

Stock, C. (2022). Informelle Lernverlaufsdiagnostik im Rahmen eines Lesetrainings bei Kindern einer Sprachheilschule. In M. Gebhardt, D. Scheer & M. Schurig (Hrsg.), *Handbuch der sonderpädagogischen Diagnostik. Grundlagen und Konzepte der Statusdiagnostik, Prozessdiagnostik und Förderplanung* (S. 805-828). Regensburg: Universitätsbibliothek. <https://doi.org/10.5283/epub.53149>

# Informelle Lernverlaufsdiagnostik im Rahmen eines Lesetrainings bei Kindern einer Sprachheilschule

*Claudia Stock*

## 1 Einführung in die Thematik

Um den individuellen Lernweg und -erfolg von Schulkindern im Laufe eines Schuljahres festzustellen, individuelle Fördermöglichkeiten zu erarbeiten und damit die Entwicklungspotentiale der Kinder optimal zu nutzen, wird in den letzten Jahrzehnten vermehrt die Curriculumbasierte Messung (CBM) bzw. Lernverlaufsdiagnostik (LVD) diskutiert, die eine Form der formativen Lernfortschrittsmessung darstellt, also begleitend zur Leistungsvermittlung verläuft (z. B. Klauer, 2006, Voß & Hartke, 2017, Walter, 2008). Darüber hinaus ermöglicht sie den Lehrenden die Evaluation ihrer eigenen Lehrmethoden (vgl. Klauer, 2006, S. 17). Möglichkeiten der Anwendung der CBM auf dem Gebiet des Lesens im sprachheilpädagogischen Kontext in Kombination mit einem Lesetraining werden in diesem Artikel vorgestellt.

## 2 CBM als Methode der Leistungserfassung

Im Rahmen der in der schulischen Praxis immer mehr geforderten Integration und Inklusion von Kindern mit Behinderungen ist es erforderlich, diese Kinder speziell auf ihrem Lernweg zu begleiten und ihnen spezielle Förderpläne anzubieten. Wie Voß und Hartke (2014, S. 83f.) sowie Walter (2008, S. 63) in ihren Überblicksartikeln berichten, erkannten dies bereits vor ca. 30 bis 40 Jahren Deno und Mirkin (1977) sowie Deno (1985), die unter diesen Gesichtspunkten ein Verfahren zur Erfassung von Schülerleistungen im sonderpädagogischen Kontext (speziell Lernbehinderung und Erziehungsschwierigkeiten) entwickelten. Anhand regelmäßiger Leistungserhebungen sowie deren statistischer und graphischer Aufarbeitung wurde es möglich, den Sonderpädagog\*innen die Leistungsentwicklung der Schüler\*innen visuell schneller zu verdeutlichen, sodass diese ihre Fördermaßnahmen flexibler anpassen und ihren Unterricht optimieren konnten. In späteren Studien konnte gezeigt werden, dass das Vorgehen der CBM sich auch auf den Regelschulkontext erweitern ließ (vgl. Voß & Hartke, 2014, S. 84). So wurde die CBM als formative Lernfortschrittsmessung in den USA nach der Einführung des 2001 verabschiedeten No-Child-Left-Behind-Gesetzes aufgegriffen, um sicherzustellen, dass möglichst

alle Schüler – auch die Leistungsschwachen – die vorgegebenen Jahrgangsnormen in der Mathematik und in der Leseleistung erreichten (vgl. Klauer, 2006, S. 16), und stellt im Rahmen des ebenfalls in den USA verbreiteten Response-to-Intervention-Ansatz (RTI) ein Vorgehen dar, das »...Maßnahmen zur Systematischen Evaluation von Schülerleistungen und deren Entwicklung...« mit der Förderung dieser Schüler\*innen verbindet (Voß & Hartke, 2014, S. 83).

Parallel dazu entwickelte sich die Debatte um das sog. »mastery learning« nach Bloom (1974, zitiert nach Maier, 2010), wonach der Lernende »...durch häufige Leistungsüberprüfungen und schülerspezifisches Feedback ... bis zu einem gewissen Grad jedes Lernziel erreichen kann...« (Maier, 2010, S. 299), wenn das weitere Vorgehen des Lehrenden an den aktuellen Wissens- und Leistungsstand des Lernenden angepasst wird. So wird auch geschlussfolgert, dass ein Grund für die Wirksamkeit der CBM in dezidierten und regelmäßigen Leistungsrückmeldungen an die Schüler\*innen liegt, bei denen die Aufmerksamkeit der Kinder auf den eigenen Lernfortschritt gelenkt wird und bewusst nicht die Benotung der Leistung im Vordergrund steht. Von dieser Vorgehensweise sollen gemäß verschiedener empirischer Studien v.a. leistungsschwächere Kinder profitieren (vgl. Crooks, 1988, Black & Wiliam, 1998, beide zitiert nach Maier, 2010, S. 300). Ein weiterer Grund ist die Ausrichtung des Unterrichts am Leistungsstand der Schüler\*innen.

Das Kernstück der CBM machen die regelmäßigen, begleitenden, immer gleich aufgestellten und am Lehrplan orientierten, möglichst standardisierten Messungen aus. Dadurch lassen sich Lernfortschritte, aber auch -stagnationen feststellen, die umso schneller von der Lehrkraft aufgegriffen werden können, je engmaschiger die Testungen durchgeführt werden. Üblicherweise erfolgen die Messungen einmal wöchentlich bis einmal monatlich (Voß und Hartke, 2014, S. 87; Fuchs & Fuchs, 1993, zit. nach Klauer, 2006, S. 18). Dabei ist zu beachten, dass zu jedem Messzeitpunkt die gleiche Leistung mit der gleichen Schwierigkeit erhoben wird. Dies lässt sich in Anlehnung an Fuchs (2004) durch zwei Vorgehensweisen sicherstellen:

Die erste wird als »general outcome measures« (GOM) bezeichnet. Hierbei werden Aufgaben verwendet, die das »Zusammenspiel ausgeprägter Teilkompetenzen« (Voß & Hartke, 2014, S. 85; Deno et al. 1982, zit. nach Klauer 2006, S. 17) erfassen und so den Leistungsstand des Kindes erheben. Als Beispiel wird in der Literatur die Leseflüssigkeit genannt, die u.a. die Buchstaben-, Silben- und Wortkenntnis erfasst (ebd.), indem man die Kinder über den gesamten Messzeitraum aus ein und demselben Text (z. B. ein Kinderbuch) laut vorlesen lässt. Eine weitere Möglichkeit besteht im deutschen Sprachraum darin, die Kinder ein Blatt mit Silben vorlesen zu lassen (vgl. Jungjohann et al., 2017).

Die zweite Vorgehensweise – die »mastery measures« (MM) – zeichnen sich dadurch aus, dass einzelne Teilkompetenzen eines Fachgebietes erhoben werden (vgl. Voß & Hartke, 2014, S. 86). Bei Erstellung der Aufgaben orientiert man sich am Lehrplan der Kinder, d.h. bspw. an den Zielen, die am Ende des Schuljahres erreicht werden sollten. Zur Umsetzung dieser Vorgehensweise für Kinder mit sprachheilpädagogischem Hintergrund wurde für diese Studie kurze Texte mit Multiple-Choice-Aufgaben verwendet. Anhand der Anzahl der richtig beantworteten Fragen lässt sich der Lesefortschritt feststellen.

### 3 Leseerwerb

Bereits seit über einhundert Jahren befasst sich die deutsch- und englischsprachige Forschung mit dem Leseprozess und dabei auftretenden Problemen. Meringer und Mayer veröffentlich-

ten z. B. bereits 1895 ein umfassendes Werk über Sprech- und Lesefehler und erste Überlegungen zu deren Entstehungsursachen, die Wissenschaftler wie Ranschburg (1916) dazu anregte, diese Fehler experimentell im Zusammenhang mit der Legasthenie zu betrachten (vgl. Stock, 2005, S. 3,7, 17ff.). Eine Betrachtung der Problematik im Rahmen der in dieser Zeit aufkommenden Sprachheilschulen und -klassen erfolgte aber nicht, da sich diese vor allem mit der gesprochenen Sprache bei Kindern mit Agrammatismus, bei Stammlern (Lispler, Näslern), bei Stimmkranken und Stotterern befassten. Ziel war es, dass diese Kinder »...die Sprache ... ästhetisch einwandfrei und phonetisch richtig...« verwenden und damit »... sozial brauchbar, also grammatisch richtig, mühelos und eindeutig zu verstehen, in nichts auffällig...« erscheinen konnten (Hansen, 1929, S. 76-77). Siebzig Jahre Später liegt der Schwerpunkt in der Legasthenieforschung auch auf der Schriftsprache. So formulieren Klicpera und Gasteiger-Klicpera (1998, S. 5): »Von Anfang an hatte ... das Erlernen des Lesens und Schreibens etwas mit der rationalen Lebensführung, mit dem beruflichen Handeln und dem Fortschritt der Güterproduktion, aber auch mit der Gestaltung und der Teilhabe am gesellschaftlichen Leben zu tun.« In dieser Tradition versteht sich dieser Artikel, der den auch für Kinder mit sprachheilpädagogischem Hintergrund bedeutsamen, aber erschwerten Leseerwerb durch die Mittel der CBM unterstützen will.

Welche Erkenntnisse zum Leseerwerb liegen mittlerweile vor? In seinem Situationsmodell des Lesens unterscheidet Lenhard (2013, S. 15, Abb. 1) zwischen allgemeinen kognitiven Fähigkeiten wie z. B. das Sprachverständnis, die Arbeitsgedächtniskapazität (vgl. Baddeley & Hitch, 1974, Abb. 3), die Geschwindigkeit des Zugriffs auf das semantische Lexikon im Langzeitgedächtnis (vgl. Warnke, 1992) oder die phonologische Bewusstheit (vgl. Mannhaupt & Jansen, 1989) als Grundlage für den Leseerwerb, hierarchieniederen Prozessen, zu denen das Rekodieren und Dekodieren im Sinne des indirekten Leseweges nach Coltheart (1978, S. 199, 211f., Abb. 2) sowie die Propositionsbildung zwischen den Wörtern auf Satzebene und das Erkennen von Kohäsionsmitteln zwischen Sätzen, sowie letztendlich hierarchiehohe Prozesse gehören, unter die man das Textverständnis inklusive Inferenzbildung unter Einbeziehung des Vorwissens zählen kann.

Dabei wurde einerseits berücksichtigt, dass entsprechend der Ergebnisse der LOGIK-Studie (Schneider et al., 1989) Kinder mit einer geringen phonologischen Bewusstheit im weiteren Sinne (pB i.w.S.), welche sich nach Marx (2007, S. 44) v.a. auf das Segmentieren von Silben und das Erkennen von Reimen und Wörtern bezieht, zu Beginn der Schulzeit im zweiten Schuljahr ein geringes Leseverständnis aufweisen. Andererseits wurden die Erkenntnisse der Doppel-Defizit-Hypothese nach Wolf und Bowers (1999) einbezogen, wonach v.a. Kinder mit Problemen im phonologischen Arbeitsgedächtnis im späteren Verlauf ihrer Schulzeit eine grundlegende Leseschwäche zeigen. Mit Bezug zum Modell von Baddeley & Hitch (1974, Abb. 3) ist hier vor allem die phonologische Schleife betroffen, die weniger Einheiten aufnehmen kann und zudem eine geringere Wiederholungsgeschwindigkeit aufweist.

