHNEESchuhhasen als 71 BEUtetiere enden 72 un wir können ALLgemein sagen; 73 je !MEHR! räuber in dem geBIET sind, 74 desto !WE!:niger beutetiere gibt es. 75 und !DES! is genau diese ABwärtsschwankung, 76 die wir HIER sehen; 77 wenn es !VIE!le beutetiere gibt (.) 78 und die anzahl der LUCHse MITsteigt, 79 dann SINKT die anzahl der hasen AUCH wieder, 80 (.) WEIL=s eben (.) VIEle (.) ähm (.) Hasen 81 gibt, 82 die als BEUtetiere enden. 83 ja, 84 und wenn wir WEIter denken; 85 jetzt gibt es nur noch WE:nig beutetiere, 86 was passiert denn mit den LUCHsen? (.) 87 es werden AUCH weniger werden, (-) 88 denn die LUCHse finden nur noch GANZ schwer 89 NAhrung (.) 90 und DA:mit (.) SINKT die anzahl der RÄUÜber 91 in diesem geBIET 92 und wir können ALLgemein sagen; 93 Je !WE!:niger beutetiere, (.) 94 je weniger HA:sen, 95 desto weniger LUCHse. 96 und im Anschluss (.) der LETZte schritt, 97 ihr könnt=s euch DENken: 98 je !WE!niger luchse es gibt, 99 desto MEHR hasen überLE:ben wieder, 100 es werden NICH so viele gefRESsen. 101 also; 102 Je !WE!:niger räuber, 103 desto !MEHR! beutetiere gibt es. 104 und DES is eben die ABhängigkeit, 105 die ihr dort O:ben seht (.) 106 und die SCHWANKungen. 107 SPANNend wird=s jetz, 108 wenn wir uns überLEgen, 109 was in ZEHN oder zwanzig JAHren mit den.../ 110 mit der ANzahl der HAsen passiert. (.) 111 [denn] HIER sind wir (-) 112 ja, RECHT weit unten, 113 das heißt, 114 ungefähr HIER in dem zyklus 115 und die ANzahl wird über die NÄCHSten jahre 116 hin wieder STEIgen.</p>	<p>L bringt links neben dem Bild eine Wortkarte mit der Aufschrift an: <i>Je mehr Beute, desto mehr Räuber.</i></p> <p>L bringt rechts neben dem Bild eine Wortkarte mit der Aufschrift an: <i>Je mehr Räuber, desto weniger Beute.</i></p> <p>L zeigt auf die Abwärtskurve in dem Schaubild.</p> <p>L heftet ein Bild mit zwei Luchsen in der Mitte des Pfeils an, der von den wenigen zu den vielen Hasen führt.</p> <p>L bringt rechts neben dem Bild eine Wortkarte mit der Aufschrift an: <i>Je weniger Beute, desto weniger Räuber.</i></p> <p>L bringt links neben dem Bild eine Wortkarte mit der Aufschrift an: <i>Je weniger Räuber, desto mehr Beute.</i></p> <p>L deutet auf das Ende des Schaubildes.</p>
---	--

Ök: Warum gibt es Schwankungen bei der Anzahl der Erdkröten?



Die Vignette befindet sich auf der beigelegten CD unter dem Dateinamen „Ökologie_komplex“.

Nr.	Gesprochener Text	Art der Visualisierung: Bilder und Textkarten an der Tafel; Schaubild auf Beamer, das permanent oberhalb der Tafel sichtbar bleibt 
01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51	<p>un was vielleicht n bisschen überRASCHT, (.) ist, dass es NICHT immer GLEICH viele erdkröten in diesem gebiet gibt, !SON!dern eben mal sehr VIELe, wenn ihr hier kuckt; Neunzehnhundert (.)VIERunneunzig oder FÜNFunneunzig gibt es n GANZ großen AUSSchlag. da hat man sehr, sehr VIELe erdkröten gefunden und nur WE:nige jahre später, so zweiTAUsend, sin es DEUTlich weniger. (.) und genau DAMit wollen wir uns jetzt beschäftigen, also (.) wa!RUM! es diese schwankungen gibt, welche (.) U:Rsache oder welche GRÜNde denn da daHINTER stecken, dass es mal MEHR und mal weniger gibt. (.) und DA:zu (.) müssen wir uns (.) die LEbensweise der erdkröte etwas genaUER anschauen. (.) und zwa:r; (-) ihr WISST sicherlich noch: die erdkröte ist ein amPHI:B (.) und beNÖ:tigt deshalb (.) ein gemÄ:Bigtes KLIma, also nicht ZU viele heiße tage und ausreichend FEUCHtigkeit oder NIEderschlag. (.)</p> <p>Das heißt; gibt es !ZU:! viele heiße TAge in einem sommer, würde die anzahl der erdkröten SINKen. (.) s also eine mögliche URsache. (.) AU:Berdem (-) braucht die erdkröte stehende geWÄSser, (.) denn sie ist ein amPHI:B .../ gehört zu den FRöschen oder KRöten, (.) und (.) legt ihre Eier in stehende geWÄSser, aus denen sich dann KAULquappen entwickeln und DANN eben die erwachsenen ERDkröten. gibt es NICHT genügend stehende geWÄSser, (.) SINKT die anzahl an erdkröten (-) beträchtlich. (-) ja, (.) dann braucht die erdkröte natürlich auch NA:hrung und das sind inSEKten, SCHNEcken und WÜRmer.</p>	<p>L deutet bei <i>hier</i> auf die entsprechende Stelle im Schaubild. L zeigt auf die Jahreszahl.</p>  <p>L heftet das Bild einer Erdkröte an die Tafel.</p> <p>L zeichnet eine horizontale Linie vom Bild nach links und heftet eine Gedankenblase mit dem Wort <i>benötigt</i> über die Linie. L heftet an das linke Ende einen Zettel mit dem Begriff <i>GemäBigtes Klima</i>.</p> <p>L heftet einen Zettel mit dem Begriff <i>Stehende Gewässer</i> über den ersten Zettel und verbindet den Zettel durch eine Linie mit der vorherigen. L deutet auf <i>Stehende Gewässer</i>.</p> <p>L heftet einen Zettel <i>mit Nahrung</i> (z. B. <i>Insekten, Würmer, Schnecken, ...</i>) unter <i>GemäBigtes Klima</i> und verbindet die Begriffe wie zuvor mit einer Linie. L hebt die linke Hand.</p>

52	(.)	L lässt die Hand wieder sinken.
53	wenn es davon geNÜgend gibt,	
54	STEIGT die anzahl an erdkröten,	L zieht eine horizontale Linie vom ersten
55	gibt es NICHT genug,	Bild nach rechts und bringt darüber eine
56	SINKT sie,	Gedankenblase mit <i>wird bedroht durch</i> an.
57	weil eben viele verHUNGern. (.)	L heftet einen Zettel mit dem Begriff
58	es gibt aber noch ME:HR faktoren, (--)	<i>Fressfeinde</i> an das rechte Ende der Linie.
59	denn die erdkröte wird auch beDROHT (--)	L deutet auf <i>Fressfeinde</i> und dann
60		auf das Bild der Erdkröte.
61	und zwar zum EInen (-) durch natürliche	
62	FRESSfeinde.	
63	das sind RATten oder größere VÖ:gel.	
64	wenn es davon VIE:le gibt,	L bringt einen Zettel mit <i>Menschen</i> (z. B.
65	werden !NOCH! mehr ERDKröten gefressen	<i>Straßenbau</i>) über <i>Fressfeinde</i> an und
66	und die anzahl der erdkröten wird SINKen.	verbindet ihn mit der rechtet Ausganglinie.
67	(-)	
68	dann sind AUCh wir MENsChen (.) ne !SEHR!	
69	große !UR!sache (.) DA:für,	
70	dass die ANzahl an erdkröten SINKT,	
71	denn,	L deutet auf das Bild der Erdkröte.
72	wenn wir zum beispiel STRAßen bauen	L deutet auf <i>Stehende Gewässer</i> .
73	und damit den !WE:G! abschneiden von ihrem	
74	EIGentlichen lebensraum	
75	und dem (.) geWÄSser,	
76	in dem sie ABlaicht	
77	- in dem sie ihre EIer legt -	
78	dann wird es (.) WENiger erdkröten geben.	L bringt einen Zettel mit <i>Schadstoffe</i> (z. B.
79	(.)	<i>Dünger, Insektizide, ...</i>) unter <i>Fressfeinde</i>
80	und ein (.) hier LETZter faktor sind	an und zeichnet eine Linie von dort zur
81	SCHADstoffe, (-)	rechtet Ausganglinie. L deutet auf
82	häufig aus der LANDwirtschaft,	<i>Schadstoffe</i> .
83	zum beispiel DÜNger oder insektiZi:de.	L deutet auf das Bild der Erdkröte.
84	also, insektiZIde sind mittel !GE:!gen	
85	insekten (.)	
86	und die sind eben AUCh giftig für die	
87	ERDKröte	L deutet auf das erste Diagramm.
88	und reduZIERen die anzahl an erdkröten. (-)	
89	ja, ihr MERKT schon;	
90	des sin schon ganz schön !VIE!le faktoren,	
91	die hier zuSAMmen(.)kommen	
92	und (.) die ANzahl der erdkröten	
93	beEINflüssen können. (-)	
94	u:nd dIEse faktoren tragen eben gemeINSam	
95	dazu bei,	
96	ob die (.) erdkröte GU:te bedingungen (.)	
97	vorfindet	
98	und sich verMEHren kann,	
99	als=zum bespiel,	L deutet auf <i>Nahrung</i> und <i>Gemäßigtes</i>
100	wenn AUSreichend NA:Hrung da is	<i>Klima</i> .
101	und der sommer nicht zu viele HEIße tage	
102	hatte, (-)	
103	a:ber auch (.) dazu BEItragen,	
104	dass es !WE!niger erdkröten gibt,	L deutet erst auf <i>Nahrung</i> ,
105	wenn es zum beispiel zwar ausreichend	dann auf <i>Menschen</i> .
106	NAHrung gibt,	
107	!A:!ber (-) ne neue STRAße gebaut wurde. (.)	
108	und !DAS! macht es ganz schön SCHWIERig, (.)	
109	ABzusehen,	L deutet auf die entsprechende Stelle im
110	warUM denn jetzt gerade (.) bis zweiTAUsend	Diagramm.
111	die anzahl allmählich abgesunken is,	
112	!O:!der (-) es macht es NOCH schwieriger,	
113	zu sagen,	L deutet über das rechte Ende des
114	was in ZEHN oder ZWANzig jahren denn mit der	Diagramms hinaus.
115	ANzahl passieren wird. (.)	
116	also des is GAR=nich so einfach. (.)	
117	und das wollen wir uns jetzt nochmal geNAUer	
118	anschauen.	

Pj: Wie kommt ein Vitamin-A-Mangel zustande?




Die Vignette befindet sich auf der beigelegten CD unter dem Dateinamen „Physiologie_linear“.


Nr.	Gesprochener Text	Art der Visualisierung:
<p>01 okay, Süper. (.) 02 03 04 05 06 also vitamin-A-mangel kann zu 07 WACHstumsstörungen und NACHTblindheit 08 führen, !WEIL! es eben wichtig für den 09 SEHprozess is 10 und an wachstum und erneuerung der HAUT 11 beteiligt is. (.) 12 was wir uns jetzt noch NICH angekuckt haben, 13 is, 14 wie denn überhaupt so=n mangel zuSTANDe 15 kommt, 16 und das möchte=ich euch kurz an der Tafel 17 zeigen. (--) 18 19 recht offenSICHTlich is, 20 denke ich, 21 dass, (.) 22 wenn wir keine vitamine mit der nahrung ZU 23 uns nehmen, (.) 24 n=vitaMINmangel entstehen kann, 25 weil wir eben vitamine nicht selbst 26 HERstellen können, 27 sondern eben mit der NAHRung aufnehmen 28 müssen. 29 und im FALL von vitamin A: sin eben WICHTige 30 lebensmittel, 31 wie ihr schon, ähm, rausgefunden habt; 32 EI:gelb, karotte und spiNA:T. 33 das heißt, (.) 34 das, 35 was wir !ZU! uns nehmen 36 - also die erNÄHRung - (.) 37 is=n ganz wichtiger faktor bei nem, äh .../ 38 bei der erklärung, 39 wie=n vitamin-a-mangel zustande kommt. (.) 40 41 es gibt allerdings noch ne ZWEIte 42 möglichkeit. 43 nämlich, (.) wenn wir diese lebensmittel 44 zwar ESSen, 45 !A:!ber (.) die vitamine gar=nich AUFnehmen 46 können. 47 und DA müssen wir uns die verDAUung noch=n 48 bisschen genauer anschauen. 49 das is der NÄCHste SCHRITT. (--) 50 51 und verDAUung heißt ja, 52 die nahrung zu zerKLEInern 53 und diese KLEI:nen bestandteile der NAHRung 54 dann IN den körper aufzunehmen. (.) 55 und das passiert eben vor ALlem im (.) 56 DÜNNDarm (.) 57 und das is hier=n AUSSchnitt. (--) 58 59 also das is die dünndarmWAND 60 und das is ein ANliegendes BLU:Tgefäß, 61 in das dann die nährstoffe AUFgenommen 62 werden können. (.) 63 und bei vitamin A gibt es nun eine</p>	<p>Art der Visualisierung: Tafelbild mit Wort- und Bildkarten</p> <p>Schaubild, das oberhalb der Tafel mit dem Beamer projiziert ist:</p>  <p>L deutet auf die Spalte <i>Wichtig für ...</i></p> <p>L klebt obiges Bild links an die Tafel.</p> <p>L zeichnet einen Blitz unter das Bild und klebt einen Zettel mit der Aufschrift <i>Ernährung</i> darunter.</p>  <p>L deutet auf das Bild.</p> <p>L klebt folgendes Schaubild links neben die Nahrung und zeichnet einen Pfeil von links nach rechts.</p>	

<p>64 beSONderheit, 65 DENN (-) vitamin A is ein sogenanntes 66 FETTLösliches vitaMIN. 67 das HEIßT, 68 es löst sich WUNderbar in FETT, 69 aber SEHR sehr schlecht in WASSer. (.) 70 nur unser DA:RM is ein wässriges milIEU. 71 das HEIßT, 72 das vitaMIN verTEILT sich nur recht 73 schlecht, 74 bekommt dadurch SEHR wenig kontakt zu 75 unserer DARMSchleimhaut 76 und kann DAdurch auch nur ganz wenig 77 AUFgenommen werden 78 und !VIEL! wird einfach dann ausgeschieden. 79 (.) 80 und !DES!wegen (.) wird vitamin A (.) als 81 FETTLösliches vitamin (-) in einem KLEI:nen 82 fetttröpfchen, 83 das man mizELLE nennt, 84 ähm, transportIERT (.) 85 !UND! da es in fett besser löslich is, 86 bekommt es mehr konTAKT zur darmwand, (.) 87 KANN auch in diese kleinen AUSstülpungen der 88 darmwand geLANGen 89 !UND! es wird mehr in die blutbahn 90 ABgegeben. (.) 91 und wenn mehr ABgegeben werden kann, (-) 92 dann steht auch MEHR (.) für diese 93 WACHstumprozesse, 94 für die es wichtig is, 95 zur verfügung oder E:ben für den SEHprozess. 96 (.) 97 un DA is schon des ZWEite verSTECKT gewesen. 98 nämlich; (.) 99 100 101 102 103 104</p>		<p>L klebt links daneben folgendes Schaubild und zeichnet einen weiteren Pfeil von links nach rechts.</p> <p>L klebt einen blauen Kreis mit der Aufschrift <i>Vitamin A</i> links oben neben den Darmausschnitt.</p> <p>Deutet auf den blauen Kreis, dann auf den Darmausschnitt.</p> <p>L unterlegt den blauen Kreis mit einem größeren gelben.</p>
<p>105 wenn wir zu wenige (.) FETTE aufnehmen, 106 die den transport begünstigen (.) 107 oder die AUFnahme begünstigen, (.) 108 DANN kann eben AUCh=n mangel entstehen. (-) 109 110 111 UND (.) über ne ausgewogene ERnährung 112 - des heißt, (.) 113 vitaMINreich, aber auch (.) f.../ mit FETTen - 114 können wir so=n vitamin-a-spiegel, 115 der geSUND is, 116 ganz gut erreichen. 117 und des is eigentlich auch des 118 nächstestichwort.</p>		<p>L deutet auf die beiden Kreise, dann auf</p>  <p>L zeichnet einen Blitz durch den Pfeil, der vom Darm zum Darmausschnitt führt, und klebt einen Zettel mit der Aufschrift <i>fehlende Fette</i> darunter. Deutet auf das linke Ende des Tafelbilds.</p> 

Pk: Wie kommt ein Folsäuremangel zustande?

Die Vignette befindet sich auf der CD unter dem Dateinamen „Physiologie_komplex“.

Nr.	Gesprochener Text	Art der Visualisierung: Tafelbild mit Wort- und Bildkarten; Tabelle mit Informationen zu Vitamin D und Folsäure, die permanent oberhalb der Tafel sichtbar bleibt:												
<p>01 okay Super. (.) 02 a:also ein FO:lsäuremangel kann zu BLU:Tarmut 03 und fehlbildungen beim baby im MUTterleib 04 führen, (.) 05 !WEIL! es (.) wichtig ist für die bildung 06 roter BLUTkörperchen 07 und die herstellung von dnA. (-) 08 was wir noch NICHT besprochen haben, 09 is, 10 !WIE! es denn eigentlich zu so einem MANGel 11 kommen kann, 12 und des möcht ich euch kurz an der tafel 13 zeigen. (.) 14 15 ja, was wahrscheinlich recht 16 offen!SICHT!lich is, (.) 17 ist, 18 dass, wenn wir uns folsäureARM ernähren 19 - also zum beispiel DIEse lebensmittel 20 !NICHT! zu uns nehmen - (.) 21 ein folsäure!MAN!gel entstehen kann, 22 denn wir können ja vitamine nicht selbst 23 HERstellen 24 und !MÜS!sen sie mit der nahrung aufnehmen. 25 (.) 26 das !HEI:!ßt; 27 die erNÄHrung (--) ist ein ganz 28 !WE!sentlicher faktor, 29 wenn wir verSUCHen zu erklären, 30 WIE es denn zu einem folsäuremangel kam. (.) 31 aber es is nicht die EINzige ursache 32 und DESwegen müssen wir uns noch die 33 weiteren SCHRITte der folsäure im körper 34 anschauen. (.) 35 und (.) der nächste SCHRITT ist eben die 36 verDAUung. (--) 37 also verDAUung heißt ja, die nahrung zu 38 zerKLEInern 39 und die kleineren beSTANDteile dann (.) IN 40 den körper aufzunehmen. 41 und diese AUfnahme 42 des (.) geschieht vor allem HIER im dünndarm 43 (.) 44 und (.) das ist ein AU:sschnitt des 45 dünndarms.--) 46 47 48 49 also hier die dünndarmSCHLEIMhaut 50 un ein anliegendes BLU:Tgefäß, (.) 51 in das DANN zum beispiel die folsäure (.) 52 ÜBERgeht 53 und im KÖRper transportiert werden kann. 54 und (.) n=großer ANteil dieser folsäure (.)</p>		<p>Art der Visualisierung: Tafelbild mit Wort- und Bildkarten; Tabelle mit Informationen zu Vitamin D und Folsäure, die permanent oberhalb der Tafel sichtbar bleibt:</p> <table border="1" data-bbox="997 465 1417 584"> <thead> <tr> <th>Vitamin</th> <th>Wichtig für...</th> <th>z.B. enthalten in...</th> <th>Mangel führt zu...</th> </tr> </thead> <tbody> <tr> <td>Vitamin D</td> <td>Wachst- und Zahnbildung</td> <td>Fische, Eigelb, Pilze, Leber</td> <td>Knochenerweichung, Zahnschmerzen</td> </tr> <tr> <td>Folsäure</td> <td>Bildung roter Blutkörperchen, Herztätigkeit</td> <td>Spinat, Spargel, Tomaten, Spinat, Gemüse</td> <td>Schwäche, Müdigkeit, Blässe</td> </tr> </tbody> </table> <p>L deutet auf die entsprechende Spalte in der Tabelle. L deutet auf die Spalte <i>Wichtig für</i>. L Heftet Bild an die Tafel und deutet darauf.</p>  <p>L Bringt einen Zettel mit <i>Ernährung</i> unter dem Bild an und zeichnet einen Blitz dazwischen.</p>  <p>L bringt Schaubild rechts neben dem Darmmodell an und zeichnet einen Pfeil vom Darmmodell zum Darmausschnitt. L deutet auf die entsprechenden Bestandteile. L heftet einen blauen Kreis mit der Aufschrift <i>Folsäure</i> oben links neben den Darmausschnitt. L bringt Schaubild rechts neben dem Darmausschnitt an und zeichnet einen Pfeil vom Darmausschnitt zur Leber.</p> 	Vitamin	Wichtig für...	z.B. enthalten in...	Mangel führt zu...	Vitamin D	Wachst- und Zahnbildung	Fische, Eigelb, Pilze, Leber	Knochenerweichung, Zahnschmerzen	Folsäure	Bildung roter Blutkörperchen, Herztätigkeit	Spinat, Spargel, Tomaten, Spinat, Gemüse	Schwäche, Müdigkeit, Blässe
Vitamin	Wichtig für...	z.B. enthalten in...	Mangel führt zu...											
Vitamin D	Wachst- und Zahnbildung	Fische, Eigelb, Pilze, Leber	Knochenerweichung, Zahnschmerzen											
Folsäure	Bildung roter Blutkörperchen, Herztätigkeit	Spinat, Spargel, Tomaten, Spinat, Gemüse	Schwäche, Müdigkeit, Blässe											

