

**Kritische Auseinandersetzung mit bestehenden Verfahren zur
Erfassung des Schriftspracherwerbs und Entwicklung eines neuen
Lesetests: Eine empirische Untersuchung mit Schüler:innen der
ersten Klasse**

Schriftliche Hausarbeit

Zur Erlangung des akademischen Grades Master of Education
im Studiengang Lehramt für sonderpädagogische Förderung

Vorgelegt von

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

Fakultät Kulturwissenschaft
Institut für Diversitätsstudien
Technische Universität Dortmund

[06.02.2025]

Erstgutachterin: Prof. Dr. Barbara Mertins

Zweitgutachterin: Katrin Odermann

Danksagung

An dieser Stelle spreche ich meinen herzlichen Dank an alle aus, die mich bei der Erstellung meiner Masterarbeit begleitet und unterstützt haben.

Mein besonderer Dank gilt meiner Betreuerin, Prof. Dr. Barbara Mertins für die große Unterstützung, die Expertise und das mir entgegengebrachte Vertrauen während des gesamten Prozesses. Ebenfalls danke ich Katrin Odermann von Herzen für das immer offene Ohr und die enge Begleitung. Danke, dass Ihr meine Begeisterung für das Thema *Schriftspracherwerb* auf besondere Weise geweckt habt. Der wertvolle Austausch mit Euch wird mich weit über diese Arbeit hinaus begleiten.

Genauso danke ich dem gesamten Team der *psycholinguistics laboratories* für die Ermutigungen und stetigen Anregungen, die das Projekt vorangetrieben haben. Die Entwicklung des Tests war ein gemeinsamer Prozess, der durch die wertvollen Impulse der Schriftspracherwerbs-AG maßgeblich geprägt wurde – ein herzliches Dankeschön an alle Beteiligten für Eure Mühe und Euren Einsatz. Ohne Euch wäre die Arbeit nicht die, die sie heute ist! Mein besonderer Dank gilt Jana Sievering, deren sorgfältige zweite Datenauswertung die Aussagekraft dieser Studie maßgeblich bereichert hat. Ein herzliches Dankeschön an Sarah Dresch für die Überprüfung der Frequenzen der Items und die Bereitschaft, mir bei jedem Anliegen zur Seite zu stehen.

Auch möchte ich mich bei Anna, Charline, Marinus und Anna bedanken. Ihr wart mir eine große Unterstützung während meines Schreibprozesses. Mit Euren hilfreichen Gedankenanstößen, Korrekturen und aufbauenden Worten habt Ihr nicht nur zum Gelingen dieser Arbeit beigetragen, sondern auch dafür gesorgt, dass selbst anstrengende Phasen von Lachen und guter Laune begleitet wurden. Danke, dass Ihr an meiner Seite wart!

Abschließend möchte ich mich an meine Freund:innen und meine Familie wenden: Auch wenn diese Zeit nicht nur von Höhen geprägt war, wart Ihr stets an meiner Seite. Danke für Euer unerschütterliches Verständnis, für das gemeinsame Feiern jedes kleinen und großen Erfolgs und für Eure beständige Ermutigung.

Herne, den 29.01.2025

I. Gliederung

II. Abbildungsverzeichnis	iv
III. Tabellenverzeichnis	iv
1. Einleitung	1
2. Theoretischer Hintergrund	4
2.1 Schriftspracherwerb	4
2.1.1 Eine Begriffsdefinition.....	4
2.1.2 Phonologische Bewusstheit	6
2.1.3 Orthografische Kompetenz.....	9
2.1.4 Semantik.....	12
2.2 Erfassung des Schriftspracherwerbs in der ersten Klasse.....	15
3. Entwicklung der Fragestellung und der Hypothesen.....	18
4. Analyse bestehender Verfahren zur Erfassung des Schriftspracherwerbs	20
4.1 Leseverständnistest für Erst- bis Sechstklässler (ELFE 1-6)	20
4.1.1 Kritische Diskussion	21
4.2 Basiskompetenzen für Lese- und Rechtschreibleistungen (BAKO 1-4).....	24
4.2.1 Kritische Diskussion	25
4.3 Inventar zur Erfassung der Lesekompetenzen im 1. Schuljahr (IEL-1)	28
4.4 Zwischenfazit	31
5. Entwicklung des neuen Lesetests für die erste Klasse	32
5.1 Zielsetzung und Anwendungsbereiche	32
5.2 Schritte bei der Testkonstruktion	36
5.3 Aufbau des Lesetests.....	37
5.4 Durchführung des Tests	46
5.4.1 Durchführungsinstruktionen.....	47
5.5 Kodierungssystem für die Auswertung	51
6. Methode	54
6.1 Zielsetzung der Studie	54
6.2 Stichprobe.....	55
6.3 Vorgehen	56
7. Darstellung der Ergebnisse	58
8. Diskussion.....	64
9. Fazit	73
IV. Literaturverzeichnis	v
V. Anhangsverzeichnis	v

II. Abbildungsverzeichnis

Abbildung 1: ELFE 1-6 (Lenhard & Schneider 2006a), Subtest 1 (Wortverständnistest), Übungsitems	21
Abbildung 2: ELFE 1-6 (Lenhard & Schneider 2006a), Subtest 1 (Wortverständnistest), Items 23, 27 und 30.....	22
Abbildung 3: ELFE 1-6 (Schneider & Lenhard 2006a), Subtest 2 (Satzverständnistest), Item 1	23
Abbildung 4: ELFE 1-6 (Lenhard & Schneider 2006a), Subtest 3 (Textverständnistest), Übungsitem 1	23
Abbildung 5: BAKO 1-4 (Stock et al. 2003a), Subtest 5 (Lautkategorisierung).....	25
Abbildung 6: BAKO 1-4 (Stock et al. 2003a), Subtest 2 (Vokalersetzung)	26
Abbildung 7: BAKO 1-4 (Stock et al. 2003a), Subtest 4 (Phonemvertauschung), Items 42-44	27
Abbildung 8: Mittelwerte der einzelnen Subtests im Vergleich	59
Abbildung 9: Boxplot (Gesamtpunktzahl ohne Subtest 2)	60
Abbildung 10: Standardabweichungen der linguistischen Bereiche der Subtests im Vergleich.....	61
Abbildung 11: Beispiel einer Proband:innenantwort in dem dritten Subtest.....	67

III. Tabellenverzeichnis

Tabelle 1: Durchführungsinstruktionen.....	48
Tabelle 2: Vergleich der erreichten Punktzahlen in den Subtests einzelner Proband:innen...	62

1. Einleitung

Learning to read is 'a psycholinguistic guessing game' (Browne Rego & Bryant 1993, S. 1).

Das Zitat verdeutlicht, dass es sich bei dem Schriftspracherwerb um einen komplexen und dynamischen Prozess handelt, bei dem Kinder kontinuierlich Hypothesen über die Verbindung zwischen gesprochener und geschriebener Sprache entwickeln und überprüfen (Bredel et al. 2017). Diese Vielschichtigkeit macht es notwendig, den Schriftspracherwerb nicht nur als einen rein mechanischen und linearen Prozess zu charakterisieren, sondern diesen vielmehr als ein Zusammenspiel kognitiver, sprachlicher und psychologischer Fähigkeiten aufzufassen. Was für kompetente Leser:innen aufgrund vielfacher automatisierter Vorgänge ohne viel Anstrengung passiert, stellt für Leseanfänger:innen oft ein mühsames Unterfangen dar (Marx 2007).

Der Schriftspracherwerb ist für das gesamte Leben von großer Bedeutung und ist heutzutage vielleicht sogar wichtiger denn je: Durch die fortlaufende Verlagerung von Lebensbereichen in digitale Welten wird dies nicht weniger, sondern gewinnt immer mehr an Bedeutung (Lenhard 2019). Dadurch geht mit der Kompetenz im Bereich der Schriftsprache ein Teilhabeaspekt einher: „Moderne Gesellschaften sind literale Gesellschaften“ (Rosebrock & Scherf 2022 S. 11), sodass für eine Teilhabe am gesellschaftlichen, kulturellen und politischen Leben eine gewisse Kompetenz vorausgesetzt wird (Lenhard 2019).

Der Schriftspracherwerb stellt eine Grundlage für den schulischen Erfolg für nahezu alle Fächer dar. Sprache ist die Basis für die Wissensaufnahme und -verarbeitung (Schneider 2017; Lenhard 2019). Die Relevanz geht folglich über die Deutschnote hinaus, denn „wer nicht flüssig und sinnentnehmend lesen kann, tut sich auch damit schwer, Anweisungen für die Arbeit in naturwissenschaftlichen Bereichen zu verstehen und Textaufgaben im Fach Mathematik angemessen zu bearbeiten“ (Schneider 2017, S. 15). In dem Kontext des schulischen Einflusses auf den Schriftspracherwerb ist die Grundschulzeit besonders hervorzuheben. Kinder machen in der Grundschule oft ihre ersten bewussten Erfahrungen mit der Schriftsprache, indem sie sich aktiv mit dieser auseinandersetzen. Der Schriftspracherwerb basiert demnach zu der Zeit meist erstmals auf einem didaktischen Konzept, das aufeinander aufbauende Schritte und Methoden integriert. Aufgrund dessen wird in dieser Arbeit das Ende der ersten Klasse in den Mittelpunkt der Betrachtung gerückt, das in der Wissenschaft häufig unberücksichtigt bleibt.

Der Schriftspracherwerb verläuft nicht bei allen Kindern gleich (Marx 2007). Unterschiedliche Individuen können in verschiedenen Bereichen des Erwerbsprozesses unterschiedliche Stärken und Schwächen zeigen. Deshalb ist es von großer Bedeutung, die individuellen Schüler:innen und deren derzeitigen Kompetenzen im Bereich der Schriftsprache gezielt in den Blick zu nehmen. Vor diesem Hintergrund wird in der vorliegenden Arbeit ein Test entwickelt,

der verschiedene Dimensionen des Schriftspracherwerbs berücksichtigt. Hauptziel der Arbeit ist es, ein Instrument zu schaffen, welches nicht nur die phonologische Bewusstheit, sondern auch orthografische Kompetenzen sowie das Leseverständnis auf Wort-, Satz-, und Textebene abbildet. Ferner sollen in einer nachfolgenden Studie die Kompetenzen in der ersten Klasse in den Fokus gerückt werden, indem der entwickelte Test erstmals eingesetzt und einer eingehenden Evaluierung unterzogen wird.

Dafür wird im Theorieteil (Kapitel 2) die Grundlage zum Schriftspracherwerb erläutert. Zunächst erfolgt eine Begriffsdefinition des Begriffs Schriftspracherwerb. Schwerpunkte des Kapitels 2.1 sind die für die Testentwicklung relevanten Aspekte, welche die phonologische Bewusstheit, die orthografische Kompetenz sowie die Semantik umfassen. In diesem Zuge werden sowohl zugrundeliegende Konzepte dargestellt als auch der aktuelle Forschungsstand und die daraus entstehende Bedeutung für den Schriftspracherwerb. Im Kapitel 2.2 wird die Relevanz einer Erfassung des Schriftspracherwerbs in der ersten Klasse erläutert und auf gängige Vorgehensweisen verwiesen.

Aus diesen theoretischen Grundlagen werden in Kapitel 3 die Forschungsfragen sowie die daraus folgenden Hypothesen entwickelt und ausgeführt. Anschließend liegt der Fokus in Kapitel 4 auf der kritischen Auseinandersetzung mit bereits bestehenden Verfahren zur Erfassung des Schriftspracherwerbs, in dessen Zuge der *ELFE 1-6* (Lenhard & Schneider 2006a) (Kapitel 4.1), der *BAKO 1-4* (Stock et al. 2003a) (Kapitel 4.2) sowie der *IEL-1* (Diehl & Hartke 2012a) (Kapitel 4.3) jeweils kurz dargelegt sowie kritisch diskutiert werden. Das Kapitel schließt mit einem Zwischenfazit (Kapitel 4.4) ab, in dem die zentralen Erkenntnisse für die Weiterentwicklung des eigenen Tests zusammengefasst werden.