Zudem ist nach Marx (2007, S. 39) anzunehmen, dass die grammatischen Fähigkeiten und das Hörverstehen grundlegende Fähigkeiten für den Schriftspracherwerb darstellen. Auch nach Rost und Hartmann (1992) zeigt das Hörverstehen eine nicht zu vernachlässigende Verknüpfung von 26,6% mit dem Leseverstehen. In diese Richtung weisen auch das bereits beschriebene Situationsmodell des Lesens nach Lenhard (2019, S. 15) und das Modell des Textverstehens nach Kintsch (1977, S. 380). In diesem Zusammenhang stehen zudem verschiedene Stufenmodelle des Schriftspracherwerbs aus den 1980er Jahren, wie bspw. das in Abbildung 4 vereinfacht wiedergegebenen Stufenmodell nach Frith (1985) oder jenes nach Günther (1986), wonach sich Leseanfänger überwiegend auf dem Weg vom Übergang aus der phonetischen Stufe

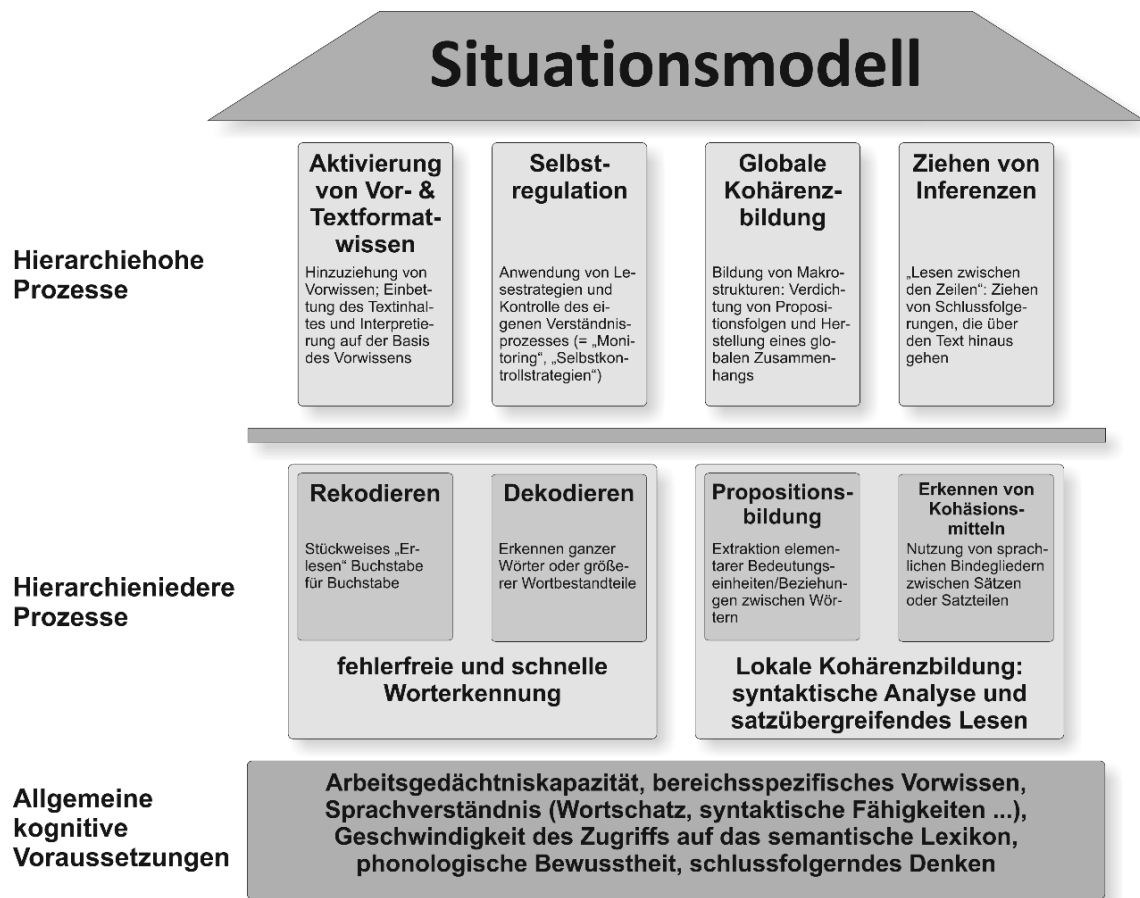


Abbildung 1: Situationsmodell des Lesens nach Lehnard (2019, S. 15). Abdruck mit Genehmigung des Verlags W. Kohlhammer

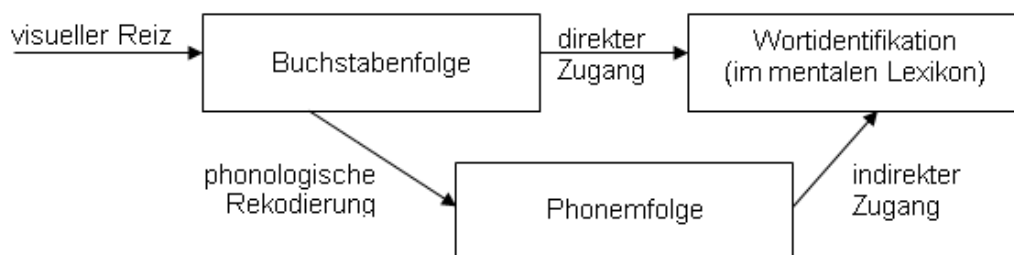


Abbildung 2: Deutsche Fassung des Modells der zweifachen Zugangswege beim Worterkennen von Coltheart (1978, S. 199, 211f.) in Anlehnung an Klicpera und Gasteiger-Klicpera (1998, S. 18)

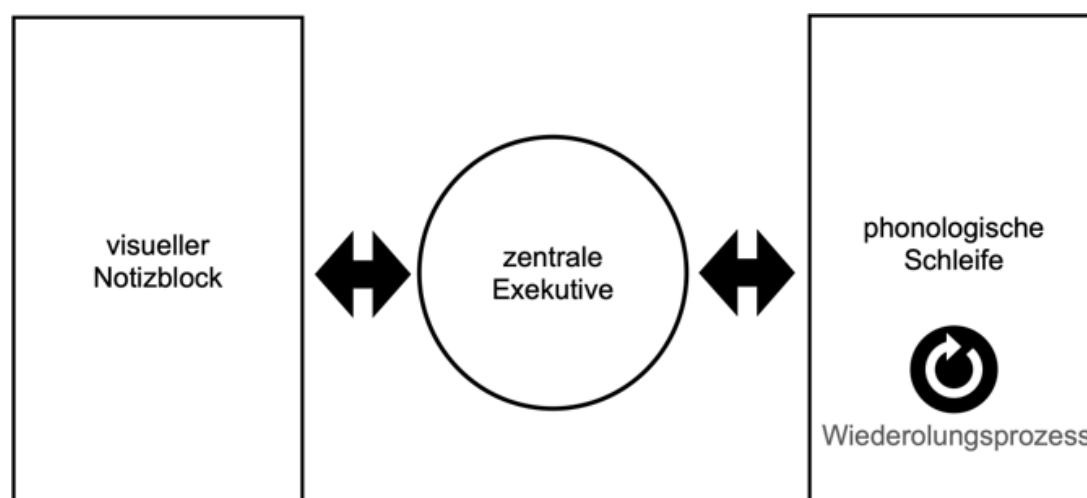


Abbildung 3: Übersetztes und leicht modifiziertes Arbeitsgedächtnis-Modell nach Baddeley und Hitch (1974) in Anlehnung an Baddeley (2014, S. 46)

Tabelle 1: Stichprobe der CBM-Testung

Test-Stichprobe	Jungen	Mädchen	gesamt
1-A	7	6	13
1-B	9	4	13
gesamt	16 (61,5%)	10 (38,5%)	26

in die alphabetische befinden, in der sie dann die Verbindung zwischen Buchstaben und Lauten erwerben, um lautgetreu lesen zu können. Im weiteren Grundschulalter gehen die Kinder dann langsam zur orthographischen Stufe über, in der die Erkennung ganzer Silben oder kurzer Wörter möglich wird und langsam auch verstehendes Lesen (vgl. Frith, S. 308f., 311; Günther, 1986, S. 33, 41f.)

## 4 Methodisches Vorgehen und Prätestung

Die Studie steht in der Tradition ähnlicher Studien mit Grundschulkindern (vgl. bspw. Anderson et al., 2020) und fand ab dem Schuljahr 2018/19 für zwei aufeinanderfolgende Schuljahre an einer Sprachheilschule statt. Es nahmen insgesamt zwei erste, später zweite Klassen der Schule teil, was einer Stichprobengröße von insg. 26 Kindern entsprach. Aus Tabelle 1 sind die Verteilungen pro Klasse zu entnehmen, wobei ersichtlich wird, dass die Jungen mit 61,5% deutlich häufiger vertreten sind als die Mädchen.

Aufgrund der Erkenntnisse zum Leseerwerb wurde zu Beginn der Studie im November 2018 als standardisierte Verfahren der TEPHOBE (Mayer, 2016) zur Erfassung der pB i.w.S. und der Benennungsgeschwindigkeit (RAN) als Maß für die Verarbeitungsgeschwindigkeit der phonologischen Schleife nach Baddeley (2014, S. 46ff) sowie der MuSePro (Berg, 2015) für die Überprüfung der grammatischen Fähigkeiten der Kinder durch Studentinnen der Sprachheilpädagogik der Universität Würzburg durchgeführt. Der TEPHOBE beinhaltet zur Erfassung der pB i.w.S. für die betrachtete Altersgruppe die vier Subtests Onse-Rime-Synthese, Phonemsynthese, Reimen und Anlautkategorisierung, deren richtige Lösungen als Summenwert zur Bestimmung der pB heran-

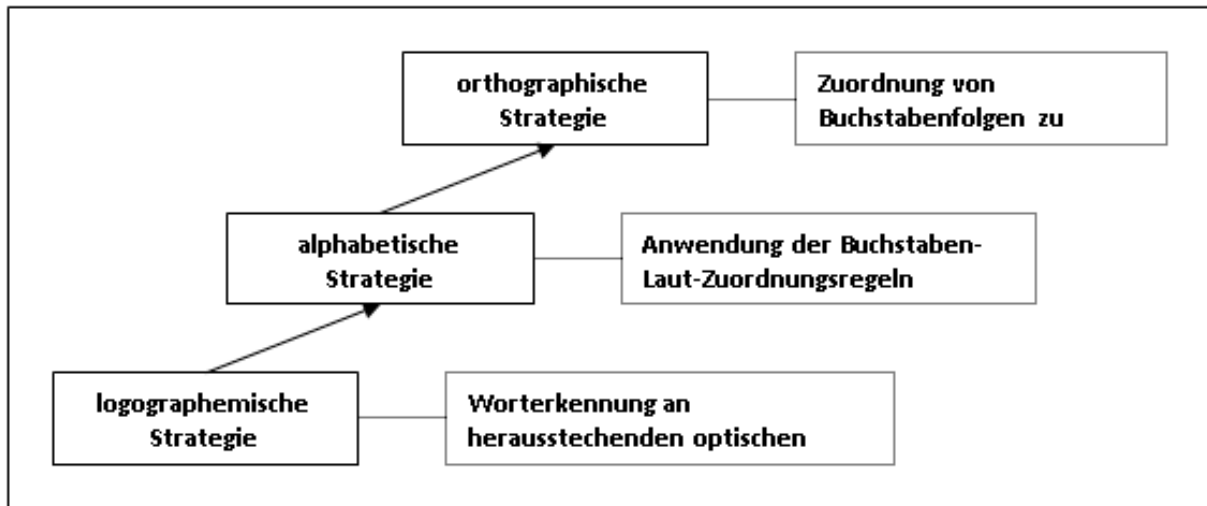


Abbildung 4: Vereinfachtes Entwicklungsmodell des Schriftspracherwerbs von Frith (1985, S. 311) und Günther (1986, S.33) nach Stock et al. (2003, S. 10) in Anlehnung an Schulte-Körne (2001, S. 12)

gezogen werden (Mayer, 2016, S. 6). Zur RAN-Erfassung stehen für Vorschulkinder die Subtests Objekte und Farben und für Erstklässler die Subtests Buchstaben, Zahlen und Farben zur Verfügung (ebd.), die jeweils einzeln betrachtet werden müssen, um die Benennungsgeschwindigkeit der Kinder einzustufen. Aufgrund der zu erwartenden Wortschatzprobleme bei Kindern mit sprachheilpädagogischem Hintergrund wurden alle vier Subtests durchgeführt. Wie erwartet zeigte sich, dass nur zwölf Kinder über Buchstabenkenntnisse verfügten, die eine Durchführung des Subtests ermöglichten, sowie jeweils zwei Kinder die Farb- bzw. die Objektbenennung nicht beherrschten. Dies erklärt die unterschiedlichen Stichprobengrößen in Tabelle 2. Der MuSePro überprüft die grammatischen Fähigkeiten der Kinder anhand der fünf Subtests Verbzweitstellungsregel im Hauptsatz, Subjekt-Verb-Kongruenz, Akkusativmarkierung, Dativmarkierung und Verbendstellungsregel im Nebensatz (vgl. Berg, 2015, S. 11).

