

*Markus Gebhardt**Kirsten Diehl**Andreas Mübling*

Online-Lernverlaufsmessung für alle Schülerinnen und Schüler in inklusiven Klassen

Zusammenfassung

Mit der Entwicklung zu einer inklusiven Schule sind erhebliche Veränderungen verbunden. Eine zentrale Herausforderung für Lehrkräfte der Allgemeinen Schule und Förderschullehrkräfte stellt dabei der gemeinsame Unterricht von Kindern mit sonderpädagogischen und ohne sonderpädagogischen Förderbedarf dar. Wie gelingt es, Leistungen und Leistungsentwicklungen der Kinder valide einzuschätzen? Wie können Lernlücken oder Stagnationen frühzeitig erkannt werden und Lernentwicklungen durch angemessene Förderung unterstützt werden? Eine mögliche Antwort auf diese Fragen, bietet die Lernverlaufsdagnostik. In diesem Beitrag wird die Onlineplattform LEVUMI vorgestellt. Sie bietet rechnergestützt kurze, leicht zu handhabende Messverfahren an – sog. curriculumbasierte Tests, mittels derer Lehrkräfte Einblick in Lernverläufe der Schülerinnen und Schüler im Bereich Lesen (Anfangsunterricht) erhalten. Die Ergebnisse der Lernfortschrittsdiagnostik bietet die Grundlage für eine darauf abgestimmte Förderung. Zudem erhält die Lehrkraft Rückmeldung zur Wirksamkeit ihres pädagogischen Handelns. Die bisherigen statistischen Überprüfungen der LEVUMI Testverfahren in Hinblick auf die Reliabilität sind positiv und unterstützen die weitere Arbeit und Entwicklung der Onlineplattform LEVUMI.

Bedarf der Lernverlaufsdagnostik in der Praxis

Ziel des inklusiven Unterrichts ist es, Schülerinnen und Schüler mit sonderpädagogischen und ohne sonderpädagogischen Förderbedarf (SPF) gemeinsam zu unterrichten und zu fördern. Hierbei liegt die Herausforderung darin, die unterschiedlichen schulischen Lern- und Leistungsvoraussetzungen der Schülerinnen und Schüler zu berücksichtigen und optimal zu unterrichten. Im Fokus des inklusiven Unterrichts steht ein individualisierter Unterricht (Gebhardt, Schwab, Krammer, Gasteiger-Klicpera & Sälzer, 2014), in dem den Schülerinnen und Schülern individuell und kriteriumsorientiert die Leistungsstände rückgemeldet werden. Eine Rückmeldung aufgrund des Leistungsvergleichs mit der Klasse (soziale Bezugsnorm) ist vor allem bei leistungsschwachen Schülerinnen und Schülern wenig sinnvoll, da dies die Motivation und auch das Selbstkonzept mindert (Elbaum, 2002).

Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf im Bereich Lernen erlernen Inhalte des regulären Curriculums verzögert und haben schulische Leistungen, die meist ein bis zwei Schuljahre hinter jenen ihrer Klassenkameraden liegen (Wocken & Gröhlich, 2007; Kocaj, Kuhl, Kroth, Pant & Stanat, 2014). Schülerinnen und Schüler mit einem sonderpädagogischen Förderbedarf im Bereich Lernen erzielen Lernfortschritte, die Leistungsdifferenzen zu den Klassenkameraden ohne sonderpädagogischen Förderbedarf werden dennoch immer größer (Blackorby, Chorost, Garza & Guzman, 2003; Gebhardt, Zehner & Hessels, 2014). Dieser Effekt ist auch in Klassen der allgemeinen Schule zwischen leistungsstarken und leistungsschwachen Kindern bekannt (Becker, Lütke, Trautwein & Baumert, 2006). Während einige Schülerinnen und Schüler im regulären Unterricht Lernfortschritte erzielen, schlagen die gewählten Methoden bei anderen Mitschülerinnen und Mitschülern nicht an. In der Forschung werden diese Kinder als sog. Non Responder (Fuchs, Fuchs & Compton, 2004) bezeichnet. Dies tritt vor allem bei

Schülerinnen und Schülern mit einem Risiko für einen sonderpädagogischen Förderbedarf im Bereich Lernen in der Grundschule auf. Im aktuellen Schulsystem wird für diese Schülerinnen und Schüler selten eine präventive Förderung angeboten. Zusätzliche Fördermaßnahmen erfolgen, sobald ein sonderpädagogischer Förderbedarf festgestellt wurde. Eine solche Feststellung erfolgt erst, wenn die Lernlücke der Non Responder im Vergleich zu den durchschnittlichen Schülerinnen und Schülern in der Grundschule mehr als einer Standardabweichung entspricht. Man spricht in diesem Fall von der „Wait to Fail Problematik“ (Huber & Grosche, 2012). Um rechtzeitig präventive Maßnahmen im Unterricht zu ergreifen, hat sich das Konzept der Lernverlaufsdagnostik als effektiv erwiesen (Stecker, Fuchs & Fuchs, 2005). Aus diesem Grund ist die Lernverlaufsdagnostik in neueren Unterrichts- und Förderkonzepten, wie zum Beispiel dem Ansatz Response to Intervention (RtI) (Huber & Grosche, 2012; Mahlau, Diehl, Voß & Hartke, 2011), ein zentraler Baustein.

Das Ziel der Lernverlaufsdagnostik ist es, Lernverläufe von Schülerinnen und Schülern über einen längeren Zeitraum hinweg darzustellen und den Lernfortschritt an Lehrkräfte rückzumelden. Anhand individueller Rückmeldungen über Leistungsstände und -entwicklungen einzelner Schülerinnen und Schüler, Schülergruppen oder der ganzen Klasse erhalten die Lehrkräfte eine wichtige Informationsbasis für pädagogische Entscheidungen (Blumenthal, Kuhlmann & Hartke, 2014). Die Lernverlaufsdagnostik basiert auf dem Ansatz der formativen Evaluation. Durch die kontinuierliche Rückmeldung während des Lernprozesses kann die Lehrkraft die Fördermaßnahmen bei Bedarf ändern. Formative Bewertungssysteme dienen der Beantwortung folgender Fragen: Was hat die Schülerin oder der Schüler im gegenwärtigen Unterricht mit den aktuell praktizierten pädagogischen Maßnahmen bereits gelernt? In welchem Bereich bestehen noch Schwierigkeiten? Wie kann die Lehrkraft diese Schwierigkeiten durch intensivere oder andere Fördermaßnahmen aufheben (Diehl & Hartke, 2011)?