Die Entwicklung des neuen Tests für eine Erfassung des Schriftspracherwerbs explizit für die erste Klasse steht im Mittelpunkt des Kapitel 5. Es werden elementare Zielsetzungen und die intendierten Anwendungsbereiche erläutert (Kapitel 5.1). In Kapitel 5.2 werden die Schritte der Testkonstruktion dargestellt. Anschließend werden die Subtests inklusive deren zu testenden Zielkompetenzen kleinschrittig beschrieben (Kapitel 5.3). Die spezifischen Überlegungen zu der Durchführung des Tests stehen im Kapitel 5.4 im Mittelpunkt der Betrachtung, in welchem Kontext auch die Durchführungsinstruktionen in den Blick genommen werden. Die Auswertung des Tests basiert auf einem Kodierungssystem, welches sich mit der Vergabe der Punkte auseinandersetzt und im nachfolgenden Kapitel 5.5 ausgeführt wird.

In Kapitel 6 liegt der Fokus auf der detaillierten Beschreibung der Forschungsmethodik, einschließlich der Zielsetzung der Studie (Kapitel 6.1), der Auswahl der Stichprobe (Kapitel 6.2) sowie dem konkreten Vorgehen (Kapitel 6.3). Nachfolgend werden in Kapitel 7 die aus der Studie gewonnenen Ergebnisse präsentiert. Schließlich werden die Ergebnisse dieser Arbeit mit Blick auf die jeweiligen Forschungsfragen und Hypothesen diskutiert (Kapitel 8). In einem

Schlusskapitel wird ein Fazit gezogen (Kapitel 9), in dem alle erlangten Erkenntnisse kompakt zusammengeführt und einer kritischen Betrachtung unterzogen werden. Ferner werden potenzielle Herausforderungen und Grenzen dieser Studie aufgezeigt und daraufhin Anregungen für weiterführende Fragestellungen formuliert, die als Grundlage für zukünftige Forschungsprojekte dienen können.

2. Theoretischer Hintergrund

Um eine einheitliche Basis für diese Arbeit zu schaffen, müssen die bisherigen Erkenntnisse über den Schriftspracherwerb ausgeführt werden. Wenngleich sich zu den Themen zahlreiche Aspekte und Perspektiven anführen lassen, wird im Rahmen dieser Arbeit nur auf die für die Beantwortung der Fragestellung und die Zielsetzung relevanten eingegangen. Dafür wird der Schriftspracherwerb in der ersten Klasse besonders hervorgehoben und darüber hinaus die Bedeutung der Erfassung des Schriftspracherwerbs herausgestellt.

2.1 Schriftspracherwerb

Die folgende Arbeit bezieht sich auf die fundamentalen Erkenntnisse über den Schriftspracherwerb mit besonderem Blick auf die phonologische Bewusstheit, die orthografische Kompetenz sowie die Semantik. Daher werden nachfolgend grundlegende Begriffe definiert, in einen Kontext gesetzt und der jeweilige aktuelle Forschungsstand dargestellt.

2.1.1 Eine Begriffsdefinition

In der Wissenschaft existieren zahlreiche Begriffe, die sich mit der Entwicklung des Lesens und Schreibens auseinandersetzen. Der Grund, weshalb dieser Arbeit der Begriff *Schriftspracherwerb* zugrunde liegt, findet sich in dem fundamentalen Verständnis, dass Kinder die Schriftsprache systematisch erwerben, nicht ausschließlich lernen. Bredel et al. (2017) legen dies dahingehend aus, dass „der Erwerb ohne eine direkte Steuerung durch eine Lehrperson abläuft“, während „es sich beim Lernen um einen gesteuerten Prozess des Wissensaufbaus“ (S. 71) handelt. An dieser Stelle lässt sich allerdings nicht in absoluten Kriterien sprechen, denn beide Prozesse, der ungesteuerte sowie der gesteuerte Erwerb, finden im Kontext der Schriftsprache statt und beeinflussen sich gegenseitig (Bredel et al. 2017). Auf der einen Seite wird der Lernprozess von außen wie beispielsweise von Lehrkräften oder Erziehungsberechtigten geplant und zielgerichtet gesteuert (Bredel et al. 2017). Auf der anderen Seite bedeutet Schriftspracherwerb auch Spontanität und Alltäglichkeit: Der ungesteuerte Erwerb geht nicht mit einer systematischen Anleitung einher, sondern stellt die individuelle Lebenswelt der Kinder in den Mittelpunkt der Betrachtung. Kinder erschließen sich unter anderem selbstständig die Schriftsprache durch ihre Erfahrungen „mithilfe innerer Hypothesen- und Regelbildungen“ (Bredel et al. 2017, S. 72) sowie ihrer intrinsischen Motivation und Neugier. Mehlem (2024) präzisiert dies, indem er herausstellt, dass der Schriftspracherwerb „als aktiver Konstruktionsprozess verstanden [wird], bei dem die Lernenden sich mit ihrer Umwelt auseinandersetzen und sich deren Strukturen schrittweise, über mehr oder weniger angemessene kognitive Konstrukte und Schemata, erschließen“ (S. 9). Während durch ungesteuerte Erfahrungen Grundlagen entwickelt werden können, die den gesteuerten Erwerb erleichtern, können gesteuerte Lernprozesse Anstöße für das ungesteuerte Lernen geben.

Des Weiteren geht mit dem Begriff der Schriftsprache eine ganzheitliche Perspektive einher. Es wird in den Vordergrund gerückt, dass es sowohl um das Lesen als auch um das Schreiben geht. Oftmals wird lediglich ein Teilaspekt aufgegriffen. Dies ist beispielsweise bei dem Begriff des *Schrifterwerbs* der Fall, welcher oft synonym zum Schriftspracherwerb verwendet wird. Hierbei wird der technische Aspekt, also die reine Fähigkeit, Buchstaben, Wörter und Sätze zu entziffern und zu verschriften, viel stärker in den Fokus gerückt. Der Begriff Schriftspracherwerb impliziert im Gegensatz dazu einen vielschichtigen Blick:

Was die Kinder auf dem Weg zu kompetenten Lesern und Schreibern lernen müssen, ist ja nicht nur ein besonderes Zeichensystem, sondern zusätzlich die Beherrschung besonderer Strukturmerkmale, die es im Gesprochenen so nicht gibt, wie bestimmte graphische Gesetzmäßigkeiten, satz- und textstrukturierende Mittel und ein besonderer – eben schriftsprachlicher – Stil. (Bredel et al. 2017, S. 73)

Der Schriftspracherwerb kann als ein Prozess angesehen werden, in dem Kinder einerseits lernen gesprochene Sprache in schriftliche Sprache umzusetzen und andererseits die schriftliche Sprache in gesprochene Sprache umzusetzen. Allerdings kommt es als Folge dessen fälschlicherweise zu der Annahme, Lesen und Schreiben seien klar voneinander trennbare Kompetenzen. Schröder-Lenzen (2004) stellt in dem Kontext heraus, dass der Schriftspracherwerb sich „in einem komplexen Erwerbsprozess [vollzieht], indem sich produktive (schreiben) und rezeptive (lesen) Aneignungsformen wechselseitig stützen“ (S. 43).

Darüber hinaus geht mit dem Begriff der Schriftsprache ein Entwicklungscharakter einher, der den Erwerb der Schriftsprache als einen dynamischen Prozess porträtiert, da er sich über einen langen Zeitraum erstreckt und sich nicht nur auf die Vermittlung bestimmter Fertigkeiten beschränkt. Zusätzlich impliziert der Begriff, dass der Erwerb der Schriftsprache nicht isoliert stattfindet, sondern eingebettet ist in soziale, kulturelle und schulische Kontexte. Wenn früher Begrifflichkeiten wie *Erstlesen* verwendet wurden, wurde indes „suggestiert, Lesen würde erst mit Schuleintritt beginnen“ (Schröder-Lenzen 2004, S. 29). In heutigen wissenschaftlichen Diskursen ist jedoch sicher belegt, dass der Schriftspracherwerb bereits vor Schuleintritt beginnt (Bredel et al. 2017).

Es wurden vielfache Versuche unternommen, den Schriftspracherwerb anhand eines Stufenmodells zu systematisieren und zu definieren.¹ Auch wenn diese Modelle inklusive festgelegter Reihenfolge der Phasen das Ziel verfolgen „anhand der registrierten Schriftsprachleistungen den Entwicklungsstand zu bestimmen und den weiteren Entwicklungsverlauf vorherzusagen“ (Schneider 2017, S. 20), muss festgehalten werden, dass der Erwerb nicht modular

¹ Aufgrund des begrenzten Umganges dieser Arbeit, können diese Modelle nicht detailliert erläutert werden. Das dreistufige Modell von Frith (1986) zählt jedoch zu den bekanntesten Modellen zum Schriftspracherwerb und wird trotz seines Entstehungszeitraums weiterhin häufig verwendet.

einhergeht, sondern dass verschiedene Prozesse ineinandergreifen. In der Wissenschaft besteht jedoch Einigkeit darüber, dass der Schriftspracherwerb auf basalen, niedrigeren Prozessen beruht, die das Entschlüsseln von Wörtern und Sätzen umfassen, sowie auf höheren Prozessen, bei denen Schriftsprache mit dem individuellen Vorwissen und Weltwissen verknüpft wird (Schneider 2017).

Was aus diesen Ausführungen bereits kenntlich wird, ist die Komplexität, die mit dem Schriftspracherwerb einhergeht. Was aus der Perspektive erfahrener Leser:innen selbstverständlich erscheint, da Prozesse bereits automatisiert und internalisiert wurden, ist hochkomplex und vielschichtig. Doch was bedeutet das konkret? Welche spezifischen Anforderungen werden an die Schüler:innen gestellt, wenn sie lesen oder schreiben sollen? Antworten auf diese Fragen werden in den nächsten Kapiteln näher beleuchtet.

2.1.2 Phonologische Bewusstheit

Da die *phonologische Bewusstheit* einen Teilbereich des *metalinguistischen Bewusstseins* darstellt, muss zunächst der Begriff des metalinguistischen Bewusstseins kurz definiert werden. Hierunter wird allgemein die Fähigkeit verstanden „seine eigenen Sprachverarbeitungsprozesse (beim Verstehen und bei der Produktion) bewusst zu kontrollieren und zu planen“ (Reber 2017, S. 34). Es bezieht sich auf die Fähigkeit über Sprache nachzudenken, sie zu analysieren und ihre Struktur unabhängig von dem Inhalt zu erkennen. Auch wenn dies nahezu alle Aspekte der Sprache betreffen kann, stellt das Erkennen und Manipulieren von Phonemen in diesem Kontext eine große Relevanz dar (Reber 2017). Diese genannten Kompetenzen fallen unter die phonologische Bewusstheit. Längsschnittstudien zeigen, dass sich eine frühzeitige Entwicklung des metalinguistischen Bewusstseins bereits vor dem Schuleintritt positiv auf die Lese- und Schreibfähigkeiten in der Grundschule auswirkt, was sich unter anderem in einer schnelleren Wortdekodierung und einer gesteigerten Effizienz bei der Informationsverarbeitung äußert (Schneider 2017).

Der Terminus der *phonologischen Bewusstheit* kann übergeordnet als die „Fähigkeit zur Erkennung der Lautstruktur der gesprochenen Sprache“ (Schneider 2017, S. 36) definiert werden. Wie an dem Wort bereits deutlich wird, steht die Phonologie, welche ein Teilgebiet der Linguistik darstellt, im Vordergrund. In diesem Bereich wird sich mit der Funktion und Struktur von Phonemen innerhalb eines Sprachsystems beschäftigt. Als Phoneme werden indes „die kleinsten bedeutungsunterscheidenden Einheiten der Sprache“ (Marx 2007, S. 23) bezeichnet.

Der Erwerb der phonologischen Bewusstheit beginnt bereits im Vorschulalter (Mehlem 2024). Infolgedessen wird die phonologische Bewusstheit vielfach als eine bedeutsame Vorläuferfertigkeit des Schriftspracherwerbs definiert (Schröder-Lenzen 2004). Auch Marx (2007) stellt

im Kontext eines Versuchs der Systematisierung von Vorläuferfertigkeiten die phonologische Bewusstheit als Teilbereich der phonologischen Informationsverarbeitung heraus.