Anhand dieser beiden Testverfahren wurde in beiden Klassen die Trainingsgruppe (TG) ausgewählt. Dabei wurde vor allem auf unterdurchschnittliche Leistungen ( $T < 30$ ) in der p.B.i.w.S. und in den Ran-Subtests fokussiert und unterdurchschnittliche Werte im MuSePro als Entscheidungshilfe bei uneindeutigen TEPHOBE-Ergebnissen herangezogen. Es ergaben sich die aus der Abbildung 5 zu entnehmenden Anfangsergebnisse. Die pB i.w.S. im TEPHOBE fiel nur bei neun Kindern unterdurchschnittlich bis schwach aus, in der Benennungsgeschwindigkeit zeigten über alle vier Subtests hinweg insgesamt elf Kinder auffällige Werte. Deutlich wird aber auch die vergleichsweise große Anzahl an Kindern, die beim Benennen von Farben und Zahlen, aber auch in der pB überdurchschnittliche Leistungen zeigten ( $T > 60$ ). Für den MuSePro werden die pro Subtest keine T-Werte bestimmt, sondern anhand von im Handbuch festgelegten Punktzahlen festgestellt, ob Förderbedarf besteht. Es zeigen sich Häufungen des Förderbedarfs in alle Subtests (rote Balken), aber vor allem im Dativ- und im Nebensatzerwerb, was für die Schüler\*innen einer Sprachheilschule ein eher zu erwartendes Ergebnis darstellt. Aus der Klasse A wurden aufgrund dieser Daten fünf Kinder (Nr. 1, 2, 6, 10 und 12) sowie aus der Klasse B vier Kinder (Nr. 3, 10, 12 und 13) für die das Lesetraining und die darauffolgende und monatlich stattfindenden CBM-Testungen ausgewählt (später als Trainingsgruppe, TG bezeichnet). Die Kontrollgruppe (KG) umfasste die restlichen 17 Kinder.

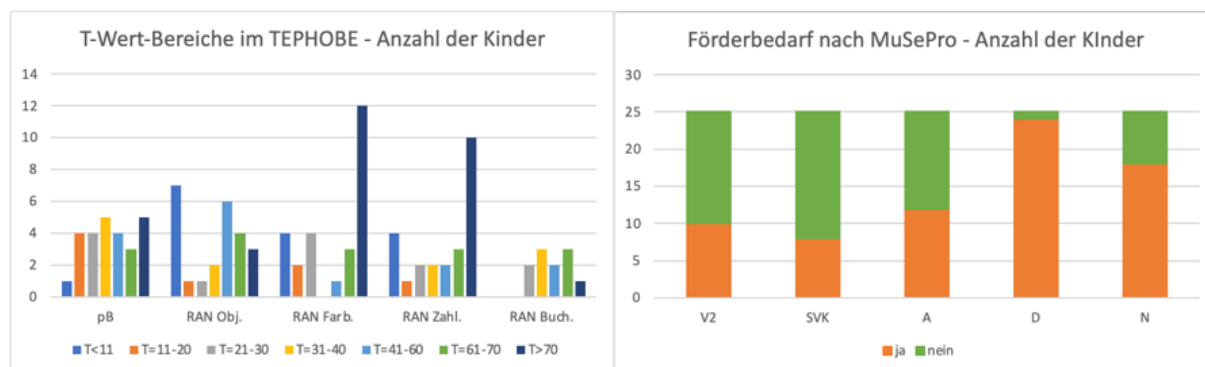


Abbildung 5: Ergebnisse der Prätestung mit dem TEPHOBE (Mayer, 2016) und dem MuSePro (Berg, 2015), (angegeben ist jeweils die Anzahl der Kinder). Anm: Subtests TEPHOBE: pB = phonologische Bewusstheit, Obj. = RAN-Objekte, Farb. = RAN-Farben, Zahl. = RAN-Zahlen, Buch. = RAN-Buchstaben, Subtests MuSePro: V-2 = Verbzweitstellung, SVK = Subjekt-Verb-Kongruenz, A = Akkusativerwerb, D = Dativerwerb, N = Nebensatzerwerb

Zur Untermauerung der deskriptiven Analyse wurden die Daten der Vortestung t-Test auf Gruppenunterschiede unterzogen, die in Tabelle 2 zu sehen sind. Dabei wird ersichtlich, dass die Kinder der TG sowohl in der pB als auch in allen Subtests der Benennungsgeschwindigkeit signifikant schwächer abschnitten als die Kinder der KG, während sich die Kinder der beiden Gruppen in den grammatischen Fähigkeiten überwiegend nicht signifikant unterschieden. Hier zeigt sich nur für den Subtest Dativerwerb ein signifikanter Unterschied bei durchschnittlich besserer Leistung der TG und für den Subtest Akkusativerwerb eine Tendenz zu Signifikanz besserer Leistung der KG.

## 5 Trainingsergebnisse

### 5.1 Leseleistung im Verlauf der ersten Klassenstufe

Die CBM-Testungen wurden von Anfang Dezember 2018 bis Anfang Juli 2019 einmal im Monat von den Studentinnen unter Anwendung des GOM-Verfahrens mit von den Lehrkräften selbst erstellten Leseteppichen (vgl. Abb. 6) durchgeführt. Die Leseteppiche entsprachen dem Voranschreiten der Lehrkräfte im Anfangsunterricht und bildeten Graphem-Phonem-Korrespondenzen und Silben ab. In den jeweils dazwischenliegenden Wochen (mit Ausnahme der Schulferien) trainierten sie mit den Kindern das Lesen in Kleingruppen mit diesen Leseteppichen.

Mit dieser Methode sollte im Laufe eines Schuljahres die Graphem-Phonem-Korrespondenz sowie das Erkennen von Silben über eine vertiefte Einspeicherung im Langzeitgedächtnis gestärkt werden. Diese Leseleistung ist dem Rekodieren innerhalb der hierarchieniedereren Prozesse nach Lenhard (2019, S. 15f.) und der indirekten Leseroute nach Coltheart (1978, S. 211ff.) zuzuordnen, da sich die Leseanfänger erst am Beginn der alphabetischen Stufe nach Frith (1985, S. 308) oder Günther (1986, S. 33) befinden. Es sollte also der Grundstein für später darauf aufbauende hierarchiehöhere Leseprozesse gelegt werden.

Tabelle 2: t-Test-Kennwerte der Prätestung mit dem TEPHOBE (Mayer, 2016) und dem MuSePro (Berg, 2015) gerechnet über die T-Werte *Anm.:* Die Berechnung erfolgte für den TEPHOBE anhand der T-Werte und für den MuSePro anhand der Rohwerte. Subtests MuSePro: V-2 = Verbzweitstellung, SVK = Subjekt-Verb-Kongruenz, A = Akkusativerwerb, D = Dativerwerb, N = Nebensatzerwerb  
n.s. = nicht signifikant

Subtest	Gruppen	N	m	s	t	df	p
TEPHOBE pB	TG	9	43,50	2,88	-3,231	21,819	<0.01
	KG	17	50,72	8,32			
TEPHOBE RAN-Objekte	TG	8	37,95	12,64	-2,771	22	<0.05
	KG	16	50,30	8,99			
TEPHOBE RAN-Farben	TG	9	41,51	14,62	-3,504	24	<0.01
	KG	17	58,34	9,85			
TEPHOBE RAN-Zahlen	TG	7	40,69	13,39	-2,906	7,781	<0.05
	KG	17	56,38	7,76			
TEPHOBE RAN-Buchstaben	TG	3	43,63	1,38	-4,724	9,882	<0.01
	KG	9	51,36	4,21			
MuSePro V-2	TG	9	8,67	2,06	-0,771	23	n.s.
	KG	16	9,19	1,33			
MuSePro SVK	TG	9	15,56	20,52	0,939	8,034	n.s.
	KG	16	9,13	1,26			
MuSePro A	TG	9	8,67	1,94	2,035	23	<0.10
	KG	16	6,44	3,54			
MuSePro D	TG	9	4,67	1,73	2,954	23	<0.01
	KG	16	2,38	1,93			
MuSePro N	TG	9	6,44	3,94	-0,396	23	n.s.
	KG	16	7,00	3,01			

S	l	m	A	t
l	T	s	o	a
lo	ma	si	ta	li
sa	mi	so	la	mo

Abbildung 6: Beispielhafter Ausschnitt aus einem der verwendeten Leseteppiche



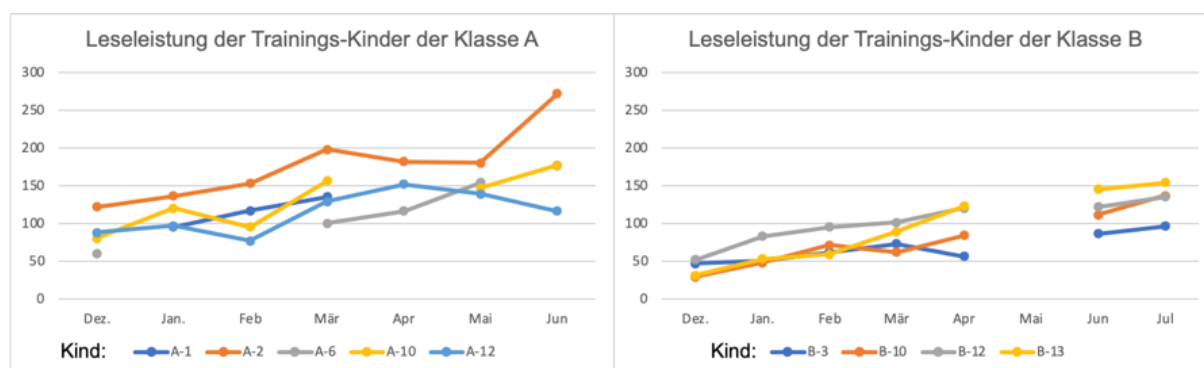


Abbildung 7: CBM-Lesekurven der richtig gelesenen Buchstaben und Silben pro zwei Minuten in den Klassen A und B im ersten Schuljahr (Messzeitraum Dezember 2018 bis Juni/Julii 2019)

## 5.2 CBM-Lese-flüssigkeitstestung in der ersten Klassenstufe

Die Kinder hatten konkret die Aufgabe, die abgebildeten Buchstaben bzw. Silben so schnell und fehlerfrei wie möglich innerhalb von zwei Minuten vorzulesen. Dabei wurde erfasst, wie viele Buchstaben und Silben insgesamt gelesen wurden und wie viele davon falsch waren. Aus der Differenz ergab sich dann die Anzahl der richtig gelesenen Buchstaben und Silben. Das Ziel war zu erkennen, inwieweit die Lesekurve im Laufe des Schuljahres ansteigen würde, da zu erwarten war, dass mit vertiefter Einspeicherung der Buchstaben und Silben auch die Lesegeschwindigkeit steigen würde, weil die Kinder den nach Coltheart (1978) direkten und schnelleren Leseweg wählen konnten und die Leseinheiten nicht mehr mühsam rekodieren mussten. Die Darstellung der individuellen Lesekurven pro Kind ergibt sich aus Abbildung 7. Dabei ist zu beachten, dass in der Klasse A im Juli 2019 und in der Klasse B im Mai 2019 eine CBM-Testung nicht stattfinden konnte sowie dass einige Kinder Fehltag aufgrund von Krankheiten o.ä. aufweisen. Die Klassen werden getrennt voneinander dargestellt, weil in jeder Klasse ein anderer Leseteppich verwendet wurde und damit die Klassen untereinander nicht vergleichbar sind.