<p>55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122</p>	<p>gelangt zuNÄCHST mal in die LE:ber, denn !DORT! (.) kann sie ZWischengespeichert werden (-) und DANN bei beDARF dann ehen auch wieder FREIgegeben werden, ZUM beispiel, wenn BLUT gebildet wird O:der WACHStumsprozesse stattfinden. (-)</p> <p>ja, und wenn man sich dieses ganze system den ganzen KÖRper ansieht, dann entdeckt man schon noch EINige weitere MÖGlichkeiten, wie so ein MANGel erklärt werden kann. (-) zum beispiel HIER bei der verDAUung. (-) WENN wir eine geSTÖRte verdauung haben, also zum beispiel durch einen magen-DARM infekt O:der wenn wir zu viel STRESS haben, denn auch die stresshorMO:ne beeinflussen unsere verdauung, DANN kann nich mehr genug FOLSäure aufgenommen werden. (.) oder !HIER!, (-) direkt beim Ü:bergang der folsäure in den blutkreislauf. (.) da wechselwirken relativ VIELe medikaMente, !ZUM! beispiel die pille zur verHÜTung. (.) oder hier bei der LEber. (-) wenn die leber gerade damit beSCHÄFTigt is, ein GIFT, zum beispiel alkohol, abzubauen, !DANN! funktioniert des SPEIchern nicht mehr so gut !UND! auch die ABgabe. (.) und hier zuLETZT, (-) WENN wir einen erhöhten beDARF haben.</p> <p>WENN also gerade SEHR viel BLUT gebildet werden muss O:der gerade ein WACHStumsschub stattfindet, dann reicht uns so=ne norMA:le ernährung gar=nich mehr aus, sondern wir müssen besonders VIEL folsäure zuführen. (.) ja und ihr SEHT schon; NEben der ernÄHrung gibt es noch ganz viele ANDere faktoren, DIE einen mangel erklären können. (.) und (.) selbst (.) diese VIElen faktoren .../ des war auch nur eine !AUS!wahl von mir. des heißt, es gibt noch viele WEITere. (.) was daZUKommt, is, dass=n MANGel meist nicht nurch EI:nen isoLIERTen faktor (.) erklärt werden kann, SONdern es (.) HÄUfig ne kombinaTION von vielen ist. (.) und DES führt uns eigentlich auch zu unserem nächsten...//</p>	<p>L bringt Schaubild rechts neben der Leber an und zeichnet einen Pfeil von der Leber zum Schaubild.</p>  <p>L zeichnet Darmmodell und bringt einen zettel mit <i>gestörte Verdauung</i> an.</p> <p>L zeichnet einen Blitz durch den Pfeil vom Darmmodell zum Darmausschnitt</p> <p>L bringt einen Zettel mit <i>Medikamente</i> an. L deutet auf die Leber und zeichnet daneben einen Blitz. L bringt einen Zettel mit <i>Alkoholkonsum</i> an.</p> <p>L zeichnet einen Blitz unter den letzten Pfeil und bringt einen Zettel mit <i>erhöhter Bedarf</i> an. L deutet auf <i>Blut</i>.</p> <p>L deutet auf <i>Ernährung</i>.</p> <p>L deutet auf die anderen Faktoren.</p> <p>L legt die rechte Hand auf die Tafel. L zeigt einen Finger.</p> <p>L deutet mehrmals auf die Tafel.</p>
---	--	--

B) Kategorienhandbuch

Das nachfolgende Kategorienhandbuch mit den Codierhinweisen ist in MAXQDA generiert worden und war über die Funktion „Memo“ neben dem Kategoriensystem pro Kategorie einsehbar.

Allgemeine Anmerkung zum Vorgehen bei der Codierung:

Aufgrund der vielen Subkategorien ist es sinnvoll, zunächst den Bezug zu einer Oberkategorie herzustellen.

Anmerkung zur Beschreibung der Kategorien:

Die Äußerungen von Teilnehmenden können positiv oder negativ formuliert sein (z. B. adressatengerechte Erklärung/die Erklärung wurde den Adressat:innen nicht gerecht). Die Beschreibungen beziehen sich für eine Erhöhung der Lesbarkeit nur auf die positiven Formulierungen.

1. Adressatenorientierung

1.1. Adressatenorientierung global

Inhaltliche Beschreibung:

In „Adressatenorientierung global“ werden ausschließlich globale Aussagen gesammelt, die keine spezifische Anpassung an die Adressat:innen beispielsweise hinsichtlich des Vorwissens oder Vortellungen den Blick nehmen.

Beispiele für die Anwendung des Codes:

- „schülergerechte Erklärung“ (Stu – NAAN13 – a)
- „nicht auf Schüler bezogen (eher auf Ärzte) (Stu – MOAL18 – d)

1.2. Didaktische Reduktion

Inhaltliche Beschreibung:

Der Begriff der didaktischen Reduktion wird in der Biologiedidaktik auf fünf Ebenen beschrieben:

- f. auf sprachlicher Ebene,
- g. auf Ebene der Darstellungen,
- h. auf Ebene der naturwissenschaftlichen Arbeitsweisen,
- i. auf sektoraler und
- j. auf struktureller Ebene

Unter dieser Kategorie sollen globale Äußerungen mit direktem Bezug zur didaktischen Reduktion („Die Erklärung ist angemessen didaktisch reduziert“) sowie Aussagen in Bezug auf die inhaltliche Eingrenzung (also sektorale und strukturelle didaktische Reduktion) gefasst werden. Äußerungen, die unter dieser Kategorie subsumiert werden, können aufgrund von dafür notwendigem Professionswissen nur von Personen der Metaperspektive (Studierende, Lehrkräfte und Didaktiker:innen) kommen.

Beispiele für die Anwendung des Codes:

- „sehr starke didaktische Reduktion“ (Did – MAPA898 – a)
- „Informationsgehalt entsprechend eingegrenzt“ (Did – BEHE12 – c)
- „m. M. nach unzulässige Vereinfachung Transkription / Translation“ (Leh – SOFR10 – a)
- „sehr gelungene Reduktion auf das Wesentliche“ (Leh – MAED33 – a)

Abgrenzung zu anderen Kategorien:

Äußerungen, die sich explizit darauf beziehen, dass das natürliche Phänomen durch die sektorale Reduktion nicht mehr angemessen in der Erklärung repräsentiert wird, werden unter „Fach – Phänomenorientierung“ codiert.

Äußerungen, bei denen sich Teilnehmer:innen darauf beziehen, dass nichts Unnötiges, sondern nur Wesentliches für die Sache in der Erklärung aufgegriffen wurde und damit nicht der Prozess des sektoralen und strukturellen Reduzierens im Vordergrund steht, werden unter „Struktur – Fokussierung“ codiert.

1.2.1. Passung des Informationsgehalts

Inhaltliche Beschreibung:

Die Erklärung kann als ausführlich oder detailliert beschrieben werden. Häufig werden auch weitere Informationen gewünscht oder einzelne Aspekte hätten genauer erklärt werden sollen. Die Beurteilung hängt nicht nur vom Fachgegenstand oder von den Adressat:innen, sondern im Sinne der didaktischen Reduktion von beiden Seiten ab.

Beispiele für die Anwendung des Codes:

- „detailliert“ (SuS – SIAN10a – b)
- „oberflächlich“ (Leh – IDER37 – d)
- „keine Vertiefung“ (Leh – INPE49 – b)

1.2.2. Passung des Schwierigkeitsgrads

Inhaltliche Beschreibung:

Mit diesem Code sollen allgemeine Aussagen zum Schwierigkeitsgrad der Erklärung codiert werden. Auch darunter zu fassen sind beispielsweise Aussagen über den Abstraktionsgrad in Bezug auf die Fähigkeiten von Schüler:innen dieser Altersstufe.

Beispiele für die Anwendung des Codes:

- „Schön erklärt, aber Niveau für eine 11. Jgst vielleicht etwas zu niedrig angesetzt.“ (Leh – CHMI75 – a)
- „11. Klasse wird ein wenig unterfordert“ (Stu – MAWE12 – d)
- „einfach erklärt“ (SuS – ANAN23 – a)

1.3. Berücksichtigung des Vorwissens

Inhaltliche Beschreibung:

Mit dieser Kategorie werden Segmente codiert, die sich auf die Vorwissensangepasstheit der fachlichen Erklärinhalte durch die erklärende Person beziehen, also darauf, inwieweit das Vorwissen der Schüler:innen in Bezug auf den Erklärgegenstand/das Phänomen korrekt eingeschätzt wurde.

Beispiele für die Anwendung des Codes:

- „relativ viel Hintergrundwissen wird vorausgesetzt“ (Leh – INPE49 – c)
- „auch auf Grundwissen wurde nochmals eingegangen“ (SuS – ANAN23 – c)
- „jedoch waren teilweise Dinge dabei, wie zum Beispiel die Flimmerhärschen, die nicht genauer erklärt wurden“ (SuS – CHKA89 – a)
- - „Hab vorher schon was über das Thema gewusst und da hat mir ihre Erklärung irgendwie nicht ganz so geholfen“ (SuS – ANIN13 – c)

Abgrenzung zu anderen Kategorien:

Auch die Berücksichtigung des sprachlichen Vorwissens ist ein relevantes Kriterium. Dies wird unter „Verständliche Sprache“ bzw. der entsprechenden Subkategorie gefasst.

1.3.1. Übererklärung

Inhaltliche Beschreibung:

Wenn das Vorwissen unterschätzt wird und die erklärende Person zu weit ausführt, was eigentlich bekannt ist oder leicht von den Schüler:innen selbst erschlossen werden kann, gilt eine Erklärung als „übererklärt“. Sie unterfordert ihre Zuhörer:innen und ist dadurch nicht aufmerksamkeitsgenerierend, sodass ein Lernen erschwert wird und Schüler:innen mitunter verärgert oder frustriert werden. Es werden nur Äußerungen von Schüler:innen mit diesem Code versehen.

Beispiele für die Anwendung des Codes:

- „sehr gut verständlich, aber das Thema kann man sich auch selbst erschließen. Diese Erklärung ist eher für Grundschüler (SuS – SAST36a – e)
- „wiederholung und sehr langsames Reden sorgt für Langeweile, man kommt sich (auch teilweise bei vorherigen Videos) vor als wäre man schwer von Begriff (SuS – EVJO38 – b)

Abgrenzung zu anderen Kategorien:

Äußerungen, die sich lediglich auf eine Unterschätzung des Vorwissens beziehen, ohne, dass Schüler:innen emotional /verärgert äußern, dass sie die Erklärung selbst als nutzlos empfinden oder denken, die Lehrkraft „verkauft sie für dumm“, werden mit „Adressatenbezug – Berücksichtigung des Vorwissens“ codiert.

Die Abgrenzung zur Kategorie „Schwierigkeitsgrad“ erfolgt über die Person, die die Äußerung vornimmt – nur Studierende, Lehrkräfte und Didaktiker:innen können die Übererklärung nicht mit einer (emotionalen) Betroffenheit einschätzen. Schüler:innen, deren Aussagen ebenfalls keine Abwertung oder einen negativen emotionalen Bezugspunkt aufweisen, werden mit dem Code Schwierigkeitsgrad versehen.

1.3.2. Erklärtempo

Inhaltliche Beschreibung:

Mit dieser Kategorie werden Segmente codiert, wenn sich Teilnehmer:innen auf die Dichte der Erklärung beziehen. Sie wird als unangemessen beurteilt, wenn zu viele Inhalte/Informationen in zu kurzer Zeit aufgegriffen oder zu wenige Inhalte auf zu lange Zeit hin ausgedehnt werden. Entscheidend ist der Aspekt der Zeit für die Zuordnung.

Beispiele für die Anwendung des Codes:

- „Lehrkraft erklärt langsam“ (Stu – MOAL18 – a)
- „Evtl. Abstraktionsebene vom Gen/Portein/zur Funktion im Immunsystem zu schnell“ (Did – MOBE36 – d)
- „zu viel auf einmal (SuS – BERO 16 – c)

Abgrenzung zu anderen Kategorien:

Äußerungen, die sich auf das Erklären von zu vielen Details beziehen, werden unter „Struktur – Fokussierung auf Wesentliches“ codiert.

1.4. Berücksichtigung von (Fehl-)Vorstellungen

Inhaltliche Beschreibung:

Mit dieser Kategorie werden Äußerungen codiert, die sich darauf beziehen, dass die erklärende Person mögliche Fehlvorstellungen von Schüler:innen (nicht) berücksichtigt hat. Das heißt, gezielt in der Erklärung auf mögliche Fehlvorstellungen eingegangen oder / und sie durch die Ausführungen nicht bestärken.

Beispiele für die Anwendung des Codes:

- „fachlich kritisch: Gen wird vergegenständlicht --> Gen sieht anders aus --> fördert fachlich nicht korrekte Alltagsvorstellungen (z. B. Gene als Gegenstände / Dinge; verändertes Gen = veränderter Phänotyp = krank)“ (Did – ELER15 – a)
- „im Gegenteil es scheint graphisch größer und somit schwerer verteilbar--> ggf. unangemessene Vorstellung“ (Did – ELER15 – c)
- „+ Gen wird kontextualisiert; Gene wirken im System, nicht allmächtig“ (Did – ELER15 – d)

1.5. Interessante und motivierende Gestaltung

Inhaltliche Beschreibung:

Mit dieser Kategorie werden alle Äußerungen codiert, die sich auf die interessante Aufbereitung des Inhalts sowie die Interessantheit des Inhalts selbst für die Schüler:innen beziehen. Auch Äußerungen zur Erzeugung von Motivation, sich weiterhin damit zu beschäftigen, werden hiermit codiert sowie allgemeinere Aussagen

Beispiele für die Anwendung des Codes:

- „hat nicht die Motivation angeregt“ (Stu – BRUD10 – d)
- „Erklärung im allgemeinen allerdings etwas trocken“ (SuS – BAFE23 – d)

Abgrenzung zu anderen Kategorien:

Das Adjektiv monoton kann sowohl für die Gestaltung der Erklärung in inhaltlicher Sicht als auch für die Gestaltung des Sprechausdrucks verwendet werden. Hier ist zu differenzieren – Äußerungen, die sich auf die stimmliche Gestaltung beziehen, werden unter der Kategorie „Prosodische Merkmale – Variation der Stimme“ gefasst.

1.6. Lebensweltbezug

Inhaltliche Beschreibung:

Mit dieser Kategorie werden alle Segmente codiert, die der Erklärung oder den darin verwendeten Beispielen, Bildern und Modellen etc. einen Lebenswelt- bzw. Alltagsbezug zuschreiben oder diesen bemängeln.

Beispiele für die Anwendung des Codes:

- „+ Anknüpfen an Immunsystem --> Anknüpfen an Ebene, die SuS bekannt sein sollte (Lebensweltbezug)“ (Did – ELER15 – d)
- „Erklärung wofür man diesen Stoff im alltäglichen Leben braucht“ (SuS – MAWI3 – a)
- „Außerdem fehlt hier eine Motivation für was man dieses Wissen im Alltag benötigt oder welche Menschen mit diesem Wissen forschen bzw. arbeiten.“ (SuS – MAWI3 – b)

1.7. (Aktiver) Einbezug der Rezipient:innen

Inhaltliche Beschreibung:

Mit diesem Code werden alle Segmente codiert, die sich auf das Unterrichtsprinzip der Schüleraktivität beziehen, also den Einbezug der Schüler:innen in die Konstruktion der Erklärung fordern.

Beispiele für die Anwendung des Codes:

- „Zwar im Sinne der Erklärung, aber in Realität Problematisierung sowie gesamter Erarbeitungsprozess als Erklärung?“ (Did – MOBE36 – e)
- „man könnte als Schüler kaum mitarbeiten“ (SuS – GUDA10 – d)
- „durch Fragen denkt man mit, die zwischendurch gestellt wurden“ (SuS – TAGE15 – e)

2. Strukturiertheit

Unter dieser Hauptkategorie sollen alle Aspekte der Strukturiertheit einer Erklärung zusammengefasst werden. Durch die Merkmale der Fokussierung auf Wesentliches, eines logischen Aufbaus, eines schrittweisen Vorgehens, der Zusammenfassung wesentlicher Punkte sowie der Klarheit des Erklärgegenstandes im Sinne einer Zielangabe werden die Rezipient:innen durch die Erklärung gewissermaßen geleitet. Es wird Orientierung im Stoffgebiet geschaffen, sodass Zusammenhänge deutlich werden können.

2.1. Strukturiertheit global

In „Strukturiertheit global“ werden ausschließlich globale Aussagen gesammelt, die keine spezifische Facette der Strukturiertheit adressieren, sondern sich holistisch auf die gelungene oder nicht gelungene Strukturiertheit beziehen.

Beispiele für die Anwendung des Codes:

- „Die Erklärung ist sehr gut Strukturiert“ (Stu – GARA38 – d)
- „inhaltlich strukturiert erklärt“ (Did – BEHE29 – a)

2.2. Fokussierung auf Wesentliches

Inhaltliche Beschreibung:

Mit diesem Code werden Segmente codiert, die sich darauf beziehen, dass sich die Erklärung auf das Wesentliche, Nötige oder Wichtige konzentriert.

Beispiele für die Anwendung des Codes:

- „[...] aber zu viel drum herum geredet“ (SuS – ANMA27 – c)
- „Das Thema wurde ohne unnötigen Ausschweifungen auf den Punkt gebracht.“ (SuS – EVMI30 – a)
- „zu weit ausgeholt (Ampibien usw.)“ (Stu – ANAN85 – b)

Abgrenzung zu anderen Kategorien:

Segmente, die sich auf den Aspekt der Auswahl bzw. Reduktion der Inhalte beziehen, werden unter „Adressatenbezug – Didaktische Reduktion“ codiert.

2.3. Logischer Aufbau

Inhaltliche Beschreibung:

Hiermit werden Segmente codiert, die die Ordnung der Teilschritte der Erklärung als sinnvoll bzw. nicht sinnvoll beschreiben. Der Aufbau gewährleistet Stringenz und bildet den Kernaspekt von Strukturiertheit. Oft wird dies als roter Faden beschrieben.

Beispiele für die Anwendung des Codes:

- „Schritt für Schritt [...] erklärt, sodass man gedanklich gut folgen und Kausalkette nachvollziehen kann“ (Did – MAWE02 – a)
- „Diagramm sollte erst nach der Einführung zu Ansprüchen/Bedrohungen kommen“ (Leh – SOFR10 – b)
- „aber die Erklärung war manchmal durcheinander“ (SuS – NANA23 – b)

2.4. Schrittweise Erklärung

Inhaltliche Beschreibung:

Mit diesem Code werden Äußerungen versehen, die sich auf den schrittweisen Aufbau der Erklärung beziehen. Auch Äußerungen die hervorheben, dass die Schritte innerhalb der Erklärung erkennbar oder dem Verständnis zuträglich sind, werden hiermit codiert.

Beispiele für die Anwendung des Codes:

- „Schritt für Schritt und anschaulich erklärt, sodass man gedanklich gut folgen und Kausalkette nachvollziehen kann“ (Did – MAWE02 – a)
- „schritt für schritt erklärt“ (SuS – IRKL11 – e)

2.5. Zusammenfassung zentraler Punkte

Inhaltliche Beschreibung:

Zur Strukturiertheit einer Erklärung zählt auch das gezielte Zusammenfassen von zentralen Punkten (z. B. zum Ende der Erklärsequenz).