Innerhalb der phonologischen Bewusstheit wird zwischen der *phonologischen Bewusstheit im engeren Sinne* und der *phonologischen Bewusstheit im weiten Sinne* differenziert. Während sich Ersteres auf das Erkennen und Verarbeiten von größeren Einheiten in der gesprochenen Sprache wie beispielsweise Silben bezieht, wird bei der phonologischen Bewusstheit im engeren Sinne der Fokus auf die Fähigkeit einzelne Phoneme wahrzunehmen und zu manipulieren gelegt (Schneider 2017). Bei der phonologischen Bewusstheit im weiten Sinne geht es beispielsweise um das Erkennen von Reimen, das Unterscheiden von vollständigen und unvollständigen Sätzen, dem Erkennen von Wörtern in Sätzen sowie das Segmentieren von Silben in Wörtern (Schneider 2017). Für diese Arbeit ist die Ebene der Silben hervorzuheben, weshalb im Folgenden hierauf der Fokus gelegt wird. Die Silbe ist eine grundlegende Einheit der gesprochenen Sprache und „leichter zugänglich als Einheiten wie Laute, Wörter oder Sätze“ (Bredel et al. 2017, S. 30). Demnach unterstützt die Fähigkeit, Silben zu identifizieren und zu erkennen, das Verständnis für die Lautstruktur von Wörtern und hilft bei dem Erlernen der phonologischen Prinzipien, die für den Schriftspracherwerb notwendig sind. Im Gegensatz zu Phonemen ist die Silbe eine größere lautliche Einheit, „die in der Regel aus mehreren lautlichen Segmenten besteht [...]“ und „die kleinste Einheit, die tatsächlich isoliert ausgesprochen werden kann“ (Mehlem 2024, S. 41). Vor allem bei dem Erlernen der Orthografie ist die Silbe relevant, da sie einer natürlichen Unterteilung von Wörtern dient und das Erkennen von Mustern in der Rechtschreibung, wie beispielsweise die korrekte Verwendung von Vokalen und Konsonanten in verschiedenen Silbenstrukturen, unterstützt.

Die phonologische Bewusstheit im engeren Sinne stellt die Voraussetzung für das Erlernen der Phonem-Graphem-Korrespondenz dar (Bosse 2015), welche im folgenden Kapitel zu der orthografischen Kompetenz näher erläutert wird. An dieser Stelle gilt es jedoch bereits festzuhalten, dass die phonologische Bewusstheit im engeren Sinne die Lernenden dazu befähigt, die einzelnen Phoneme in Wörtern klar zu identifizieren. Es wird demnach die notwendige Wahrnehmung und Unterscheidung von Phonemen gefördert, die zur korrekten Anwendung der Phonem-Graphem-Korrespondenz beim Schriftspracherwerb unerlässlich ist.

Beide Faktoren sind folglich entscheidend für den Schriftspracherwerb. Die phonologische Bewusstheit im weiten Sinne bildet jedoch die Grundlage für die im engeren Sinne, wenngleich an dieser Stelle betont werden muss, dass es sich nicht um strikt nacheinander ablaufende Prozesse handelt, sondern um solche, die ineinander übergehen (Schneider 2017). Während sich die phonologische Bewusstheit im weiten Sinne bereits vor Schuleintritt entwickelt, entwickelt sich die phonologische Bewusstheit im engeren Sinne meist erst durch den aktiven Umgang mit der Schriftsprache (Reber 2017). Die erste Klasse ist hierbei besonders relevant,

wobei insbesondere die ersten Wochen hervorgehoben werden müssen, denn in dieser Zeit entwickelt sie sich aufgrund der aktiven Auseinandersetzung mit der Schriftsprache schnell (Marx 2007).

Dass eine starke Korrelation zwischen der phonologischen Bewusstheit und dem Schriftspracherwerb besteht, wurde vielfach belegt. Auch Schröder-Lenzen (2004) stellt die Prognosekraft der phonologischen Bewusstheit für den Schriftspracherwerb heraus. Diese Aussage unterstreicht die Untersuchung von Schneider und Näslund (1993), die unter anderem den Einfluss unterschiedlicher Faktoren wie beispielsweise des Arbeitsgedächtnisses und des phonologischen Bewusstseins auf den Schriftspracherwerb in den Blick nahmen. Bei den Proband:innen handelte es sich zunächst um Kindergartenkinder, wobei diese nachfolgend nach den ersten zwei Grundschuljahren in den Bereichen der Wortdekodierung, des Leseverstehens sowie der orthografischen Kompetenzen getestet wurden (Schneider & Näslund 1993). Obgleich die Ergebnisse zeigen, dass alle getesteten Faktoren einen signifikanten Einfluss auf den Schriftspracherwerb hatten, konnte das phonologische Bewusstsein als stärkste Einflusskategorie bestimmt werden (Schneider & Näslund 1993).

Ähnliche Ergebnisse liefert die Studie von Hogan et al. (2005), welche ebenfalls eine Längsschnittstudie mit Kindergartenkindern durchführten, die in der zweiten und vierten Klasse erneut getestet wurden. Die phonologische Bewusstheit wurde bei allen Messzeitpunkten in den Blick genommen, wobei im Kindergarten zusätzlich die Identifikation von Buchstaben und in der zweiten und vierten Klasse explizit das phonetische Dekodieren sowie das Lesen von Wörtern getestet wurde (Hogan et al. 2005). Wie Schneider und Näslund (1993) haben Hogan et al. (2005) unter anderem herausgestellt, dass die phonologische Bewusstheit ein Prädiktor für die Kompetenz im Schriftspracherwerb in der zweiten Klasse ist. Zudem wurde eine wechselseitige Beziehung zwischen der phonologischen Bewusstheit und dem Lesen von Wörtern festgestellt, was die Annahme bestätigt, dass es sich bei dem Schriftspracherwerb nicht um klar abgrenzbare und sequenziell nacheinander ablaufende Bereiche handeln kann (Hogan et al. 2005).

Milankov et al. (2021) sind ebenfalls der Frage nachgegangen, inwiefern die phonologische Bewusstheit mit der Entwicklung im Kontext des Schriftspracherwerbs zusammenhängt. Die Stichprobe umfasste Grundschüler:innen von der ersten bis zur dritten Klasse. Das Ergebnis bestand darin, dass die phonologische Bewusstheit als eine der Hauptursachen für spätere Leseschwierigkeiten definiert werden konnte (Milankov et al. 2021). Grundlage für diese Annahme fand sich darin, dass Kinder mit Leseschwierigkeiten bei Aufgaben, die sich mit der phonologischen Bewusstheit beschäftigen, häufig schlechter abschnitten als Kinder ohne Leseschwierigkeiten (Milankov et al. 2021).

Zusammenfassend lässt sich feststellen, dass die phonologische Bewusstheit eine zentrale Rolle im Schriftspracherwerb einnimmt. Sie bildet nicht nur die Grundlage für die Fähigkeit, die Lautstruktur der Sprache zu erkennen, sondern stellt auch eine essenzielle Voraussetzung für den Erwerb der Phonem-Graphem-Korrespondenz dar. Die zahlreichen empirischen Studien unterstreichen die enge Verknüpfung zwischen der phonologischen Bewusstheit und dem Schriftspracherwerb.

2.1.3 Orthografische Kompetenz

Die Rolle des Schreibens, insbesondere die des orthografisch korrekten Schreibens, bleibt in dem öffentlichen Diskurs im Kontext des Schriftspracherwerbs oft unberücksichtigt, wobei „das deutsche Schriftsystem [...] am Leser orientiert ist“ (Bredel et al. 2017, S. 3). Das verdeutlicht den direkten Zusammenhang. Dornheim (2008) führte eine Längsschnittstudie durch, in der sie Kinder im Kindergarten bis zum Ende der zweiten Klassen begleitet und an vier Messzeitpunkten getestet hat. Im Zuge ihrer Hauptforschungsfrage nach Prädiktoren für die Rechenleistung, geht sie unterschiedlichen Faktoren auf den Grund. Hierbei betrachtete sie unter anderem die schriftsprachlichen Leistungen, wobei ein korrelativer Zusammenhang zwischen den Lese- und Rechtschreibkompetenzen herausgestellt werden konnte (Dornheim 2008).

Vettori et al. (2023) sind der Frage nach dem Zusammenhang zwischen der lexikalischen Kompetenz und der Lese- und Rechtschreibkompetenz nachgegangen. An der Studie nahmen Kinder der zweiten bis fünften Klasse teil, welche eine Vokalaufgabe, um die lexikalischen Fähigkeiten zu testen, eine Textleseaufgabe für die Erfassung von Lesegenauigkeit und -geschwindigkeit sowie ein Textdiktat umfasste, in dem die Orthografie im Fokus stand (Vettori et al. 2023). Es konnte herausgestellt werden, dass es starke Zusammenhänge zwischen allen Bereichen gibt, wobei die orthografische Kompetenz eine Schlüsselrolle einnimmt (Vettori et al. 2023).

Retelsdorf und Köller (2014) stellen ebenfalls die Frage nach wechselseitigen Effekten zwischen der orthografischen Kompetenz und der Lesekompetenz, wobei sie das Leseverstehen in den Mittelpunkt ihrer Studie stellen. Die Studie basiert auf einer Stichprobe von Schüler:innen der Sekundarstufe I (Retelsdorf & Köller 2014). Die Annahme einer Interdependenz konnte verifiziert werden. Während Vettori et al. (2023) die orthografischen Kompetenzen als einen zentralen Prozess definiert haben, um den herum die Lesekompetenz und die lexikalischen Kompetenzen aufgebaut werden, stellen Retelsdorf und Köller (2014) allerdings den Einfluss des Leseverstehens auf die orthografischen Kompetenzen als größer heraus. Unzureichende orthografische Kompetenzen können sich folglich sowohl auf das Schreiben als auch auf das Lesen auswirken und umgekehrt. Darüber hinaus kann die Orthografie ein sinnvoller, zusätzlicher Anhaltspunkt dafür sein, wie Kinder auf die Schriftsprache blicken. Nach Schröder-

Lenzen (2004) gelten Rechtschreibfehler „als entwicklungsnotwendig und das eigenaktive Üben einer phasentypischen Zugriffsweise auf Schrift als unhintergebar“ (S. 39).

Allerdings fangen Kinder nicht von Beginn an orthografisch korrekt zu schreiben. Der Weg dorthin ist langandauernd und komplex. In diesem Kontext ist die *Phonem-Graphem-Korrespondenz* als eine Schlüsselkompetenz für den Schriftspracherwerb zu nennen, welche eine der ersten Grundlagen für eine orthografische Kompetenz darstellt. Die Phonem-Graphem-Korrespondenz beschreibt die systematische Zuordnung zwischen den Phonemen und deren schriftlicher Darstellung durch Buchstaben, genauer gesagt durch Grapheme (Schröder-Lenzen 2004). Grapheme sind nicht mit Buchstaben gleichzusetzen. Marx (2007) konkretisiert dies: „Ein Graphem ist der mit einem Phonem korrespondierende Buchstabe bzw. die korrespondierende Buchstabengruppe. Grapheme können also auch mehrere Buchstaben enthalten“ (S. 23). Das bedeutet, dass „ein Laut [...] durch verschiedene Buchstaben und Buchstabenkombinationen orthografisch korrekt geschrieben werden kann“ (Schröder-Lenzen 2004, S. 50). Eine Einsicht in diesen Zusammenhang und deren Umsetzung ist demnach eine unabdingbare Schwelle für den Schriftspracherwerb und zeichnet sich durch eine große Komplexität aus. Obwohl es regelmäßige Beziehungen auf der Grundlage einer intuitiven Phonem-Graphem-Korrespondenz gibt, wie etwa bei dem Wort <Oma> /oma/, besteht eine deutlich größere Anzahl an Phonemen im Deutschen, verglichen mit der Zahl an verfügbaren Graphemen, sodass „Rechtschreibprobleme [] sich daraus [ergeben], dass verschiedene Buchstaben ein und denselben Laut und verschiedene Laute ein und denselben Buchstaben darstellen können“ (Schneider 2017, S. 26). So besteht auf der einen Seite eine komplexe Phonem-Graphem-Korrespondenz, beispielsweise bei dem Phonem [f], welches durch die Grapheme <f>, <v> und <ph> realisiert werden kann. Auf der anderen Seite gibt es auch Grapheme, welche durch mehrere Phoneme artikuliert werden können, wie beispielsweise das Graphem <s>. In dem Wort *Sonne* wird es als [z] und in dem Wort *Haus* als [s] artikuliert. Dieses linguistische Phänomen bezeichnet „die Tatsache [], dass ein Buchstabe bzw. eine Buchstabenkombination je nach Wortkontext ganz unterschiedlich ausgesprochen wird“ (Schröder-Lenzen 2004, S. 53). Die daraus entstehende Bedeutung der orthografischen Kompetenz wird an der folgenden Ausführung von Bredel et al. (2017) deutlich:

Wir hören nur wenige der Wortzwischenräume, man kann keine zwei [t] in *Matte* hören, wir hören weder ein [e]/[ə] noch einen r-Laut in *Männer*, großgeschriebene Buchstaben ‚klingen‘ nicht anders als kleingeschriebene und wir hören keinen Unterschied zwischen *dass* und *das*. Die Reihe ließe sich fast beliebig fortsetzen. Die Schrift ist keine Abbildung der Lautung; sie ist vielmehr eine Abbildung von Grammatik. (S. 22)

Es wird deutlich, dass das lautgetreue Schreiben eine Grundlage für viele deutsche Wörter ist, allerdings schnell an die Grenzen kommt und für eine korrekte Orthografie nicht ausreichend

ist. Orthografie als Regelwerk zur normgerechten Schreibung geht über diese lautgetreue Schreibweise hinaus und setzt Aspekte wie beispielsweise das Groß- und Kleinschreiben oder das Dehnungszeichen in den Mittelpunkt der Betrachtung. In diesem Kontext ist auch die Silbenstruktur von erneut großer Relevanz (Schröder-Lenzen 2004). Bei Wörtern, die eine Konsonantendopplung aufweisen, dient diese der Kennzeichnung einer kurzen betonten Vokalsilbe, wie beispielsweise in dem Wort *rennen* (del Toro & Thomé 2021). Dieses orthografische Prinzip zeigt, dass die Rechtschreibung nicht nur phonetischen Gesetzmäßigkeiten folgt, sondern auch durch morphologische und silbische Strukturen geprägt ist.