In beiden Klassen zeigten sich der erwartete Anstieg in der Leseflüssigkeit, wobei aber auch deutliche Niveau- und Verlaufsunterschiede zwischen den Kindern pro Klasse sichtbar werden. So zeigte bspw. das Kind A-2 einen starken Anstieg in der Anzahl der richtig gelesenen Wörter pro zwei Minuten, hier hat sich die Anzahl der gelesenen Wörter mehr als verdoppelt, wobei das Kind schon zu Beginn der Testung eine vergleichsweise hohe Leseflüssigkeit aufwies. Dagegen stagniert das Kind A-12 auf niedrigem Niveau und zeigt nur eine Verbesserung um 28%. In einem solchen Fall ist es ratsam, entweder das didaktische Vorgehen in der Lesevermittlung zu überdenken und anzulassen, und/oder die Anzahl der Trainingseinheiten zu erhöhen.

Aufgrund der verschiedenen Fehlzeiten der Kinder war es nicht sinnvoll, zur Feststellung der Wirksamkeit des Vorgehens eine Messwiederholungsanalyse über alle Testzeitpunkte zu rechnen, da sich so die ohnehin schon kleine Gruppe an Testkindern noch mehr reduzieren würde. Ein Messwiederholungsanalysen als Vergleich des ersten und letzten Messzeitpunktes pro Klasse zeigen aber, dass in beiden Klassen die Leseflüssigkeit der Trainingskinder innerhalb des ersten Schuljahres signifikant angestiegen ist (vgl. Tab. 3).

Tabelle 3: Messwiederholungsanalyse der CBM-Lese-flüssigkeit für beide Klassen im ersten Schuljahr

Klasse	Testzeitpunkt	N	m	s	F	df	p
A	Dez.'19	5	89,80	22,88	20,783	1	<0.01
	Juni '20	5	179,20	57,56			
B	Dez '19	4	41,75	12,09	29,329	1	<0.01
	Juli '20	4	130,50	24,53			

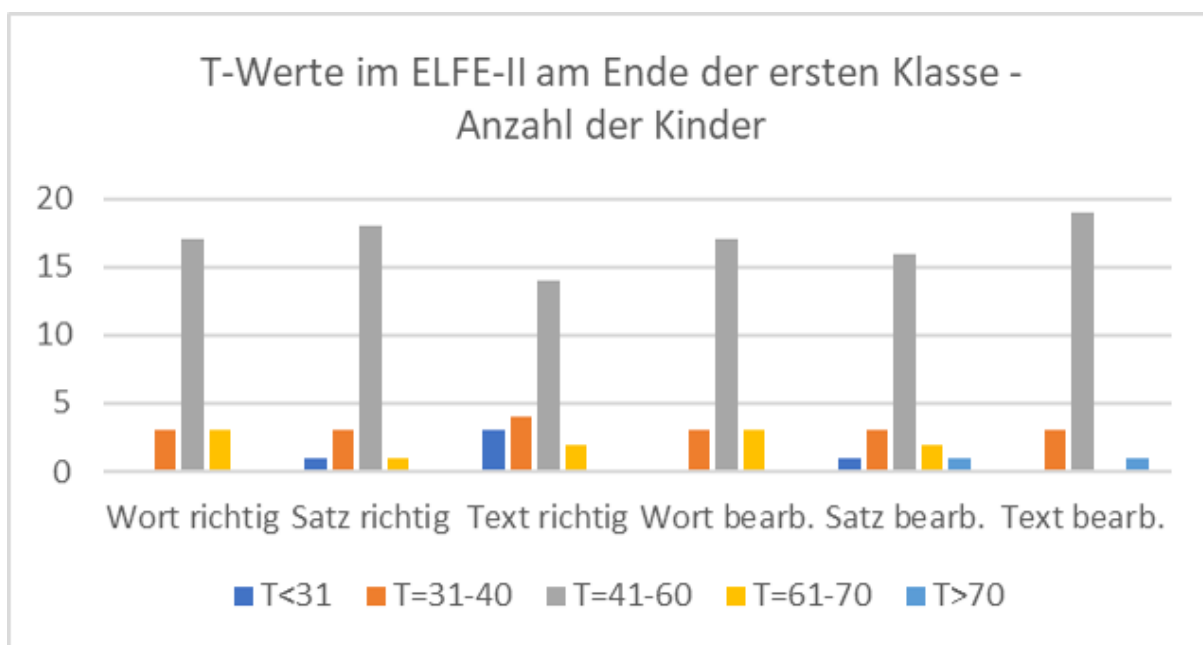


Abbildung 8: Ergebnisse des ELFE II (Lenhard et al. 2017)

### 5.3 Standardisierte Statustestung am Ende der ersten Klassenstufe

Um die Leseleistung der Kinder richtig einzustufen und nicht dem als Bezugsgruppen- oder Fischteich-Effekt bekannten Effekt (vgl. Marsh, 2005, S. 120) zu erliegen, wurde am Ende des ersten Schuljahres der standardisierte Lesetest ELFE II von Lenhard et al. (2017) durchgeführt, mit dem das Wort- Satz- und Textverständnis der Kinder mit Bezug zu den richtig gelesenen und zu den insgesamt bearbeitete Aufgaben gemessen werden kann, was der reinen Lesegeschwindigkeit ohne Betrachtung der falsch vs. richtig gelesenen Wörter entspricht. Die Ergebnisse sind Abbildung 8 zu entnehmen.

Es zeigt sich, dass am Ende des ersten Schuljahres nur wenige Kinder unterdurchschnittliche Leistungen im Wort-, Satz- und Textverständnis zeigen, das betrifft für die richtig gelesene Aufgaben je nach Subtest drei bis sieben Kinder und für die bearbeiteten Aufgaben je Subtest drei bis vier Kinder. Unter diesen Kindern befanden sich auch solche aus der KG, die Kinder aus der TG überwogen jedoch. Die überwiegende Anzahl der Kinder (ca. 61% bis 78% – je nach Subtest) zeigten durchschnittliche Leistungen und jeweils ein bis drei Kinder aus der KG lagen im überdurchschnittlichen Bereich. Wenn man beachtet, dass der ELFE II weit überwiegend an Kindern aus Regelschulen normiert wurde, hier aber Ergebnisse von Kindern einer Sprachheilschule vorgestellt werden, sind das gute Ergebnisse, die sich auch in der Anzahl der bear-

Tabelle 4: t-Test-Kennwerte der ELFE-II-Subtests (Lenhard et al., 2017) am Ende des ersten Schuljahres berechnet über die T-Werte der TG und der KG

Subtest	Gruppen	N	m	s	t	df	p
Wortverständnis richtig	TG	9	45,89	5,28	-2,178	21	<0.05
	KG	14	53,36	9,32			
Satzverständnis richtig	TG	9	42,56	7,42	-2,868	21	<0.01
	KG	14	51,86	7,69			
Textverständnis richtig	TG	9	42,78	6,92	-1,436	21	n.s.
	KG	14	48,93	11,53			
Wortverständnis bearbeitet	TG	9	45,22	5,59	-2,681	21	<0.05
	KG	14	53,86	8,52			
Satzverständnis bearbeitet	TG	9	40,89	8,05	-2,789	21	<0.05
	KG	14	52,07	11,12			
Textverständnis bearbeitet	TG	9	42,33	3,64	-2,046	21	<0.10
	KG	14	49,00	9,26			

bearbeiteten Aufgaben, also in der Lesegeschwindigkeit widerspiegeln. Hier zeigen je nach Subtest nur 3 bis 4 Kinder (meist aus der TG) ein zu langsames Lesetempo, ein bis drei Kinder aus der KG waren überdurchschnittlich und ca. 70% bis knapp 83% lagen im durchschnittlichen Leistungsbereich. Dabei werden die Gruppenunterschiede bis auf die Skalen im Textverständnis bei überwiegend hohen Effektstärken (hoch)signifikant mit jeweils besseren Leistungen in der KG (s. Tab. 4).

Die Korrelationen zwischen den bearbeiteten Aufgaben und den richtigen Lösungen in den Subtests sind Tabelle 5 zu entnehmen. Anhand der hohen und (hoch)signifikanten Korrelationen mit überwiegend mittleren und hohen Effekten lässt sich ableiten, dass Kinder im Wort- und Satzverständnistest gleichermaßen gut abschneiden. Zudem gibt es einen engen Zusammenhang zwischen Lesegeschwindigkeit und der Anzahl richtiger Lösungen, woraus sich auf eine hohe Lesegenauigkeit schließen lässt. D.h. die meisten Kinder begehen wenige Fehler. So entspricht auf dem Gebiet der bearbeiteten Textaufgaben der T-Wert von ca. 42 (in der TG, s. Tab. 4) einer absoluten Anzahl von drei bis vier bearbeiteten Aufgaben, der entsprechende T-Wert von ca. 43 auf dem Gebiet der richtig gelösten Textaufgaben einer absoluten Anzahl von zwei bis drei richtig gelöste Aufgaben. Die entsprechenden absoluten Werte in der KG bewegen sich in der Größenordnung von sieben bearbeiteten und ca. vier richtig gelösten Aufgaben. Vor allem bei den Trainingskindern wird hier die Strategie sichtbar, dass sie vor allem auf die Lesegenauigkeit fokussiert haben.