Entgegen einer summativen Evaluation bietet die formative Evaluation durch frühzeitiges Eingreifen die Möglichkeit, den Lernerfolg zu steigern. Durch die graphische Darstellung des Lernverlaufs beispielsweise können die Lehrkräfte positive/negative Entwicklung oder Stagnation des Lernfortschritts ablesen und so gezielte Übungen einsetzen. Insgesamt ist der Einsatz von Lernverlaufsdagnostik im Unterricht für alle Schülerinnen und Schüler hilfreich (Förster & Souvignier, 2014; Walter, 2013). Eine valide, kleinschrittige, am Curriculum orientierte Abbildung von Fähigkeiten bietet eine sehr gute Grundlage für die Passung von Unterrichtszielen und -inhalten in heterogenen Lerngruppen (Diehl & Hartke, 2012; Käter, Käter, Martenstein & Hillenbrand, 2016).

Unter dem Namen „Curriculum Based Measurements“ (CBM) entwickelten Stanley Deno und seine Mitarbeiter in den 70er Jahren die Methode, den Lernverlauf von leistungsschwächeren Schülerinnen und Schülern häufiger und dafür detaillierter zu messen (siehe Klauer, 2014). Mit Hilfe von curriculumbasierten Tests (Deno, 2003) können die Lernfortschritte der Schülerinnen und Schüler in Lernbereichen zeitnah beurteilt werden. Die Kinder lösen innerhalb einer zeitlichen Vorgabe von meist nur wenigen Minuten möglichst viele Aufgaben eines Tests zu einem Kompetenzbereich, wie z. B. Lesen oder Grundrechenoperationen. Die Messungen werden in regelmäßigen Abständen (monatlich bis wöchentlich) mittels paralleler Testversionen durchgeführt. Tests mit parallelen Versionen sind so konstruiert, dass sie entweder aus einem gemeinsamen Itempool (Aufgabenpool) gezogen werden, oder aus ähnlich schweren Items (Zum Beispiel Testversion A: $2 + 3 = _$ oder B: $3 + 2 = _$) aufgebaut sind.

Nach Fuchs (2004) kann man Lernverlaufsdagnostik nach den Ansätzen des Curriculum Sampling oder der robusten Indikatoren konstruieren. Beim Curriculum Sampling werden die am Schuljahresende geforderten Kompetenzen in Teilmengen unterteilt. Jeder Paralleltest enthält dann dieselbe Anzahl der vorab definierten Aufgabentypen (Voß & Hartke, 2014). Die Aufgaben werden zu allen Inhalten des Lehrplans gestellt und anhand der Rohwerte verrechnet (Fuchs, Fuchs, Hamlett & Stecker, 1990). Beim Ansatz der robusten Indikatoren werden Typen

Lernverlaufsdagnostik

Konstruktion von Lernverlaufsdagnostik

von Aufgaben gesucht, die die geforderte Kompetenz möglichst gut repräsentieren und hoch mit den relevanten Leistungen korrelieren. Für die Lesekompetenz hat sich beispielsweise das Messen der Leseflüssigkeit durch ein einminütiges lautes Lesen als ein robuster Indikator gezeigt (Reschly, Busch, Betts, Deno & Long, 2009; Wayman, Wallace, Wiley, Ticha & Espin, 2007).

Des Weiteren kann man den Test so konstruieren, dass er entweder eine gesamte Kompetenz bzw. Leistung (general outcome measure), oder nur eine Teilkompetenz bzw. einen engen Leistungsbereich (Skill based measure) misst. Der Vorteil eines engen Leistungsbereichs (beispielsweise Addieren im Hunderterraum ohne Zehnerübergang) ist, dass die Lehrkraft aus dem Testwert der Schülerinnen und Schüler direkt die Intervention ableiten kann (Hosp, Hosp & Howell, 2007). Der Nachteil des engen Leistungsbereichs ist, dass die Lernverlaufsmessung keine Langzeitmessung mehr erfassen kann (Fuchs, 2004) und der spezifische Testbereich durch die Lehrkraft ausgewählt werden muss. Daher schlägt Fuchs (2004) vor, dass Lernverlaufsmessungen multidimensional sein sollten. Dies kann jedoch dazu führen, dass die Lehrkräfte die Testwerte der Schülerinnen und Schüler aufgrund ihrer komplexen Zusammensetzung nicht mehr interpretieren können. Eine Ableitung einer Intervention ist zum Beispiel bei einer Lernverlaufsmessung, die die Leistung im Grundrechnen in der Grundschule misst, nicht mehr allein über den Gesamtestwert möglich, sondern erst bei Betrachtung weiterer qualitativer Auswertungen der nicht gelösten Aufgaben.

Souvignier, Förster und Salaschek (2014) haben ihre Lernverlaufsmessung in der Onlineumgebung quop (quop.de) curriculumsnah nach Klassenstufe und dem Ansatz der robusten Indikatoren konstruiert. Dies hat für Lehrkräfte der Grundschule den Vorteil, dass sie für den Bereich Lesen und Mathematik jeweils einen Test zur Verfügung haben, an dem sie für das ganze Schuljahr die Lernfortschritte ihrer Schülerinnen und Schüler ablesen können. Der Nachteil dieses auf Klassenstufen bezogenen Systems ist, dass alle Schülerinnen und Schüler mit demselben Test gemessen werden und nach dem regulären Lehrplan unterrichtet werden müssen. Schülerinnen und Schüler, die im schulischen Inhalt zurückhängen oder nach einem anderen Lehrplan unterrichtet werden, haben in einem solchen System nur die Möglichkeit, eine Lernverlaufsdagnostik der früheren Schulstufe zu benutzen.