Seymour et al. (2003) sind im Kontext des *COST A4*, eines Projektes der Europäischen Union, der Frage nach dem Zusammenhang zwischen der Komplexität des sprachspezifischen schriftsprachlichen Systems und schriftsprachlichen Kompetenzen von Schüler:innen nachgegangen. Es wurden die Buchstabenkenntnisse, das Lesen hochfrequenter Wörter sowie von Pseudowörtern bei einer Stichprobe von Kindern der ersten, zweiten und dritten Klasse getestet (Seymour et al. 2003). Die Studie zeigte, dass die Regularität einer Orthografie einen entscheidenden Einfluss auf die Entwicklung der Lesefähigkeit hat. Kinder, die das Lesen in Sprachen mit unregelmäßigen Schreibsystemen wie Englisch, Dänisch oder Französisch erlernten, hatten am Ende der ersten Klasse deutliche Schwierigkeiten: Sie lasen Wörter weniger sicher und konnten Pseudowörter nur selten korrekt erkennen (Seymour et al. 2003). Im Gegensatz dazu bewältigten Schüler:innen, die eine regelmäßige Orthografie lernten, das Lesen von Wörtern und Pseudowörtern nahezu fehlerfrei (Seymour et al. 2003). Zu den Sprachen mit einer transparenten Graphem-Phonem-Korrespondenz zählten beispielsweise das Finnische, das Griechische und das Italienische. Die deutsche Orthografie befand sich im Vergleich mit den europäischen Sprachsystemen in einem Mittelfeld, wobei nahezu jedes schriftsprachliche System Irregularitäten aufweist (Seymour et al. 2003). Die deutsche Sprache wurde als orthografisch regelmäßig herausgestellt, da die Graphem-Phonem-Korrespondenz im Vergleich zu tiefen Orthografien wie im Englischen konsistenter ist. Allerdings wurde die hohe Komplexität der Silbenstruktur in der deutschen Sprache hervorgehoben, welche den Leseerwerb beeinflusst, da diese den Dekodierungsprozess erschwert (Seymour et al. 2003). Wenngleich demnach festgehalten werden kann, dass orthografische Kompetenzen für die Entwicklung von schriftsprachlichen Kompetenzen von Bedeutung sind, zeigt dies auch, dass Wörter innerhalb des Deutschen unterschiedlich komplex sein können, beispielsweise aufgrund der silbischen oder phonologischen Struktur (Seymour et al. 2003).

Die Phonem-Graphem-Korrespondenz kann als Basis für die orthografische Kompetenz gesehen werden und die lauttreue Schreibweise als einen ersten Schritt. Schneider (2017) betont in dem Zusammenhang, dass „der Erwerb orthografischen Wissens [...] schneller [verläuft], wenn Kinder dazu imstande sind, die Buchstabenfolgen in Wörtern simultan zu erfassen, sich

die visuelle Aufmerksamkeit also gleichzeitig auf alle Teilkomponenten eines gegebenen Wortes erstreckt“ (S. 76). Der Übergang von der lautgetreuen zur orthografischen Schreibung vollzieht sich auf der Basis von regelbasiertem Lernen und einer stetig steigenden Verinnerlichung und Automatisierung der Phonem-Graphem-Korrespondenz.

Im Laufe des Schriftspracherwerbs werden immer mehr Wörter inklusive deren Bedeutung, Aussprache und grammatische Eigenschaften gespeichert. Sobald diese Faktoren für ein Wort eine Rolle spielen, führt dies zu einer Aktivierung des *mental Lexikons*, sodass diese Informationen schneller abgerufen werden können. Im nachfolgenden Kapitel wird die Bedeutung des mentalen Lexikons näher beleuchtet, dennoch muss an dieser Stelle bereits festgehalten werden, dass dieses ebenfalls eine Rolle bei der Orthografie spielt. Auch wenn das Verständnis der Phonem-Graphem-Korrespondenz sowie das Erlernen orthografischer Regelmäßigkeiten im Vordergrund stehen, werden im mentalen Lexikon unter anderem orthografische Informationen gespeichert. Marx (2007) beschreibt diesen gesamten Vorgang wie folgt:

Für den Schreibanfänger ist es als nächstes erforderlich, die zu den Lauten passenden Buchstaben zu finden, ihre Form zu erinnern und auf Papier zu bringen [...]. Für Wörter, die nicht lautgetreu verschriftet werden, müssen zudem Rechtschreibregeln beherrscht werden. Auch das Rechtschreiben fällt natürlich leichter, wenn sich der Sichtwortschatz der Kinder vergrößert und demzufolge die Schreibweise von immer mehr Wörtern oder Wortteilen abgerufen werden kann. (S. 22)

Das bedeutet, dass ein phonologischer Zugriff im Sinne der Phonem-Graphem-Korrespondenz zu einer lautorientierten Verschriftlichung führt, während der lexikalische Zugriff durch den direkten Abruf eines Schreibschemas aus einem inneren orthografischen Lexikon zu einer orthografisch richtigen Schreibweise führt (Schröder-Lenzen 2004).

Zusammenfassend kann gesagt werden, dass die Orthografie Erkenntnisse über den Schriftspracherwerb liefern kann. Bredel et al. (2017) präzisieren dies, indem sie herausstellen, dass „auch der Leseerwerb davon profitiert, wenn den Kindern Gelegenheit gegeben wird, das System der Schrift zu entdecken und zu nutzen“ (S. 3), was erneut die Wechselhaftigkeit des Lesens und Schreibens herausstellt.

2.1.4 Semantik

Nach der Wortebene folgt die Satz- und anschließend die Textebene, „denn wenn die Hürde des Wortlesens genommen ist, wartet eine neue Herausforderung auf sie: das Satzlesen, also die Verknüpfung von Wörtern zu Wortgruppen und Sätzen“ (Bredel et al. 2017, S. 152). Insbesondere bei der Satz- und Textebene kommt eine bedeutsame Komponente dazu: das semantische Bewusstsein. Das semantische Bewusstsein bezeichnet die Fähigkeit, sich aktiv mit den Bedeutungen von Wörtern, Sätzen und Texten auseinanderzusetzen (Gutzmann 2020). Es ist

ein Teilbereich des metalinguistischen Bewusstseins und bezieht sich darauf, die semantischen Aspekte der Sprache bewusst wahrzunehmen, zu analysieren und reflektiert zu nutzen (Gutzmann 2000).

Dies bezeichnet unter anderem den Prozess, in dem „Mikrostrukturen [...] durch die semantische Verarbeitung von Wortgruppen zu Bedeutungseinheiten“ (Scheerer-Neumann 2023, S. 105) entstehen. Um die Bedeutung von Wörtern zu entschlüsseln und diese in Beziehung setzen zu können, müssen folglich Wörter oder Wortgruppen miteinander verbunden werden, die im Kontext eines Satzes spezifische Bedeutungen entwickeln. Das Verständnis von Sätzen geht über das Erkennen der Wörter hinaus und baut auf dem Herstellen von Beziehungen zwischen den einzelnen Satzbausteinen auf (Marx 2007). Hierbei werden unter anderem grammatische Kompetenzen benötigt, um diese Informationen miteinfließen zu lassen. Werden die erforderlichen Kompetenzen auf der Textebene betrachtet, so werden das Herstellen von satzübergreifenden Zusammenhängen und deren Bedeutungen relevant (Marx 2007). Mit zunehmender Komplexität wird das Arbeitsgedächtnis mehr beansprucht, da bereits Gelesenes behalten werden muss.

Die Rolle der Semantik ist für den Schriftspracherwerb nicht zu unterschätzen. Um einen Satz oder einen Text verstehen zu können, ist „ein bloßes Verbinden der Laute nicht ausreichend“ (Marx 2007, S. 20). Marx (2007) stellt in diesem Zusammenhang die Bedeutung des Verstehens für die akustische Artikulation von Wörtern heraus, da „die richtige Aussprache oft erst nach Erkennen der Bedeutung möglich ist“ (S. 20). Aufgrund linguistischer Phänomene wie der bereits dargestellten Phonem-Graphem-Korrespondenz, klingen Laute im Kontext eines Wortes anders als bei einer isolierten Aussprache. An dieser Stelle wird das mentale Lexikon relevant. Dieses ist entscheidend für die Identifikation und das Verstehen von Wörtern während des Lesens, denn es ermöglicht gelesene Wörter mit gespeicherten Bedeutungen, Kontextinformationen und semantischen Netzwerken zu verbinden (Harden 2023). Das mentale Lexikon kann als ein zentraler Speicher im Gehirn angesehen werden, welcher das Wissen über Wörter und deren Eigenschaften umfasst. Wenn Kinder lernen, jedem Buchstaben einen Lautwert zuzuordnen, diesen zu artikulieren und Buchstabenfolgen zu synthetisieren, resultieren daraus zunächst Wortvorformen, die im nächsten Schritt mit einem Eintrag im mentalen Lexikon der Lesenden abgeglichen und verknüpft werden müssen (Bredel et al. 2017). Wenn ein Wort gehört, gelesen oder gedacht wird, führt dies zu einer Aktivierung im mentalen Lexikon. Diese Aktivierung umfasst nicht nur die Laut- oder Schriftform des Wortes, sondern auch die semantischen Informationen, wie die Bedeutung, die möglichen Kontexte der Verwendung und die Verknüpfung mit verwandten Konzepten. Harden (2023) stellt heraus, „dass sowohl beim Sprechen als auch beim Hören immer mehrere Kandidaten, aus denen dann der jeweils passendste ausgewählt wird, in irgendeiner Form aktiviert werden“ (S. 26). Diese Bestimmung

basiert auf drei Aspekten: „Wahrnehmung (was hat man eigentlich gehört), Häufigkeit (je häufiger, desto wahrscheinlicher) und Kontext (bestimmte Elemente sind in bestimmten Kontexten entweder völlig ausgeschlossen oder doch sehr unwahrscheinlich)“ (Harden 2023, S. 25f.). Darüber hinaus unterstützt das mentale Lexikon den Prozess der Bedeutungserschließung bei unbekannten Wörtern. Dies geschieht durch kontextuelle Hinweise, semantische Assoziationen oder morphologische Zerlegung (Harden 2023). Somit fungiert das mentale Lexikon als eine dynamische Schnittstelle, die es ermöglicht, semantische Informationen effizient zu nutzen und Sprache als bedeutungstragendes System zu verstehen. Gute Lesende zeichnen sich aufgrund dessen vor allem durch die Fähigkeit aus, Wörter und deren Bedeutung rasch und automatisiert aus dem mentalen Lexikon abzurufen (Schneider 2017).