Beispiele für die Anwendung des Codes:

- „hier fehlte allerdings die Zusammenfassung!“ (Stu – ELIL12 – c)

2.6. Zielangabe

Inhaltliche Beschreibung:

Ein Segment wird mit diesem Code versehen, wenn sich die Äußerung auf die Klarheit der Zielangabe bzw. die genaue oder ungenaue Bestimmung des Erklärgegenstandes bezieht. Es geht dabei um das Vorhandensein einer Zielangabe.

Beispiele für die Anwendung des Codes:

- „Was genau soll hier gezeigt werden? Ist die Erkenntnis, dass eine Erklärung der Populationsgröße der Erdkröte schwer ist, der zentrale Punkt?“ (Did – ANTH74 – b)

Abgrenzung zu anderen Kategorien:

Bezieht sich die Äußerung darauf, dass das vorhandene bzw. wahrgenommene Erklärziel nicht mit dem darauffolgenden Erklärinhalt zusammenpasst, wird das mit der dieser Kategorie untergeordneten Kategorie „Übereinstimmung von Erklärziel und Erklärinhalt“ codiert.

2.6.1. Übereinstimmung von Erklärziel und Erklärinhalt

Inhaltliche Beschreibung:

Segmente werden mit dieser Kategorie codiert, wenn sich Äußerungen auf die Übereinstimmung von Erklärziel (z. B. in Form einer Frage, Problemstellung etc.) und den Inhalten in der Erklärung beziehen.

Beispiele für die Anwendung des Codes:

- „Anfangsfrage wurde nicht beantwortet“ (SuS – ULHE15 – b)
- „aber es wurde nicht erklärt, weshalb die Anzahl der Erdkröten dadurch sinkt.“ (SuS – EDUD31 – b)
- „irgendwie ist die frage nicht wirklich beantwortet worden“ (Stu – NAAN 13 – d)

3. Repräsentationen

Unter dieser Hauptkategorie sollen alle Aspekte der Erklärung verstanden werden, die bei den Rezipient:innen mentale Bilder und sachgerechte Vorstellungen vom Erklärgegenstand entstehen lassen können. Beispielsweise kann dies der Lehrkraft durch die Verwendung von Metaphern, Analogien, Vergleiche, Realien und Bilder gelingen. Diese müssen von der Lehrkraft sinnvoll in die Erklärung eingebettet werden, um dem Verständnis zuträglich zu sein. Gegebenenfalls werden diese selbst zu Erklärgegenständen.

3.1. Repräsentationen global

In „Repräsentationen global“ werden ausschließlich globale Aussagen gesammelt, die keine spezifische Facette der Repräsentationen adressieren, sondern sich holistisch auf den gelungenen oder nicht gelungenen Einsatz von Repräsentationen beziehen. Weiterhin werden Äußerungen damit codiert, die global auf die gute Veranschaulichung innerhalb der Erklärung verweisen.

Beispiele für die Anwendung des Codes:

- „sehr gut veranschaulicht“ (Did – FRGE13 – b)
- „aber man könnte es noch weiter veranschaulichen“ (SuS – ESJO22 – f)

3.2. Einsatz bildlicher Visualisierung

Inhaltliche Beschreibung:

Der Code wird vergeben, wenn in Äußerungen auf das Vorhandensein bildlicher Repräsentationen verwiesen wird, die den Erklärinhalt visualisieren. Es muss klar werden, dass die Erklärung nicht allgemein anschaulich ist, sondern das Vorhandensein einer Visualisierung einen Beitrag zur Bewertung geleistet hat.

Beispiele für die Anwendung des Codes:

- „gut, dass es an bildern veranschaulicht wurde, man kann es sich besser vorstellen“ (Sus – SAFR9 – c)
- „Erklärung wird mit Abbildungen unterstützt, die nach und nach an Tafel angebracht werden. Dadurch kann der Erklärung sehr gut gefolgt werden.“ (Leh – GEUL11 – a)
- „Gut mit Bildern unterlegt“ (Stu – MATH01 – a)

Abgrenzung zu anderen Kategorien:

Entscheidend für die Zuordnung ist, dass eine konkrete Nennung bzw. ein Bezug zu konkreten Visualisierungen vorkommt (z. B. Bilder, grafische Darstellung ...). Codiert wird nicht mit diesem Code, wenn allgemein eine Anschaulichkeit der Erklärung als globale Aussage genannt wird („Globale Aussagen – Anschaulich“).

3.2.1. Erklärbedürftigkeit der Grafik

Inhaltliche Beschreibung:

Ein Segment wird dieser Kategorie zugeordnet, wenn sich die Äußerung darauf bezieht, dass es weitere Erklärungen zur Veranschaulichung selbst oder Teilen dieser geben hätte müssen, um ein Verstehen zu fördern, oder wenn – invers dazu – diese Erklärungen gegeben wurden und es positiv Erwähnung findet.

Beispiele für die Anwendung des Codes:

- „ich finde, man hätte die Achsen des Diagramms erklären müssen: X-Achse - Jahr, Y-Achse - Anzahl der Erdkröten“ (Leh – INWO07 – b)
- „Die Bedeutung des Blitzes = Ursache für Vitamin A-Mangel könnte noch erwähnt werden.“ (Did – KAGÜ08 – c)

3.2.2. Fachliche Angemessenheit

Inhaltliche Beschreibung:

Codiert werden Segmente, die sich auf die fachliche Angemessenheit der Visualisierung beziehen. Das schließt auch Äußerungen zur Begünstigung bzw. Verhinderung fachlich falscher Vorstellungen ein.

Beispiele für die Anwendung des Codes:

- „Unterschied Aufnahme Vitamin A ohne Fettröpfchen wird zwar erläutert; Visualisierung zeigt aber nicht, dass Vit. A in Fettröpfchen besser aufgenommen wird/sich besser verteilt“ (Did – ELER15 – c)
- „(Optisch ist die Micelle sehr groß, um zwischen die Zotten zu gelangen.)“ (Did – KAGÜ08 – c)
- „Gute Anschauung für ökologische Interdependenzen“ (Did – MAWE02 – b)
- „Durch die gestrichelten Pfeile und die Erläuterungen wird ein Kausalzusammenhang suggeriert“ (Stu – CHNO01 – d)

3.2.3. Sichtbarkeit der Visualisierung

Inhaltliche Beschreibung:

Code wird verwendet, wenn sich Äußerungen auf die Sichtbarkeit der Visualisierungen, sowohl der Schrift als auch der Bilder und eingesetzten Farben, für die gesamte Klasse beziehen.

Beispiele für die Anwendung des Codes:

- „– Abbildungen zu klein – beachte Sichtbarkeit in den hinteren Reihen“ (Stu – JOAN04 – a)
- „aber das mit dem Beamer projizierte konnte man überhaupt nicht lesen.“ (Sus – GUDA10 – c)

3.2.4. Verständlichkeit der Visualisierung

Inhaltliche Beschreibung:

Mit dieser Kategorie werden Segmente codiert, die auf Gestaltungskriterien für Visualisierungen verweisen, die zum Beispiel ein Entnehmen von Informationen und Lernen aus bzw. mit der Visualisierung erleichtern. Das beinhaltet zum Beispiel die Anordnung der Teile in Leserichtung, die Beschriftung der Teile in räumlicher Nähe zu diesen, das Hervorheben wesentlicher Informationen.

Beispiele für die Anwendung des Codes:

- „keine klare Beschriftung der Teile“ (Did – INMI35 – a)
- „Beschriftung der weißen Pfeile fehlt“ (Sus – CHRO28 – f)
- „es war schon ganz ok veranschaulicht aber zum einen finde ich wäre in dem fall eine Tabelle oder eine Kennzeichnung mit plus und minus übersichtlicher gewesen“ (Sus – TAST6 – b)

3.2.5. Ästhetische Gestaltung

Inhaltliche Beschreibung:

Unter dieser Kategorie werden Segmente gesammelt, die die eingesetzten Visualisierungen als (un-)ästhetisch oder (un-)ansprechend und Ähnliches beschreiben.

Beispiele für die Anwendung des Codes:

- „die Visualisierung war noch attraktiver“ (Did – KAGÜ08 – d)
- „weniger laminiertes Papier Zeichnungen an die Tafel mit bunter Kreide“ (SuS – DAAN8 – d)

Abgrenzung zu anderen Kategorien:

Wenn die farbliche Gestaltung zur besseren Unterscheidung von Unterschieden angesprochen wird, wird das Segment nicht mit diesem Code, sondern mit „Repräsentationen – Einsatz bildlicher Visualisierungen – Verständlichkeit der Visualisierung“ codiert.

3.3. Einsatz von Analogien, Metaphern u. Vergleichen

Inhaltliche Beschreibung:

Ein Segment wird mit diesem Code codiert, wenn es sich auf den Einsatz von Analogien, Metaphern oder Vergleichen bezieht. Die sprachlichen Mittel können eingesetzt werden, um bei den Schüler:innen den Aufbau mentaler Bilder zu unterstützen.

Abgrenzung zu anderen Kategorien:

Mit Vergleichen sind hier explizit solche außerhalb des Faches, also beispielsweise aus der Lebenswelt der Schüler:innen, gemeint. Davon abzugrenzen ist die hier untergeordnete Kategorie „Themenspezifisch: Vergleich bei Mukoviszidose“ – hier wird innerhalb der Erklärung zur Entstehung von Mukoviszidose ein Vergleich zwischen gesunden und erkrankten Menschen gezogen. Der Vergleich erfolgt somit innerhalb des Faches.

3.3.1. Themenspezifisch: Vergleich bei Mukoviszidose

Inhaltliche Beschreibung:

Im Video zur Erklärung der Entstehung von Mukoviszidose wird ein innerfachlicher Vergleich zwischen gesunden und erkrankten Personen herangezogen. Segmente, die diesen Vergleich adressieren, werden unter dieser Kategorie gesammelt.

Beispiele für die Anwendung des Codes:

- „Gegenüberstellen von Prozess bei Muko-Erkrankten und gesunden Menschen“ (Did – ELER15 – a)
- „gut mit Vergleich bei einem gesunden Menschen“ (SuS – STMA6 – a)

Abgrenzung zu anderen Kategorien:

Vergleiche von fachlichen Inhalten mit extrafachlich Ähnlichem – im Sinne einer Übertragung – werden mit der Kategorie „Repräsentationen – Analogien, Metaphern und Vergleiche“ codiert.

3.4. Einsatz von Beispielen

Inhaltliche Beschreibung:

In dieser Kategorie werden Segmente gesammelt, die sich auf den Einsatz von Beispielen beziehen.

Beispiele für die Anwendung des Codes:

- „gutes Beispiel mit den Räubern und der Beute“ (SuS – DOMA40 – e)
- „mit Beispielen“ (Stu – CAGE02 – f)

Abgrenzung zu anderen Kategorien:

Nicht mit diesem Code werden Segmente codiert, die sich auf die ausreichende Einbindung oder Ausführung des Beispiels in die Erklärung beziehen („Repräsentationen – Beispiele – Ausreichende Bezugnahme auf das Beispiel“).

3.4.1. Ausreichende Bezugnahme auf das Beispiel

Inhaltliche Beschreibung:

Mit diesem Code werden Segmente versehen, die sich auf die Bezugnahme des eingeführten Beispiels bzw. der konkreten Daten (z. B. im Diagramm) beziehen. Sie kann ausreichend oder nicht ausreichend sein.

Beispiele für die Anwendung des Codes:

- „immer wieder ein Rückbezug zur Graphik“ (Stu – JOAN04 – e)
- „Erklärung gut aber der Zusammenhang zum Diagramm muss noch mehr erklärt werden. [...]“ (SuS – MAWI3 – b)
- „Schwankungen nicht direkt am Beispiel erklärt“ (Did – HEMA24 – b)

3.5. Gelungene Verbindung der Repräsentationen

Inhaltliche Beschreibung:

In dieser Kategorie werden Segmente gesammelt, die sich darauf beziehen, wie gut die gewählten Repräsentationsformen von der Lehrkraft innerhalb der Erklärung miteinander verbunden wurden. Auch die Einbindung der Repräsentationsformen in das Gesagte muss stimmig sein und für Zuhörer:innen einen Mehrwert zum Verstehen bieten. Die Sprache ist auch eine Repräsentationsform.

Beispiele für die Anwendung des Codes:

- „gut: Repräsentationsformen systematisch aufeinander bezogen (Übersetzung von Diagramm mit linearer Zeitachse in zyklisches Diagramm)“ (Did – EVWO09 – e)
- „Klare Zuordnung Abbildungen zu Diagramm“ (Leh – GEUL11 – e)
- „Dopplung der Erklärung auf mehreren Kanälen (Liniendiagramm sowie Fussdiagramm) sowie immer wieder Rückbezug zwischen beiden Darstellungen“ (Did – MOBE36 – e)
- „Abbildungen unterstreichen Gesagtes“ (Leh – GEUL11 – c)

4. Sprech- und Körperausdruck

Die Oberkategorie bringt alle Äußerungen zusammen, die sich auf den nonverbalen Anteil der Kommunikation (in Abgrenzung zu den sprachlichen/verbalen Anteilen der Kommunikation) beziehen. Der nonverbale Anteil setzt sich aus paraverbalen und extraverbalen Mitteln zusammen bzw. kann durch diese beschrieben werden: Paraverbale Mittel können durch den auditiven Kanal wahrgenommen werden. Sie

sind wiederum gliederbar in artikulatorische (=Artikulation) und prosodische Merkmale (= z. B. Stimmklang, Tempo, Sprechpausen). Die extraverbalen Mittel werden über den visuellen Kanal aufgenommen und umfassen allgemein den Körperausdruck (Körperhaltung und -bewegung, Gestik, Mimik, Blickrichtung und Blickkontakt).

4.1. Artikulation

Inhaltliche Beschreibung:

Mit diesem Code werden Segmente versehen, die sich auf die Aussprache/Artikulation der Person beziehen.

Beispiele für die Anwendung des Codes:

- „gute Artikulation“ (Leh – ROAN05 – d)
- „deutlich gesprochen“ (SuS – EVJO38 – d)

4.2. Prosodische Merkmale

Inhaltliche Beschreibung:

Hierunter werden Segmente gefasst, die sich auf angenehme oder nicht angenehme Merkmale der Stimme beziehen, jedoch kein spezifisches Merkmal nennen. Spezifische Merkmale sind der Stimmklang, die Sprechgeschwindigkeit, die Sprechpausen und die Variation/Modulation der Stimme. Die Verwendung des Wortes „Stimme“ muss kritisch geprüft werden, ob tatsächlich stimmliche Merkmale gemeint sind.

Beispiele für die Anwendung des Codes:

- „wirkt wie abgelesen“ (SuS – KAUW3 – c)
- „Gleichmäßige Stimme“ (Leh – DAWI04 – e)

4.2.1. Stimmklang

Inhaltliche Beschreibung:

Hiermit werden Äußerungen codiert, die den Stimmklang adressieren. Laien können beispielsweise eine Stimme als heiser, rau, kratzig, schön, dunkel, hell oder nasal beschreiben.

Abgrenzung zu anderen Kategorien:

Abzugrenzen ist die Modulation der Stimme durch Heben und Senken oder Lautstärkenvariationen. Sie werden unter „Variation der Stimme“ gefasst.

4.2.2. Sprechgeschwindigkeit

Inhaltliche Beschreibung:

Hierunter werden Aussagen gesammelt, die sich auf die Sprechgeschwindigkeit beziehen. Sie kann als zu schnell, zu langsam, zu monoton oder zu sprunghaft beschrieben werden.

Beispiele für die Anwendung des Codes:

- „Lehrkraft spricht langsam“ (Leh – GEUL11 – a)
- „es könnte etwas schneller gesprochen werden“ (SuS – TAGE15 – b)

4.2.3. Sprechpausen

Inhaltliche Beschreibung:

Sprechpausen können auch Denkpausen für die Zuhörenden sein, wenn sie an der richtigen Stelle gesetzt werden. Sie fallen vor allem auf, wenn sie als störend erlebt werden, weil sie den Denkfluss/das Mitdenken behindern, sie also an unangemessenen/unnatürlichen Stellen gesetzt werden. Mit diesem Code werden Segmente codiert, die darauf Bezug nehmen, dass Sprechpausen angemessen oder unangemessen, störend oder hilfreich waren.

4.2.4. Variation der Stimme

Inhaltliche Beschreibung:

Eine Variation der Stimme (z. B. hinsichtlich der Lautstärke, der Modulation) wirkt angenehm bzw. aufmerksamkeitsgenerierend auf Adressat:innen. Äußerungen, die auf die Umsetzung der Variation der Stimme verweisen, werden mit diesem Code versehen.

Beispiele für die Anwendung des Codes:

- „keine Akzente bei der Lautstärke, Höhe gesetzt“ (Stu – MAWE12 – a)
- „sehr monotone Stimme“ (Stu – JOAN04 – a)
- „Variation mit der Stimme ein wenig besser, deshalb gerade noch die Note 3“ (Stu – MAWE12 – c)

4.3. Sprechflüssigkeit

Inhaltliche Beschreibung:

Hiermit werden Äußerungen codiert, die sich auf die Sprechflüssigkeit/den Redefluss beziehen. Dieser kann mit Stottern, vielen Satzabbrüchen, Wortwiederholungen, Versprechern gestört sein. Auch Hästitationspartikel (Ähm etc.) beeinflussen die Wahrnehmung.

Beispiele für die Anwendung des Codes:

- „Zwischendurch stockte die erklärende Person ein bisschen, der Redefluss wurde somit unterbrochen“ (Stu – ELIL12 – e)
- „ein paar kleine Versprecher“ (Stu – BREHE95 – a)

4.4. Zugewandter Körperausdruck

Inhaltliche Beschreibung:

Die Körperhaltung und Gestik der erklärenden Person sollten den Zuhörer:innen zugewandt und offen sein, sodass diese zum Zuhören animiert werden. Auch der Blickkontakt und eine freundliche Mimik können für das Zuhören zuträglich sein. Mit diesem Code werden Segmente codiert, die darauf hindeuten, dass sich die Adressat:innen durch den Körperausdruck angesprochen fühlen bzw. fühlen können.

Beispiele für die Anwendung des Codes:

- „evtl mehr Augenkontakt“ (Leh – ROAN05 – d)

4.5. Unterstützender Körperausdruck

Inhaltliche Beschreibung:

Mit diesem Code werden alle Äußerungen versehen, bei denen der Körperausdruck inklusive Mimik und Gestik in dem Sinn adressiert wird, dass das Gesagte unterstützt oder ein Verstehen des Gesagten gehemmt wird.

Beispiele für die Anwendung des Codes:

- „Mimik und Gestik gut / Bewegungen wirken unruhig“ (Did – BEHE12 – b)
- „nicht gut Gestikuliert“ (SuS – CLHE21 – d)

5. Verständliche Sprache

5.1. Sprache global

Inhaltliche Beschreibung:

Unter dieser globalen Kategorie werden alle Äußerungen gefasst, die sich global und/oder unspezifisch auf die Sprache der erklärenden Person beziehen. Global und unspezifisch können auch bedeuten, dass keine Ebene der Sprache adressiert wird, also kein Bezugspunkt erkennbar ist, ob sich die Äußerung auf Wort-, Satz- oder Textebene bezieht.

Beispiele für die Anwendung des Codes:

- „verständliche Ausdrucksweise“ (Stu – MARU06 – d)
- „zu viel gehobene Sprache“ (SuS – BIDZ13 – f)

5.2. Wortschatz

Inhaltliche Beschreibung:

Mit diesem Code werden alle Segmente codiert, die auf die Wortwahl der erklärenden Person verweisen. Diese kann der Situation („bildungssprachlich“) und für die Jahrgangsstufe angemessen und damit verständlich sein, also keine Verständnishürde für den Inhalt darstellen oder nicht.

Beispiele für die Anwendung des Codes:

- „gute Wortwahl“ (Did – STHE54 – c)
- „Ausdruck an manchen Stellen etwas komisch“ (Stu – NAAN13 – b)

- „Verwendet zu oft Wörter wie tun - außerdem sind die Wörter wie kucken absolut nicht schön ausgedrückt“ (Stu – MAWE12 – e)

Abgrenzung zu anderen Kategorien:

Abzugrenzen sind hier explizite Verweise auf den Fachwortschatz („Sprache – Umgang mit Fachworten“) und auf den präzisen Ausdruck der erklärenden Person („Sprache – (Un-)präziser Ausdruck“).