Die nicht-signifikanten Korrelationen bei überwiegend mittleren Effekten zu den insgesamt bearbeiteten Aufgaben im Textverständnistest sind zum einen mit der geringen Stichprobengröße, zum anderen aber auch mit dem jungen Lesealter der Testkinder erklärbar, da sich die Kinder am Ende des ersten Schuljahres nach dem Modell von Lenhard (2019, s. Abb. 1) noch überwie-

Tabelle 5: Korrelationen der ELFE-II-Subtests (Lenhard et al., 2017) am Ende des ersten Schuljahres

Anm.: – Signifikanzniveaus der Korrelationskoeffizienten: \*  $p < 0.05$ ; \*\*  $p < 0.01$  – Effektstärke  $r \geq 0.1$  – klein,  $r \geq 0.3$  – mittel,  $r \geq 0.5$  – groß

ELFE II	(2)	(3)	(4)	(5)	(6)
(1) Wortverständnis richtig	.788**	.538**	.964**	.506*	.343
(2) Satzverständnis richtig		.768**	.837**	.883**	-.134
(3) Textverständnis richtig			.583**	.794**	-.697
(4) Wortverständnis bearbeitet				.644**	.234
(5) Satzverständnis bearbeitet					-.407
(6) Textverständnis bearbeitet					

Tabelle 6: Vorhersage der Leseleistung am Ende der 1. Klasse durch die pB und RAN (N=23)

ELFE II	TEPHOBE	R2	korr. R2	F	p	$\beta$	T	p
Wortverst.	pB	0.467	0.414	8,764	.002	.437	2,566	.018
	RAN					0.416	2,442	.024
Satzverst.	RAN	0.486	0.434	9,448	.001	0.577	3,449	.003
	pB					0.26	1,556	.135
Textverst.	pB	0.17	0.087	2,051	.155	0.236	1,109	.281
	RAN					0.278	1,310	.205

gend auf der Stufe der hierarchieniederen Prozesse der Worterkennung und lokalen Kohärenzbildung auf Satzebene befinden und sich die hierarchiehoheren Prozesse, die das Textverständnis ermöglichen, erst noch erschließen und festigen müssen. Zu diesen Erkenntnissen passen die Ergebnisse dreier Regressionsanalysen, die Tabelle 6 zu entnehmen sind. Die mit dem ELFE (Lenhard et al., 2017) erfassten richtigen Lösungen im Wort- und Satzverständnis am Ende der ersten Klasse ließ sich signifikant durch die Subtests des TEPHOBE (Mayer, 2016) vorhersagen. Dabei wurden für das Wortverständnis sowohl die pB als auch die Benennungsgeschwindigkeit als Prädiktoren signifikant, für das Satzverständnis nur die Benennungsgeschwindigkeit. Die Textverständnisleistung konnte nicht signifikant vorhergesagt werden.

## 5.4 Leseleistung in der zweiten Klassenstufe

Mit den Testkindern wurde auch im darauffolgenden Schuljahr 2019/20 weitergearbeitet. Geplant waren die üblichen monatlichen CBM-Testungen ab November 2019 bis Jun/Juli 2020 sowie die wöchentlichen Trainingseinheiten. Ein ähnliches Vorgehen zeigen Jungjohann et al. (2021). Aufgrund der Corona-Pandemie und der damit verbundenen Schulschließungen musste das CBM-Training sowie die -Testungen im Februar bis März 2020 beendet werden. Die Lehrkräfte führten am Ende des Schuljahres den ELFE II (Lenhard et al, 2017) selbst durch, sodass die Studentinnen das Schulgebäude nicht betreten mussten.

Im Rahmen der CBM-Testungen wurde zwischen der Erfassung der Leseflüssigkeit (das bereits bekannte GOM-Vorgehen) und des Leseverständnisses (MM-Vorgehen) unterschieden. Die Leseflüssigkeit wurde durch zweiminütiges lautes Vorlesen aus einem der beiden Bücher von Nöstlinger (2013 und 2015 – Geschichten vom Franz) erfasst, wobei die studentischen Testlei-

terinnen Beginn und Ende der gelesenen Textpassagen sowie die Anzahl der falsch gelesenen Wörter notiert wurde. So konnte wiederum die Anzahl der richtig gelesenen Wörter ermittelt und zudem sichergestellt werden, dass die Kinder in der darauffolgenden Testung oder beim Training dort weiterlesen konnten, wo sie aufgehört hatten. Das Training erfolgte mit denselben Büchern von Nöstlinger (2013, 2015) in Kleingruppen.

Die Messung des Leseverständnisses wurde in der B-Klasse durch das stille Lesen von kurzen Geschichten, die von einer Lehrkraft der Schule bereitgestellt und durch single-choice-Fragen ergänzt. Zu jeder Geschichte gab es je sechs Fragen mit vier Antwortalternativen, die bei richtiger Beantwortung mit je einem Punkt bewertet wurden. Dieses Vorgehen entspricht dem oben beschriebenen MM-Verfahren.

## **5.5 Standardisierte Testung am Beginn der zweiten Klassenstufe**

Am Anfang der zweiten Klassenstufe wurden zudem zwei standardisierte Testverfahren – die WLLP-R (Schneider et al., 2011) und der Leseflüssigkeitstest aus dem SLRT II (Moll & Landerl, 2014) – an insgesamt 20 Kindern durchgeführt, um festzustellen, ob sich Lernrückschritte durch die Sommerferien eingestellt hatten. Diese Vorhersage gelang gut (s. Tab. 7), wobei vor allem durch den Wortverständnistest die Leistungen in der WLLP-R und im Pseudowortlesen des SLRT-II vorhergesagt werden konnten und aus dem Satzverständnistest das Wortlesen im SLRT II. Das Textverständnis lieferte wiederum keinen signifikanten Beitrag. Die Ergebnisse lassen sich so erklären, dass sowohl für das Pseudowortlesen als auch für das Wortlesen im ELFE II wie im WLLP-R auch die indirekte Leseroute nach Coltheart (1978, s. Abb. 2) ausreichend ist, während das Satzverständnis umso besser ausfällt, je schneller Wörter als Ganzes erkannt werden (direkte Route nach Coltheart, 1978, s. Abb. 2).

Die deskriptiven Ergebnisse sind in Abbildung 9 als T-Werte dargestellt, die nach Liener & Raatz (1998, S. 410) aus den Prozentrangwerten ermittelt wurden. Es wird ersichtlich, dass sowohl in der Leseflüssigkeit als auch in der mit der WLLP-R gemessenen Lesegeschwindigkeit die meisten Kinder eine durchschnittliche Leistung zeigen. Nur wenige Kinder müssen als unterdurchschnittlich (PR <25) bis schwach (PR <11) eingestuft werden. Insgesamt handelt es sich jeweils um maximal fünf Kinder: zwei in der WLLP-R, fünf im Wortlesen und vier im Pseudowortlesen des SLRT-II, wobei diese Kinder sowohl der TG als auch der KG angehörten. Dagegen schneiden bis zu neun Kinder mit einem PR >75 überdurchschnittlich ab. Da diese beiden Testverfahren nicht Förderschulen normiert wurden und somit ein Vergleich mit Regelschulkindern erfolgen musste, sind auch diese Ergebnisse als überwiegend positiv zu bewerten.

Ebenso ließen sich die Ergebnisse im Lesen am Anfang der zweiten Klasse durch die Benennungsgeschwindigkeit am Beginn der ersten Klasse vorhersagen. Die  $p_B$  (s. jeweils TEPHOBE, Mayer, 2016) spielte hier keine signifikante Rolle. Dies entspricht den Ergebnissen der Doppel-Defizit-Hypothese nach Wolf und Bowers (1999).

## **5.6 CBM-Leseflüssigkeitstestung in der zweiten Klassenstufe**

Aufgrund der gewonnenen Erfahrungen aus der ersten Klassenstufe und auf Bitten der Lehrkräfte wurden nicht nur alle CBM-Trainingskinder, sondern auch die Kinder der KG in die Testungen mit einbezogen und regelmäßig in der Leseflüssigkeit getestet. Die Ergebnisse finden sich in Abbildung 10, wiederum getrennt nach Klasse.

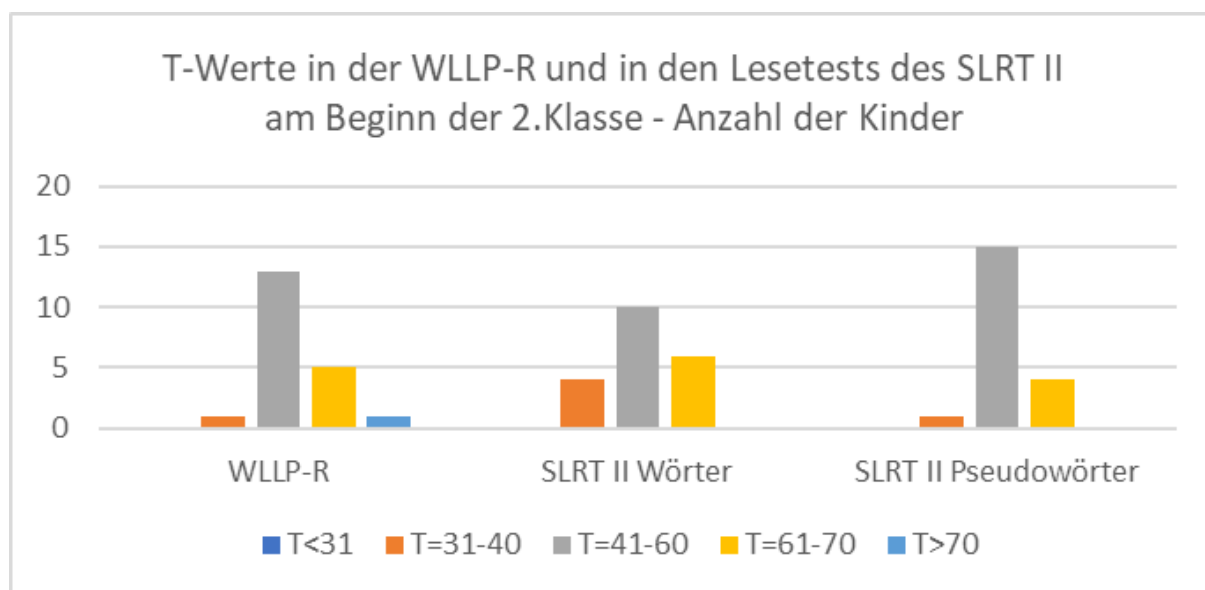


Abbildung 9: T-Werte in der WLLP-R (Schneider et al., 2011) und in den Lesetests des SLRT II (Moll & Landerl, 2014)

Tabelle 7: Vorhersage der Leseleistung am Beginn der 2. Klasse durch pB und RAN (N=20) sowie aus den ELFE-Subtests (N=19) der ersten Klassenstufe

Anm.: Die Unterschiede in den Stichprobenzahlen ergeben sich daraus, dass nicht immer alle Kinder an allen Testzeitpunkten anwesend waren.