Für schwache Lerner und Schülerinnen und Schüler mit einem sonderpädagogischen Förderbedarf lohnt sich daher eher ein System, bei dem die Lehrkraft zwischen verschiedenen Niveaustufen eines robusten Indikators wählen kann. Durch diese Flexibilität wird aber in Kauf genommen, dass eine klassenbezogene Auswertung mittels einer Jahrgangstufennorm nur eingeschränkt möglich ist. Dieser Ansatz findet sich beispielsweise in den CBM im Projekt Rügener Inklusionsmodell (RIM) (Voß, 2014). Hier gibt es Vorschläge, wann in welcher Jahrgangstufe welcher Test mit welchem Niveau durchgeführt werden kann, die Lehrkraft wählt jedoch selbst nach Bedarf aus. Für inklusive Klassen mit einer hohen Leistungsheterogenität ist dieser Vorgang vorzuziehen. Er setzt aber voraus, dass die Lehrkraft ein hohes Interesse für Diagnostik und Lernentwicklung hat, denn es besteht die Gefahr, dass die Lehrkraft ohne ausreichendes Wissen falsche Tests auswählt. Werden beispielsweise immer nur zu leichte Tests verwendet, kann das aktuelle Niveau nicht gemessen werden.

CBM werden überwiegend nach der klassischen Testtheorie (KTT) konstruiert (Deno, 2003). Dies ist aber aufgrund der Überprüfung von Dimensionalität, Messinvarianz und Änderungssensibilität nicht mehr der aktuelle Stand der Wissenschaft (Klauer, 2014, Wilbert, 2014; Wilbert, & Linnemann, 2011). Aus diesem Grund verwendet dieses Forschungsprojekt zur Skalierung der Tests die Konzepte der Item Response Theory (IRT) (Rost, 2004). Dies bedeutet, dass versucht wird, die Tests für Schülerinnen und Schüler mit sonderpädagogischen und ohne sonderpädagogischen Förderbedarf oder Migrationshintergrund fair zu konstruieren. Sollte in Analysen festgestellt werden, dass dies nicht der Fall ist, werden die Tests bzw. deren einzelne Items geändert.

Projekt: Lernverlaufsdagnostik mit der Onlineplattform LEVUMI

Ziel der Onlineplattform LEVUMI (Lern-Verlaufs-Monitoring) ist es einerseits, die Forschung zur Lernverlaufsdagnostik zu vertiefen und andererseits, der Schulpraxis ein praktikables Online-instrument zur Lernverlaufsmessung anzubieten. Die Plattform bietet aktuelle Lernverlaufstests, sowie deren Auswertung mit quantitativen und qualitativen Daten an. Die Benutzung ist kostenfrei und setzt nur einen aktuellen Browser (empfohlen wird Firefox) mit Internetzugang voraus. Die Installation eines Programms ist nicht nötig. Alle Daten werden auf Servern der TU München und der Universität Kiel gesichert.

Eine computergestützte Lernverlaufsdagnostik bietet gegenüber einer papierbasierten Version mehrere Vorteile: So werden der Lehrkraft organisatorische Aufgaben abgenommen und es bestehen verschiedene Möglichkeiten der Darstellung und Auswertung, die sofort angeboten werden. Die Umsetzung mittels einer Online Plattform bietet ebenso Vorteile gegenüber einer Version, die auf einem zu installierenden Programm besteht: Es gibt praktisch keine speziellen Voraussetzungen, die die Rechnerausstattung der Schule erfüllen muss. Die Daten sind für die Lehrkraft auf allen Rechnern (auch zu Hause oder mobil auf Smartphone oder Tablet) verfügbar. Die Daten sind sicherer abgelegt, als es bei einer lokalen Installation auf einem Rechner der Fall wäre. Die Administratoren der Internet-Seite können, falls nötig, unkompliziert eingreifen und bei Bedienungsfehlern oder -problemen unterstützen. Die Plattform kann unkompliziert um weitere Funktionen und Tests erweitert und aktualisiert werden. Neue Erkenntnisse aus Datenanalysen und Feedback der Nutzer kommen so allen Nutzern zu Gute. Für die Forschung hat die Onlineplattform den Vorteil, dass sie direkt auf die Daten zugreifen kann und die Tests sowie deren Auswertungsmöglichkeiten jederzeit verbessern kann.

Jeder Schüler erhält einen eigenen Test mit unterschiedlich angeordneten Items. Dadurch werden mehrere Paralleltests gebildet und es kann in der Forschungsuntersuchung festgestellt werden, ob die einzelnen Tests der Schüler pro Messzeitpunkt gleich schwer waren. Wichtig ist hierbei für die Testkonstruktion, dass alle Aufgaben eines Tests für den einzelnen Schüler und die Schülerin entweder gleich schwer sind, oder die Schwierigkeit der Aufgaben bei der Ziehung berücksichtigt wird. Ansonsten kann es passieren, dass ein Schüler oder eine Schülerin in einem Test weniger Aufgaben löst, als in einem vorherigen Test, nicht weil er oder sie weniger kann, sondern weil der Test einen höheren Anteil an schwierigeren Aufgaben hat. Dieses Problem kennen Lehrkräfte aus der Unterrichtspraxis beispielsweise bei selbstkonstruierten Diktaten. Nur wenn die Voraussetzung erfüllt ist, dass alle Paralleltests gleich schwierig sind, dürfen die Ergebnisse der einzelnen Messzeitpunkte in einem Graphen als Lernverlaufsmessung zusammengefasst und interpretiert werden (Klauer, 2014).

Die Onlineplattform LEVUMI ist so gestaltet, dass die Lehrkraft zuerst ihre Klasse mit allen Schülerinnen und Schülern anlegt und dann aus einer Auswahl Tests wählt. Für die Lernverlaufsmessung gibt es eine Klassenansicht und eine Schüleransicht mit je einem Graphen zur Anzahl und Quote der gelösten und nicht gelösten Aufgaben, sowie einer Auflistung der Aufgaben.