5.3. Umgang mit Fachworten

Inhaltliche Beschreibung:

Dieser Code wird vergeben, wenn in einem Segment explizit auf den Umgang oder den Einsatz von Fachworten verwiesen wird: auf die Bekanntheit der genutzten Fachworte, auf die (fehlende) Erklärung oder Paraphrase.

Beispiele für die Anwendung des Codes:

- „was bedeutet Fett-reich“ (Did – MAPA89 – c)
- „keine Note 1, da ggf. zu viele neue Fachbegriffe für Sus“ (Leh – INWO07 – a)
- „und ich fand auch gut dass immer bei Wörtern die manche Schüler eventuell nicht kennen (z. B. Insektizide) auch erklärt wurden.“ (SuS – GUDA10 – b)

5.4. Satzbau

Inhaltliche Beschreibung:

Außerungen werden in diese Kategorie sortiert, wenn Teilnehmer:innen auf den Satzbau verweisen. Bedingt angemessen wären beispielsweise Nominalisierungen, häufige Nebensatzkonstruktionen und zu lange Sätze, die das Verstehen behindern können.

Beispiele für die Anwendung des Codes:

- „knappe und einfache Sätze“ (Did – STHE 54 – c)
- „Sätze teilweise recht lang“ (Leh – FRCH02 – d)
- „etwas umständliche Je mehr...desto weniger... Folgerungen“ (SuS – KAYÜ3 – e)

5.5. (Un-)Präziser Ausdruck

Inhaltliche Beschreibung:

Mit diesem Code werden Segmente codiert, die auf einen präzisen Ausdruck der erklärenden Person referieren, der der Klarheit der Erklärung dient und keinen (unangemessenen) Deutungsspielraum lässt, also nicht schwammig oder vage ist. Codiert wird damit ebenfalls das Gegenteil. (Findeisen schreibt dazu: „Vage Formulierungen wie ‚es wäre möglich‘, ‚üblicherweise‘, ‚meistens‘ [...] beeinträchtigen die Klarheit der Erklärung. Hierdurch wird die Kommunikation negativ beeinflusst. Außerdem könnten die Lernenden diese Formulierungen als Defizit im Fachwissen bzw. unzureichende Vorbereitung der Lehrperson interpretieren“ (S. 72))

Beispiele für die Anwendung des Codes:

- „klare Sprache“ (Did – ro37 – a)
- „Fachsprache nicht immer präzise“ (Leh – ASBR09 – a)
- „aber Erklärungen selbst schwammig formuliert“ (Stu – NGAI04 – e)
- „viel und wenig - relative Begriffe“ (Leh – ANMA14 – b)

Abgrenzung zu anderen Kategorien:

Abzugrenzen sind hier die Verwendung von Fachwörtern oder von Wortschatz, der Schüler:innen nicht geläufig ist, wodurch das Verstehen erschwert wird („Sprache – Fachwörter“/„Sprache – Wortschatz“). Abzugrenzen sind auch das allgemeine Sprachniveau und die Satzstruktur, die unangemessen und deshalb eine Hürde zum Verstehen sein kann („Sprache – Sprachniveau und Satzbau“).

5.6. Hervorheben von Wichtigem

Inhaltliche Beschreibung:

Hiermit werden Segmente codiert, die auf ein sprachliches Hervorheben von wichtigen Punkten und Zusammenhängen in der Erklärung abzielen. Die Rezipient:innen werden mit sprachlichen Mitteln auf besonders Relevantes hingewiesen.

6. Fachliche Aspekte

Neben der Berücksichtigung des Adressaten muss die Lehrkraft auch den Gegenstand, der in der Erklärung Lerngegenstand sein soll, in angemessener Weise präsentieren. Die hierunter gesammelten Äußerungen beziehen sich allgemein auf eine fachliche korrekte und vollständige Darstellung und nehmen im Speziellen mit Fokus auf das Erkenntnisinteresse zur unterschiedlichen Wahrnehmung von komplexen und linearen Erklärungen mit weiteren Subkategorien diese in den Blick.

6.1. Fach global

Inhaltliche Beschreibung:

Hiermit werden nur diejenigen Segmente codiert, die eine klare Referenz auf fachliche Aspekte darstellen, aber keiner Subkategorie zugeordnet werden können.

Beispiele für die Anwendung des Codes:

- „Wissenschaftsorientierung“ (Did – GEHA01 – d)

6.2. Korrektheit

Inhaltliche Beschreibung:

Mit diesem Code werden alle Begründungen versehen, die die Erklärung als frei von inhaltlichen Mängeln und somit als sachlich richtig oder korrekt beschreiben oder das Gegenteil aussagen (sachlich nicht richtig oder inkorrekt). Dabei können sich die benannten Fehler auch auf falsche Begrifflichkeiten beziehen.

Beispiele für die Anwendung des Codes:

- „fachlich kritisch: Vermischung Ebene Genotyp und Phänotyp“ (Did – ELER15 – a)
- „Fehler in der Folgerichtigkeit: Zitat: Vitamin A kann zu Nachblindheit führen, weil es wichtig beim Sehvorgang und bei der Erneuerung der Haut ist.“ (Stu – JOAN04 – c)
- „Fachliche Fehler bei Verwendung des Wortes Gen“ (Stu – MAWE12 – d)

6.3. Fachspezifische Erklärungsform

Inhaltliche Beschreibung:

Als Begründung für die Bewertung kann von den Teilnehmer:innen auch das Kriterium der Korrektheit der fachspezifischen Erklärungsform (z. B. funktionalen, teleologischen, mechanistischen, kausalen) genannt werden. Bedingung für die Codierung ist das explizite Nennen der Erklärungsform.

Beispiele für die Anwendung des Codes:

- „Guter Überblick aber einige der Erklärungen bezogen sich nicht auf Ursachen, sondern auf Zwecke, bspw. deswegen wird es in Mizellen transportiert (oder so ähnlich). Das könnte den Raum für teleologische Denkweisen eröffnen.“ (Did – ANTH74 – c)
- „Einbettung von Vitamin A in eine Fettmicelle: Die Erklärung klingt finalistisch, d. h. es klingt, als ob zielgerichtet Fettmicellen gebildet würden, die nun auf Vitamin-A-Fang ausgehen.“ (Leh – MAED33 – c)

6.4. Berücksichtigung von Nichtwissen

Inhaltliche Beschreibung:

Mit diesem Code werden alle Äußerungen codiert, die sich darauf beziehen, dass in der Erklärung auch auf Zusammenhänge, Mechanismen und Faktoren verwiesen wird, die noch nicht bekannt sind. Es ist also nicht das persönliche Nichtwissen von Adressat:innen gemeint, sondern ein gesamtwissenschaftliches Nichtwissen.

Beispiele für die Anwendung des Codes:

- „gut: Unsicherheit bei Erklärungen einbezogen (statt positivistischer richtig/falsch-Dichotomie)“ (Did-EVWO09-b)
- „Es wird deutlich, dass es in der Wissenschaft noch viel zu erforschen gibt und man noch längst nicht alles weiß“ (Stu – BRHE95 – d)
- „- gut: Unsicherheit bei Erklärungen im naturwissenschaftlichen Erkenntnisgewinnungsprozess wird hervorgehoben, insbesondere unter Einbezug aktueller Forschung“ (Did – EVWO09 – d)

Abgrenzung zu anderen Kategorien:

Segmente, die eine unklare und vage Sprache betreffen, werden mit „Sprache – (un-)präziser Ausdruck“ codiert.

6.5. Vagheit/Eindeutigkeit

Inhaltliche Beschreibung:

Ein Segment wird dann mit diesem Code codiert, wenn es sich auf die (unzureichende) Eindeutigkeit oder (zu große) Vagheit der Zusammenhänge bezieht.

Beispiele für die Anwendung des Codes:

- „aber für Kinder eventuell zu diffus da nur potentielle Faktoren genannt werden“ (Did – MAWE02 – b)
- „Es sollten z. B. einzelne Ausschläge zumindest ein bisschen erklärt werden, was heutzutage sicher möglich ist, und nicht nur gesagt werden, dass es sehr schwierig ist das Diagramm zu erklären.“ (SuS – MAWI3 – b)
- „nicht konkret“ (Leh – IDER37 – d)
- „Die Erklärung sagt nur aus, dass Faktoren zu multipler Sklerose führt (das kann man sich denken)“ (Stu – SUJO15 – d)

Abgrenzung zu anderen Kategorien:

Wenn sich Äußerungen auf schwammige, vage, unpräzise Wortwahl beziehen, die unangemessenen Deutungsspielraum offenlassen, wird das Segment mit „Sprache – (Un)präziser Ausdruck“ codiert.

6.6. Gefühl der (Un-)Sicherheit

Inhaltliche Beschreibung:

Mit diesem Code werden Äußerungen codiert, die sich auf eine mögliche Verunsicherung der Schüler:innen durch wenig eindeutige Zusammenhänge und viele Unbekannte beziehen.

Abgrenzung zu anderen Kategorien:

Wird explizit die Wortwahl angesprochen, die unangemessenen Deutungsspielraum offenlässt und eben nicht klar bzw. präzise ist, wird das Segment mit „Sprache – (Un-)präziser Ausdruck“ codiert.

6.7. Faszination und Interessiertheit

Inhaltliche Beschreibung:

Mit diesem Code werden Äußerungen codiert, die deutlich machen, dass der ausgewählte Fachinhalt Faszination und Interessiertheit bei den Adressat:innen entstehen lässt bzw. lassen kann.

6.8. Phänomenorientierung

Inhaltliche Beschreibung:

Mit diesem Code werden Äußerungen codiert, die sich auf die angemessene Abbildung des Phänomens mit seinen Charakteristika bezieht. Eine Eigenschaft biologischer Phänomene ist zum Beispiel deren Komplexität (Multikausalität).

Beispiele für die Anwendung des Codes:

- „weniger gut: monokausale Erklärung an sehr einfachem Modell; Komplexität natürlicher Räuber-Beute-Zusammenhänge [...] wird wenig deutlich“ (Did – EVWO09 – e)
- „Zusammenspiel Faktoren wird deutlich“ (Leh – ASBR09 – b)
- „Es wird sehr deutlich, wie komplex das System ist.“ (Stu – BRHE95 – f)

6.9. Prinzipienorientierung

Inhaltliche Beschreibung:

Mit diesem Code werden Segmente codiert, die sich auf den Verweis fachlicher Prinzipien beziehen. Diese sind beispielsweise in Basiskonzepten oder Erschließungsfeldern zusammengefasst. Das Aufgreifen/Herausstellen der Prinzipien kann zur Übertragbarkeit auf weitere Beispiele verhelfen.

Beispiele für die Anwendung des Codes:

- „möglich auch Hinweis auf Basiskonzepte bzw. Erschließungsfelder“ (Did – ANTH74 – a)
- „Möglichkeit des systemischen Denkens nicht genutzt“ (Did – MAPA89 – b)
- „Je - Desto - Beziehungen kommen wenig zum tragen“ (Stu – JOAN04 – b)

- „Also expliziter Bezug zur Regelmäßigkeit natürlicher Phänomene möglich aber nicht genutzt“ (Did – ANTH74 – e)

Abgrenzung zu anderen Kategorien:

Von Schüler:innen schwerlich einzuschätzen (infrage kommende Segmente sollten genau auf Passung zu anderen Kategorien hin überprüft werden).

6.10. Vollständigkeit

Inhaltliche Beschreibung:

Mit diesem Code werden alle Äußerungen zur Vollständigkeit oder zur Nichtvollständigkeit der Erklärung codiert. Darunter ist zu verstehen, dass in der Erklärung nichts vernachlässigt wird, was für eine korrekte Vorstellung über das Phänomen aus fachlicher Sicht notwendig gewesen wäre (oder dass etwas vernachlässigt wird). Die Adressat:innen spielen bei den Äußerungen keine explizite Rolle, sondern der Fachgegenstand steht im Fokus.

Beispiele für die Anwendung des Codes:

- „Nicht klar ersichtlich, warum Micelle besseren Kontakt zur Darmschleimhaut im wässrigen Milieu haben soll.“ (Did – HEMA24 – c)
- „Wichtigkeit bei Schwangerschaft fehlte. In wie weit trägt Empfängnisverhütung mit der Pille zur guten/schlechten Aufnahme bei?“ (Stu – ELUL12 – f)
- „und der Unterschied zwischen gleichläufiger und gegenläufiger Beziehung wird nicht herausgestellt“ (Leh – MAED33 – e)

Abgrenzung zu anderen Kategorien:

Äußerungen, bei denen es darum geht, dass die Inhalte entweder zu schnell aneinandergereiht oder zu zerdehnt werden, indem zu den einzelnen Verstehenselementen zu viel ausgeführt wird, werden mit Erklärtempo codiert.

Äußerungen beziehen sich in dieser Kategorie auf Dinge, die nicht gesagt wurden, aber für ein korrekteres Verständnis über den Fachgegenstand notwendig wären. In Abgrenzung zur Kategorie „Adressatenorientierung – Berücksichtigung des Vorwissens“ bei dem es darum geht, dass angesprochene Dinge nicht weit genug expliziert oder erklärt wurden, sodass sie von Schüler:innen verstanden werden können.

6.10.1. „Fehlende“ Inhalte, die über das Erklärziel hinausgehen

Inhaltliche Beschreibung:

Mit diesem Code werden alle Segmente codiert, bei denen sich – meist Schüler:innen – weitere Inhalte wünschen, die nicht direkt das Erklärziel, also die anfangs gestellte Frage, betreffen, sondern darüber hinausgehen.

Beispiele für die Anwendung des Codes:

- „keine Folgen der Erkrankung“ (SuS – OPFA31 – c)
- „und es wäre noch interessant gewesen die mangel Erscheinungen zu erfahren“ (SuS – TAST6 – d)
- „Das Protein kann nicht mehr in die Membran eingebaut werden, aber was geschieht damit dann?“ (Stu – BASIO2 – a)

6.11. Zusammenhänge werden deutlich

Inhaltliche Beschreibung:

Segmente werden mit diesem Code codiert, wenn Teilnehmer:innen in Ihrer Begründung anführen, dass in der Erklärung das Ursache-Wirkungs-Gefüge deutlich wird, dass die Ursachen mit den jeweiligen Folgen gut in Verbindung gebracht und Zusammenhänge deutlich aufgezeigt werden.

Beispiele für die Anwendung des Codes:

- „Das Je ... desto war gut nach zu vollziehen“ (SuS – CHRO28 – e)
- „schlüssige Zusammenhänge“ (SuS – JOCH4 – c)
- „ansonsten sehr anschauliches Wirkungsgefüge“ (Stu – JOAN04 – f)
- „Bekanntere Faktoren werden aber nur (teilweise) erwähnt, nicht erklärt.“ (Stu – ELUL12 – d)

7. Globale Äußerungen

Globale Wirkungen/Eindrücke der Erklärung (meist als einzelne Stichworte), die durch mehrere Aspekte bzw. Kriterien realisiert werden können. Die Äußerungen werden von den Teilnehmer:innen nicht genauer zum Beispiel durch eine Begründung oder einen Bezugspunkt in der Erklärung spezifiziert. In der Oberkategorie werden keine Äußerungen gesammelt, sondern ausschließlich die Codes der Unterkategorien genutzt.

7.1. Gut Erklärt

Beispiele für die Anwendung des Codes:

- „Grundsätzlich finde ich es gut erklärt“ (Stu – NAAN13 – c)
- „Stoff gut erklärt“ (SuS – MAWI3 – a)

7.2. Verständlich

Inhaltliche Beschreibung:

Verständlichkeit wird als globale Aussage codiert, da sie durch mehrere Merkmale einer Erklärung oder Handlungen bzw. Umsetzungen der Lehrkraft erreicht werden kann (z. B. Wortwahl, Aufbau und Anordnung der Verstehenselemente, ...). Ebenfalls codiert werden Äußerungen von Schüler:innen, dass sie die Erklärung verstanden hätten.

Beispiele für die Anwendung des Codes:

- „man hat es schlussendlich verstanden“ (SuS – EDUD31 – a)
- „Insgesamt gut verständliche Erklärung“ (Stu – SUNO37 – a)

7.3. Klar

Beispiele für die Anwendung des Codes:

- „Klare Erläuterung“ (Did – GEHA01 – f)
- „klar erklärt“ (SuS – TAGE15 – a)

7.4. Nachvollziehbar

Abgrenzungen:

Wenn gleichzeitig beispielsweise die Struktur angesprochen wird, wird das Segment nicht mit diesem Code codiert, sondern „Strukturiertheit – logischer Aufbau“ oder „Strukturiertheit – Schrittweise Erklärung“ verwendet.

7.5. Anschaulich

Abgrenzung zu anderen Kategorien:

Sobald Teilnehmer:innen spezifischer auf die eingesetzten Bilder rekurrieren, wird der Code nicht vergeben, sondern „Repräsentationen – Einsatz bildlicher Visualisierungen“.

7.6. Übersichtlich

Beispiele für die Anwendung des Codes:

- „Übersichtlich“ (SuS – ANST15 – d)
- „übersichtliche Erklärung“ (Leh – ELED03 – a)
- „übersichtlich dargestellt“ (Leh – ASBR09 – b)

Abgrenzung zu anderen Kategorien:

Segmente werden nicht mit dieser Kategorie codiert, wenn gleichzeitig die eingesetzten Visualisierungen erwähnt werden, die übersichtlich gestaltet sind (z. B. Diagramm, Mindmap).

7.7. Sachlich

Beispiele für die Anwendung des Codes:

- „sachlich erklärt“ (SuS – BIRE38 – a)
- „sachliche Informationen“ (SuS – GAMA44 – a)

7.8. Informativ

Beispiele für die Anwendung des Codes:

- „informativ erklärt“ (SuS – CHBA9 – a)
- „der Unterricht was sehr informativ“ (SuS – CARO9 – d)

8. Lehrperson

Inhaltliche Beschreibung:

Mit diesem Code werden Segmente versehen, die sich auf die Wirkung der Lehrperson und nicht diejenige der Erklärung selbst beziehen. Dabei werden Eigenschaften der Person adressiert, die sie als Persönlichkeit betreffen.

Beispiele für die Anwendung des Codes:

- „hat sich viel mühe gegeben“ (SuS – TAST6 – a)
- „und damit der Eindruck entsteht,dass die Lehrerin selbst nicht 100% weiß, was genau Sache ist“ (SuS – CEAM05 – d)
- „Zu wenig Humor.“ (SuS – KAYÜ3 – b)

9. Merkbarkeit und Lernbarkeit

Inhaltliche Beschreibung:

Mit dieser Kategorie werden alle Segmente codiert, die eine Äußerung über die Lernbarkeit oder Merkbarkeit des erklärten Schulstoffes anhand des Gesagten, der Repräsentationen und der Visualisierungen bzw. des Tafelanschiebs als möglichen Hefteinträge betreffen.

Beispiele für die Anwendung des Codes:

- „nicht besonders gut zu behalten[...]“ (SuS – CHJO26 – a)
- „[...] und man kann es sich auch leichter merken und vergisst es nicht so schnell.“ (SuS – SAFR9 – b)
- „Falls mitschreiben erwünscht ist -> zu wenige Angaben (s.u.) oder Beschriftungen (Dünndarmausschnitt etc. Leber (Zwischenspeicher) (links oben: Der Folsr. Stoffwechsel; links unten: Mögliche Ursachen für Mangelerscheinungen)“ (Stu – INST04 – f)

10. Sonstiges

Mit dem Code „Sonstiges“ werden Segmente codiert, die keiner der bestehenden Kategorien zugeordnet werden können. Einige der Äußerungen beziehen sich dabei auf das Studiendesign selbst oder sind deskriptive Bemerkungen.