Im Vergleich zu der bereits dargestellten Studienlage zum phonologischen Bewusstsein, gibt es nur wenige Studien, welche sich mit dem semantischen Bewusstsein im Kontext des Schriftspracherwerbs auseinandersetzen. Eine von diesen ist die Längsschnittstudie von Frost et al. (2005), welche die Beziehung zwischen vorschulischen semantischen Fähigkeiten wie Wortschatz, Verständnis und Satzbau, dem phonologischen Bewusstsein sowie das Leseverstehen untersuchten. Die Stichprobe bestand aus Kindern und Jugendlichen, welche in regelmäßigen Abständen im Alter von drei bis 16 getestet wurden (Frost et al. 2005). Die Ergebnisse verdeutlichten, dass sowohl semantische Faktoren und das Interesse an Büchern im Alter von drei Jahren als auch phonologische Faktoren im Alter von sechs Jahren signifikante Prädiktoren für die Leseentwicklung im Alter von 16 Jahren waren (Frost et al. 2005). Zudem zeigte sich ein wechselnder Haupteffekt zwischen semantischen und phonologischen Variablen auf die Leseentwicklung. Während die phonologische Bewusstheit im Alter von sechs Jahren den stärksten Einfluss auf die Lesefähigkeiten zu Beginn der zweiten Klasse hatte, dominierten bei späteren Messzeitpunkten die semantischen Variablen als zentraler Einflussfaktor (Frost et al. 2005).

Die Relevanz des semantischen Bewusstseins für die Lesekompetenz wird auch durch die Langzeitstudie von Browne Rego und Bryant (1993) deutlich. Die Untersuchung zeigt, dass syntaktische und semantische Fähigkeiten Kindern helfen, sowohl mehrsilbige als auch fremdsprachige Wörter zu verstehen und zu lesen. Die Studie begleitete fünf- bis sechsjährige Kinder über ihr erstes Schuljahr. Im ersten Abschnitt wurden drei Tests durchgeführt, um das syntaktische und semantische Bewusstsein der Kinder zu bewerten (Browne Rego & Bryant 1993). Fünf Monate später, im zweiten Abschnitt, wurde ein Lesetest durchgeführt, bei dem die Kinder zunächst isolierte Wörter und anschließend Wörter, eingebettet in einen Kontext, lasen. Die Ergebnisse zeigten, dass Kinder mit höheren Punktzahlen bei den syntaktischen und semantischen Aufgaben des ersten Teils später deutlich mehr kontextgebundene Wörter lesen konnten als Kinder mit niedrigeren Punktzahlen (Browne Rego & Bryant 1993). Die

Autoren schlossen daraus, dass Kinder, die ein stark ausgeprägtes Bewusstsein für semantische und syntaktische Kategorien besitzen, diese effektiv als Orientierungshilfen beim Lesen nutzen können (Browne Rego & Bryant 1993).

Wird von Leseverständnis gesprochen, wird zwischen dem *lokalen Verständnis* und dem *globalen Verständnis* unterschieden. Das lokale Verständnis legt den Fokus auf das Entschlüsseln und das inhaltliche Verknüpfen der Bedeutung einzelner Wörter, Satzteile oder Sätze (Rosebrock & Nix 2015). Bredel et al. (2017) stellen in diesem Kontext die Bedeutung der Interpunktion in den Mittelpunkt der Betrachtung, da grammatische Faktoren hierbei eine ausschlaggebende Rolle einnehmen. Im Zuge des Erkennens der Bedeutung eines Satzes müssen Wortgruppen zu sinnvollen Einheiten verknüpft und direkte Zusammenhänge verstanden werden. Die Entwicklung der Kompetenz, Wörter und Sätze sinnvoll zu dekodieren und grammatische sowie semantische Strukturen zu erkennen, hat eine immense Relevanz für den Schriftspracherwerb (Marx 2007). Bei dem globalen Verständnis wird das Erfassen der „inhaltlichen Gesamtvorstellung des Textes, die aus dem Zusammenschließen von erfahrungsba- sierten Vermutungen zum Thema (»top down«) und ersten Verarbeitungsschritten der Semantik des Textes (»bottom up«) entsteht“ (Rosebrock & Nix 2015, S. 18), relevant. Es lässt sich demnach als ein Zusammenspiel charakterisieren, indem top-down-Prozesse helfen, den Text im größeren Kontext zu interpretieren, während bottom-up-Prozesse sicherstellen, dass alle Details des Textes korrekt aufgenommen werden. Dies erfordert das Verknüpfen von Informationen aus verschiedenen Teilen eines Textes, um ein kohärentes Gesamtbild zu erstellen. Schneider (2017) betont, „dass der Leser Texte nicht passiv verarbeitet, sondern für sich selbst aktiv Bedeutungen konstruiert“ (S. 77). Diese Aussage wird durch Scheerer-Neumann (2023) weitergeführt, indem daraus gefolgert wird, dass unterschiedliche Leser:innen zu unterschiedlichen Bedeutungen kommen können. Die aktive Auseinandersetzung mit dem Text, einschließlich der Verbindung zu dem individuellen Vor- und Weltwissen, ist hierfür entscheidend (Marx 2007). Das lokale und das globale Verständnis hängen demnach eng zusammen, wobei das lokale Verständnis als Voraussetzung für das globale Verständnis interpretiert werden kann. Globale Zusammenhänge können nur dann erkannt werden, wenn die Bedeutung einzelner Sätze verstanden wird, weshalb es sich um sehr komplexe Vorgänge handelt. Für Kinder der ersten Klasse steht im Kontext des Schriftspracherwerbs das lokale Verständnis im Vordergrund, während das globale Verständnis mit zunehmender schriftsprachlicher Kompetenz gestärkt wird (Rosebrock & Nix 2015).

2.2 Erfassung des Schriftspracherwerbs in der ersten Klasse

Aus den vorherigen Ausführungen wird deutlich, dass es sich bei dem Schriftspracherwerb um einen sehr komplexen Prozess handelt, welcher aus vielen Teilprozessen besteht. Wenn Kinder lesen und schreiben lernen, müssen sie all diese Schritte erlernen (Bredel et al. 2017).

Das folgende Kapitel setzt sich mit der Frage auseinander, warum es von zentraler Bedeutung ist, den Schriftspracherwerb in der ersten Klasse zu erfassen. Des Weiteren werden unterschiedliche Vorgehensweisen für die Erfassung erklärt.

Grundlegend lässt sich festhalten, dass die Schule für den Schriftspracherwerb eine immense Bedeutung hat. Bredel et al. (2017) stellen heraus, dass wenngleich in diesem Kontext häufig der eigenaktive Anteil des Gesamtprozesses jedes Individuums in den Mittelpunkt der Betrachtung gerückt wird, der Einfluss von äußeren Faktoren nicht zu unterschätzen ist. Folglich ist die Schule die Institution, welche sich auf der einen Seite aktiv zum Ziel setzt, an dem Schriftspracherwerb teilzuhaben (QUA-LiS NRW 2024b) und auf der anderen Seite einen großen Teil des alltäglichen Lebens der Schüler:innen einnimmt. Darüber hinaus kommen die Kinder mit den unterschiedlichsten Vorerfahrungen in die Schule. Insbesondere bei den Kindern, die aus einem für den Schriftspracherwerb weniger förderlichen Umfeld kommen, ist die Grundschule von noch größerer Bedeutung (Bredel et al. 2017). In diesem Zusammenhang kann sich zu Nutzen gemacht werden, dass „die Bereitschaft, sich mit den Mustern der Schrift auseinanderzusetzen, sich also ein Wissensgerüst aufzubauen, das die unabdingbare Basis für alle weiteren Lernprozesse darstellt, [] nie größer [ist] als in dieser Zeit“ (Bredel et al. 2017, S. 2). Die Kinder haben zu Schulbeginn große Kompetenzen im mündlichen Sprachgebrauch und verschiedenste Arten impliziten Wissens über Sprache auf unterschiedlichen Ebenen erworben (Bredel et al. 2017). Mit der Beschulung wird das explizit sprachliche Wissen vertieft. Es lässt sich demnach festhalten, dass die ersten Schuljahre „wesentliche Weichenstellungen für den Erwerb des Lesens und Schreibens und damit auch für den Schulerfolg“ (Schneider 2017, S. 73) beinhalten.

Daraus lassen sich bestimmte Ansprüche an die Kompetenzen der Lehrkräfte ableiten. Diese nehmen im Bildungsprozess eine zentrale Rolle ein. An sie wird die Anforderung gestellt, verschiedene Verfahren zur Erfassung des Schriftspracherwerbs angemessen sowohl anwenden als auch die Ergebnisse interpretieren zu können. Die KMK (2022b) hat in den Standards für die Ausbildung von Lehrkräften verschiedene Kompetenzbereiche definiert. Eine von diesen nennt sich *Beurteilen*, in dem es sowohl um das Erfassen der Lernvoraussetzungen der Schüler:innen als auch um die daraus resultierende Wahl der individuellen Förderung inklusive der Beratung von den Lernenden und deren Erziehungsberechtigten geht. Dies zeigt die Bedeutung dieses Themenbereichs für die Qualität schulischer Bildung auf.

Der Grund warum diese Kompetenzen im Lehramt in dieser Form hervorgehoben werden stützt sich schwerpunktmäßig auf die Heterogenität der Schüler:innenschaft. Aufgrund derer ist es relevant, dass Lehrkräfte sich mit den individuellen Lernvoraussetzungen der Schüler:innen auseinandersetzen, um den Unterricht differenziert planen zu können (Helmke 2015). Dadurch können zusätzlich Förderbedarfe erkannt werden. Eine schulische Herausforderung

liegt in dem Anliegen individuelle Unterstützungsbedarfe von Schüler:innen rechtzeitig zu erkennen und dem *wait-to-fail-Prinzip* entgegenzuwirken, welches aussagt, dass Förderungen erst dann begonnen werden, wenn die Wahrnehmungs- und Belastungsgrenze der Lehrkraft zu einem gewissen Maß überschritten wird (Huber et al. 2013). Dies impliziert jedoch eine späte Förderung, was bedeuten kann, dass sich bestimmte Strukturen oder Verhaltensmuster bei den Kindern bereits verfestigt haben. Daraus lässt sich der Wunsch nach einer möglichst frühzeitigen Erkennung eskalierender Lernverläufe sowie nach einer angemessenen Förderung ableiten (Huber et al. 2013). Rosebrock und Nix (2015) betonen, dass „nur Lehrende, die die Leseleistung ihrer Schüler(innen) beobachten und diese Beobachtungen vor dem Horizont eines Lesekompetenzbegriffs bewerten können, [i]n der Lage [sind], das Lesen erfolgreich zu unterstützen“ (S. 13). Auf der einen Seite wird durch diese Ausführung die Relevanz einer differenzierten Ermittlung des individuellen Leistungsstandes in dem Bereich des Schriftspracherwerbs in den Fokus gerückt, während auf der anderen Seite der Aspekt des dafür notwendigen fachfundierten Wissens eingebracht wird. Dies betont die damit einhergehenden und bereits erwähnten Anforderungen an die Kompetenzen der Lehrkräfte.

Neben der dadurch entstehenden besseren Planbarkeit von Lernprozessen, kann auch der Erfolg vorangegangener pädagogischer Bemühungen überprüft werden (Lenhard 2019). Dabei kann eine Erfassung des Schriftspracherwerbs entscheidende Erkenntnisse liefern, um den Unterricht gezielt weiterzuentwickeln (Rosebrock & Nix 2015). Hierbei wird zwischen einer einmaligen Erfassung und einer Erhebung, welche den Lernverlauf erfassen kann, unterschieden. Klauer (2014) differenziert in diesem Zuge zwischen der summativen und formativen Evaluation. Erstere zielt darauf ab, den aktuellen Lernstand zu einem bestimmten Zeitpunkt zu erfassen. Hierbei liegt der Fokus auf dem Ergebnis am Schluss einer Lernperiode (Klauer 2014). Die formative Evaluation hingegen verfolgt das Ziel Veränderungen über einen Zeitraum zu dokumentieren, um Entwicklungen sichtbar zu machen (Klauer 2014).