Um die Lesekompetenz zu erfassen, kann man einerseits die Geschwindigkeit des Lesens, die sogenannte Leseflüssigkeit, und andererseits das sinnentnehmende Lesen prüfen. Insbesondere in den ersten Schuljahren hat sich die Messung der Leseflüssigkeit im Rahmen der Eine-Minute-Lese-Aufgabe als ein wichtiger Prädiktor für die Lesefertigkeit gezeigt (Landerl, Wimmer & Moser, 1997; Fuchs, Fuchs, Hosp & Jenkins, 2001). Während die phonologischen Kodierungsprozesse insbesondere in frühen Phasen des Lesenlernens (Beginn in der Vorschulzeit) eine entscheidende Bedeutung haben, nimmt mit zunehmender Leseerfahrung die Bedeutung orthographischer Vergleichsprozesse zu (Frith, 1986). Für gute Leser der zweiten Klasse bietet sich daher der Vergleich von Wörtern und Pseudowörtern an. Des Weiteren gibt es eine Tradition in der Verwendung von Eine-Minute-Lese-Aufgaben als curriculumbasierte Tests (Stecker, Fuchs & Fuchs, 2005; Walter, 2009; 2013; Diehl & Hartke, 2012). Auch die Erfahrung und Ergebnisse des Rügener Inklusionsmodells (RIM) mit curriculumbasierten Messverfahren im Bereich Lesen in der Schuleingangsphase untermauern die Wahl von Einer-Minute-Lese-Aufgaben (Diehl, 2014;

Konstruktion der Tests

Leseflüssigkeitstests in LEVUMI

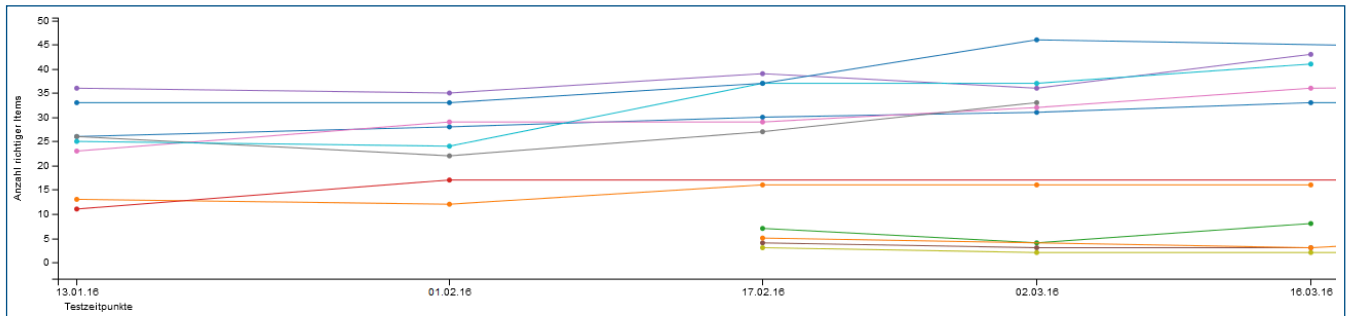
Diehl & Hartke, 2012; Diehl & Hartke, 2011). Daher haben wir uns entschieden, in LEVUMI zuerst lehrerzentrierte Einzeltests für die Messung der Leseflüssigkeit zu implementieren.

Die Tests zur Leseflüssigkeit orientieren sich am Lehrplan der Grundschule und dem Kieler Leseaufbau (KLA) (Dummer-Smoch & Hackethal, 2007). Der Kieler Leseaufbau basiert auf Erkenntnissen der Legasthenie (Lese-Rechtschreib-Störung)-Therapie und ist insbesondere bei Schülerinnen und Schülern mit Lernschwierigkeiten wirksam (Groth, Hasko, Bruder, Kunze & Schulte-Körne, 2013; Kuhl, Euker, & Ennemoser, 2015). Die Konzeption des Kieler Leseaufbaus ist kleinschrittig und klar strukturiert. Aktuell liegen auf der Lernplattform LEVUMI für jede Testreihe je vier Niveaustufen (N1 bis N4) vor. Für die Lernverlaufsdagnostik muss der Unterricht nicht auf den Kieler Leseaufbau umgestellt werden, wir empfehlen aber, dieses Konzept als Strukturierungshilfe einzusetzen. Insbesondere für Kinder mit Lernschwächen und in strukturierter Förderung bietet es sich an, nach den Schwierigkeitsstufen des Kieler Leseaufbaus zu arbeiten.