- „Schade, dass man das Video nicht zurückspulen kann, da ich eine Sekunde mal unaufmerksam war und ein Wort nicht richtig verstanden habe.“ (Leh – CHMI75 – a)
- „Erklärungen von ähnlichen Themen in den einzelnen Klassen sehr repetitiv (nicht zwingend schlecht!)“ (Stu – SUNO37 – f)
- „Erklärung der Populationsdynamik unter Bezugnahme auf Lotta-Volterra“ (Did – ANTH74 – e)
- „wichtiges Thema“ (SuS – UTAN26 – c)

C) Ergebnisse der Augenscheininvalidierung

Zur Absicherung der Augenscheininvalidität bezüglich der Trennschärfe der fächerübergreifenden Aspekte Adressatenorientierung, Strukturiertheit, sprachliche Verständlichkeit, Sprech- und Körperausdruck sowie der Wirkung der erklärenden Person wurden Lehramtsstudierende, Promovierende und Professor:innen der empirischen Bildungsforschung ($N = 98$) der Universität Regensburg befragt. Sie wurden mit zwei Texten instruiert (gerahmte Texte auf den folgenden Seiten; „Validierung des Fragebogens der FALKE-Studie“ und „Kurzbeschreibung der Konstrukte“). Die Tabelle mit den Ergebnissen der Zuordnung findet sich nach diesen Instruktionstexten.

Validierung des Fragebogens der FALKE-Studie

In der Studie FALKE („Fachspezifische Lehrerkompetenzen im Erklären“) untersuchen elf Didaktiken der Universität Regensburg im Verbund mit der Sprach- und der Sprechwissenschaft empirisch Aspekte guten Erklärens. Dazu wurden bereits pro Unterrichtsfach je sechs kurze videografierte Lehrerklärungen in einen Online-Fragebogen implementiert, den dann jeweils Fachdidaktiker/innen, aber auch Lehrkräfte, Studierende und Schüler/innen bearbeitet haben, indem sie für jedes gezeigte Video bestimmte Qualitätskriterien guten Erklärens beurteilt haben.

Sie müssen diese Videos nun NICHT ein weiteres Mal ansehen und beurteilen, sondern sollen uns vielmehr dabei helfen, die abgefragten Kriterien (siehe unten „Kriterien im Fragebogen“) theoretisch „zuzuordnen“. In der Online-Fragebogenstudie wurden Studienteilnehmenden pro Video dazu die untenstehenden Aussagen gestellt, die sie für jedes Video von 1 = „stimme voll zu“ bis 6 = „stimme überhaupt nicht zu“ beurteilen sollten.

Diese zu beurteilenden Aussagen sollten theoretisch fünf übergeordnete Konstrukte abbilden (siehe „Kurzbeschreibung der Konstrukte“). Wir wollen nun gerne von Ihnen wissen, welchem der fünf auf der nächsten Seite beschriebenen Konstrukte Sie die untenstehenden Fragen jeweils am ehesten zuordnen würden. Manche der Aussagen sind sehr offensichtlich einem der Konstrukte zuzuordnen, bei manchen dürfte Ihnen die Zuordnung etwas schwerer fallen.

Bitte setzen Sie aber auf jeden Fall in jeder Zeile ein Kreuz. Wenn Sie der Meinung sind, eine Aussage betrifft mehrere der fünf Konstrukte, machen Sie Ihr Kreuz bitte bei dem Konstrukt, zu dem die Aussage Ihrer Meinung nach „am besten“ passt. Orientieren Sie sich dabei bitte gegebenenfalls an der folgenden Kurzbeschreibung der Konstrukte.

Bitte beachten Sie, dass es durchaus möglich ist, dass eines der Konstrukte beispielsweise nur durch drei Aussagen abgebildet wird, ein anderes aber durch sieben Kriterien.

Kurzbeschreibung der Konstrukte

Die Strukturiertheit einer Erklärung beinhaltet einerseits die Konzentration auf die wesentlichen Aspekte des zu Erklärenden und andererseits die Anordnung dieser Aspekte in einer, dem Inhalt angemessenen, Reihenfolge. Das Nennen des Themas zu Beginn und das Aufzeigen der Anordnung der einzelnen Aspekte unterstützen die Strukturiertheit.

Unter Adressatenorientierung versteht man die Anpassung der Erklärung und des Erklärprozesses an die jeweiligen Rezipient/innen, um Verstehensprozesse zu ermöglichen und den Aufbau von Wissen zu unterstützen. Die Einschätzung der kognitiven, motivationalen und affektiven Charakteristika der Zielgruppe (z. B. Vorwissen, Schülervorstellungen, mentale Fähigkeiten, Interesse, Motivation, ...) ist die Voraussetzung für adressatengerechte didaktische Überlegungen zur Komplexitätsreduktion oder Visualisierung.

Der Sprech- und Körperausdruck der erklärenden Lehrkraft umfasst zum einen akustisch-auditiv wahrnehmbare Merkmale gesprochener Sprache melodischer, dynamischer, temporaler und artikulatorischer Art sowie die Stimmqualität. Zum anderen werden darunter körperausdrucksbezogene Spezifika wie Mimik und Blickverhalten, Gestik, Körperhaltung, -spannung und -orientierung sowie die Bewegung und das Verhalten im Raum gefasst.

Die sprachliche Verständlichkeit einer Erklärung erfordert, dass man einerseits Ausdrucksweisen verwendet, die das inhaltliche Verstehen für die Zuhörenden unterstützen, und andererseits Ausdrucksweisen vermeidet, die das Verstehen erschweren. Wichtige sprachliche Ausdrucksmittel sind verständliche Wörter und Begriffe, eine angemessene Satzlänge sowie gezielte Hervorhebungen, um die Aufmerksamkeit der Zuhörenden auf Wichtiges zu lenken.

Unter der Persönlichkeitswirkung verstehen wir Aspekte, die theoretisch eigentlich keine ureigenen Aspekte guten Erklärens sind, von denen wir aber vermuten, dass sie ebenfalls eine Rolle für die Beurteilung einer in einem Video gezeigten Erklärung spielen dürften. Diese Aspekte lassen sich aber eher der Persönlichkeit einer Lehrkraft im Allgemeinen zuordnen.

Ergebnisse der Befragung zur Zuordnung der Items zu den fünf Aspekten

Die nachstehende Tabelle zeigt die relativen Häufigkeiten für die Zuordnung der Items zu den fünf fächerübergreifenden Aspekten guten Erklärens, die im Fragebogen implementiert werden sollten.

Tabelle: Relative Häufigkeiten der Zuordnung der Items zu den Aspekten durch Lehramtsstudierende, Promovierende und Professor:innen der empirischen Bildungsforschung (N = 98) von der Universität Regensburg

Items	Aspekte	ST	AD	SE	SW	PE
Schüler/innen konnten der Erklärung mit ihrem Wissen gut folgen.		.19	.69	.01	.10	.01
Bei manchen Wörtern wissen die Schüler/innen eventuell nicht, was sie bedeuten.		.01	.26	.06	.66	.01
Die Lehrkraft hatte eine angenehm klingende Stimme.		.00	.01	.78	.05	.15
Die Erklärung passte für Schüler-/innen dieser Jahrgangsstufe.		.03	.92	.00	.04	.01
Die Lehrkraft hatte eine deutliche Aussprache.		.00	.00	.57	.43	.00
Manche Sätze hat die Lehrkraft zu lang gemacht.		.14	.04	.11	.70	.01
Die Lehrkraft hat die Sprechpausen so gesetzt, dass man gut folgen konnte.		.15	.01	.40	.42	.03
Die Erklärung berücksichtigte die Lebenswelt der Schüler/innen.		.01	.87	.02	.00	.10
Die Lehrkraft wirkte in diesem Video auf mich sympathisch.		.00	.00	.10	.03	.86
Die Lehrkraft hatte einen Körperausdruck, der die Erklärung unterstützt hat.		.00	.01	.91	.03	.05
Die Erklärung hatte einen roten Faden.		1.00	.00	.00	.00	.00
Die Art und Weise, wie die Lehrkraft spricht, gefällt mir.		.00	.00	.47	.14	.38
Die Lehrkraft hat sich den Schüler/innen zugewendet.		.00	.38	.44	.07	.10
Die Lehrkraft wirkte in diesem Video auf mich begeistert.		.00	.02	.19	.02	.77
Es ist klar, was genau die Lehrkraft erklären wollte.		.65	.03	.00	.31	.01
Die Lehrkraft hatte eine angemessene Sprechgeschwindigkeit.		.00	.02	.56	.41	.01
Manchmal hat die Lehrkraft extra gesagt, dass etwas sehr wichtig ist.		.61	.17	.09	.13	.00
Die Lehrkraft wirkte in diesem Video auf mich natürlich.		.00	.00	.09	.02	.89
Für Schüler/innen war die Erklärung zu schwierig.		.01	.86	.00	.10	.03
Die Erklärung war gut strukturiert.		.98	.02	.00	.00	.00
Die Lehrkraft hat die Erklärung Schritt für Schritt aufgebaut.		.98	.01	.00	.01	.00
Die Lehrkraft hat sich gut verständlich ausgedrückt.		.00	.01	.20	.79	.00
Die Lehrkraft hat zu viele Einzelheiten erklärt.		.51	.21	.02	.21	.05
Die Lehrkraft hatte eine abwechslungsreiche Sprechweise.		.00	.01	.84	.14	.01
Die Erklärung war adressatengerecht.		.01	.96	.01	.02	.00
Die Erklärung war angemessen didaktisch reduziert.		.38	.54	.02	.05	.01

Bem.: AD: Adressatenorientierung, ST: Strukturiertheit, SE: Sprech- und Körperausdruck, SW: Sprachliche Verständlichkeit, PE: Persönlichkeitswirkung

D) Deskriptive und psychometrische Kennwerte auf Itemebene pro Statusgruppe und Videovignette

Schüler:innen

Tabelle: Übersicht zu Mittelwerten und Standardabweichungen der einzelnen Items sowie Itemtrennschärfen der in der Skala berücksichtigten Items und Cronbachs Alpha als Maß für die interne Konsistenz der Skalen für die Statusgruppe der Schüler:innen ($N_{max} = 122$)

			G _l	G _k	Ö _l	Ö _k	P _l	P _k
			M	M	M	M	M	M
			(SD)	(SD)	(SD)	(SD)	(SD)	(SD)
			r _{it}	r _{it}	r _{it}	r _{it}	r _{it}	r _{it}
Skala/ temabkürzung/Itemtext								
Ad_	Adressatenorientierung	<i>Cronbachs Alpha:</i>	0,54	0,38	0,29	0,16	0,53	0,55
sch	Für Schülerinnen und Schüler war die Erklärung zu schwierig.*		2,43 (1,22)	2,40 (1,31)	1,88 (1,24)	2,07 (1,20)	2,45 (1,24)	2,38 (1,22)
			0,27	0,04	0,04	-0,13	0,19	0,16
lbw	Die Erklärung berücksichtigte die Lebenswelt der Schülerinnen und Schüler.		4,22 (1,32)	4,12 (1,44)	4,50 (1,27)	4,24 (1,32)	3,36 (1,42)	3,49 (1,33)
			-0,04	-0,09	-0,17	-0,12	0,13	0,16
pas	Die Erklärung passte für Schülerinnen und Schüler dieser Jahrgangsstufe.		2,69 (1,12)	2,69 (1,03)	2,58 (1,12)	2,58 (1,01)	2,58 (1,01)	2,76 (1,12)
			0,60	0,42	0,26	0,25	0,37	0,43
red	Die Erklärung war angemessen didaktisch reduziert.		---	---	---	---	---	---
			---	---	---	---	---	---
wwi	Schülerinnen und Schüler konnten der Erklärung mit ihrem Wissen gut folgen.		2,37 (1,22)	2,36 (1,14)	1,87 (1,04)	1,94 (0,87)	2,29 (1,08)	2,50 (1,20)
			0,53	0,42	0,49	0,24	0,56	0,63
glo	Die Erklärung war adressatengerecht.**		---	---	---	---	---	---
			---	---	---	---	---	---
St_	Strukturiertheit	<i>Cronbachs Alpha:</i>	0,74	0,69	0,63	0,71	0,75	0,74
was	Es ist klar, was genau die Lehrkraft erklären wollte.		2,13 (1,06)	2,31 (1,10)	1,97 (0,89)	2,22 (1,07)	2,25 (0,98)	2,35 (1,10)
			0,63	0,61	0,55	0,65	0,55	0,58
rot	Die Erklärung hatte einen roten Faden.		2,37 (1,06)	2,51 (1,08)	2,18 (1,02)	2,42 (1,17)	2,43 (0,98)	2,47 (1,10)
			0,66	0,60	0,49	0,64	0,63	0,67
bau	Die Lehrkraft hat die Erklärung Schritt für Schritt aufgebaut.		2,02 (0,97)	2,23 (0,87)	1,98 (0,91)	2,24 (1,05)	2,31 (0,88)	2,33 (0,93)
			0,67	0,62	0,54	0,69	0,67	0,60
ein	Die Lehrkraft hat zu viele Einzelheiten erklärt.*		3,04 (1,14)	3,26 (1,21)	3,27 (1,32)	3,17 (1,23)	3,12 (1,21)	3,31 (1,29)
			0,20	0,07	0,01	0,06	0,29	0,25
glo	Die Erklärung war gut strukturiert.**		2,30 (0,97)	2,56 (1,04)	2,20 (0,91)	2,48 (1,06)	2,39 (0,88)	2,57 (1,10)
			---	---	---	---	---	---
Se_	Sprech- und Körperausdruck	<i>Cronbachs Alpha:</i>	0,78	0,69	0,77	0,88	0,85	0,80
sti	Die Lehrkraft hatte eine angenehm klingende Stimme.		2,12 (0,76)	2,24 (0,73)	2,13 (0,80)	2,23 (0,81)	1,95 (0,58)	2,07 (0,78)
			0,46	0,34	0,46	0,77	0,60	0,56
aus	Die Lehrkraft hatte eine deutliche Aussprache.		1,65 (0,78)	1,59 (0,63)	1,66 (0,63)	2,21 (1,06)	1,88 (0,71)	1,81 (0,74)
			0,59	0,27	0,35	0,57	0,35	0,42
ges	Die Lehrkraft hatte eine angemessene Sprechgeschwindigkeit.		1,95 (0,87)	2,46 (1,05)	2,23 (0,97)	2,51 (0,94)	2,10 (0,96)	2,19 (0,82)
			0,38	0,33	0,47	0,75	0,70	0,49
pau	Die Lehrkraft hat die Sprechpausen so gesetzt, dass ich gut folgen konnte.		2,26 (0,90)	2,50 (1,07)	2,19 (1,03)	2,67 (1,34)	2,10 (0,79)	2,32 (0,78)
			0,51	0,53	0,53	0,59	0,66	0,56

abw	Die Lehrkraft hatte eine abwechslungsreiche Sprechweise.		3,51 (1,20)	3,77 (1,25)	3,17 (1,45)	3,02 (1,30)	2,79 (1,07)	3,10 (1,33)
kor	Die Lehrkraft hatte einen Körperausdruck, der die Erklärung unterstützt hat.		0,48 (0,83)	0,45 (1,14)	0,60 (1,19)	0,62 (1,14)	0,63 (1,01)	0,62 (0,95)
zug	Die Lehrkraft hat sich den Schülerinnen und Schülern zugewendet.		0,64 (0,88)	0,52 (1,29)	0,63 (1,19)	0,62 (1,13)	0,65 (0,82)	0,71 (1,10)
glo	Die Art und Weise, wie die Lehrkraft spricht, gefällt mir.**		0,42 (0,72)	0,42 (1,05)	0,40 (0,98)	0,70 (1,05)	0,66 (0,68)	0,39 (0,94)
Sw_	Sprachliche Verständlichkeit	<i>Cronbachs Alpha:</i>	---	---	---	---	---	---
			0,43	0,45	0,48	0,50	0,37	0,57
ver	Die Lehrkraft hat sich gut verständlich ausgedrückt.		2,04 (0,91)	2,17 (0,88)	1,86 (0,82)	2,07 (0,84)	2,12 (0,85)	2,11 (0,88)
wic	Manchmal hat die Lehrkraft extra gesagt, dass etwas sehr wichtig ist.**		0,08 (1,31)	0,11 (1,37)	0,11 (1,32)	0,22 (1,41)	0,05 (1,28)	0,27 (1,21)
bed	Bei manchen Wörtern wissen die Schülerinnen und Schüler eventuell nicht, was sie bedeuten.		---	---	---	---	---	---
			2,75 (1,39)	2,70 (1,29)	2,10 (1,37)	2,11 (1,22)	2,73 (1,39)	2,70 (1,39)
zul	Manche Sätze hat die Lehrkraft zu lang gemacht.*		0,36 (1,15)	0,42 (1,13)	0,41 (1,22)	0,42 (1,18)	0,32 (1,18)	0,40 (1,11)
			0,44	0,39	0,52	0,34	0,39	0,49
Sy_	Persönlichkeitswirkung	<i>Cronbachs Alpha:</i>	0,84	0,68	0,74	0,86	0,89	0,78
sym	Die Lehrkraft wirkte in diesem Video auf mich sympathisch.		2,23 (1,07)	2,34 (1,03)	2,26 (0,97)	2,37 (1,18)	2,21 (1,16)	2,40 (1,08)
beg	Die Lehrkraft wirkte in diesem Video auf mich begeistert.		0,79 (1,26)	0,55 (1,25)	0,62 (1,22)	0,77 (1,26)	0,76 (1,27)	0,67 (1,12)
nat	Die Lehrkraft wirkte in diesem Video auf mich natürlich.		0,62 (1,37)	0,47 (1,26)	0,59 (1,08)	0,74 (1,28)	0,78 (1,36)	0,62 (1,07)
			0,68	0,44	0,47	0,71	0,84	0,58
Fa_	Fachspezifische Items	<i>Cronbachs Alpha:</i>	0,76	0,72	0,69	0,72	0,80	0,77
erf	Nach dieser Erklärung möchten Schüler_innen noch mehr darüber erfahren.		3,95 (1,39)	3,87 (1,40)	4,17 (1,27)	4,20 (1,38)	3,79 (1,34)	3,94 (1,36)
fas	Die Erklärung ist für Schüler_innen faszinierend.		0,69 (1,36)	0,54 (1,29)	0,51 (1,31)	0,55 (1,38)	0,65 (1,37)	0,57 (1,40)
nat	Die Erklärung zeigt, wie es in der Natur wirklich ist.		0,53 (1,13)	0,67 (1,25)	0,56 (0,99)	0,64 (1,17)	0,68 (1,16)	0,66 (1,25)
zan	Die Anzahl der Zusammenhänge ist für Schüler_innen angemessen.		0,48 (0,87)	0,39 (1,12)	0,40 (1,07)	0,36 (1,06)	0,60 (1,13)	0,52 (1,10)
			0,53	0,44	0,44	0,51	0,55	0,53
hil	Die Erklärung hilft Schüler_innen andere Beispiele zu verstehen.		---	---	---	---	---	---
aus	Die Erklärung zeigt nur einen kleinen Ausschnitt auf die Frage „Thema der Erklärung“.**		---	---	---	---	---	---

Bem.: Einheitliche Skalierung aller Items (1 = Stimme voll zu; 2 = Stimme zu; 3 = Stimme eher zu; 4 = Stimme eher nicht zu; 5 = Stimme nicht zu; 6 = Stimme gar nicht zu); *M*: Mittelwert; *SD*: Standardabweichung; *N_{max}*: aufgrund fehlender Werte bei einzelnen Videos den Berechnungen pro Statusgruppe maximal zugrunde liegende Personenanzahl; *: Item wird in der zugehörigen Skala *rekodiert* berücksichtigt; **: Item wird in der zugehörigen Skala *nicht* berücksichtigt; ---: Item ist in der Skala der jeweiligen Statusgruppe *nicht* enthalten.