Lenhard (2019) fasst dies passend zusammen: „Letztlich stellen diese nichts anderes als den Versuch dar, Leistungserhebungen, wie sie jeden Tag in der Schule erfolgen, objektiv und zuverlässig zu gestalten.“ (S. 76). Es wird zwischen standardisierten und informellen Verfahren unterschieden (Lenhard 2019). Die informelle Erfassung wird im Folgenden aufgrund ihrer Irrelevanz für diese Arbeit nicht näher beleuchtet, da es sich hierbei „um die normalen Instrumente der Schulleistungserhebung wie Zensuren aus Schularbeiten, Bewertung von Hefteinträgen und die Beurteilung mündlicher Leistungen im Unterricht“ (Lenhard 2019, S. 73) handelt und beispielsweise Fehleranalyse oder qualitative Beobachtungen beinhaltet. Standardisierte Verfahren zeichnen sich durch drei Kriterien aus: Objektivität, Reliabilität und Validität (Marx 2007). Ein Verfahren ist objektiv, wenn die Ergebnisse unabhängig von der durchführenden Person ausgewertet und interpretiert werden (Bredel et al. 2017). Wenn das Verfahren bei

wiederholter Anwendung unter gleichen Bedingungen zu den gleichen Ergebnissen führt, erfüllt es das Kriterium der Reliabilität (Bredel et al. 2017). Ein Verfahren ist darüber hinaus valide, wenn es tatsächlich das misst, was es zu messen vorgibt (Bredel et al. 2017). Im Kontext des Schriftspracherwerbs fallen hierunter einige Testverfahren, welche sich zum Ziel gesetzt haben, objektive, reliable und valide Aussagen über den Entwicklungsstand der Lese- und Schreibkompetenz zu ermöglichen.

Wenngleich diese Instrumente weiterhin vor allem in der Schulpraxis eingesetzt werden, steigt die Kritik an der Methodik fortlaufend. So wird bemängelt, dass die Tests teilweise auf veralteten Normen basieren und zudem auf nicht repräsentativen Stichproben bauen (Lenhard 2019). Des Weiteren weisen einige Testverfahren eine zu lange Testdauer auf und zeichnen sich durch eine niedrige Reliabilität aus (Lenhard 2019). Es können auch Deckeneffekte auftreten, die dann entstehen, wenn viele Testpersonen die maximale Punktzahl erreichen. Diese mindern die Aussagekraft des Tests, da er keine ausreichende Differenzierung im oberen Leistungsbereich erlaubt (Lenhard 2019). Gleichzeitig ist im Gegensatz dazu bei einigen Tests die Komplexität zu hoch, dass Bodeneffekte auftreten können, wenn viele Proband:innen niedrige Punkte erreichen (Marx 2007). Bredel et al. (2017) führen darüber hinaus noch aus, dass der einigen Tests zugrundeliegende Lesebegriff in der modernen Leseforschung nicht mehr vertreten wird. Viele standardisierte Tests legen zudem oft ihren Fokus auf Lesegenauigkeit und -geschwindigkeit, was „die Entwicklung der Lesekompetenz jedoch noch nicht hinreichend“ (Schneider 2017, S. 22) beschreibt. Es wird deutlich, dass die genannten Kritikpunkte die Aussagekraft diverser Testinstrumente erheblich einschränkt.

3. Entwicklung der Fragestellung und der Hypothesen

Empirisch werden im Kontext der Frage des Schriftspracherwerbs häufig nur die Klassenstufen ab dem dritten Schuljahr in den Blick genommen (z.B. Vettori et al. 2023; Retelsdorf & Köller 2014). Das führt dazu, dass die erste Klasse oft unberücksichtigt bleibt. Selbst die großen internationalen Schulleistungsstudien, die sich mit der Lesekompetenz beschäftigen, konzentrieren sich auf ältere Schüler:innen. Während die PISA-Studie die Kompetenzen von 15-Jährigen untersucht (OECD 2024), nimmt die IGLU-Studie die Lesekompetenz von Schüler:innen der vierten Klasse in den Blick (McElvany et al. 2023). Gerade in der Grundschule zeigt sich jedoch eine erhebliche Leistungsfluktuation, da die Entwicklung der Schriftsprache zwischen der ersten und vierten Klasse besonders dynamisch und schnell verläuft. Dies wird beispielsweise in der Studie von Byrd et al. (2024) deutlich, in der Erst- und Zweitklässler:innen Schreibaufgaben absolvieren sollten. Auch wenn sie den Fokus ihrer Forschungsfrage auf das Identifizieren von Dialektmerkmalen legten, zeigten die Ergebnisse den bereits vorhandenen großen Unterschied in der schriftsprachlichen Kompetenz zwischen diesen beiden Klassenstufen auf (Byrd et al. 2024). Die Teilnehmenden der zweiten Klasse erreichten erheblich

bessere Ergebnisse im Vergleich zu den Schüler:innen der ersten Klasse (Byrd et al. 2024). Dies lässt die Schlussfolgerung zu, dass die Erkenntnisse über den Schriftspracherwerb in anderen Klassenstufen nicht ohne Weiteres auf die erste Klasse übertragen werden können.

Wenn die erste Klasse in Studien Einhalt findet, dann lediglich als Teil einer Längsschnittstudie, in dessen Zuge der Zeitraum der ersten Klasse eine Teilerhebung darstellt, wie auch vielfach in den bereits erläuterten Studien (z.B. Schneider & Näslund 1993; Frost et al. 2005; Dornheim 2008). Die Planung einer Längsschnittstudie als Basis einer Untersuchung des Schriftspracherwerbs ist grundsätzlich als sinnvoll zu bewerten, wenn das Ziel verfolgt wird, Entwicklungsverläufe aufzuzeigen und Korrelationen zu identifizieren (Klauer 2014). Allerdings ergibt sich eine Lücke im Forschungsstand, wenn nicht auch die konkreten Ergebnisse aus der ersten Klasse detailliert beleuchtet werden. Ohne eine differenzierte Darstellung des aktuellen Entwicklungsstandes in der Schuleingangsphase bleibt ein zentraler Aspekt des Schriftspracherwerbs an der Oberfläche oder gar unbeachtet.

Diese Herausforderung wird auch bei einer Betrachtung der didaktischen Konzepte augenscheinlich. Wenngleich Kompetenzerwartungen für die Grundschule differenziert werden, indem welche für das Ende der Schuleingangsphase und welche für das Ende der vierten Klasse formuliert werden (QUA-LiS NRW 2024a), bestehen keine Konkretisierungen für das Ende der ersten Klasse.

Vor diesem Hintergrund muss sich die Frage gestellt werden, wie der Schriftspracherwerb einer ersten Klasse erfasst werden soll, woraufhin die Antwort oft in standardisierten Tests gefunden wird. Dass standardisierte Tests jedoch an einigen Stellen Schwächen aufweisen und zunehmend Gegenstand intensiver Kritik sind, wurde bereits vielfach thematisiert (Lenhard 2019; Bredel et al. 2017; Schneider 2017; Marx 2007). Daher scheint eine Überarbeitung oder Neuentwicklung im Einklang mit aktuellen wissenschaftlichen Erkenntnissen erforderlich. Auf dieser Grundlage werden folgende Forschungsfragen und Hypothesen abgeleitet:

Forschungsfrage 1: Erfassen bestehende Lesetests tatsächlich die intendierten Aspekte der Leseleistung?

Hypothese 1: Methodische Schwächen führen dazu, dass bestehende Lesetests nicht die intendierten Aspekte der Lesekompetenz erfassen.

Forschungsfrage 2: Ist der neu entwickelte Lesetest geeignet, um den Schriftspracherwerb zu erfassen?

Hypothese 2: Der entwickelte Lesetest ist methodisch dazu geeignet, die Kompetenzen von Schüler:innen in der ersten Klasse in den Bereichen der phonologischen Bewusstheit, der orthografischen Kompetenz sowie der Semantik zu erfassen.

4. Analyse bestehender Verfahren zur Erfassung des Schriftspracherwerbs

Um den Forschungsfragen nachgehen zu können, werden in diesem Kapitel drei standardisierte Testverfahren in den Mittelpunkt der Betrachtung gerückt, die sich zum Ziel setzen den Schriftspracherwerb zu erfassen. Hierbei wurde sich für die Instrumente *ELFE 1-6* (Lenhard & Schneider 2006a), *BAKO 1-4* (Stock et al. 2003a) und *IEL-1* (Diehl & Hartke 2012a) entschieden, da es sich dabei um in der Praxis vielfach eingesetzte Tests handelt. Die Tests werden jeweils kurz erläutert und anschließend einer kritischen Betrachtung unterzogen.

4.1 Leseverständnistest für Erst- bis Sechstklässler (ELFE 1-6)

Ein in der Schulpraxis häufig eingesetzter Test, welcher den Fokus auf das Leseverständnis legt, ist der *Leseverständnistest für Erst- bis Sechstklässler (ELFE 1-6)* (Lenhard & Schneider 2006a). Hierbei handelt es sich um einen normierten und standardisierten Leseverständnistest für die erste bis sechste Klasse. Obwohl er zu Beginn lediglich für die erste bis vierte Klasse konzipiert wurde, ist er mit Hilfe einer verkürzten Bearbeitungszeit auch in der fünften und sechsten Klasse anwendbar. Laut den Autoren müsse man in diesem Kontext den Test eher als Screening-Verfahren sehen (Lenhard & Schneider 2006b).

Der Test betrachtet basale Lesestrategien sowie die Kompetenzen, die benötigt werden, um Sätze und Texte verstehen zu können (Lenhard & Schneider 2006b). Er besteht aus drei Subtests, wobei es einen weiteren, ausschließlich für die Computerversion erstellten Subtest gibt, welcher sich mit einer Schwellenmessung der Worterkennungsgeschwindigkeit beschäftigt.² Bei dem ersten Subtest handelt es sich um einen Wortverständnistest, in dem die Fähigkeiten zum Dekodieren und Synthetisieren von Lauten in den Mittelpunkt der Betrachtung gerückt werden. Hierfür wird den Proband:innen ein Bild präsentiert, dem vier Wörter nebenstehend zugeordnet sind. Die Kinder sollen das Wort unterstreichen, welches zu dem Bild passt. Bei dem zweiten Subtest geht es schwerpunktmäßig um das Satzverständnis mit besonderem Blick auf syntaktische Fähigkeiten. Die Proband:innen lesen einen Satz, in dem sich an einer Stelle für das richtige Wort aus multiplen Möglichkeiten entschieden werden muss. Der dritte Subtest beschäftigt sich mit dem Textverständnis. Hierbei wird die Kompetenz zum Identifizieren von Informationen, das satzübergreifende Lesen und das schlussfolgernde Denken in den Mittelpunkt der Betrachtung gerückt. Den Proband:innen wird ein unterschiedlich langer Text präsentiert, den sie lesen und anschließend dazugehörige Fragen mit Hilfe von vier Antwortmöglichkeiten beantworten sollen.

Insgesamt beinhaltet der ELFE 1-6 120 Items, wobei an dieser Stelle anzumerken ist, dass die Items nach Schwierigkeit gestaffelt sind und es für die jeweiligen Subtests

² Da dieser Subtest für einen Einsatz in der Papierversion ausgeschlossen ist und demnach auch für die Entwicklung des neuen Tests zur Erfassung des Schriftspracherwerbs irrelevant ist, wird dieser nachfolgend nicht genauer betrachtet.

Zeitbeschränkungen gibt. Damit ist intendiert, dass nicht alle Items bearbeitet werden (Lenhard & Schneider 2006b).