LEVUMI bietet zur Messung der Leseflüssigkeit Testreihen für Silben, Wörter und Pseudowörter. Die Silben- und Wörtertests wurden im Rahmen des Rügener Inklusionsmodells (RIM) in Verantwortung von Frau Diehl entwickelt und für die Onlinetestung adaptiert. Für diesen Bereich liegen in der Testentwicklung und Evaluation umfangreiche Vorarbeiten vor (Diehl, 2009; Diehl, Hartke & Knopp, 2009; Diehl, 2011; Diehl & Hartke, 2011; Diehl & Hartke, 2012). Dem RIM wird der US-amerikanische Response to Intervention-Ansatz zugrunde gelegt. Die Kernelemente im RTI-Ansatz sind die Mehrebenenprävention, die Evidenzbaisierung und das Monitoringsystem (National Center on Response to Intervention, 2010, S. 1–2, S. 6). Im Rahmen des Beobachtungssystems (Monitoring) von Lernprozessen für datenbasierte Förderentscheidungen orientieren sich die im RIM entwickelten Silben- und Wörterlesetests an den Schwierigkeitsstufen des Kieler Leseaufbaus (Dummer-Smoch, 2002) und werden als Paper Pencil Tests durchgeführt. Der Deutschunterricht im RIM passt sich in der Auswahl des Leselehrwerks sehr an die Kieler Konzeption an (Tolkmitt, 2005). Anfängliche Bedenken der Lehrkräfte, die CBM in ihren Unterricht zu integrieren, konnten schnell überwunden werden. Die Vorteile einer zeitnahen Rückmeldung zu Lernfortschritten der Schülerinnen und Schüler als Grundlage für eine adaptive Förderung und die Erfahrung, darüber hinaus auch Feedback zum eigenen pädagogischen Handeln zu bekommen, haben anfängliche Bedenken der Handhabung zurückgestellt (Voß & Hartke, 2014). Die Verfahren wurden sehr gut von den Lehrkräften im RIM angenommen. Die kleinschrittige systematische Aufbereitung der CBM in mehrere Niveaustufen in den Silben- und Wörtertests half den Lehrkräften, auf unterschiedliche Lernentwicklungen und damit verbundene Herausforderungen im inklusiven Unterricht und die individuellen Bedürfnisse der Schülerinnen und Schüler einzugehen. Beim Silbenlesen erproben sich Kinder in der Synthese zweier Laute. Sie lernen, das Prinzip der Lautschrift zu verstehen. Silbenlesen, eine Schwierigkeiten vermeidende Form des Lesens, bildet die Grundlage für das Wortlesen. Im LEVUMI ist sie damit die einfachste Lesestufe. Der Schüler bzw. die Schülerin übt und automatisiert zu Beginn des Leseerwerbs das lautierete Lesen. Die Testreihe Silbenlesen (Beispiel N2 hau, del, che) ist besonders für Anfänger im Leseerwerb in den ersten beiden Klassenstufen und für langsame Leser geeignet. Bei der Testreihe Wörterlesetest (Beispiel N2 Hafan, Maler, Rosen) wird neben dem synthetischen Lesen auch das analoge Ganzwortlesen (lexikalisches Lesen) getestet. Der Pseudoworttest im LEVUMI wurde neu aus den Silbentests gebildet. Pseudowörter sind sinnfreie Wörter und finden in der deutschen Sprache keine Verwendung. Über basale Leseflüssigkeit hinaus kann mit den Pseudowörtern das nichtlexikalische Lesen der Kinder (Dekodierfähigkeit) erfasst werden. Kinder, die in den Pseudowörterlesetests gute Ergebnisse erreichen, zeigen damit, dass sie orthographische Merkmale in den Fokus nehmen und diese für den Leseprozess erfolgreich nutzen können. Die schwierigsten Tests sind damit die Pseudowörterlesetests (Beispiel N2 teifen, mufer, tafei). Diese wurden aus den Silben der Silbenlesetests der jeweiligen Niveaustufe und nach der einfachsten Wortstruktur Konsonant-Vokal-Konsonant-Vokal-(Konsonant) (KV(KV(K))) gebildet. LEVUMI bietet zur Interpretation des Lernerfolgs eine Klassenübersicht (Abbildung 1) und eine Schülerübersicht für jeden Test an. Neben den Graphen bekommen die Lehrkräfte für jede Messung

qualitative und quantitative Ergebnisse. Pro Testzeitpunkt erhalten sie die richtig gelösten Items, die falsch gelösten Items, die Anzahl der richtig gelösten Items, die Anzahl der falsch gelösten Items und die Lösungswahrscheinlichkeit in Prozent. So kann man erkennen, bei welchen Buchstaben ein Schüler bzw. eine Schülerin Schwierigkeiten hat.

Abbildung 1:
Klassenansicht eines Tests



Die Pilotstudie wurde gemeinsam mit den Diagnose- und Förderklassen Arbeitskreis Schwaben (Bayern) im Schuljahr 2015/2016 durchgeführt. Im gemeinsamen Austausch wurde ein Lehrerhandbuch verfasst (Gebhardt, Diehl & Mühlhng, 2016), das frei im Internet zur Verfügung steht. Alle interessierten Lehrkräfte bekamen Zugang zum Onlinesystem. Anhand der Rückmeldungen der Lehrkräfte wurden Fehler gefunden und behoben, sowie eine leichtere Niveaustufe N0 eingeführt, die aus einer Auswahl von Buchstaben von N1 besteht. Ebenso ist eine Druckmöglichkeit der Schüler- und Klassenprofile ermöglicht worden. Insgesamt ist der Diagnose- und Förderklassen Arbeitskreis mit der Handhabung des Programms zufrieden, auch wenn noch nicht alle Wünsche realisiert werden konnten.

Ergebnisse der Pilotstudie

Die Auswertung der Pilotdaten erfolgt mithilfe von GNU-R. Wir stellen hier exemplarisch die Ergebnisse für eine Niveaustufe des Silbenlesetests vor. Mittels des R Pakets pairwise (Heine, 2015) wurden Raschmodelle zur Veränderungsmessung (Rost, 2004) für den Silbentest N3 berechnet. Aufgrund der geringen Stichproben von 90 Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf Lernen bei der ersten Testung (328 Tests gesamt) bei 116 Silben mit zufälliger Ziehung sind die Aussagen jedoch nur vorläufig. Insgesamt zeigen die Analysen für den Silbentest N3 reliable Ergebnisse des Pilotversuchs (WLE Reliabilität zwischen 0,82 und 0,89).

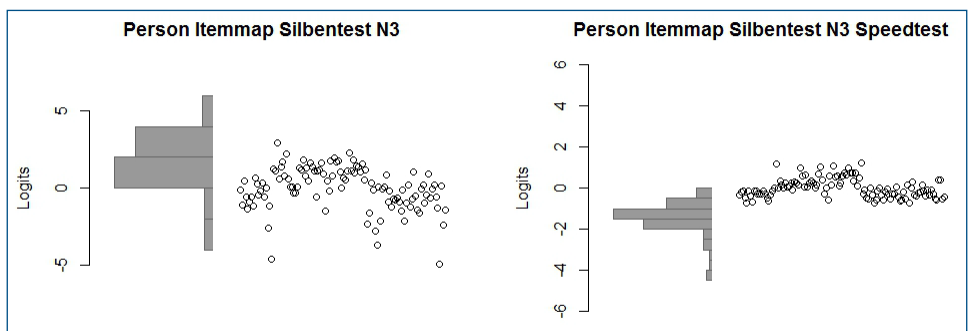


Abbildung 2:
Person-Item-Map des Raschmodells bei der ersten Testung als Power- und Speedtest