Studierende

Tabelle: Übersicht zu Mittelwerten und Standardabweichungen der einzelnen Items sowie Itemtrennschärfen der in der Skala berücksichtigten Items und Cronbachs Alpha als Maß für die interne Konsistenz der Skalen für die Statusgruppe der Studierenden ($N_{max} = 44$)

			G_I	G_K	\bar{O}_I	\bar{O}_K	P_I	P_K
			M	M	M	M	M	M
			(SD)	(SD)	(SD)	(SD)	(SD)	(SD)
Skala/ Itemabkürzung/Itemtext			r_{it}	r_{it}	r_{it}	r_{it}	r_{it}	r_{it}
Ad_	Adressatenorientierung	<i>Cronbachs Alpha:</i>	0,70	0,81	0,62	0,38	0,86	0,84
sch	Für Schülerinnen und Schüler war die Erklärung zu schwierig.*		2,32 (0,88)	2,77 (1,02)	1,93 (0,86)	2,20 (1,03)	2,81 (1,01)	2,88 (1,15)
lbw	Die Erklärung berücksichtigte die Lebenswelt der Schülerinnen und Schüler.		0,26 3,50 (0,93)	0,52 3,35 (1,00)	0,23 3,33 (1,02)	-0,10 3,32 (0,93)	0,58 3,09 (1,04)	0,59 2,57 (1,09)
pas	Die Erklärung passte für Schülerinnen und Schüler dieser Jahrgangsstufe.		0,26 2,45 (1,01)	0,47 2,60 (0,85)	0,09 2,79 (1,01)	0,14 2,80 (1,03)	0,54 2,42 (0,93)	0,64 2,38 (0,88)
red	Die Erklärung war angemessen didaktisch reduziert.		0,53 2,41 (1,02)	0,58 2,70 (1,01)	0,47 2,51 (1,01)	0,32 2,41 (0,95)	0,77 2,30 (0,94)	0,60 2,40 (0,96)
wvi	Schülerinnen und Schüler konnten der Erklärung mit ihrem Wissen gut folgen.		0,63 2,23 (0,77)	0,68 2,42 (1,01)	0,65 1,88 (0,79)	0,37 1,95 (0,65)	0,77 2,30 (0,89)	0,65 2,38 (1,06)
glo	Die Erklärung war adressatengerecht.**		0,61 2,34 (0,94)	0,74 2,65 (1,09)	0,47 2,49 (1,14)	0,25 2,55 (0,95)	0,76 2,35 (0,92)	0,71 2,38 (1,01)
			---	---	---	---	---	---
St_	Strukturiertheit	<i>Cronbachs Alpha:</i>	0,79	0,85	0,72	0,80	0,77	0,85
was	Es ist klar, was genau die Lehrkraft erklären wollte.		1,80 (0,80)	2,07 (0,99)	1,86 (0,86)	1,91 (0,77)	1,86 (0,86)	1,98 (0,72)
rot	Die Erklärung hatte einen roten Faden.		0,68 1,73 (0,90)	0,75 2,12 (1,01)	0,48 1,67 (0,64)	0,63 2,16 (0,99)	0,62 1,93 (0,80)	0,70 1,81 (0,74)
bau	Die Lehrkraft hat die Erklärung Schritt für Schritt aufgebaut.		0,69 1,55 (0,79)	0,80 1,95 (0,82)	0,60 1,56 (0,70)	0,78 1,95 (0,89)	0,74 2,00 (0,79)	0,71 1,81 (0,71)
ein	Die Lehrkraft hat zu viele Einzelheiten erklärt.*		0,67 2,59 (1,25)	0,75 2,84 (1,17)	0,54 2,95 (1,51)	0,63 3,09 (1,24)	0,63 2,95 (1,13)	0,74 3,26 (1,29)
glo	Die Erklärung war gut strukturiert.**		0,33 1,93 (0,90)	0,42 2,21 (0,89)	0,11 2,05 (0,98)	0,35 2,30 (0,82)	0,23 2,09 (0,90)	0,50 2,10 (0,98)
			---	---	---	---	---	---
Se_	Sprech- und Körperausdruck	<i>Cronbachs Alpha:</i>	0,84	0,59	0,80	0,84	0,91	0,78
sti	Die Lehrkraft hatte eine angenehm klingende Stimme.		1,52 (0,70)	2,00 (1,11)	1,91 (1,07)	1,81 (0,83)	1,74 (0,86)	1,95 (1,00)
aus	Die Lehrkraft hatte eine deutliche Aussprache.		0,63 1,56 (0,93)	0,14 1,36 (0,58)	0,48 1,23 (0,43)	0,54 1,67 (0,62)	0,77 1,41 (0,57)	0,51 1,36 (0,58)
ges	Die Lehrkraft hatte eine angemessene Sprechgeschwindigkeit.		0,49 1,67 (0,96)	0,28 1,95 (1,09)	0,60 1,77 (0,97)	0,45 2,07 (1,17)	0,80 1,63 (0,84)	0,40 1,82 (1,01)
pau	Die Lehrkraft hat die Sprechpausen so gesetzt, dass ich gut folgen konnte.		0,72 1,74 (1,02)	0,29 2,36 (0,95)	0,44 2,14 (0,83)	0,70 3,04 (1,56)	0,59 1,93 (1,04)	0,50 2,00 (0,87)
abw	Die Lehrkraft hatte eine abwechslungsreiche Sprechweise.		0,53 2,67 (1,07)	0,29 3,55 (1,30)	0,70 2,73 (1,12)	0,76 2,74 (1,13)	0,65 2,11 (1,09)	0,78 2,55 (1,30)
			0,61	0,62	0,75	0,65	0,79	0,66

kor	Die Lehrkraft hatte einen Körperausdruck, der die Erklärung unterstützt hat.		2,33 (0,96)	2,86 (1,21)	2,18 (0,91)	2,56 (1,01)	1,78 (1,01)	2,45 (1,10)
zug	Die Lehrkraft hat sich den Schülerinnen und Schülern zugewendet.		0,79 (0,89)	0,48 (1,02)	0,51 (1,13)	0,63 (0,85)	0,68 (0,70)	0,61 (0,72)
glo	Die Art und Weise, wie die Lehrkraft spricht, gefällt mir.**		0,39 (1,05)	0,17 (1,14)	0,26 (1,21)	0,48 (1,24)	0,72 (0,95)	0,20 (1,22)
Sw_	Sprachliche Verständlichkeit	<i>Cronbachs Alpha:</i>	0,62	0,69	0,67	0,53	0,63	0,73
ver	Die Lehrkraft hat sich gut verständlich ausgedrückt.		1,82 (0,84)	2,00 (0,95)	1,70 (,91)	1,86 (0,82)	1,79 (0,74)	1,88 (0,92)
wic	Manchmal hat die Lehrkraft extra gesagt, dass etwas sehr wichtig ist.**		0,44 (1,31)	0,47 (1,01)	0,46 (1,01)	0,28 (1,08)	0,35 (1,14)	0,48 (1,13)
bed	Bei manchen Wörtern wissen die Schülerinnen und Schüler eventuell nicht, was sie bedeuten.		0,40 (1,20)	0,50 (1,12)	0,43 (1,13)	0,30 (1,21)	0,56 (1,30)	0,57 (1,10)
zul	Manche Sätze hat die Lehrkraft zu lang gemacht.*		0,43 (1,07)	0,56 (1,30)	0,56 (1,07)	0,49 (1,25)	0,46 (1,12)	0,61 (1,04)
Sy_	Persönlichkeitswirkung	<i>Cronbachs Alpha:</i>	0,73	0,69	0,76	0,82	0,87	0,82
sym	Die Lehrkraft wirkte in diesem Video auf mich sympathisch.		0,47 (0,79)	0,57 (1,23)	0,67 (1,14)	0,58 (0,92)	0,77 (0,64)	0,75 (1,16)
beg	Die Lehrkraft wirkte in diesem Video auf mich begeistert.		0,28 (0,96)	0,51 (1,37)	0,50 (1,22)	0,71 (0,87)	0,64 (0,93)	0,51 (1,19)
nat	Die Lehrkraft wirkte in diesem Video auf mich natürlich.		0,32 (1,28)	0,43 (1,10)	0,59 (1,17)	0,72 (1,22)	0,81 (0,95)	0,76 (1,16)
Fa_	Fachspezifische Items	<i>Cronbachs Alpha:</i>	0,78	0,83	0,78	0,72	0,83	0,87
erf	Nach dieser Erklärung möchten Schüler_innen noch mehr darüber erfahren.		0,52 (1,15)	0,67 (1,04)	0,58 (0,97)	0,64 (0,99)	0,71 (1,01)	0,80 (0,91)
fas	Die Erklärung ist für Schüler_innen faszinierend.		0,73 (0,99)	0,61 (1,04)	0,60 (1,01)	0,59 (1,06)	0,71 (1,00)	0,71 (1,03)
nat	Die Erklärung zeigt, wie es in der Natur wirklich ist.		0,52 (1,04)	0,57 (1,20)	0,56 (1,22)	0,20 (0,88)	0,40 (1,02)	0,53 (0,93)
zan	Die Anzahl der Zusammenhänge ist für Schüler_innen angemessen.		0,53 (0,84)	0,65 (1,01)	0,51 (0,97)	0,49 (0,81)	0,79 (0,85)	0,74 (0,94)
hil	Die Erklärung hilft Schüler_innen andere Beispiele zu verstehen.		0,50 (1,04)	0,59 (1,05)	0,50 (0,90)	0,53 (0,85)	0,51 (0,88)	0,73 (0,89)
aus	Die Erklärung zeigt nur einen kleinen Ausschnitt auf die Frage „Thema der Erklärung“.**		0,50 (1,20)	0,59 (1,17)	0,50 (1,19)	0,53 (1,22)	0,51 (1,12)	0,73 (1,23)

Bem.: Einheitliche Skalierung aller Items (1 = Stimme voll zu; 2 = Stimme zu; 3 = Stimme eher zu; 4 = Stimme eher nicht zu; 5 = Stimme nicht zu; 6 = Stimme gar nicht zu); *M*: Mittelwert; *SD*: Standardabweichung; *N_{max}*: aufgrund fehlender Werte bei einzelnen Videos den Berechnungen pro Statusgruppe maximal zugrunde liegende Personenanzahl; *: Item wird in der zugehörigen Skala *rekodiert* berücksichtigt; **: Item wird in der zugehörigen Skala *nicht* berücksichtigt; ---: Item ist in der Skala der jeweiligen Statusgruppe *nicht* enthalten.

Lehrkräfte

Tabelle: Übersicht zu Mittelwerten und Standardabweichungen der einzelnen Items sowie Itemtrennschärfen der in der Skala berücksichtigten Items und Cronbachs Alpha als Maß für die interne Konsistenz der Skalen für die Statusgruppe der Lehrkräfte ($N_{max} = 29$)

			G _i	G _k	Ö _i	Ö _k	P _i	P _k
			M	M	M	M	M	M
			(SD)	(SD)	(SD)	(SD)	(SD)	(SD)
			r _{it}	r _{it}	r _{it}	r _{it}	r _{it}	r _{it}
Skala/	Itemabkürzung/	Itemtext						
Ad_	Adressatenorientierung	<i>Cronbachs Alpha:</i>	0,64	0,60	-0,19	0,48	0,84	0,80
sch	Für Schülerinnen und Schüler war die Erklärung zu schwierig.*		2,24 (1,02)	2,36 (0,99)	1,79 (0,69)	2,15 (0,95)	2,52 (1,01)	2,46 (1,07)
lbw	Die Erklärung berücksichtigte die Lebenswelt der Schülerinnen und Schüler.		0,23 3,52 (0,95)	0,22 4,25 (1,18)	-0,31 4,29 (1,08)	0,09 3,67 (1,07)	0,68 3,19 (1,08)	0,48 3,11 (1,07)
pas	Die Erklärung passte für Schülerinnen und Schüler dieser Jahrgangsstufe.		0,36 3,14 (1,09)	0,27 3,36 (1,13)	-0,21 2,82 (1,22)	0,17 3,11 (1,12)	0,65 2,89 (1,16)	0,52 2,79 (0,96)
red	Die Erklärung war angemessen didaktisch reduziert.		0,59 2,97 (1,24)	0,42 3,29 (1,12)	0,24 2,79 (1,10)	0,28 3,00 (0,96)	0,57 2,74 (1,13)	0,63 2,86 (0,97)
wwi	Schülerinnen und Schüler konnten der Erklärung mit ihrem Wissen gut folgen.		0,49 2,24 (0,83)	0,59 2,43 (0,96)	0,49 1,75 (0,70)	0,55 1,89 (0,69)	0,63 2,33 (1,07)	0,71 2,25 (0,89)
glo	Die Erklärung war adressatengerecht.**		0,34 2,83 (1,14)	0,41 3,25 (1,04)	-0,20 2,68 (1,16)	0,22 2,81 (1,15)	0,64 2,74 (1,06)	0,56 2,68 (1,09)
			---	---	---	---	---	---
St_	Strukturiertheit	<i>Cronbachs Alpha:</i>	0,74	0,78	0,83	0,92	0,90	0,92
was	Es ist klar, was genau die Lehrkraft erklären wollte.		1,79 (0,86)	2,36 (1,03)	1,71 (0,71)	2,30 (1,38)	2,19 (0,96)	1,75 (0,70)
rot	Die Erklärung hatte einen roten Faden.		0,61 1,76 (0,87)	0,61 2,21 (0,92)	0,73 1,61 (0,83)	0,89 2,22 (1,19)	0,87 2,07 (0,87)	0,83 1,82 (0,77)
bau	Die Lehrkraft hat die Erklärung Schritt für Schritt aufgebaut.		0,68 1,66 (0,72)	0,60 1,93 (0,77)	0,84 1,57 (0,69)	0,88 1,89 (0,89)	0,78 1,81 (0,79)	0,87 1,75 (0,70)
ein	Die Lehrkraft hat zu viele Einzelheiten erklärt.*		0,59 2,48 (1,24)	0,70 3,11 (1,37)	0,76 1,96 (0,74)	0,84 3,07 (1,30)	0,66 2,41 (1,15)	0,83 2,54 (1,26)
glo	Die Erklärung war gut strukturiert.**		0,16 1,79 (0,77)	0,36 2,36 (0,91)	0,36 1,79 (0,96)	0,67 2,11 (0,80)	0,82 1,96 (0,85)	0,59 1,82 (0,82)
			---	---	---	---	---	---
Se_	Sprech- und Körperausdruck	<i>Cronbachs Alpha:</i>	0,72	0,90	0,89	0,90	0,87	0,94
sti	Die Lehrkraft hatte eine angenehm klingende Stimme.		1,25 (0,45)	1,58 (0,69)	1,68 (0,75)	1,33 (0,65)	1,33 (0,65)	1,68 (0,82)
aus	Die Lehrkraft hatte eine deutliche Aussprache.		0,33 1,17 (0,39)	0,77 1,74 (1,15)	0,73 1,53 (0,70)	0,65 1,67 (0,78)	0,55 1,33 (0,49)	0,87 1,58 (0,84)
ges	Die Lehrkraft hatte eine angemessene Sprechgeschwindigkeit.		0,62 1,75 (0,87)	0,79 2,00 (1,16)	0,68 1,89 (0,66)	0,55 2,17 (1,19)	0,62 2,08 (1,00)	0,85 1,89 (0,81)
pau	Die Lehrkraft hat die Sprechpausen so gesetzt, dass ich gut folgen konnte.		0,60 2,00 (1,04)	0,71 2,42 (1,54)	0,79 1,95 (0,85)	0,80 2,42 (1,24)	0,81 2,00 (1,04)	0,81 1,95 (0,85)
abw	Die Lehrkraft hatte eine abwechslungsreiche Sprechweise.		0,48 2,83 (1,12)	0,61 3,26 (1,33)	0,74 2,74 (0,93)	0,80 2,92 (1,08)	0,73 2,33 (0,89)	0,84 2,95 (1,18)
			0,44	0,62	0,69	0,72	0,55	0,76

kor	Die Lehrkraft hatte einen Körperausdruck, der die Erklärung unterstützt hat.		2,75 (0,97)	2,58 (1,07)	2,47 (1,12)	2,42 (1,17)	1,83 (0,72)	2,47 (1,12)
zug	Die Lehrkraft hat sich den Schülerinnen und Schülern zugewendet.		1,92 (1,00)	2,11 (1,05)	2,42 (0,96)	2,33 (1,16)	1,75 (0,97)	2,16 (0,90)
glo	Die Art und Weise, wie die Lehrkraft spricht, gefällt mir.**		0,71 (0,87)	0,64 (0,96)	0,67 (0,82)	0,74 (1,29)	0,63 (1,00)	0,76 (0,86)
Sw_	Sprachliche Verständlichkeit	<i>Cronbachs Alpha:</i>	0,57	0,38	0,67	0,64	0,74	0,73
ver	Die Lehrkraft hat sich gut verständlich ausgedrückt.		1,62 (0,68)	1,79 (0,69)	1,71 (0,76)	1,89 (0,89)	1,81 (0,83)	1,82 (0,77)
wic	Manchmal hat die Lehrkraft extra gesagt, dass etwas sehr wichtig ist.**		0,17 (1,18)	0,35 (1,28)	0,44 (1,30)	0,50 (1,22)	0,56 (1,14)	0,40 (1,17)
bed	Bei manchen Wörtern wissen die Schülerinnen und Schüler eventuell nicht, was sie bedeuten.		0,63 (1,38)	0,13 (1,18)	0,49 (0,88)	0,42 (0,85)	0,60 (1,41)	0,70 (1,22)
zul	Manche Sätze hat die Lehrkraft zu lang gemacht.*		0,54 (1,31)	0,12 (0,98)	0,52 (1,14)	0,42 (1,06)	0,52 (1,13)	0,65 (1,14)
Sy_	Persönlichkeitswirkung	<i>Cronbachs Alpha:</i>	0,67	0,74	0,76	0,84	0,84	0,84
sym	Die Lehrkraft wirkte in diesem Video auf mich sympathisch.		0,54 (1,42)	0,61 (1,07)	0,74 (0,83)	0,86 (0,75)	0,72 (0,78)	0,67 (1,17)
beg	Die Lehrkraft wirkte in diesem Video auf mich begeistert.		0,41 (1,27)	0,49 (1,05)	0,33 (1,18)	0,64 (1,27)	0,67 (1,00)	0,59 (1,30)
nat	Die Lehrkraft wirkte in diesem Video auf mich natürlich.		0,40 (1,17)	0,62 (1,24)	0,66 (1,05)	0,54 (1,17)	0,70 (0,94)	0,86 (1,17)
Fa_	Fachspezifische Items	<i>Cronbachs Alpha:</i>	0,72	0,75	0,59	0,58	0,81	0,74
erf	Nach dieser Erklärung möchten Schüler_innen noch mehr darüber erfahren.		0,29 (0,92)	0,53 (1,04)	0,42 (0,92)	0,32 (1,06)	0,43 (1,06)	0,31 (0,88)
fas	Die Erklärung ist für Schüler_innen faszinierend.		0,64 (0,98)	0,66 (1,24)	0,33 (1,08)	0,50 (0,98)	0,72 (1,09)	0,55 (1,10)
nat	Die Erklärung zeigt, wie es in der Natur wirklich ist.		0,23 (1,03)	0,21 (1,22)	0,13 (1,43)	0,21 (0,97)	0,51 (1,13)	0,61 (1,15)
zan	Die Anzahl der Zusammenhänge ist für Schüler_innen angemessen.		0,58 (1,06)	0,56 (1,13)	0,53 (1,32)	0,49 (1,35)	0,72 (1,23)	0,56 (1,18)
hil	Die Erklärung hilft Schüler_innen andere Beispiele zu verstehen.		0,72 (0,98)	0,62 (0,91)	0,33 (0,94)	0,23 (0,88)	0,66 (0,93)	0,55 (0,98)
aus	Die Erklärung zeigt nur einen kleinen Ausschnitt auf die Frage „Thema der Erklärung“.**		0,72 (1,11)	0,62 (1,03)	0,33 (1,28)	0,23 (1,11)	0,66 (1,11)	0,55 (1,12)

Bem.: Einheitliche Skalierung aller Items (1 = Stimme voll zu; 2 = Stimme zu; 3 = Stimme eher zu; 4 = Stimme eher nicht zu; 5 = Stimme nicht zu; 6 = Stimme gar nicht zu); *M*: Mittelwert; *SD*: Standardabweichung; *N_{max}*: aufgrund fehlender Werte bei einzelnen Videos den Berechnungen pro Statusgruppe maximal zugrunde liegende Personenanzahl; *: Item wird in der zugehörigen Skala *rekodiert* berücksichtigt; **: Item wird in der zugehörigen Skala *nicht* berücksichtigt; ---: Item ist in der Skala der jeweiligen Statusgruppe *nicht* enthalten.