Lese-Test	TEPHOBE	R2	korr. R2	F	p	$\beta$	T	p
WLLP-R	RAN	-0.369	0.334	10,547	.004	.608	3,248	.004
	pB					0.333	1,845	.083
SLRT II – Wort	RAN	0.303	0.265	7,843	.012	0.551	2,800	.012
	pB					0.222	1,104	.285
SLRT II – Pseudow.	RAN	0.328	0.29	8,773	.008	0.572	2,962	.008
	pb					0.292	1,524	.146
Lese-Test	ELFE II	R2	korr. R2	F	p	$\beta$	T	p
WLLP-R	Wort	0.758	0.744	53,340	<.001	0.871	7,303	<.001
	Satz					.109	0,334	.742
	Text					0.321	2,038	.058
SLRT II – Wort	Satz	0.753	0.738	51,822	<.001	0.868	7,199	<.001
	Wort					-0.153	-0,467	.646
	Text					-0.044	-0,222	.827
SLRT II – Pseudow	Wort	0.601	0.578	25,648	<.001	0.775	5,064	<.001
	Satz					0.038	0,091	.928
	Text					-0.032	-0,142	.889

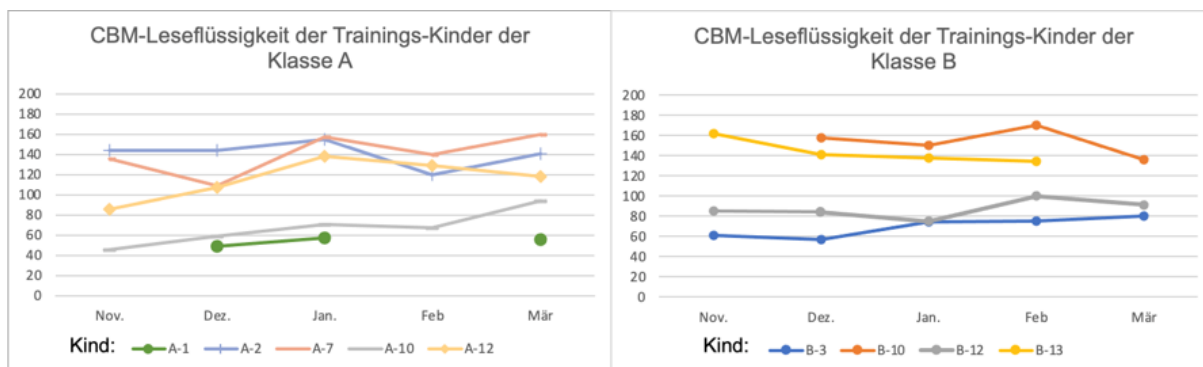


Abbildung 10: CBM-Lesekurven der richtig gelesenen Wörter pro zwei Minuten (Leseflüssigkeit) in den Trainingsgruppen der Klassen A und B im zweiten Schuljahr (Messzeitraum November 2019 bis März 2020)

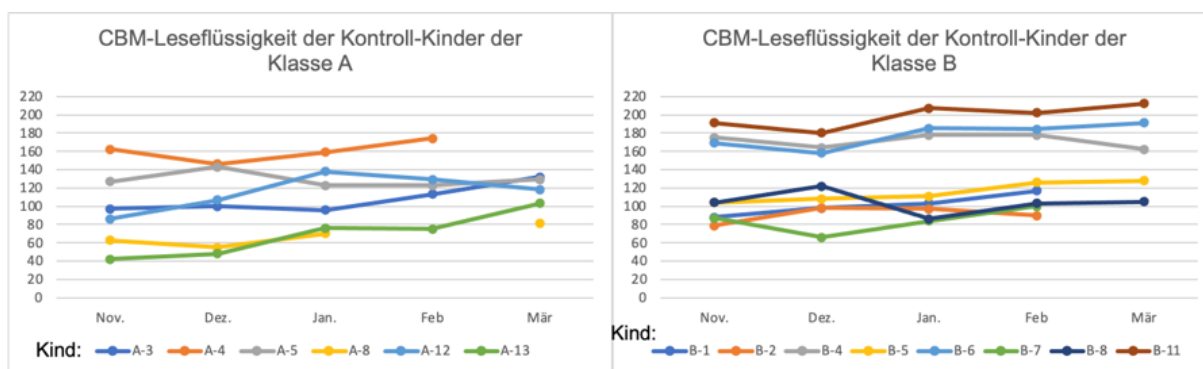


Abbildung 11: CBM-Lesekurven der richtig gelesenen Wörter pro zwei Minuten (Leseflüssigkeit) in den Kontrollgruppen der Klassen A und B im zweiten Schuljahr (Messzeitraum November 2019 bis März 2020)

Sie zeigen, dass vier Kinder einen Zuwachs in der Leseflüssigkeit um ca. 20 bis 50 Wörter in zwei Minuten verzeichnen konnten (A-7, A-10, A-12, B3). Zwei weitere Kinder stagnierten (z. B. A-1, B-12) mit einem Zuwachs um nur ca. 10 Wörter. Die drei Kinder (A-2, B-10, B-13), die scheinbar einen Rückschritt machten und im März weniger Wörter richtig lasen als im November, befanden sich von Beginn an auf einem so hohem Niveau, das die meisten anderen Kinder auch im März nicht erreichten, sodass davon ausgegangen werden muss, dass diese Kinder an ihre obere Leistungsgrenze gestoßen sind.

Es kann nur spekuliert werden, welche Fortschritte in den fehlenden vier Schulmonaten hätten erreicht werden können, aber da nur zwei Kinder auf niedrigem Lesetempo stagnierten, kann das vorgestellte Vorgehen als zielführend angesehen werden. Vor allem, weil auch die Kinder der KG ohne Intervention ähnliche Verläufe zeigten (vgl. Abb. 11), wengleich häufiger auf höherem Leistungsniveau. So kann davon ausgegangen werden, dass der überwiegende Teil der Kinder, die bei Schuleintritt als Risikokinder getestet wurden, durch regelmäßige Testungen und daran anknüpfende Interventionen zumindest auf dem Leistungsniveau der Klasse gehalten werden konnten, was die Wirksamkeit des Vorgehens unterstreicht.

Eine Messwiederholungsanalyse aufgrund der individuellen Fehlzeiten der Kinder berechnet über den ersten und letzten Messzeitpunkt bestätigt diese Annahme (siehe Tabelle 8). Wie im

Tabelle 8: Messwiederholungsanalyse der Leseflüssigkeitsentwicklung im zweiten Schuljahr;  $F(1;??)=5,785$ ,  $p < 0.05$ , Kovariate (Gruppe) nicht signifikant.

Gruppe	N	Nov.'20 [M (s)]	Mär.'21 [M (s)]
TG	7	106,00 (47,87)	113,43 (24,87)
KG	11	121,55 (47,80)	140,18 (41,46)

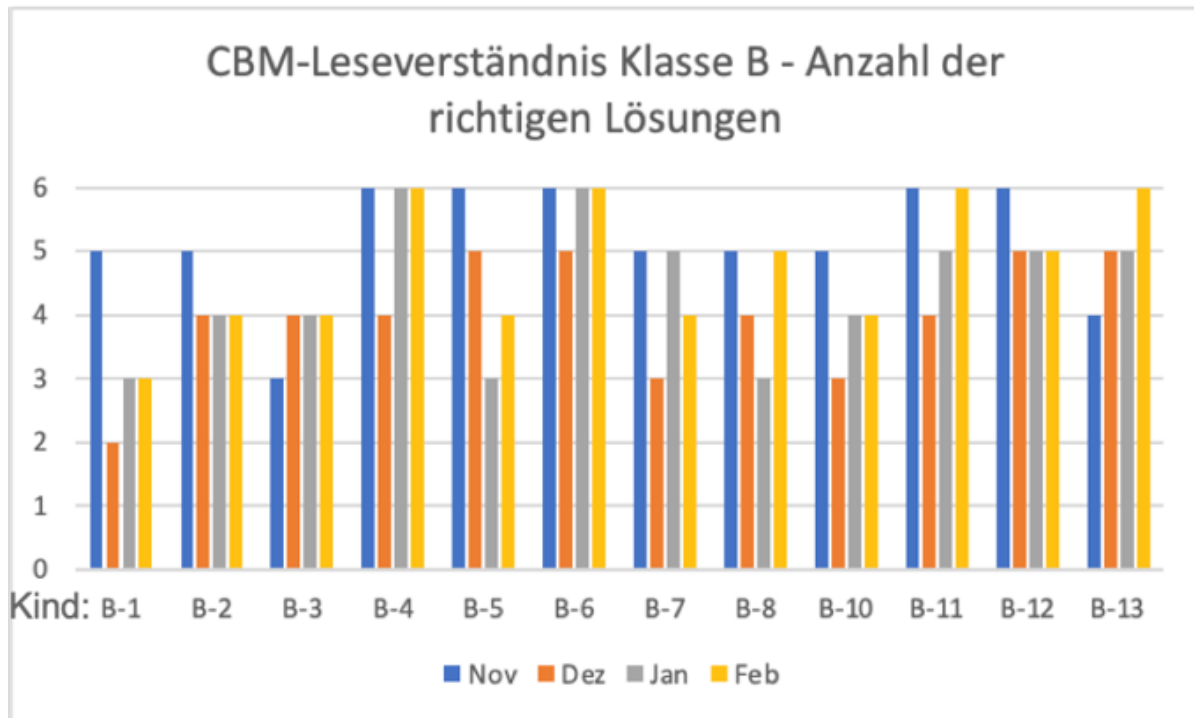


Abbildung 12: Ergebnisse der CBM-Leseverständnistests im zweiten Schuljahr (Messzeitraum November 2019 bis Februar 2020)

ersten Schuljahr wird der Zuwachs an Leseflüssigkeit signifikant, ein Gruppeneffekt hingegen nicht, sodass davon ausgegangen werden kann, dass sie TG und KG nicht in (der Entwicklung) der Leseflüssigkeit unterscheiden.

## 5.7 CBM-Leseverständnistestung in der zweiten Klassenstufe

Der Vergleich der TG mit der KG der Klasse B im Leseverständnis von November 2019 (Text 1) bis Februar 2020 (Text 4) zeigt ähnliche Ergebnisse (vgl. Abb. 12). Die maximale Anzahl an richtigen Lösungen erreichten insgesamt sechs Kinder an ein bis drei Testtagen. Zwei dieser Kinder gehörten der TG, vier der KG an. Das Kind mit der geringsten Anzahl an richtigen Lösungen (B-1) gehörte ebenfalls zur KG.

Eine Messwiederholungsanalyse (Tabelle 9) über alle vier Messzeitpunkte hinweg zeigt für die Entwicklung des Leseverständnisses nur die Tendenz zur Signifikanz. Der Gruppenunterschied wird wiederum nicht signifikant. Aufgrund dieser Datenlage könnte nun behauptet werden, dass die Kinder in allen Gruppen und über alle Messzeitpunkte gleich gut gelesen haben, unter der Annahme, dass auch die Schwierigkeit der CBM-Tests gleich war.



Tabelle 9: Messwiederholungsanalyse der Leseverständnisentwicklung im zweiten Schuljahr von November 2019 (=Text 1) bis Februar 2020 (=Text 4);  $F(3;??)=2,901$ ,  $p < 0.10$ , Kovariate (Gruppe) nicht signifikant.

Gruppe (N)	Text 1 [M (s)]	Text 2 [M (s)]	Text 3 [M (s)]	Text 4 [M (s)]
TG (4)	4,50 (0,54)	4,25 (0,96)	4,50 (0,58)	4,75 (0,96)
KG (8)	5,50 (1,29)	4,00 (0,93)	4,38 (1,30)	

Tabelle 10: Skalenwerte der vier Lesetexte zum CBM-Leseverständnis berechnet über 12 Kinder

Skalenwerte	Text 1	Text 2	Text 3	Text 4
M	5,17	4,00	4,42	4,75
SD	0,94	0,97	1,08	1,06
mittlere Itemschwierigkeit	0,86	0,67	0,74	0,79

Tabelle 10 ist zu entnehmen, dass die Mittelwerte der richtigen Lösungen nicht stark voneinander abweichen und sich die Standardabweichungen überschneiden. Auch die mittleren Itemschwierigkeiten der Fragen zu den Texten liegen in einem vergleichbaren Bereich und zeigen auf, dass die Fragen eher als leicht zu lösen einzustufen waren. Mit Blick auf die Stichprobe war dies auch beabsichtigt. Die Testgenauigkeit liegt über alle vier Texte hinweg bei einem Wert von  $\alpha = .63$  und bei einer Paralleltestreliabilität von  $r_{tt} = .75$ , was keinem zufriedenstellenden Wert entspricht, da nur ca. 40% der Testvarianz aufklärt werden. Dies ist auch der Kürze des Tests geschuldet, der über die vier Monate hinweg aus nur 24 Items bestand. Lienert & Raatz (1998, S. 209f.) beschreiben in einem ähnlichen Beispiel, dass aus einer Erhöhung der Itemzahl auf 50 ein »...einigermaßen hinreichend reliables...« Testverfahren resultieren würde (Lienert & Raatz, 1998, S. 210). Wäre die Testdurchführung wie geplant bis einschließlich Juli 2020 möglich gewesen, hätten die Daten von 30 zusätzlichen Fragen aus fünf Testmonaten hinzugezogen werden können. Bei einer analogen Anwendung der von den beiden Autoren vorgestellten Formel würde eine als gut einzustufenden Reliabilitätswerte von  $\alpha = .79$  und  $r_{tt} = .87$ . Auch wenn diese Berechnungen nur aus Schätzung zu werten sind, weil sie zudem auf einer kleinen Stichprobe basiert, zeigen diese Überlegungen, dass das gewählte Vorgehen auf einen gewinnbringenden Pfad führt.