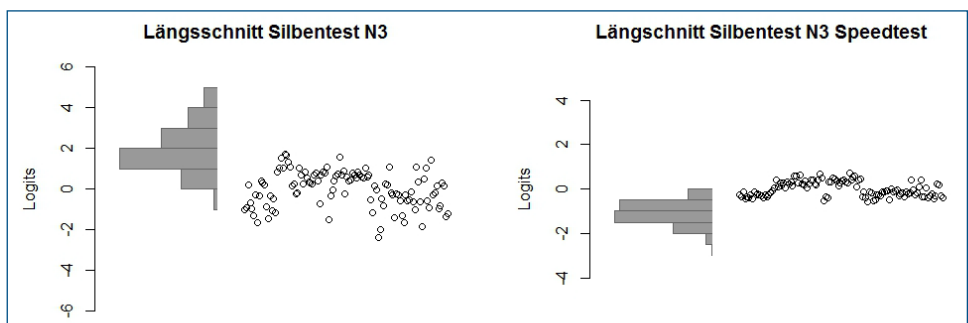


Abbildung 3:
Person-Item-Map des Raschmodells bei allen Messungen als Power- und Speedtest

Schlüsselwörter

Lernverlaufsmessung, Curriculum Based Measurement CBM, Lesen, Online Messung

Abstract

The development of inclusive schools is creating significant changes in current school systems. One key challenge for teachers in general schools is teaching combined classes that include students with and without special educational needs. How can teachers accurately judge students' achievements and learning development? How can gaps and stagnation in learning be detected early and suitable support given to foster learning progression? One possible solution to these problems is curriculum based measurement. In this article, we present the web platform LEVUMI. It offers short, easy to handle tests that offer teachers insights into the learning progressions of their students' reading abilities. The results of these tests can then be used as the basis for carefully coordinated support. Additionally, teachers receive feedback about their pedagogical actions. The reliability evaluations of LEVUMI are positive and support the future work and development of the platform.

Keywords

Curriculum Based Measurement CBM, Reading, Online Assessment

In den Abbildungen 1 und 2 sind in den Person-Item-Maps die ermittelten Personenparameter und Itemparameter anhand der Logitstransformierung verbunden. In Abbildung 2 wurde nur der erste Testzeitpunkt der Schüler skaliert, während in Abbildung 3 alle Messzeitpunkte berücksichtigt wurden. Des Weiteren wurden die Analysen einmal nur für die erhaltenen Items als Powertests und einmal für alle möglichen Items als Speedtest ausgewertet. Die Personenparameter sind hierbei links als Histogramm und die Items rechts als einzelne Punkte (Scatterplot) dargestellt. Die Auswertung zeigt, dass die Items relativ leicht zu lösen sind. Die Annahme, dass die Silben ähnlich schwer sind, gelingt besser für die Auswertung aller Messzeitpunkte als in der Auswertung für den ersten Messzeitpunkt. Dies kann aufgrund der geringen Stichprobe an Ziehungseffekten liegen. In weiteren Auswertungen müssen dieses beachtet und gegebenenfalls Items gelöscht oder die Ziehung in verschiedene Schwierigkeitsbereiche unterteilt werden. Nachdem der Test als Speedtest ausgegeben wird, muss man beachten, ob die Itemschwierigkeiten angemessen sind. Insgesamt erweisen sich die Items im Speedtest für die Personen als relativ schwierig. Bei der Auswertung aller Messzeitpunkte haben die Schülerinnen und Schüler höhere Personenkennwerte als bei ihrer ersten Testung und nähern sich den Itemschwierigkeiten durch einen messbaren Lernerfolg an. Somit kann in den ersten Analysen des Piloten sichergestellt werden, dass der Silbentest N3 ausreichend schwierig ist, um den Lernerfolg für Schülerinnen und Schüler mit einem sonderpädagogischen Förderbedarf im Bereich Lernen über die Zeit zu

messen. Anhand weiterer Auswertungen wird der Test der gemessenen Lernerfolge in LEVUMI in den nächsten Jahren verbessert.

Offenes Forschungsprojekt

Mit dem Projekt LEVUMI gehen wir den Weg eines offenen Forschungsprojekts. Wir entwickeln nicht zuerst die Tests, normieren diese und publizieren in einem Verlag, sondern wir erarbeiten im engen Austausch mit Lehrerinnen und Lehrern die Tests und entwickeln diese konstant weiter. Die Tests werden auf Anwendbarkeit, Interpretierbarkeit und psychometrische Qualität geprüft, bevor eine Normierung erfolgt. Ziel ist es zudem, auf der Grundlage der Ergebnisse, den Lehrkräften Maßnahmen zur Förderung an die Hand zu geben. Im Sinne eines „whole in one-Pakets“ (Hartke, Diehl & Vrban, 2008) soll es die Lehrkraft unterstützen, Kinder mit sehr unterschiedlichen Lernvoraussetzungen in ihrer Entwicklung angemessen zu fordern und zu fördern. Daher ist das Projekt längerfristig angelegt, um die Instrumente und daran anknüpfende Empfehlungen angemessen weiter zu entwickeln. Das Projekt steht am Anfang und viele Fragen bezüglich der Umsetzung und Anwendung der Lernverlaufsdagnostik sind noch offen. Die in LEVUMI verwendeten Tests sind aus reliablen Tests entwickelt. Da jedoch die Forschung zu Onlinetestungen und Lernverlaufsmessung noch in den Kinderschuhen steckt, wird ein endgültiger Nachweis der Reliabilität und Validität der Instrumente erst im Laufe der Zeit möglich werden. Wir wollen mit LEVUMI hierzu einen möglichen Baustein für die Erforschung der Lernverlaufsdagnostik beitragen. Für interessierte Lehrkräfte ist es möglich, einen Account in LEVUMI zu bekommen, die Instrumente und Informationsmaterialien frei zu nutzen und Verbesserungen anzuregen. Um einen Account zu erhalten, schreiben Sie eine E-Mail an markus.gebhardt@tum.de.

Literatur

- Becker, M., Lüdtke, O., Trautwein, U., & Baumert, J. (2006). Leistungszuwachs in Mathematik: Evidenz für einen Schereneffekt im mehrgliedrigen Schulsystem? *Zeitschrift für Pädagogische Psychologie*, 20, S. 233–242.
- Blackorby, J., Chorost, M., Garza, N. & Guzman, A. M. (2003). *The Academic Performance of Secondary School Students with Disabilities*. In U.S. Department of Education (Hrsg.),

The Achievement of Youth with Disabilities during Secondary School. A Report from the National Longitudinal Transition Study 2 (S. 61 – 75). Menlo Park, CA: SRI International. Verfügbar unter http://www.nlts2.org/reports/2003_11/nlts2_report_2003_11_complete.pdf [19.07.2016].