Didaktiker:innen

Tabelle: Übersicht zu Mittelwerten und Standardabweichungen der einzelnen Items sowie Itemtrennschärfen der in der Skala berücksichtigten Items und Cronbachs Alpha als Maß für die interne Konsistenz der Skalen für die Statusgruppe der Didaktiker:innen ($N_{max} = 24$)

			G _I	G _K	Ö _I	Ö _K	P _I	P _K
			M	M	M	M	M	M
			(SD)	(SD)	(SD)	(SD)	(SD)	(SD)
Skala/Itemabkürzung/Itemtext			r _{it}	r _{it}	r _{it}	r _{it}	r _{it}	r _{it}
Ad_	Adressatenorientierung	<i>Cronbachs Alpha:</i>	0,76	0,64	0,69	0,84	0,78	0,57
sch	Für Schülerinnen und Schüler war die Erklärung zu schwierig.*		2,39 (1,12)	2,04 (0,86)	1,91 (0,79)	2,65 (1,23)	2,43 (0,90)	2,48 (1,12)
lbw	Die Erklärung berücksichtigte die Lebenswelt der Schülerinnen und Schüler.		0,60 4,30 (1,15)	0,32 3,83 (1,27)	0,12 4,35 (1,03)	0,74 3,78 (1,23)	0,66 2,91 (0,90)	0,41 2,96 (1,30)
pas	Die Erklärung passte für Schülerinnen und Schüler dieser Jahrgangsstufe.		0,22 2,70 (1,15)	0,36 2,50 (0,89)	0,35 2,17 (0,94)	0,26 2,65 (1,15)	0,13 2,17 (0,78)	-0,17 2,17 (0,89)
red	Die Erklärung war angemessen didaktisch reduziert.		0,66 2,74 (1,25)	0,30 2,71 (1,08)	0,59 2,17 (1,27)	0,69 2,43 (1,16)	0,68 2,26 (1,10)	0,31 2,57 (1,16)
wvi	Schülerinnen und Schüler konnten der Erklärung mit ihrem Wissen gut folgen.		0,64 2,26 (1,05)	0,45 2,25 (0,85)	0,59 1,78 (0,60)	0,75 2,30 (0,93)	0,62 2,17 (0,89)	0,47 2,39 (0,84)
glo	Die Erklärung war adressatengerecht.**		0,53 2,39 (1,20)	0,56 2,33 (1,13)	0,65 2,04 (0,88)	0,80 2,43 (1,16)	0,74 2,22 (1,00)	0,66 2,30 (0,88)
			---	---	---	---	---	---
St_	Strukturiertheit	<i>Cronbachs Alpha:</i>	0,89	0,84	0,84	0,86	0,88	0,88
was	Es ist klar, was genau die Lehrkraft erklären wollte.		0,72 1,70 (0,82)	0,74 1,96 (0,86)	0,73 1,52 (0,51)	0,82 1,96 (1,07)	0,82 1,91 (0,90)	0,53 1,78 (0,67)
rot	Die Erklärung hatte einen roten Faden.		0,95 1,65 (0,71)	0,80 1,83 (0,87)	0,79 1,39 (0,50)	0,92 2,00 (1,04)	0,83 2,00 (0,91)	0,86 1,61 (0,66)
bau	Die Lehrkraft hat die Erklärung Schritt für Schritt aufgebaut.		0,60 1,61 (0,78)	0,73 1,79 (0,93)	0,70 1,39 (0,50)	0,53 2,13 (1,01)	0,66 1,78 (0,80)	0,80 1,57 (0,59)
ein	Die Lehrkraft hat zu viele Einzelheiten erklärt.*		0,76 2,22 (1,23)	0,42 2,33 (1,05)	0,47 1,78 (0,67)	0,58 2,78 (1,45)	0,62 2,52 (1,24)	0,73 2,30 (1,02)
glo	Die Erklärung war gut strukturiert.**		0,76 1,96 (0,98)	0,42 1,63 (0,77)	0,47 1,65 (0,78)	0,58 2,26 (0,92)	0,62 2,09 (0,85)	0,73 1,96 (0,98)
			---	---	---	---	---	---
Se_	Sprech- und Körperausdruck	<i>Cronbachs Alpha:</i>	0,87	0,89	0,87	0,80	0,89	0,91
sti	Die Lehrkraft hatte eine angenehm klingende Stimme.		0,58 1,50 (0,71)	0,60 1,37 (0,50)	0,70 1,47 (0,51)	0,52 1,40 (0,52)	0,43 1,40 (0,52)	0,67 1,47 (0,61)
aus	Die Lehrkraft hatte eine deutliche Aussprache.		0,45 1,40 (0,52)	0,49 1,42 (0,51)	0,69 1,47 (0,61)	0,45 1,40 (0,52)	0,57 1,40 (0,52)	0,64 1,42 (0,61)
ges	Die Lehrkraft hatte eine angemessene Sprechgeschwindigkeit.		0,72 1,70 (0,95)	0,69 2,05 (1,22)	0,46 1,53 (0,84)	0,38 2,10 (0,99)	0,69 1,90 (1,20)	0,78 1,58 (0,77)
pau	Die Lehrkraft hat die Sprechpausen so gesetzt, dass ich gut folgen konnte.		0,83 2,00 (1,16)	0,72 2,00 (1,41)	0,55 1,53 (0,70)	0,55 2,10 (1,20)	0,84 2,40 (1,43)	0,71 1,53 (0,84)
abw	Die Lehrkraft hatte eine abwechslungsreiche Sprechweise.		0,51 2,70 (0,95)	0,70 2,47 (1,31)	0,87 1,95 (0,97)	0,60 2,50 (0,85)	0,79 2,40 (1,08)	0,84 2,37 (1,17)

kor	Die Lehrkraft hatte einen Körperausdruck, der die Erklärung unterstützt hat.		2,50 (1,27)	1,89 (0,99)	1,95 (1,08)	1,90 (0,88)	2,00 (0,82)	2,16 (1,07)
zug	Die Lehrkraft hat sich den Schülerinnen und Schülern zugewendet.		2,00 (1,16)	1,79 (0,98)	1,79 (0,92)	1,80 (0,92)	1,80 (0,63)	1,74 (0,87)
glo	Die Art und Weise, wie die Lehrkraft spricht, gefällt mir.**		0,61 1,90 (1,10)	0,78 1,68 (0,75)	0,63 1,47 (0,70)	0,41 1,90 (0,88)	0,66 1,90 (0,88)	0,79 1,63 (0,76)
Sw_	Sprachliche Verständlichkeit	<i>Cronbachs Alpha:</i>	0,62	0,44	0,53	0,68	0,70	0,63
ver	Die Lehrkraft hat sich gut verständlich ausgedrückt.		1,70 (0,70)	1,67 (0,70)	1,57 (0,73)	1,83 (0,83)	1,96 (0,83)	1,61 (0,58)
wic	Manchmal hat die Lehrkraft extra gesagt, dass etwas sehr wichtig ist.**		0,36 3,87 (1,14)	0,18 3,38 (1,38)	0,05 3,61 (1,20)	0,60 3,96 (1,02)	0,51 3,39 (1,31)	0,16 3,87 (1,29)
bed	Bei manchen Wörtern wissen die Schülerinnen und Schüler eventuell nicht, was sie bedeuten.		0,51 3,65 (1,40)	0,22 3,63 (1,25)	0,66 2,00 (0,91)	0,36 2,74 (1,21)	0,61 3,35 (1,47)	0,66 2,87 (1,18)
zul	Manche Sätze hat die Lehrkraft zu lang gemacht.*		0,46 2,83 (1,19)	0,51 2,88 (1,30)	0,51 2,13 (0,82)	0,46 3,09 (1,54)	0,46 2,65 (1,19)	0,79 2,52 (1,16)
Sy_	Persönlichkeitswirkung	<i>Cronbachs Alpha:</i>	0,61	0,56	0,65	0,68	0,67	0,63
sym	Die Lehrkraft wirkte in diesem Video auf mich sympathisch.		0,57 1,90 (0,99)	0,51 1,68 (1,00)	0,49 1,53 (0,70)	0,64 1,60 (0,52)	0,48 1,60 (0,52)	0,37 1,63 (0,90)
beg	Die Lehrkraft wirkte in diesem Video auf mich begeistert.		0,36 3,10 (0,99)	0,25 3,26 (1,20)	0,44 2,53 (1,26)	0,43 2,60 (0,97)	0,59 2,80 (0,92)	0,47 2,95 (1,13)
nat	Die Lehrkraft wirkte in diesem Video auf mich natürlich.		0,34 2,30 (1,16)	0,35 2,37 (1,38)	0,42 2,26 (1,33)	0,37 1,90 (0,74)	0,42 1,70 (0,68)	0,48 2,05 (1,22)
Fa_	Fachspezifische Items	<i>Cronbachs Alpha:</i>	0,72	0,74	0,69	0,74	0,80	0,56
erf	Nach dieser Erklärung möchten Schüler_innen noch mehr darüber erfahren.		0,52 3,61 (1,08)	0,42 2,79 (1,18)	0,50 3,48 (1,28)	0,66 3,78 (1,17)	0,53 3,74 (1,21)	0,48 3,57 (1,34)
fas	Die Erklärung ist für Schüler_innen faszinierend.		0,36 3,26 (1,18)	0,67 3,46 (1,29)	0,53 3,00 (1,21)	0,58 3,70 (1,02)	0,53 3,70 (1,06)	0,33 3,52 (1,08)
nat	Die Erklärung zeigt, wie es in der Natur wirklich ist.		0,26 3,22 (1,09)	0,33 3,17 (0,87)	0,29 3,04 (1,33)	0,22 2,96 (1,11)	0,54 3,26 (1,01)	0,07 2,91 (1,08)
zan	Die Anzahl der Zusammenhänge ist für Schüler_innen angemessen.		0,69 2,30 (0,97)	0,47 2,17 (0,64)	0,49 2,09 (1,04)	0,29 2,57 (1,41)	0,75 2,30 (0,93)	0,27 2,65 (1,27)
hil	Die Erklärung hilft Schüler_innen andere Beispiele zu verstehen.		0,63 3,30 (1,36)	0,68 3,58 (1,21)	0,40 2,26 (1,21)	0,76 2,48 (1,12)	0,58 3,26 (1,45)	0,53 2,96 (1,26)
aus	Die Erklärung zeigt nur einen kleinen Ausschnitt auf die Frage „Thema der Erklärung“.**		0,63 4,13 (1,14)	0,68 4,75 (1,11)	0,40 4,57 (1,27)	0,76 3,70 (1,26)	0,58 4,26 (1,29)	0,53 3,78 (1,35)

Bem.: Einheitliche Skalierung aller Items (1 = Stimme voll zu; 2 = Stimme zu; 3 = Stimme eher zu; 4 = Stimme eher nicht zu; 5 = Stimme nicht zu; 6 = Stimme gar nicht zu); *M*: Mittelwert; *SD*: Standardabweichung; *N_{max}*: aufgrund fehlender Werte bei einzelnen Videos den Berechnungen pro Statusgruppe maximal zugrunde liegende Personenanzahl; *: Item wird in der zugehörigen Skala *rekodiert* berücksichtigt; **: Item wird in der zugehörigen Skala *nicht* berücksichtigt; ---: Item ist in der Skala der jeweiligen Statusgruppe *nicht* enthalten.

Lehrende/Metaperspektive

Tabelle 62: Übersicht zu Mittelwerten und Standardabweichungen der einzelnen Items sowie Itemtrennschärfen der in der Skala berücksichtigten Items und Cronbachs Alpha als Maß für die interne Konsistenz der Skalen für die Statusgruppe der Lehrenden bzw. der Metaperspektive (Studierende, Lehrkräfte, Didaktiker:innen; $N_{max} = 97$)

			G_I	G_K	\bar{O}_I	\bar{O}_K	P_I	P_K
			M	M	M	M	M	M
			(SD)	(SD)	(SD)	(SD)	(SD)	(SD)
			r_{it}	r_{it}	r_{it}	r_{it}	r_{it}	r_{it}
Skala/	Itemabkürzung/	Itemtext						
Ad_	Adressatenorientierung	<i>Cronbachs Alpha:</i>	0,69	0,71	0,49	0,62	0,83	0,76
sch	Für Schülerinnen und Schüler war die Erklärung zu schwierig.*		2,31 (0,98)	2,46 (1,01)	1,88 (0,79)	2,23 (1,07)	2,63 (0,99)	2,66 (1,13)
lbw	Die Erklärung berücksichtigte die Lebenswelt der Schülerinnen und Schüler.		0,34 3,70 (1,04)	0,32 3,74 (1,18)	0,06 3,86 (1,14)	0,24 3,53 (1,03)	0,60 3,08 (1,01)	0,47 2,83 (1,15)
pas	Die Erklärung passte für Schülerinnen und Schüler dieser Jahrgangsstufe.		0,27 2,72 (1,12)	0,38 2,80 (1,01)	0,02 2,65 (1,11)	0,21 2,85 (1,09)	0,50 2,49 (1,00)	0,38 2,45 (0,93)
red	Die Erklärung war angemessen didaktisch reduziert.		0,58 2,66 (1,16)	0,49 2,87 (1,08)	0,40 2,51 (1,11)	0,42 2,59 (1,03)	0,67 2,42 (1,04)	0,54 2,58 (1,03)
wwi	Schülerinnen und Schüler konnten der Erklärung mit ihrem Wissen gut folgen.		0,59 2,24 (0,86)	0,59 2,38 (0,95)	0,58 1,82 (0,72)	0,54 2,02 (0,74)	0,68 2,28 (0,94)	0,61 2,34 (0,95)
glo	Die Erklärung war adressatengerecht.**		0,50 2,50 (1,08)	0,58 2,75 (1,13)	0,31 2,44 (1,10)	0,45 2,60 (1,06)	0,71 2,43 (0,99)	0,63 2,45 (1,01)
			---	---	---	---	---	---
St_	Strukturiertheit	<i>Cronbachs Alpha:</i>	0,80	0,83	0,78	0,85	0,83	0,88
was	Es ist klar, was genau die Lehrkraft erklären wollte.		1,77 (0,81)	2,13 (0,97)	1,73 (0,75)	2,03 (1,05)	1,97 (0,90)	1,86 (0,70)
rot	Die Erklärung hatte einen roten Faden.		0,67 1,72 (0,84)	0,71 2,07 (0,95)	0,60 1,59 (0,68)	0,77 2,14 (1,05)	0,73 1,99 (0,84)	0,71 1,76 (0,73)
bau	Die Lehrkraft hat die Erklärung Schritt für Schritt aufgebaut.		0,74 1,59 (0,76)	0,74 1,91 (0,83)	0,69 1,52 (0,65)	0,85 1,98 (0,92)	0,76 1,89 (0,79)	0,78 1,73 (0,68)
ein	Die Lehrkraft hat zu viele Einzelheiten erklärt.*		0,62 2,47 (1,21)	0,72 2,79 (1,23)	0,61 2,37 (1,26)	0,64 3,01 (1,30)	0,64 2,69 (1,18)	0,78 2,81 (1,28)
glo	Die Erklärung war gut strukturiert.**		0,36 1,90 (0,88)	0,42 2,11 (0,91)	0,23 1,87 (0,93)	0,51 2,23 (0,84)	0,49 2,05 (0,86)	0,58 1,98 (0,93)
			---	---	---	---	---	---
Se_	Sprech- und Körperausdruck	<i>Cronbachs Alpha:</i>	0,82	0,82	0,85	0,85	0,89	0,89
sti	Die Lehrkraft hatte eine angenehm klingende Stimme.		1,45 (0,65)	1,67 (0,86)	1,70 (0,83)	1,61 (0,76)	1,57 (0,76)	1,72 (0,85)
aus	Die Lehrkraft hatte eine deutliche Aussprache.		0,55 1,43 (0,76)	0,43 1,50 (0,79)	0,58 1,40 (0,59)	0,56 1,61 (0,64)	0,59 1,39 (0,53)	0,55 1,45 (0,68)
ges	Die Lehrkraft hatte eine angemessene Sprechgeschwindigkeit.		0,41 1,69 (0,92)	0,55 2,00 (1,14)	0,58 1,73 (0,84)	0,50 2,10 (1,12)	0,71 1,80 (0,96)	0,55 1,77 (0,87)
pau	Die Lehrkraft hat die Sprechpausen so gesetzt, dass ich gut folgen konnte.		0,70 1,86 (1,04)	0,55 2,27 (1,30)	0,55 1,88 (0,83)	0,66 2,69 (1,45)	0,64 2,04 (1,12)	0,56 1,83 (0,87)
abw	Die Lehrkraft hatte eine abwechslungsreiche Sprechweise.		0,59 2,71 (1,04)	0,59 3,12 (1,37)	0,68 2,48 (1,07)	0,73 2,73 (1,06)	0,70 2,22 (1,03)	0,67 2,62 (1,22)
			0,55	0,65	0,77	0,66	0,74	0,67

kor	Die Lehrkraft hatte einen Körperausdruck, der die Erklärung unterstützt hat.		2,47 (1,02)	2,47 (1,16)	2,20 (1,04)	2,39 (1,04)	1,84 (0,90)	2,37 (1,09)
zug	Die Lehrkraft hat sich den Schülerinnen und Schülern zugewendet.		0,68 (0,96)	0,69 (1,01)	0,55 (1,03)	0,70 (0,94)	0,72 (0,76)	0,60 (0,84)
glo	Die Art und Weise, wie die Lehrkraft spricht, gefällt mir.**		0,50 (1,00)	0,52 (1,01)	0,51 (1,00)	0,53 (1,21)	0,67 (0,93)	0,51 (1,02)
Sw_	Sprachliche Verständlichkeit	<i>Cronbachs Alpha:</i>	---	---	---	---	---	---
			0,60	0,56	0,64	0,59	0,67	0,72
ver	Die Lehrkraft hat sich gut verständlich ausgedrückt.		1,73 (0,76)	1,85 (0,83)	1,67 (0,82)	1,86 (0,84)	1,84 (0,78)	1,80 (0,80)
wic	Manchmal hat die Lehrkraft extra gesagt, dass etwas sehr wichtig ist.**		0,33 (1,25)	0,38 (1,18)	0,37 (1,16)	0,40 (1,13)	0,44 (1,18)	0,39 (1,22)
bed	Bei manchen Wörtern wissen die Schülerinnen und Schüler eventuell nicht, was sie bedeuten.		---	---	---	---	---	---
			3,29 (1,31)	3,24 (1,19)	2,04 (1,02)	2,39 (1,14)	3,33 (1,37)	2,91 (1,16)
zul	Manche Sätze hat die Lehrkraft zu lang gemacht.*		0,48 (1,17)	0,30 (1,22)	0,47 (1,06)	0,33 (1,29)	0,59 (1,17)	0,62 (1,12)
Sy_	Persönlichkeitswirkung	<i>Cronbachs Alpha:</i>	0,46	0,43	0,54	0,46	0,47	0,66
sym	Die Lehrkraft wirkte in diesem Video auf mich sympathisch.		0,67	0,66	0,72	0,79	0,83	0,78
beg	Die Lehrkraft wirkte in diesem Video auf mich begeistert.		1,63 (0,78)	1,93 (1,12)	1,73 (0,92)	1,86 (0,82)	1,57 (0,65)	1,83 (1,08)
nat	Die Lehrkraft wirkte in diesem Video auf mich natürlich.		0,58 (1,03)	0,53 (1,23)	0,61 (1,24)	0,65 (1,00)	0,72 (0,96)	0,61 (1,21)
Fa_	Fachspezifische Items	<i>Cronbachs Alpha:</i>	0,41	0,42	0,43	0,61	0,61	0,52
erf	Nach dieser Erklärung möchten Schüler_innen noch mehr darüber erfahren.		2,43 (1,21)	2,52 (1,23)	2,22 (1,17)	2,39 (1,13)	1,82 (0,88)	2,28 (1,18)
fas	Die Erklärung ist für Schüler_innen faszinierend.		0,44 (1,05)	0,46 (1,19)	0,54 (1,09)	0,63 (1,05)	0,69 (1,07)	0,71 (1,10)
nat	Die Erklärung zeigt, wie es in der Natur wirklich ist.		0,61 (1,07)	0,65 (1,15)	0,52 (1,37)	0,58 (0,98)	0,68 (1,05)	0,58 (1,03)
zan	Die Anzahl der Zusammenhänge ist für Schüler_innen angemessen.		0,40 (0,94)	0,44 (1,00)	0,37 (1,10)	0,24 (1,16)	0,49 (1,00)	0,42 (1,10)
hil	Die Erklärung hilft Schüler_innen andere Beispiele zu verstehen.		0,57 (1,13)	0,57 (1,06)	0,50 (0,99)	0,43 (0,92)	0,69 (1,10)	0,54 (1,01)
aus	Die Erklärung zeigt nur einen kleinen Ausschnitt auf die Frage „Thema der Erklärung“.**		0,61 (1,15)	0,58 (1,14)	0,40 (1,24)	0,48 (1,19)	0,58 (1,15)	0,61 (1,24)

Bem.: Einheitliche Skalierung aller Items (1 = Stimme voll zu; 2 = Stimme zu; 3 = Stimme eher zu; 4 = Stimme eher nicht zu; 5 = Stimme nicht zu; 6 = Stimme gar nicht zu); *M*: Mittelwert; *SD*: Standardabweichung; *N_{max}*: aufgrund fehlender Werte bei einzelnen Videos den Berechnungen pro Statusgruppe maximal zugrunde liegende Personenanzahl; *: Item wird in der zugehörigen Skala *rekodiert* berücksichtigt; **: Item wird in der zugehörigen Skala *nicht* berücksichtigt; ---: Item ist in der Skala der jeweiligen Statusgruppe *nicht* enthalten.