4.1.1 Kritische Diskussion

Wenngleich sich der ELFE 1-6 durch eine große Popularität kennzeichnen lässt, bedarf es bei einigen Aspekten eine kritische Betrachtung. Zunächst kommt die Frage nach der Auswahl der Items auf. Die Testautoren basieren ihre Begründung ausschließlich auf einen zu 98-prozentigen Einsatz des Grundwortschatzes für die Grundschule (Lenhard & Schneider 2006b). Jedoch fallen bereits bei der Betrachtung der Itemauswahl des ersten Subtests einige Items auf, die diese Aussage anzweifeln lassen. Darüber hinaus lässt sich keine Einheitlichkeit finden bezüglich semantischer, phonemischer und/oder graphemischer Unterschiede, worunter die Vergleichbarkeit der Items leidet. Zusätzlich ergibt sich die Frage nach einer klaren Zielsetzung dessen, was konkret getestet werden soll, da die Nutzung von solchen Gemeinsamkeiten oder Unterschieden diverse Kompetenzen der Proband:innen voraussetzt. Dies wird im folgenden Beispiel verdeutlicht:

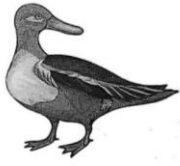
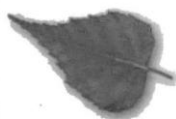
	Auto		Ente		Bein
	Flugzeug		Stein		Bild
	<u>Fahrrad</u>		Blume		Bett
	Pferd		Fisch		Blatt

Abbildung 1: ELFE 1-6 (Lenhard & Schneider 2006a), Subtest 1 (Wortverständnistest), Übungsitems

Die Auswahlmöglichkeiten des ersten Items, welches gleichzeitig als Übungsitem fungiert, lassen sich durch semantische Ähnlichkeiten charakterisieren: Es handelt sich bei jedem Begriff um ein Fortbewegungsmittel. Phonemisch und graphemisch lassen sich keine starken Zusammenhänge feststellen. Bei dem zweiten Item bestehen semantische Ähnlichkeiten zwischen den Begriffen *Ente* und *Fisch*: Es sind Tiere, die im Wasser leben. Es können unterschiedliche Phonemstrukturen sowie -anzahlen beobachtet werden: So besteht beispielsweise die Lautfolge /f/, /i/, /ʃ/ aus vier Phonemen, während /b/, /l/, /u:/, /m/, /ə/ aus fünf Phonemen zusammengesetzt ist. Zusätzlich führen die unterschiedlichen Grapheme auch zu Unterschieden auf der graphemischen Ebene. Im Gegensatz zu diesen beiden Items, unterscheiden sich die Begriffe aus dem dritten Item semantisch stark, wobei sie eine phonemische Ähnlichkeit aufweisen, da alle Wörter mit dem Phonem /b/ beginnen. Der Anfangsbuchstabe lässt auch eine graphemische Ähnlichkeit zu, was zusätzlich dadurch bestärkt wird, dass dreiviertel der Wörter vier Grapheme aufweisen. Zudem ist als richtige Antwort für dieses Item der Begriff *Blatt* bestimmt worden, wobei der Begriff *Bild* auch nicht falsch wäre, da es sich um ein Bild von einem Blatt handelt. Folglich ist dieses Item uneindeutig.

Dass es sich hierbei nicht um einen Einzelfall handelt, zeigen die nächsten Items:

 Gramm Kamm Lamm Kampf	 Blumenvase Blumensträuße Blumenladen Blumenwiese	 Marssonde Maskottchen Sandmännchen Marsmännchen
--	---	--

Abbildung 2: ELFE 1-6 (Lenhard & Schneider 2006a), Subtest 1 (Wortverständnistest), Items 23, 27 und 30

Bei diesem ersten Item sind keine semantischen Ähnlichkeiten vorhanden, allerdings ähneln sie sich phonemisch durch die Phoneme /a/ und /m/. Insbesondere die Begriffe *Gramm*, *Kamm* und *Lamm* lassen sich durch eine ähnliche Phonemstruktur kennzeichnen, sodass sie sich nur im Anlaut unterscheiden. Der Begriff *Kampf* (/k/, /a/, /m/, /pf/) weist aufgrund des Konsonantenclusters und des Frikativs eine komplexe Phonemstruktur auf und unterscheidet sich folglich von den anderen drei Wörtern. Graphemische Parallelen kommen ebenso durch die gleiche Endung <amm> der Begriffe *Gramm*, *Kamm* und *Lamm*. Zusätzlich beginnen die Begriffe *Kampf* und *Kamm* mit der Buchstabenkombination <Ka>. Zudem weisen dreiviertel der Wörter eine gleiche Länge auf. Bei dem zweiten Item bestehen wiederum semantische Ähnlichkeiten durch die Kategorie *Blume*. Durch die Komposita der Begriffe (<Blumen->) bestehen zusätzlich phonemische und graphemische Ähnlichkeiten. Die Endungen unterscheiden sich jedoch sehr stark. Zudem ist hier erneut eine Uneindeutigkeit in der richtigen Antwort zu finden: Das als korrekt vorgegebene Wort ist *Blumenvase*, allerdings ist die Angabe des Begriffs *Blumensträuße* oder *Blumenladen* nicht vollständig falsch. Trotz, dass eine Vase abgebildet ist, kann diese für Blumensträuße genutzt werden beziehungsweise kann diese in einem Blumenladen erworben werden. Besonders in der Arbeit mit Kindern ist es wichtig, auf solche potenziellen Uneindeutigkeiten zu achten, da Kinder häufig noch unterschiedliche oder unerwartete Assoziationen und Perspektiven entwickeln. Bei dem letzten Item gibt es lediglich bei den Begriffen *Marssonde* und *Marsmännchen* semantische Ähnlichkeiten, da sie semantisch durch den Planet Mars verbunden sind. Wird diese Kategorie auf Planeten erweitert, könnte hier auch eine semantische Parallele zu dem Begriff *Sandmännchen* bestehen, da dieser häufig im Kontext mit dem Mond steht. Es bestehen phonemische Ähnlichkeiten bei den Begriffen *Marssonde* und *Marsmännchen* durch die erste Silbe /maks/. Die erste Silbe bei dem Item *Maskottchen* /mas/ ähnelt dieser ersten Silben ebenfalls sehr und unterscheidet sich nur anhand eines Phonems. Die Begriffe *Sandmännchen* und *Marsmännchen* teilen sich darüber hinaus die Endung /mɛnçən/. Graphemische Ähnlichkeiten kommen durch die gleichen Phänomene auf.

Ferner wird lediglich von den Testautor:innen darauf verwiesen, dass die Items nach Schwierigkeit gestaffelt wurden (Lenhard & Schneider 2006b). Auf welchen Annahmen diese

Sortierung stattgefunden hat, ist nicht ersichtlich. Es wird deutlich, dass eine Willkür in der Auswahl der Items nicht nur zwischen den Items, sondern auch innerhalb eines Items vorzufinden ist. Hinzukommend darf nicht unbeachtet bleiben, dass, wie bereits angesprochen, die Auswahl der Wörter lediglich auf den Grundwortschatz für die Grundschule NRW basieren soll. Allerdings reicht eine Betrachtung lediglich der hier beispielhaft aufgeführten Items aus, um festzustellen, dass dies oft nicht der Fall ist. Ganz abgesehen davon wird auf eine Überprüfung der Frequenz der Wörter nicht eingegangen, obwohl die Frequenz in dem wissenschaftlichen Diskurs ein gängiger Faktor ist, um die Bekanntheit der Wörter bei den Kindern vorherzusagen. Harden (2023) stellt heraus, dass hochfrequente Worte einen schnelleren Zugriff im mentalen Lexikon bedeuten. Auch Schröder-Lenzen (2004) beschreibt die statistische Häufigkeit der Wörter als eine von drei für die Unterrichtspraxis entscheidenden Faktoren. Der Begriff *Mars-sonde* zeigt an der Stelle beispielhaft sehr ausdrucksstark, dass dies nicht in die Konzeption mit einbezogen wurde.

Diese dargestellten Uneindeutigkeiten ziehen sich auch durch die weiteren Subtests. Dies wird an dem folgenden Beispiel deutlich:

Ein Auto kann	härter	fahren als ein Fahrrad.
	heller	
	schiefer	
	schneller	
	schärfer	

Abbildung 3: ELFE 1-6 (Schneider & Lenhard 2006a), Subtest 2 (Satzverständnistest), Item 1

Die als richtig bewertete Antwort wäre *schneller*, jedoch kann an dieser Stelle berechtigterweise die Frage gestellt werden, ob ein Auto nicht auch *härter* fährt. Darüber hinaus ist an dieser Stelle anzumerken, dass dieser Subtest weniger das Satzverständnis testet, sondern vorrangig der Frage nachgeht, wo unterschiedliche Konstruktionen auftauchen. So ist der Satz syntaktisch und morpho-syntaktisch immer richtig. Bredel et al. (2017) stellen heraus, dass bei der Proband:innen bei diesem Subtest kein syntaktisches Wissen benötigen und es sich daher eher um einen Wortbedeutungstest handelt.

Auch bei Subtest Drei gibt es einige Aspekte kritisch zu hinterfragen.

Heute scheint den ganzen Tag die Sonne.	
<u>Welcher Satz stimmt?</u>	
☐ Heute ist schönes Wetter.	☐ Morgen wird es regnen.
☐ Gestern war schönes Wetter.	☐ Heute regnet es.

Abbildung 4: ELFE 1-6 (Lenhard & Schneider 2006a), Subtest 3 (Textverständnistest), Übungssitem 1

Das vorliegende Beispiel verdeutlicht eine erneute Uneindeutigkeit bei der richtigen Antwort. Es stellt sich die Frage, warum die Aussage, dass gestern schönes Wetter war oder dass es morgen regnen wird, als falsch betrachtet werden sollte. Auf den ersten Blick erscheint die Antwort eindeutig, da sie sich offensichtlich auf die Aussage im Satz bezieht. Allerdings muss bei der Konstruktion solcher Tests oder Aufgaben mit Kindern stets berücksichtigt werden, dass für sie andere Wege intuitiver sein könnten. Aus diesem Grund ist es von besonderer Bedeutung, klare und eindeutige Antwortmöglichkeiten zu geben. Zusätzlich lässt sich festhalten, dass manche Items vor allem aufgrund der Textlänge zu komplex für die erste Klasse sind und eher ab der dritten Klasse anwendbar sind (Marx 2007). Dadurch können Bodeneffekte entstehen, sodass die tatsächlichen Unterschiede in den Kompetenzen der Proband:innen nicht differenziert werden können (Marx 2007). Dies schmälert die Aussagekraft des Subtests. Zu Subtest Drei lässt sich demnach festhalten, dass es zwar wirklich das Textverständnis testet, das Design aber einige Schwierigkeiten aufweist.

Schlussfolgernd kann gesagt werden, dass sich der ELFE 1-6 durch eine hohe Willkür in vielerlei Richtungen auszeichnet. Es bleibt unklar, wie die Items ausgewählt sowie nach Komplexität gestaffelt wurden. Dadurch entsteht eine geringe Vergleichbarkeit, da Items beispielsweise im Kontext von Semantik und Phonetik unbegründet eingesetzt werden. Zudem sind einige Antwortmöglichkeiten unklar und setzen eine zu hohe Anforderung an eine erste Klasse.

4.2 Basiskompetenzen für Lese- und Rechtschreibleistungen (BAKO 1-4)

Ein weiterer Test ist der *Basiskompetenzen für Lese- und Rechtschreibleistungen (BAKO 1-4)* (Stock et al. 2003a), welcher die phonologische Bewusstheit in den Mittelpunkt der Betrachtung rückt. Es handelt sich dabei um einen standardisierten Test, welcher für die erste bis vierte Klasse konzipiert wurde. Der BAKO 1-4 ist ein Einzeltest und weist eine Länge von 30 Minuten auf. Der Schwerpunkt liegt auf der Erfassung der phonologischen Bewusstheit im engeren Sinne (Stock et al. 2003b). Die Besonderheit dieses Tests ist seine Sprachgebundenheit. Dies bedeutet, dass die Items per CD-ROM vorgegeben werden und die Proband:innen verbal ihre Antwort äußern. Jedes Item wird gesondert abgespielt. Folglich gibt es keinen Testbogen, sondern lediglich einen Protokollbogen für die testleitende Person. Der Test besteht aus sieben Subtests und umfasst insgesamt 74 Items (Stock et al. 2003a). Laut Angaben der Testautor:innen sind die Items innerhalb eines Subtests nach Schwierigkeit gestaffelt (Stock et al. 2003b). In vier der Subtests ist ein Abbruchkriterium vorgesehen, das besagt, den Subtest zu beenden, sobald drei Fehler in Folge auftreten. Mit dieser Hilfe sollen die Proband:innen nicht mit zu schweren Aufgaben konfrontiert werden.