Blumenthal, Y., Kuhlmann, K. & Hartke, B. (2014). *Diagnostik und Prävention von Lernschwierigkeiten im Aptitude Treatment Interaction- (ATI) und Response to Intervention- (RTI)-Ansatz*. In M. Hasselhorn, W. Schneider & U. Trautwein (Hrsg.), *Tests & Trends*, NF Bd. 12. Formative Leistungsdiagnostik (S. 61–82). Göttingen: Hogrefe.

Deno, S. L. (2003). Developments in curriculum-based measurement. *The Journal of Special Education*, 37 (3), S. 184–192.

Diehl, K. (2009). *Schriftspracherwerb und Lernfortschrittsdokumentation*. Saarbrücken: svh.

Diehl, K. (2011). Innovative Lesediagnostik – Ein Schlüssel zur Prävention von Lese-Rechtschreibschwierigkeiten. *Zeitschrift für Heilpädagogik*, 62, S. 164–172.

Diehl, K. (2014). *Lesekompetenzen formativ evaluieren mit dem IEL-1 – Inventar zur Erfassung der Lesekompetenzen von Erstklässlern*. In M. Hasselhorn, W. Schneider & U. Trautwein (Hrsg.), *Jahrbuch der pädagogisch-psychologischen Diagnostik* (Bd. 12, S. 145–164). Göttingen: Hogrefe.

Diehl, K. & Hartke, B. (2011). Zur Reliabilität und Validität des formativen Bewertungssystems IEL-1: Inventar zur Erfassung der Lesekompetenz von Erstklässlern. *Empirische Sonderpädagogik*, 3, S. 121–146.

Diehl, K. & Hartke, B. (2012). *IEL-1 – Inventar zur Erfassung der Lesekompetenzen von Erstklässlern*. Göttingen: Hogrefe.

Diehl, K., Hartke, B. & Knopp, E. (2009). Curriculum-Based Measurement & Leeringonderwijsvolgysteem – Konzepte zur theoriegeleiteten Lernfortschrittsmessung im Anfangsunterricht Deutsch & Mathematik. *Zeitschrift für Heilpädagogik*, 60, S. 122–130.

Dummer-Smoch, L. & Hackethal, R. (2002). *Kieler Leseaufbau*. Handbuch (6. Aufl.). Kiel: Veris.

Dummer-Smoch, L. & Hackethal, R. (2007). *Kieler Leseaufbau*. Handbuch (7. Auflage). Kiel: Veris.

Elbaum, B. (2002). The self-concept of students with learning disabilities: A meta-analysis of comparisons across different placements. *Learning Disabilities Research & Practice*, 17 (4), S. 216–226.

Förster, N., & Souvignier, E. (2014). Learning progress assessment and goal setting: Effects on reading achievement, reading motivation and reading self-concept. *Learning and Instruction*, 32, S. 91–100.

Frith, U. (1986). A developmental framework for developmental dyslexia. *Annals of Dyslexia*, 36, S. 69–81.

Fuchs, L. S. (2004). The past, present, and future of curriculum-based measurement research. *School Psychology Review*, 33 (2), S. 188–192.

Fuchs, D., Fuchs, L.S. & Compton, D. L. (2004). Identifying Reading Disabilities by Responsiveness-to-Instruction. Specifying Measures and Criteria. *Learning Disability Quarterly*, 27, S. 216–227.

Fuchs, L. S., Fuchs, D., Hamlett, C. L. & Stecker, P. M. (1990). The role of skills analysis in curriculum-based measurement in math. *School Psychology Review*, 19, S. 617–641.

Fuchs, L. S., Fuchs, D., Hosp, M. K. & Jenkins, J. R. (2001). Oral Reading Fluency as an Indicator of Reading Competence: A Theoretical, Empirical, and Historical Analysis. *Scientific Studies of Reading*, 5 (3), S. 239–256.

Gebhardt, Diehl & Mühling (2016). *Lehrerhandbuch LEVUMI*. Verfügbar unter: http://ddi.in.tum.de/levumi/assets/LEVUMI_Lehrerhandbuch-e4be8726eacb0121247626010d57c010.pdf [19.07.2016]

Gebhardt, M., Schwab, S., Krammer, M., Gasteiger-Klicpera, B. & Sälzer, C. (2014). Erfassung von individualisiertem Unterricht in der Sekundarstufe I. Eine Quantitative Überprüfung der Skala „Individualisierter Unterricht“ in zwei Schuluntersuchungen in der Steiermark. *Zeitschrift für Bildungsforschung*, 4 (3), S. 303–316.