## 5.8 Standardisierte Testung am Ende der zweiten Klassenstufe

Wie bereits im ersten Schuljahr wurde auch am Ende des zweiten Schuljahres der standardisierte Lesetest ELFE II von Lenhard et al. (2017) an insgesamt 19 Kindern durchgeführt (vgl. Abb. 13), da nicht alle Kinder zum Testzeitpunkt anwesend waren. Es zeigt sich wiederum, dass am Ende des zweiten Schuljahres nur wenige Kinder unterdurchschnittliche Leistungen im Wort-, Satz- und Textverständnis aufweisen. Sowohl für die richtig gelesene Aufgaben als auch für die bearbeiteten Aufgaben sind dies je nach Subtest drei bis vier Kinder sowohl aus der TG als auch der KG (jeweils gleich verteilt). Die überwiegende Anzahl der Kinder (je nach Subtest ca. 63% bis 74%) zeigten durchschnittliche Leistungen und wiederum ein bis drei Kinder (alle aus der KG) lagen im überdurchschnittlichen Bereich. Auch hier sind mit Verweis auf die Normierungstichprobe des ELFE II die Leistungen der Kinder als überwiegend beachtlich hervorzuheben. Die Gruppenunterschiede werden am Ende des zweiten Schuljahres in allen ELFE-Skalen mit

Tabelle 11: t-Test-Kennwerte der ELFE-II-Subtests (Lenhard et al., 2017) am Ende des zweiten Schuljahres berechnet über die T-Werte der TG und der KG

Subtest	Gruppen	N	m	s	T	df	p
Wortverständnis richtig	TG	5	42,60	4,93	-2,664	11	<0.05
	KG	8	56,75	11,07			
Satzverständnis richtig	TG	5	40,60	5,22	-1,802	11	<0.10
	KG	8	50,25	11,09			
Textverständnis richtig	TG	5	42,20	9,09	-2,122	11	<0.10
	KG	8	52,75	8,50			
Wortverständnis bearbeitet	TG	5	44,40	4,83	-3,551	11	<0.01
	KG	8	58,62	8,02			
Satzverständnis bearbeitet	TG	5	41,40	4,10	-2,993	11	<0.05
	KG	8	52,75	7,74			
Textverständnis bearbeitet	TG	5	39,80	6,72	-2,907	11	<0.05
	KG	8	54,37	9,78			

Tabelle 12: Korrelationen der ELFE-II-Subtests am Ende des zweiten Schuljahres

Anm.: – Signifikanzniveaus der Korrelationskoeffizienten: \* p<0.05; \*\* p<0.01, T<0.10 – Effektstärke r $\cong$ 0.1 – klein, r $\cong$ 0.3 – mittel, r $\cong$ 0.5 – groß

ELFE II	(2)	(3)	(4)	(5)	(6)
(1) Wortverständnis richtig	.892**	.792**	.942**	.686**	.431
(2) Satzverständnis richtig		.900**	.850**	.770**	.492T
(3) Textverständnis richtig			.773**	.813**	.680*
(4) Wortverständnis bearbeitet				.849**	.602*
(5) Satzverständnis bearbeitet					.882**
(6) Textverständnis bearbeitet					

besseren Ergebnissen für die KG signifikant oder weisen zumindest die Tendenz zur Signifikanz auf (s. Tabelle 11).

Die am Ende des ersten Schuljahres gefundenen hohen Korrelationen zwischen den Subtests des ELFE II lassen sich ein Jahr später bestätigen (vgl. Tab. 12). Die Ergebnisse der beiden Schuljahre sind vergleichbar. Der Wort-, Satz- und Textverständnistest korrelieren hochsignifikant bei mittleren bis großen Effekten miteinander, sowohl was die bearbeiteten als auch die richtig gelösten Aufgaben betrifft. Dies lässt sich zu diesem Zeitpunkt überwiegend auch für die Skala der bearbeiteten Aufgaben im Textverständnistest feststellen, die nun nur noch mit den richtigen Lösungen im Wortverständnis nicht korreliert, aber zum Satzverständnis die Tendenz zur Signifikanz und insgesamt zumindest mittlere Effekte aufweist. Daraus lässt sich schließen, dass sich die Leseleistung der Kinder weiterentwickelt und vervollständigt hat und nun nicht mehr nur die hierarchieniederen, sondern zunehmend auch die hierarchiehohe Leseprozesse nach dem Modell von Lenhard (2019, vgl. Abb. 1) beherrscht werden.

Die Leseleistung in der ELFE II (Lenhard et al., 2017) am Ende der zweiten Klasse lässt sich aus den Werten aller drei vorheriger Messzeitpunkte vorhersagen. Für das Wortverständnis sind hier v.a. die pB und RAN (TEPHOBE, Mayer, 2016) zu Beginn und die Wortverständnisleistung am Ende der ersten sowie das Wortlesen mit der WLLP-R (Schneider et al., 2011) zu Beginn

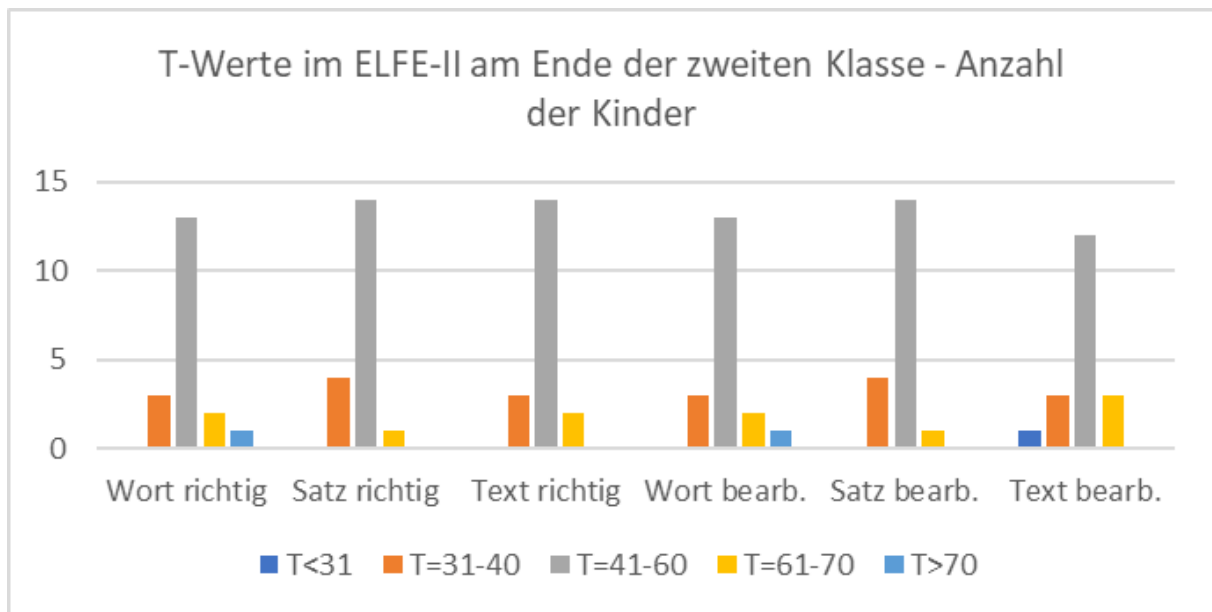


Abbildung 13: Ergebnisse des ELFE II (Lenhard et al. 2017)

der zweiten Klasse signifikante Prädiktoren. Für das Satz- und Textverständnis wird nur die pB signifikant, nicht aber die Benennungsgeschwindigkeit. Auch hier lassen sich die Leseleistungen am Ende der zweiten Klasse aus dem Wortverständnistest der ersten Klasse vorhersagen. Zudem sind wiederum für das Satzverständnis am Ende der zweiten Klasse die Wortlesefähigkeit (SLRT II, Moll & Landerl, 2014) zu Beginn dieser Klasse und für das Textverständnis das Wortlesen mit der WLLP-R von Relevanz. Zudem werden die Korrelationen der CBM-Testleistung am Ende der zweiten Klasse mit dem Wort- (Kl.1:  $r=.754$ ,  $p<.001$ ; Kl.2:  $r=.809$ ,  $p<.01$ ), Satz- (Kl.1:  $r=.765$ ,  $p<.001$ ; Kl.2:  $r=.959$ ,  $p<.001$ ) und Textverständnis (Kl.1:  $r=.564$ ,  $p<.05$ ; Kl.2:  $r=.929$ ,  $p<.001$ ) in beiden Schuljahren hoch signifikant. Eine Regressionsanalyse konnte hier aufgrund des geringen Stichprobenumfangs nicht gerechnet werden.

## 6 Fazit

Die Studie beschreibt die Entwicklung der Leseleistungen von Kindern einer Sprachheilschule über die ersten beiden Schuljahre hinweg. Dabei konnte einerseits gezeigt werden, dass durch das regelmäßige Training, dessen Erfolg durch verschiedene CBM-Testungen regelmäßig überprüft wurde, die überwiegende Zahl der zu Beginn der ersten Klasse als Risikokinder eingestuft Kinder soweit gefördert werden konnten, dass sie den Anschluss an ihre Klassenkamerad\*innen mit besseren Lernvoraussetzungen nicht verloren haben. Somit ist dieses Vorgehen als gewinnbringend und leistungssteigernd einzustufen (vgl. auch für Viertklässler Souevignier & Förster, 2011 und als Überblick Blumenthal, in diesem Band). Bereits Jungjohann et al. (2018) konnten den Nutzen dieses Vorgehens anhand anderer Verfahren in einem Review feststellen bzw. Jungjohann et al. (2021) in ihrer Längsschnittstudie zu ähnlichen Ergebnissen gelangen. Andererseits weisen die hohen Korrelationen zwischen den CBM-Tests und dem ELFE-Test darauf hin, dass auch die Validität des GOM-Vorgehens gegeben ist. Am Vorgehen der Erfassung des Leseverständnisses im Rahmen des CBM-Verfahrens muss jedoch noch weitergearbeitet werden.