E) Prüfung auf Normalverteilung

Ergebnisse zur Prüfung auf Normalverteilung getrennt nach Statusgruppe für das Globalurteil sowie die fünf Aspekte guten Erklärens. Für die Statusgruppe der Schüler:innen anhand des Kolmogorov-Smirnov-Tests und für die Gruppen der Lehrenden anhand des Shapiro-Wilk-Tests.

Schüler:innen

Tabelle: Ergebnisse der Kolmogorov-Smirnov-Tests für Schüler:innen je Konstrukt und Videovignette

Konstrukt	Video	Video G _l	Video G _k	Video Ö _l	Video Ö _k	Video P _l	Video P _k
Globalurteil	N	122	121	119	121	120	116
	D	0,13	0,16	0,15	0,15	0,16	0,14
	p	≤ 0,01	≤ 0,01	≤ 0,01	≤ 0,01	≤ 0,01	≤ 0,01
Adressatenorientierung	N	122	121	120	121	121	120
	D	0,12	0,15	0,11	0,12	0,09	0,09
	p	≤ 0,01	≤ 0,01	≤ 0,01	≤ 0,01	0,02	0,03
Strukturiertheit	N	122	121	120	121	121	120
	D	0,12	0,11	0,11	0,10	0,14	0,12
	p	≤ 0,01	≤ 0,01	≤ 0,01	≤ 0,01	≤ 0,01	≤ 0,01
Sprech- und Körperausdruck	N	43	70	70	43	42	68
	D	0,14	0,12	0,08	0,08	0,16	0,11
	p	0,03	0,01	0,20	0,20	0,01	0,04
Sprachliche Verständlichkeit	N	122	121	120	121	121	120
	D	0,11	0,12	0,15	0,11	0,10	0,10
	p	≤ 0,01	≤ 0,01	≤ 0,01	≤ 0,01	≤ 0,01	≤ 0,01
Fachlicher Aspekt	N	122	121	120	121	121	120
	D	0,07	0,10	0,08	0,10	0,08	0,08
	p	0,20	≤ 0,01	0,04	≤ 0,01	0,05	0,06

Bem.: N: Stichprobengröße; D: Kolmogorov-Smirnovs D; p: p-Wert.

Studierende

Tabelle: Ergebnisse der Shapiro-Wilk-Tests für Studierende je Konstrukt und Videovignette

Konstrukt	Video	Video G _l	Video G _k	Video Ö _l	Video Ö _k	Video P _l	Video P _k
Globalurteil	N	50	51	51	50	50	50
	W	0,91	0,95	0,90	0,97	0,93	0,93
	p	≤ 0,01	0,03	≤ 0,01	0,25	≤ 0,01	≤ 0,01
Adressatenorientierung	N	44	43	43	44	43	42
	W	0,98	0,94	0,97	0,97	0,93	0,98
	p	0,51	0,04	0,22	0,38	0,01	0,80
Strukturiertheit	N	44	43	43	44	43	42
	W	0,92	0,96	0,96	0,97	0,97	0,95
	p	≤ 0,01	0,15	0,10	0,37	0,31	0,07
Sprech- und Körperausdruck	N	27	22	22	27	27	22
	W	0,91	0,96	0,92	0,96	0,87	0,95
	p	0,02	0,58	0,08	0,34	≤ 0,01	0,36
Sprachliche Verständlichkeit	N	44	43	43	44	43	42
	W	0,94	0,97	0,95	0,97	0,97	0,96
	p	0,02	0,41	0,04	0,37	0,22	0,19
Fachlicher Aspekt	N	44	43	43	44	43	42
	W	0,97	0,96	0,98	0,98	0,96	0,97
	p	0,30	0,14	0,56	0,49	0,12	0,25

Bem.: N: Stichprobengröße; W: Shapiros-Wilks W; p: p-Wert.

Lehrkräfte

Tabelle: Ergebnisse der Shapiro-Wilk-Tests für Lehrkräfte je Konstrukt und Videovignette

Konstrukt	Video	Video G _I	Video G _K	Video Ö _I	Video Ö _K	Video P _I	Video P _K
Globalurteil	<i>N</i>	34	33	32	34	34	32
	<i>W</i>	0,90	0,94	0,90	0,96	0,89	0,91
	<i>p</i>	≤ 0,01	0,09	≤ 0,01	0,18	≤ 0,01	0,01
Adressatenorientierung	<i>N</i>	29	28	28	27	27	28
	<i>W</i>	0,96	0,96	0,96	0,93	0,95	0,95
	<i>p</i>	0,35	0,42	0,33	0,09	0,27	0,16
Strukturiertheit	<i>N</i>	29	28	28	27	27	28
	<i>W</i>	0,92	0,97	0,91	0,92	0,89	0,91
	<i>p</i>	0,03	0,55	0,03	0,03	≤ 0,01	0,02
Sprech- und Körperausdruck	<i>N</i>	12	19	19	12	12	19
	<i>W</i>	0,91	0,87	0,94	0,84	0,90	0,87
	<i>p</i>	0,20	0,01	0,28	0,02	0,14	0,01
Sprachliche Verständlichkeit	<i>N</i>	29	28	28	27	27	28
	<i>W</i>	0,97	0,95	0,94	0,93	0,95	0,95
	<i>p</i>	0,57	0,19	0,14	0,08	0,21	0,15
Fachlicher Aspekt	<i>N</i>	29	28	28	27	27	28
	<i>W</i>	0,96	0,98	0,96	0,97	0,96	0,97
	<i>p</i>	0,40	0,72	0,29	0,64	0,38	0,48

Bem.: *N*: Stichprobengröße; *W*: Shapiros-Wilks *W*; *p*: p-Wert.

Didaktiker:innen

Tabelle: Ergebnisse der Shapiro-Wilk-Tests für Didaktiker:innen je Konstrukt und Videovignette

Konstrukt	Video	Video G _I	Video G _K	Video Ö _I	Video Ö _K	Video P _I	Video P _K
Globalurteil	<i>N</i>	30	30	29	30	30	30
	<i>W</i>	0,92	0,92	0,95	0,95	0,94	0,93
	<i>p</i>	0,03	0,03	0,22	0,18	0,08	0,04
Adressatenorientierung	<i>N</i>	23	24	23	23	23	23
	<i>W</i>	0,97	0,93	0,96	0,89	0,97	0,96
	<i>p</i>	0,67	0,12	0,52	0,01	0,65	0,49
Strukturiertheit	<i>N</i>	23	24	23	23	23	23
	<i>W</i>	0,83	0,90	0,87	0,88	0,93	0,92
	<i>p</i>	≤ 0,01	0,03	≤ 0,01	≤ 0,01	0,13	0,05
Sprech- und Körperausdruck	<i>N</i>	10	19	19	10	10	19
	<i>W</i>	0,94	0,88	0,90	0,97	0,90	0,87
	<i>p</i>	0,06	0,02	0,04	0,85	0,21	0,02
Sprachliche Verständlichkeit	<i>N</i>	23	24	23	23	23	23
	<i>W</i>	0,94	0,95	0,89	0,94	0,94	0,93
	<i>p</i>	0,18	0,31	0,02	0,16	0,14	0,11
Fachlicher Aspekt	<i>N</i>	23	24	23	23	23	23
	<i>W</i>	0,94	0,98	0,95	0,93	0,96	0,94
	<i>p</i>	0,21	0,85	0,32	0,14	0,67	0,17

Bem.: *N*: Stichprobengröße; *W*: Shapiros-Wilks *W*; *p*: p-Wert.

F) Prüfung auf Varianzhomogenität

Globalurteil

Tabelle: Ergebnisse der Tests auf Varianzhomogenität der Globalurteile zwischen den vier Statusgruppen

Video	Video G _l	Video G _k	Video Ö _l	Video Ö _k	Video P _l	Video P _k
<i>F</i>	2,17	1,65	0,86	1,68	0,96	1,26
<i>df</i>	3	3	3	3	3	3
<i>Residuen</i>	232	231	288	231	229	224
<i>p</i>	0,09	0,18	0,46	0,17	0,41	0,29

Bem.: *F*: F-Wert; *df*: Freiheitsgrade Gruppen; *Residuen*: Freiheitsgrade teilnehmende Personen; *p*: p-Wert.

Adressatenorientierung

Tabelle: Ergebnisse der Tests auf Varianzhomogenität der Adressatenorientierung zwischen den vier Statusgruppen

Video	Video G _l	Video G _k	Video Ö _l	Video Ö _k	Video P _l	Video P _k
<i>F</i>	1,10	0,24	0,95	2,29	0,67	0,40
<i>df</i>	3	3	3	3	3	3
<i>Residuen</i>	214	212	210	211	210	209
<i>p</i>	0,35	0,87	0,42	0,08	0,85	0,75

Bem.: *F*: F-Wert; *df*: Freiheitsgrade Gruppen; *Residuen*: Freiheitsgrade teilnehmende Personen; *p*: p-Wert.

Strukturiertheit

Tabelle: Ergebnisse der Tests auf Varianzhomogenität der Strukturiertheit zwischen den vier Statusgruppen

Video	Video G _l	Video G _k	Video Ö _l	Video Ö _k	Video P _l	Video P _k
<i>F</i>	0,28	0,34	1,12	1,37	0,51	0,35
<i>df</i>	3	3	3	3	3	3
<i>Residuen</i>	214	212	210	211	210	209
<i>p</i>	0,84	0,80	0,34	0,25	0,68	0,79

Bem.: *F*: F-Wert; *df*: Freiheitsgrade Gruppen; *Residuen*: Freiheitsgrade teilnehmende Personen; *p*: p-Wert.

Sprech- und Körperausdruck

Tabelle: Ergebnisse der Tests auf Varianzhomogenität des Sprech- und Körperausdrucks zwischen den vier Statusgruppen

Video	Video G _l	Video G _k	Video Ö _l	Video Ö _k	Video P _l	Video P _k
<i>F</i>	0,57	0,18	0,19	1,09	0,33	0,57
<i>df</i>	3	3	3	3	3	3
<i>Residuen</i>	88	126	126	88	87	124
<i>p</i>	0,64	0,15	0,90	0,36	0,80	0,63

Bem.: *F*: F-Wert; *df*: Freiheitsgrade Gruppen; *Residuen*: Freiheitsgrade teilnehmende Personen; *p*: p-Wert.

Verständlichkeit der Sprache

Tabelle: Ergebnisse der Tests auf Varianzhomogenität der Verständlichkeit der Sprache zwischen den vier Statusgruppen

Video	Video G _l	Video G _k	Video Ö _l	Video Ö _k	Video P _l	Video P _k
<i>F</i>	0,14	1,06	0,69	0,52	0,17	0,45
<i>df</i>	3	3	3	3	3	3
<i>Residuen</i>	124	212	210	211	210	209
<i>p</i>	0,94	0,37	0,56	0,67	0,92	0,72

Bem.: *F*: F-Wert; *df*: Freiheitsgrade Gruppen; *Residuen*: Freiheitsgrade teilnehmende Personen; *p*: p-Wert.

Fachspezifischer Aspekt

Tabelle: Ergebnisse der Tests auf Varianzhomogenität des fachspezifischen Aspekts zwischen den vier Statusgruppen

Video	Video G _l	Video G _k	Video Ö _l	Video Ö _k	Video P _l	Video P _k
<i>F</i>	2,00	0,33	0,52	2,92	1,90	1,51
<i>df</i>	3	3	3	3	3	3
<i>Residuen</i>	214	212	210	211	210	209
<i>p</i>	0,12	0,80	0,67	0,04	0,13	0,24

Bem.: *F*: F-Wert; *df*: Freiheitsgrade Gruppen; *Residuen*: Freiheitsgrade teilnehmende Personen; *p*: p-Wert.

G) Inhaltlich-strukturierende qualitative Inhaltsanalyse – absolute und relative Häufigkeiten der Nennungen

Tabelle: Absolute und relative Häufigkeiten der Kodierungen in den Kategorien

Kategorie	Statusgruppe	Absolute Häufigkeiten					Relative Häufigkeiten				
		SuS	Stu	Leh	Did	Über alle	SuS	Stu	Leh	Did	Über alle
Adressatenorientierung											
Adressatenorientierung global		7	6	1	1	15	0,62	1,53	0,40	0,25	0,69
Didaktische Reduktion		0	3	5	7	15	0,00	0,77	1,98	1,72	0,69
Passung des Informationsgehalts		25	6	9	5	45	2,20	1,53	3,57	1,23	2,06
Passung des Schwierigkeitsgrads		31	8	5	1	45	2,73	2,05	1,98	0,25	2,06
Berücksichtigung des Vorwissens		15	5	3	6	29	1,32	1,28	1,19	1,47	1,33
Übererklärung		27	6	1	0	34	2,38	1,53	0,40	0,00	1,56
Erklärtempo		41	26	5	9	81	3,61	6,65	1,98	2,21	3,71
Berücksichtigung von (Fehl-)Vorstellungen		0	2	3	6	11	0,00	0,51	1,19	1,47	0,50
Interessante u. motivierende Gestaltung		69	9	3	0	81	6,08	2,30	1,19	0,00	3,71
Lebensweltbezug		3	0	0	4	7	0,26	0,00	0,00	0,98	0,32
(Aktiver) Einbezug der RezipientInnen		13	1	13	7	34	1,15	0,26	5,16	1,72	1,56
Σ Adressatenorientierung		231	72	48	46	397	20,35	18,41	19,05	11,30	18,17
Strukturiertheit											
Strukturiertheit global		12	4	8	9	33	1,06	1,02	3,17	2,21	1,51
Fokussierung auf Wesentliches		24	3	3	2	32	2,11	0,77	1,19	0,49	1,46
Logischer Aufbau		29	11	11	6	57	2,56	2,81	4,37	1,47	2,61
Schrittweise Erklärung		5	2	0	6	13	0,44	0,51	0,00	1,47	0,59
Zusammenfassung zentraler Punkte		0	3	0	0	3	0,00	0,77	0,00	0,00	0,14
Zielangabe		1	2	4	6	13	0,09	0,51	1,59	1,47	0,59
Übereinstimmung Erklärziel – Erklärinhalt		6	4	2	2	14	0,53	1,02	0,79	0,49	0,64
Σ Strukturiertheit		77	29	28	31	165	6,78	7,42	11,11	7,62	7,55
Repräsentationen											
Repräsentationen global		18	2	2	5	27	1,59	0,51	0,79	1,23	1,24
Einsatz bildlicher Visualisierungen		187	31	20	20	258	16,48	7,93	7,94	4,91	11,81
Erklärbedürftigkeit der Grafik		3	1	1	6	11	0,26	0,26	0,40	1,47	0,50

Fachliche Angemessenheit	1	11	4	16	32	0,09	2,81	1,59	3,93	1,46
Sichtbarkeit der Visualisierung	3	4	4	3	14	0,26	1,02	1,59	0,74	0,64
Verständlichkeit der Visualisierung	66	27	13	45	151	5,81	6,91	5,16	11,06	6,91
Ästhetische Gestaltung	7	5	4	3	19	0,62	1,28	1,59	0,74	0,87
Einsatz von Analogien, Metaphern u. Vergleichen	0	0	0	0	0	0,00	0,00	0,00	0,00	0,00
Themenspezifisch: Vergleich bei Mukoviszidose	12	0	2	4	18	1,06	0,00	0,79	0,98	0,82
Einsatz von Beispielen	10	3	1	0	14	0,88	0,77	0,40	0,00	0,64
Ausreichende Bezugnahme auf das Beispiel	7	3	7	11	28	0,62	0,77	2,78	2,70	1,28
Gelungene Verbindung der Repräsentationen	8	7	2	18	35	0,70	1,79	0,79	4,42	1,60
Σ Repräsentationen	322	94	60	131	607	28,37	24,04	23,81	32,19	27,78
Sprech- & Körperausdruck										
Artikulation	7	6	2	5	20	0,62	1,53	0,79	1,23	0,92
Prosodische Merkmale	13	4	2	2	21	1,15	1,02	0,79	0,49	0,96
Stimmklang	0	0	0	1	1	0,00	0,00	0,00	0,25	0,05
Sprechgeschwindigkeit	10	7	3	0	20	0,88	1,79	1,19	0,00	0,92
Sprechpausen	0	1	0	0	1	0,00	0,26	0,00	0,00	0,05
Variation der Stimme	9	14	2	1	26	0,79	3,58	0,79	0,25	1,19
Sprechflüssigkeit	4	5	0	0	9	0,35	1,28	0,00	0,00	0,41
Zugewandter Körperausdruck	0	0	2	1	3	0,00	0,00	0,79	0,25	0,14
Unterstützender Körperausdruck	3	4	0	2	9	0,26	1,02	0,00	0,49	0,41
Σ Sprech- und Körperausdruck	46	41	11	12	110	4,05	10,49	4,37	2,95	5,03
Sprache										
Sprache global	5	5	3	5	18	0,44	1,28	1,19	1,23	0,82
Wortschatz	2	6	2	6	16	0,18	1,53	0,79	1,47	0,73
Umgang mit Fachworten	7	6	1	8	22	0,62	1,53	0,40	1,97	1,01
Satzbau	2	9	1	3	15	0,18	2,30	0,40	0,74	0,69
(Un-)Präziser Ausdruck	0	5	12	5	22	0,00	1,28	4,76	1,23	1,01
Hervorheben von Wichtigem	0	0	0	0	0	0,00	0,00	0,00	0,00	0,00
Σ Sprache	16	31	19	27	93	1,41	7,93	7,54	6,63	4,26
Fach										
Fach global	0	0	0	3	3	0,00	0,00	0,00	0,74	0,14
Korrektheit	3	13	8	20	44	0,26	3,32	3,17	4,91	2,01
Fachspezifische Erklärungsformen	0	0	1	3	4	0,00	0,00	0,40	0,74	0,18
Berücksichtigung von Nichtwissen	0	2	3	9	14	0,00	0,51	1,19	2,21	0,64
Vagheit/Eindeutigkeit	4	4	4	5	17	0,35	1,02	1,59	1,23	0,78
Gefühl der (Un-)Sicherheit	0	0	0	0	0	0,00	0,00	0,00	0,00	0,00
Faszination und Interessiertheit	0	0	0	0	0	0,00	0,00	0,00	0,00	0,00
Phänomenorientierung	0	14	6	27	47	0,00	3,58	2,38	6,63	2,15
Prinzipienorientierung	0	2	0	7	9	0,00	0,51	0,00	1,72	0,41
Vollständigkeit	1	8	11	12	32	0,09	2,05	4,37	2,95	1,46
„Fehlende“ Inhalte, die über das Erklärziel hinausgehen	16	2	2	1	21	1,41	0,51	0,79	0,25	0,96
Zusammenhänge werden deutlich	13	8	5	12	38	1,15	2,05	1,98	2,95	1,74
Σ Fach	37	53	40	99	229	3,26	13,55	15,87	24,32	10,48
Globale Aussagen										
gut erklärt	127	20	7	12	166	11,19	5,12	2,78	2,95	7,60
verständlich	86	4	1	4	95	7,58	1,02	0,40	0,98	4,35
klar	3	1	1	6	11	0,26	0,26	0,40	1,47	0,50
nachvollziehbar	8	4	1	5	18	0,70	1,02	0,40	1,23	0,82
anschaulich	52	17	12	4	85	4,58	4,35	4,76	0,98	3,89
übersichtlich	23	4	4	3	34	2,03	1,02	1,59	0,74	1,56
sachlich	5	0	2	1	8	0,44	0,00	0,79	0,25	0,37
informativ	4	1	0	2	7	0,35	0,26	0,00	0,49	0,32
Σ Globale Aussagen	308	51	28	37	424	27,14	13,04	11,11	9,09	19,41
Lehrperson	11	1	1	5	18	0,97	0,26	0,40	1,23	0,82
Merkbarkeit u. Lernbarkeit	33	7	5	0	45	2,91	1,79	1,98	0,00	2,06
Sonstiges	54	12	12	19	97	4,76	3,07	4,76	4,67	4,44
GESAMT	1135	391	252	407	2185	100,00	100,00	100,00	100,00	100,00