- Gebhardt, M., Zehner, F. & Hessels, M. G. P. (2014). Basic Arithmetical Skills of Students with Learning Disabilities in the Secondary Special Schools: An Exploratory Study covering Fifth to Ninth Grade. *Frontier Learning Research*, 2 (1), S. 50–63.
- Groth, K., Hasko, S., Bruder, J., Kunzel, S. & Schulte-Körne (2013). Interventionseffekte bei Lese-Rechtschreibstörung: Evaluation von zwei Förderkonzepten unter besonderer Betrachtung methodischer Aspekte. *Lernen und Lernstörungen*, 2, S. 161–175.
- Hartke, B., Diehl, K. & Vrban, R. (2008). *Planungshilfen zur schulischen Prävention – Früherkennung und Intervention bei Lern- und Verhaltensproblemen*. In J. Borchert, B. Hartke & P. Jogschies (Hrsg.), Frühe Förderung entwicklungsauffälliger Kinder und Jugendlicher (S. 218–234). Stuttgart: Kohlhammer.
- Heine, J.-H. (2015). *pairwise: Rasch Model Parameters by Pairwise Algorithm. R package version 0.3.1*.
- Hosp, M. K., Hosp, J. L. & Howell, K. W. (2007). *The ABCs of CBM. A practical guide to curriculum-based measurement*. New York: Guilford Press.
- Huber, C., & Grosche, M. (2012). Das response-to-intervention-Modell als Grundlage für einen inklusiven Paradigmenwechsel in der Sonderpädagogik. *Zeitschrift für Heilpädagogik*, 63 (8), S. 312–322.
- Käter, C., Käter, T., Martenstein, R. & Hillenbrand, C. (2016). Leitfadengestützte Konstruktion eines Instruments der Lernverlaufdiagnostik (CBM) im Bereich Lesen. *Zeitschrift für Heilpädagogik*, 66, S. 168–179.
- Klauer, K. J. (2014). *Formative Leistungsdiagnostik: Historischer Hintergrund und Weiterentwicklung zur Lernverlaufdiagnostik*. In M. Hasselhorn, U. Trautwein, & W. Schneider (Hrsg.), Lernverlaufdiagnostik (Vol. N. F. Band 12) (S. 1–18). Göttingen: Hogrefe.
- Kocaj, A., Kuhl, P., Kroth, A. J., Pant, H. A., & Stanat, P. (2014). Wo lernen Kinder mit sonderpädagogischem Förderbedarf besser? Ein Vergleich schulischer Kompetenzen zwischen Regel- und Förderschulen in der Primarstufe. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 66, S. 165–191.
- Kuhl, J., Euker, N. & Ennemoser, M. (2015). Förderung des lautorientierten Lesens bei Schülerinnen und Schülern mit intellektueller Beeinträchtigung. *Empirische Sonderpädagogik*, 7 (1), S. 41–55.
- Landerl, K., Wimmer, H. & Moser, E. (1997). Salzburger Lese- und Rechtschreibtest. Bern: Huber
- Mahlau, K., Diehl, K., Voss, S. & Hartke, B. (2011). Das Rügener Inklusions-Modell -Konzeption einer inklusiven Grundschule. *Zeitschrift für Heilpädagogik*, 11, S. 464–472.
- National Center on Response to Intervention (2010). *Essential Components of RTI – A Closer Look at Response to Intervention*. Washington, DC: National Center on Response to Intervention.
- Reschly, A. L., Busch, T. W., Betts, J., Deno, S. L., & Long, J. D. (2009). Curriculum-based measurement oral reading as an indicator of reading achievement: A meta-analysis of the correlational evidence. *Journal of School Psychology*, 47, S. 427–469.
- Rost, J. (2004). *Lehrbuch Testtheorie – Testkonstruktion*. Bern: Huber.
- Souvignier, E., Förster, N., & Salaschek, M. (2014). *quop: ein Ansatz internet-basierter Lernverlaufdiagnostik und Testkonzepte für Mathematik und Lesen*. In M. Hasselhorn, W. Schneider, & U. Trautwein (Hrsg.), Lernverlaufdiagnostik (Test und Trends N.F. Band 12) (S. 239–256). Göttingen: Hogrefe.
- Stecker, P. M., Fuchs, L. S. & Fuchs, D. (2005). Using Curriculum-based Measurement to Improve Student Achievement. Review of Research. *Psychology in the School*, 42 (8), S. 795–819.
- Stern, T. (2004). Neue Wege zu einer förderlichen Leistungsbewertung. Eine neue Lernkultur braucht eine neue Prüfungskultur. *Lernende Schule*, 7 (28), S. 16–21.
- Tolkmitt, P. (2005). *Lulu lernt lesen*. Heinsberg: Dieck.
- Walter, J. (2009). Theorie und Praxis Curriculumbasierten Messens in Unterricht und Förderung. *Zeitschrift für Heilpädagogik*, 60 (5), S. 162–170.
- Walter, J. (2013). *Verlaufdiagnostik sinnerfassenden Lesens (VSL)*. Göttingen: Hogrefe Schultests.

- Wayman, M., Wallace, T., Wiley, H.I., Tichá, R., & Espin, C.A. (2007). Literature synthesis on curriculum-based measurement in reading. *Journal of Special Education, 41*, S. 85–120.
- Wilbert, J. (2014). *Instrumente zur Lernverlaufsmessung Gütekriterien und Auswertungsherausforderungen*. In M. Hasselhorn, W. Schneider & U. Trautwein (Hrsg.), *Lernverlaufsdagnostik* (S. 281–308). Weinheim: Hogrefe.
- Wilbert, J., & Linnemann, M. (2011). Kriterien zur Analyse eines Tests zur Lernverlaufsdagnostik. *Empirische Sonderpädagogik, 3* (3), S. 225–242.
- Wocken, H., & Gröhlich, C. (2007). *Kompetenzen von Schülerinnen und Schülern an Hamburger Förderschulen*. In W. Bos, & M. Bonsen (Eds.), *KESS 7: Kompetenzen und Einstellungen von Schülerinnen und Schülern der Jahrgangsstufe 7* (S. 133–142). Münster: Waxmann.
- Voß, S. (2014). *Curriculumbasierte Messverfahren im mathematischen Erstunterricht – Zur Güte und Anwendbarkeit einer Adaption US-amerikanischer Verfahren im deutschen Schulsystem*. Saarbrücken: SVH.
- Voß, S. & Hartke, B. (2014). *Curriculumbasierte Messverfahren (CBM) als Methode der formativen Leistungsdiagnostik im RTI-Ansatz*. In M. Hasselhorn, W. Schneider & U. Trautwein, U. (Hrsg.). *Formative Leistungsdiagnostik* (S. 83–100). Göttingen: Hogrefe.

Prof. Dr. Markus Gebhardt
 TU Dortmund
 Fakultät für Rehwissenschaften
 Emil-Figge-Straße 50 - 44227 Dortmund
 markus.gebhardt@tum.de

Prof. Dr. Kirsten Diehl
 Europa-Universität Flensburg
 Institut für Sonderpädagogik
 Auf dem Campus 1 - 24943 Flensburg
 kirsten.diehl@uni-flensburg.de

Prof. Dr. Andreas Mühling (Juniorprofessor)
 Arbeitsgruppe Didaktik der Informatik
 Christian-Albrechts-Universität zu Kiel
 Christian-Albrechts-Platz 4 - 24118 Kiel
 andreas.muehling@informatik.uni-kiel.de