Tabelle 13: Vorhersage der Leseleistung am Ende der 2. Klasse durch pB und RAN (N=13) durch die Leseleistung am Ende der ersten (N=13) und am Beginn der zweiten Klasse (N=12)

<b>ELFE II Klasse 2</b>	<b>TEPHOBE</b>	<b>R2</b>	<b>korr. R2</b>	<b>F</b>	<b>p</b>	$\beta$	<b>t</b>	<b>p</b>
Wortverst.	pB	0.596	0.515	7,370	0.011	.502	2,419	0.036
	RAN					0.474	2,284	0.045
Satzverst.	pB	0.322	0.261	5,231	.043	0.568	2,287	0.043
	RAN					0.244	0,948	0.365
Textverst.	RAN	0.32	0.258	5,182	.044	0.566	2,276	0.044
	pB					0.428	1,837	0.096
<b>ELFE II Klasse 2</b>	<b>ELFE II Klasse 1</b>	<b>R2</b>	<b>korr. R2</b>	<b>F</b>	<b>p</b>	$\beta$	<b>t</b>	<b>p</b>
Wortverst.	Wort	0.789	0.77	41,063	<.001	0.888	6,408	<.001
	Satz					-.200	-0,864	0.408
	Text					-0.253	-1,697	0.121
Satzverst.	Wort	0.739	0.715	31,074	<.001	0.859	5,574	<.001
	Satz					-0.198	-0,763	0.463
	Text					-.133	-0,725	0.485
Textverst.	Wort	0.713	0.687	27,328	<.001	0.844	5,228	<.001
	Satz					0.279	1,048	0.319
	Text					0.149	0,775	0.456
<b>ELFE II</b>	<b>Klasse 2</b>	<b>R2</b>	<b>korr. R2</b>	<b>F</b>	<b>p</b>	$\beta$	<b>t</b>	<b>p</b>
Wortverst.	WLLP-R	0.707	0.679	24,140	0.001	0.841	4,913	0.001
	SLRT-W					0.225	0,571	0.582
	SLRT-PW					0.389	1,693	0.125
Satzverst.	SLRT-W	0.778	0.756	35,032	<.001	0.882	5,919	<.001
	WLLP-R					0.386	1,194	0.263
	SLRT-PW					-0.181	-0,555	0.593
Textverst.	WLLP-R	0.737	0.71	27,958	<.001	0.858	5,288	<.001
	SLRT-W					0.34	0,940	0.372
	SLRT-PW					-0.095	-0,376	0.716

Zudem konnte anhand von vier Messzeitpunkten die Leseleistung der Kinder mit verschiedenen standardisierten Verfahren abgebildet werden. Es zeigte sich, dass die meisten Kinder durchschnittliche Leistungen aufwiesen, aber nur wenige Kinder im unterdurchschnittlichen bzw. auch einige im überdurchschnittlichen Bereich lagen. Diese Resultate sind umso erfreulicher, wenn man beachtet, dass sich die Stichprobe ausschließlich aus Kindern mit sprachheilpädagogischem Hintergrund zusammensetzte. Zudem verdeutlichen sie, dass durch das durchgeführte Training und die regelmäßigen CBM-Testungen auch Kinder mit schwierigen Lernvoraussetzungen soweit gefördert werden konnten, dass sie durchschnittliche bis gute Leseleistungen erbringen können. Dies ist für ihre weitere Schulzeit und den damit verbundenen Wissenserwerbs von großer Bedeutung.

Allerdings ist anzumerken, dass sich aufgrund der geringen Stichproben keine KG bilden ließ, die auf dem gleichen Leistungsniveau der Vortests lag, wie die Trainingskinder. Zudem sprachen ethische Gründe dagegen, Risikokindern keine Intervention zukommen zu lassen. Dennoch ist dies eine Schwachstelle der Studie, die in weiteren Studien behoben werden sollte.

## Literatur

- Anderson, S., Jungjohann, J. & Gebhardt, M. (2020). Effects of using curriculum-based measurement (CBM) for progress monitoring in reading and an additive reading instruction in second classes. *Zeitschrift für Grundschulforschung*, 13(1), 151-166. <https://doi.org/10.1007/s42278-019-00072-5>
- Baddeley, A.D. (2014). *Essentials auf human memory*. London & New York: Psychology Press.
- Baddeley, A. & Hitch, G.J., (1974). Working memory. In G.A. Bower (Ed.). *The psychology of learning and motivation* (pp. 47-89). New York: Academic Press.
- Berg, M. (2015). *Überprüfung grammatischer Fähigkeiten bei 5- bis 8-jährigen Kindern (MuSe-Pro)*. München: Reinhardt.
- Black, P. & William, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5(1), 7–74.
- Bloom, B. (1971). *Mastery learning*. New York: Holt, Rinehart & Winston.
- Blumenthal, S. (2021). Lernverlaufsdiagnostik. In M. Gebhardt et al (Hrsg.), *Handbuch sonderpädagogische Diagnostik*.
- Coltheart, M. (1978). Lexical access in simple reading tasks. In G. Underwood (ed.). *Strategies of information processing* (pp. 151-216). London: Academic Press.
- Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research*, 58(4), 438–481.
- Deno, S.L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children*, 52, 219-232.
- Deno, S.L. & Mirkin, P.K. (1977). *Data-based program modification: A manual*. Reston V.A.: Council of Exceptional Children.

- Frith, U. (1985). Beneath the surface of developmental dyslexia. *Developmental dyslexia*, 13, 301-330.
- Fuchs, L.S. (2004). The past, present, and future of curriculum-based measurement research. *School Psychology Review*, 33, 188-192.
- Fuchs, L.S. & Fuchs, D. (1993). Formative evaluation of academic progress. How much growth can we expect? *School Psychology Review*, 22, 1-30.
- Günther, K.B. (1986). Ein Stufenmodell der Entwicklung kindlicher Lese- und Rechtschreibstrategien. In H. Brügelmann (Hrsg.). *ABC und Schriftsprache: Rätsel für Kinder, Lehrer und Forscher*. Konstanz: Faude.
- Hansen, K. (1929). *Die Problematik der Sprachheilschule in ihrer geschichtlichen Entwicklung*. Halle/Saale: Marhold.
- Jungjohann, J., Gebhardt, M., Diehl, K. & Mühling, A. (2017). Förderansätze im Lesen mit LEVUMI. <http://dx.doi.org/10.17877/DE290R-18042>
- Jungjohann, J., Gegenfurtner, A. & Gebhardt, M. (2018). Systematisches Review von Lernverlaufsmessung im Bereich der frühen Leseflüssigkeit. *Empirische Sozialpädagogik*,
- Jungjohann, J., Schurig, M. & Gebhardt, M. (2021). Pilotierung von Leseflüchtigkeits- und Leseverständnistests zur Entwicklung von Instrumenten der Lernverlaufsdagnostik. Ergebnisse einer Längsschnittstudie in der 3ten und 4ten Jahrgangsstufe. *Vierteljahresschrift für Heilpädagogik und ihre Nachbargebiete Plus*, 90, 1-19. <https://doi.org/10.2378/vhn2021.art12d>
- Kintsch, W. (1977). *Memory and cognition*. New York: Wiley & Sons
- Klauer, K.J. (2006). Erfassung des Lernfortschritts durch curriculumbasierte Messung. *Heilpädagogische Forschung*, 32 (1), 16-26.
- Klicpera, C. & Gasteiger-Klicpera (1998). *Psychologie der Lese- und Schreibschwierigkeiten. Entwicklung, Ursachen, Förderung*. Weinheim: Beltz.
- Lenhard, W. (2019). *Leseverständnis und Lesekompetenz*. Stuttgart: Kohlhammer.
- Lenhard, W., Lenhard, D. & Schneider, W. (2017). *Ein Leseverständnistest für Erst- bis Siebtklässler – Version II (ELFE II)*. Göttingen: Hogrefe.
- Lienert, G. & Raatz, U. (1998). *Testaufbau und Testanalyse (6. Aufl.)*. Weinheim: Psychologie Verlags Union
- Maier, U. (2010). Formative Assessment – Ein erfolgsversprechendes Konzept zur Reform von Unterricht und Leistungsmessung? *Zeitschrift für Erziehungswissenschaft*, 13(2), 292-308.
- Mannhaupt, G. & Jansen, H. (1989). Phonologische Bewusstheit: Aufgabenentwicklung und Leistungen im Vorschulalter. *Heilpädagogische Forschung*, 15, 50-56.
- Marsh, H.W. (2005). Der »Big-fish-little-pond«-Effekt und das akademische Selbstkonzept. *Zeitschrift für Pädagogische Psychologie*, 19, 119-127.
- Marx, P. (2007). *Lese- und Rechtschreiberwerb*. München: Schöningh.

- Mayer, A. (2016). Test zur Erfassung der phonologischen Bewusstheit und der Benennungsgeschwindigkeit (3. Aufl.). München: Reinhardt.
- Meringer R. & Mayer, K. (1895). Versprechen und Verlesen. Eine psychologisch-linguistische Studie. Stuttgart: Göschen.
- Moll, K. & Landerl, K. (2014). Lese- und Rechtschreibtest SLRT-II; Weiterentwicklung des Salzburger Lese- und Rechtschreibtests (SLRT). Bern: Huber.
- Nöstlinger, C. (2013). Schulgeschichten vom Franz. Hamburg: Oetinger.
- Nöstlinger, C. (2015). Freundschaftsgeschichten vom Franz. Hamburg: Oetinger.
- Ranschburg, P. (1916). Die Leseschwäche (Legasthenie) und die Rechenschwäche (Arithmasthenie) der Schulkinder im Lichte des Experiments. Berlin: Springer.
- Rost, D.H. & Hartmann, A. (1992). Lesen, Hören, Verstehen. Zeitschrift für Psychologie, 200, 345-361
- Schneider, W., Blanke, I., Faust, V. & Küspert, P. (2011). Würzburger Leise Leseprobe – Revision (WLLP-R). Göttingen: Hogrefe.
- Schulte-Körne, G. (2001). Lese-Rechtschreibstörung und Sprachwahrnehmung. Münster: Waxmann.
- Schneider, W. Näslund, J.C. & Weinert, F.E. (1989). Prerequisites of school achievement. In F.E. Weinert, & W. Schneider (Eds.). The Munich Longitudinal Study on the Genesis of Individual Competencies (LOGIC). Report No. 5: results of wave three (pp. 155-182). München: Max-Planck-Institut für Psychologische Forschung.
- Souvignier, E. & Förster, N. (2011). Effekte prozessorientierter Diagnostik auf die Entwicklung der Lesekompetenz leseschwacher Viertklässler. Empirische Sonderpädagogik, 3 (3), 243-255.
- Stock, C. (2005). Über phonologische und schriftsprachliche Kompetenzen deutscher Grundschüler. Hamburg: Kovač.
- Stock, C., Marx, P. und Schneider, W. (2003). Basiskompetenzen für Lese-Rechtschreibleistungen (BAKO 1-4). Ein Test zur Erfassung der phonologischen Bewusstheit vom ersten bis vierten Grundschuljahr. Göttingen: Hogrefe.
- Voß, S. & Hartke, B. (2014). Curriculumbasierte Messverfahren (CBM) als Methode der formativen Leistungsdiagnostik im RTI-Ansatz. In M. Hasselhorn, W. Schneider & U. Trautwein (Hrsg.), Lernverlaufsdagnostik (S. 83-99). Göttingen: Hogrefe.
- Walter, J. (2008). Curriculumbasiertes Messen (CBM) als lernprozessbegleitende Diagnostik: Erste deutschsprachige Ergebnisse zur Validität, Reliabilität und Veränderungssensibilität eines robusten Indikators zur Lernfortschrittsmessung beim Lesen. Heilpädagogische Forschung, 2, 62-79.
- Warnke, A. (1992). Legasthenie und Hirnfunktion. Bern: Huber.
- Wolf, M. & Bowers, P.G. (1999). The double-deficite hypothesis for the developmental dyslexias. Journal of Educational Psychology, 91, 415-438.

**Claudia Stock** seit 2001 am Lehrstuhl für Pädagogische und Entwicklungspsychologie der Universität Würzburg und seit 2011 als Dozentin und Wissenschaftlerin am dortigen Lehrstuhl für Sonderpädagogik III. Ihre Projekte befassen sich mit dem Themengebiet Legasthenie und Schriftspracherwerb, aktuell auch in Verbindung mit der Lernverlaufsdiagnostik in diesem Kontext.