Der erste Subtest nennt sich *Pseudosegmentierung*, in dem die Proband:innen die Laute eines vorgesprochenen Wortes einzeln wiedergeben sollen (Stock et al. 2003b). Im zweiten Subtest geht es um die *Vokalersetzung* und besteht aus der Aufgabe, jedes [a] in einem Wort in ein [i]

zu verwandeln und das Wort verändert verbal wiederzugeben (Stock et al. 2003b). Dieser Subtest stellt demnach die Anforderung an die Proband:innen nicht nur die zu ändernden Vokale zu identifizieren, sondern diese auch zu ersetzen. Der vierte Subtest nennt sich *Phonemvertauschung*, in dem die Aufgabe gestellt wird, dass die ersten beiden Laute eines vorgesprochenen Wortes vertauscht werden sollen (Stock et al. 2003b). Bei Subtest fünf, der *Lautkategorisierung*, sollen die Proband:innen aus mehreren Wörtern jenes herausfinden, das aufgrund des An- oder Auslautes nicht zu den anderen passt (Stock et al. 2003b). Die Kinder müssen demnach den Anfang- und Endlaut analysieren und sie von ähnlichen Lauten differenzieren. Anschließend folgt in Subtest Sechs die *Vokallängenbestimmung*, in der die Kinder entweder ein Wort mit einem kurzen Vokal aus einer Gruppe mit langen Vokalen heraushören sollen oder umgekehrt ein Wort mit einem langen Vokal aus einer Gruppe mit kurzen Vokalen (Stock et al. 2003b). Es wird die Fähigkeit zur Unterscheidung der Vokallänge erfasst. Der letzte Subtest nennt sich *Wortumkehr*, in dem die Kinder ein vorgespieltes Wort rückwärts wiedergeben sollen (Stock et al. 2003b). Damit dies gelingt, müssen sie die Laute analysieren und neu synthetisieren.

4.2.1 Kritische Diskussion

Bei einer genaueren Betrachtung des BAKO 1-4 fallen einige Punkte auf, die die Aussagekraft des Tests infrage stellen. Zunächst lässt sich festhalten, dass bei jedem Subtest eine unterschiedliche Anzahl an Items vorliegt. So besteht beispielsweise der erste Subtest aus zehn Items, während der siebte Test aus 18 Items besteht. Aufgrund dieser Tatsache lässt die Konstruktion der Subtests keinen Vergleich untereinander zu. Diese Problematik zeigt sich auch bei einer detaillierteren Betrachtung innerhalb einzelner Subtests, was nachfolgend beispielhaft an Subtest fünf dargestellt wird:

5. Lautkategorisierung	
Items	
51.	(A) Kopf – Turm – tief – Trick
52.	(A) mib – mad – nob – mof
53.	(A) pat – kut – pit – pal
55.	(E) grau – froh – Pfau – blau
56.	(E) Rock – Zeit – rot – statt
57.	(E) haum – laum – faun – gaum
58.	(E) res – basch – pos – fas
59.	(E) fük – mük – nüt – rük

Abbildung 5: BAKO 1-4 (Stock et al. 2003a), Subtest 5 (Lautkategorisierung)

Bei diesem Subtest wird die Anforderung an die Proband:innen gestellt, das Wort zu identifizieren, welches nicht zu den anderen Worten passt (Stock et al. 2003b). Innerhalb des Subtests wird zwischen Items differenziert, bei denen entweder der Anlaut oder der Auslaut

entscheidend für die richtige Auswahl ist. An dieser Stelle stellt sich jedoch die Frage, warum die Testautor:innen für diese beiden Aspekte unterschiedliche Umfänge an Items gewählt haben: Während die Version A (Anlaut) aus drei Items besteht, umfasst die Version E (Auslaut) fünf Items. Somit kann hier eine Vergleichbarkeit zwischen den beiden Aspekten nicht erfolgen.

Wie bereits bei dem ELFE 1-6 (Lenhard & Schneider 2006a) kann die Frage nach der Auswahl der Items nicht zureichend beantwortet werden. Dies kann an dem zweiten Subtest exemplarisch dargestellt werden:

2. Vokalersetzung	
Items	
14.	Mittag
15.	sugita
16.	Ananas
17.	Handballmannschaft
18.	Namenstag
19.	arast
20.	Sandra
21.	Mathematik
22.	miranale
23.	Marmelade
24.	Januar
25.	alernaka

Abbildung 6: BAKO 1-4 (Stock et al. 2003a), Subtest 2 (Vokalersetzung)

Die Proband:innen sollen jedes [a], das in einem Wort vorkommt, in ein [i] verwandeln (Stock et al. 2003b). In diesem Kontext werden auch Pseudowörter eingesetzt, wobei anzumerken ist, dass an keiner Stelle bestätigt wird, dass diese Pseudowörter auf ihr tatsächliches Dasein als Pseudowörter getestet wurden und keine zufälligen Kohärenzen zu anderen Sprachen aufweisen. Das kann vor allem bei mehrsprachigen Kindern dazu führen, dass das Lösen der Aufgabe von dem Wissen über die Sprache beeinflusst wird und sie demnach einen Vorteil haben können (van der Straten Waillet et al. 2023). Grundlegend ist der Einsatz von Pseudowörtern kein untypisches Mittel, um die Lesekompetenz ohne eine Möglichkeit auf Rückgriffe auf bereits verinnerlichte Phonem-Graphem-Zuordnungen testen zu können (Marx 2007). Es gilt jedoch stets mitzudenken, dass der Einsatz von Pseudowörtern das Arbeitsgedächtnis beansprucht (Marx 2007). Da das Arbeitsgedächtnis zu Beginn des Schriftspracherwerbs ein wesentlicher Faktor ist, „weil phonologische Informationen über einzelne Grapheme zwischengespeichert werden müssen, bis ein ganzes Wort gelesen werden kann“ (Mehlem 2024, S. 57), ist dies kein Ausschlusskriterium. Marx (2007) stellt ebenfalls heraus, dass die Aufgaben „recht gut den Anforderungen beim Lesen und Schreiben [entsprechen], da hierbei ebenfalls phonologische Bewusstheit und phonologisches Arbeitsgedächtnis gleichzeitig beansprucht werden“ (S. 91). Trotz dessen gilt es zu bedenken, dass das Arbeitsgedächtnis

Einfluss auf die Ergebnisse haben wird. Bei diesem konkreten Subtest gibt es jedoch darüber hinaus einen weiteren Aspekt, der erneut die Möglichkeit zur Vergleichbarkeit mindert: bei vier von elf Items handelt es sich um Pseudowörter. Es besteht folglich ein ungleiches Verhältnis und es kann keine eindeutige Aussage darüber getroffen werden, wie Kinder mit bekannten und unbekannten Wörtern umgehen. Zusätzlich werden unterschiedliche Wortlängen und eine Mischung aus Pseudowörtern, Eigennamen, lexikalischen Wörtern und Komposita eingesetzt, die ebenfalls unbegründet bleiben. Laut den Testautor:innen sind die Items innerhalb der Subtest nach Schwierigkeit gestaffelt, allerdings bleibt offen wie genau dies umgesetzt wurde (Stock et al. 2003b). Es lässt sich auch nicht durch die Anzahl an [a]s innerhalb eines Wortes erklären, da es nicht fortlaufend ansteigt.

Der nächste Kritikpunkt wird an dem vierten Subtest, der *Phonemvertauschung*, ersichtlich:

42.	serog	esrog
43.	iftak	fitak
44.	Löwe	ölwe

Abbildung 7: BAKO 1-4 (Stock et al. 2003a), Subtest 4 (Phonemvertauschung), Items 42-44

Die Proband:innen sollen die ersten beiden Laute vertauschen, was eine sehr komplexe Aufgabe darstellt. Diese Aufgabe setzt ein hohes Maß an phonologischer Bewusstheit voraus, insbesondere die Fähigkeit, Laute in einem Wort isoliert wahrzunehmen, zu manipulieren und anschließend korrekt wieder zusammenzufügen. Es kommt die Frage auf, warum sich an dieser Stelle nicht für eine Vertauschung der Silben entschieden wurde, da Silben deutlich zugänglicher sind als Phoneme (Bredel et al. 2017). Unabhängig davon wird bei der Betrachtung der Musterlösung schnell offenkundig, dass die Aussprache nach der Vertauschung oft sehr schwer ist (Stock et al. 2003b). Allerdings ist es für eine erfolgreiche Umsetzung des Subtests erforderlich, dass der Neologismus sowohl rhythmisch als auch prosodisch korrekt strukturiert ist, um eine natürliche und flüssige Artikulation zu ermöglichen.³

Schlussendlich lässt sich festhalten, dass sich ein verbal-sprachgebundener Test als erkenntnisbringend für eine Betrachtung des phonologischen Bewusstseins herausstellt, aber sich der BAKO 1-4 aufgrund der dargestellten Kritikpunkte nicht ohne Weiteres in der Form umsetzen lässt. Es ist essenziell, dass die Items auch nach der Manipulation weiterhin eine klare und flüssige Aussprache ermöglichen. Des Weiteren ist der begründete Einsatz von Pseudowörtern an geeigneter Stelle sinnvoll, jedoch darf nicht außer Acht gelassen werden, dass das Arbeitsgedächtnis in dem Kontext eine große Rolle einnimmt, was eine Hürde für die eindeutige Interpretation der Ergebnisse darstellen kann.

³ Zu den prosodischen Merkmalen von Sprache „gehören ihre Gliederung in Silben, Akzente im Sinne eines stärkeren Drucks bei der Artikulation, Tonhöhe, Sprechtempo und weitere“ (Mehlem 2024, S. 41).

4.3 Inventar zur Erfassung der Lesekompetenzen im 1. Schuljahr (IEL-1)

Bei dem *Inventar zur Erfassung der Lesekompetenzen im 1. Schuljahr (IEL-1)* (Diehl & Hartke 2012a) handelt es sich um einen curriculumbasierten, standardisierten Test, welcher konkret für die erste Klasse entwickelt wurde. Der IEL-1 umfasst insgesamt 135 Items und setzt den Schwerpunkt auf die phonologische Bewusstheit sowie auf das Wort-, Satz- und Textverständnis (Diehl & Hartke 2012b). Er hat eine Länge von 90 Minuten, wobei dieser in zwei Phasen (2 x 45 Minuten) durchgeführt werden soll. Dies kann entweder mit einer Pause zwischen den Testungen an einem Tag oder an zwei kurz aufeinanderfolgenden Tagen durchgeführt werden. Er kann sowohl als Gruppentest als auch als Einzeltest genutzt werden. Darüber hinaus besteht die Möglichkeit den Test als einmalige Leistungsmessung zu nutzen oder wiederholend an insgesamt drei Messzeitpunkten für eine Dokumentation des Lernfortschritts. Vor der Durchführung des Haupttests erfolgt ein dreiminütiges Screening in Form einer Eine-Minute-Lese-Aufgabe, die im Folgenden aufgrund der Irrelevanz für diese Arbeit nicht genauer betrachtet wird (Diehl & Hartke 2012b).

Der IEL-1 besteht aus sieben Subtests, wobei die einzelnen Subtests noch durch weitere Aufgaben gegliedert sind. Da es sich demnach insgesamt um 20 Subtests handelt, würde es den Rahmen sprengen all diese in ihren Einzelheiten darzustellen.⁴ Aufgrund dessen wird sich im Folgenden auf die sieben Oberkategorien bezogen und in diesem Zuge kurz dargestellt, welche Kompetenzen dort im Mittelpunkt der Betrachtung liegen. Die erste Kategorie trägt den Titel *Buchstaben-Laut-Zuordnung* und umfasst vier Subtests, bei denen spezifische Laute analysiert, identifiziert und miteinander in Beziehung gesetzt werden müssen (Diehl & Hartke 2012a). Im Subtest *Wörter in Silben segmentieren* liegt der Schwerpunkt auf der korrekten Anwendung von Silbenbögen. Hierbei sollen einerseits Silbenbögen unter vorgegebene Wörter gesetzt werden, andererseits sollen Wörter anhand gegebener Silbenbögen segmentiert werden (Diehl & Hartke 2012a). Der Subtest *Silben zu Wörtern verbinden* fordert das Zusammensetzen verschiedener Silben zu vollständigen Wörtern (Diehl & Hartke 2012a). Im Subtest *Buchstaben zu Wörtern verbinden* besteht die Aufgabe darin, fehlende Buchstaben in Wörter einzufügen (Diehl & Hartke 2012a). Ein weiterer Subtest, *Wörter zu Bildern verbinden*, verlangt, dass entweder passende Wörter mit Bildern verknüpft oder die Wörter zu den Bildern selbstständig geschrieben werden (Diehl & Hartke 2012a). Im sechsten Subtest, *Sätze*, stehen Satzstrukturen im Mittelpunkt. Hier sollen Sätze gelesen und anschließend die richtige Antwort auf eine Frage gegeben werden. Dazu gehört, den passenden Satz zu einem Bild zu finden oder ein Bild auf Grundlage eines Satzes zu vervollständigen (Diehl & Hartke 2012a). Der letzte Subtest, *Text*, widmet sich der Informationsaufnahme aus Texten. Die Aufgaben

⁴ Die Subtests, die konkret auf die Konstruktion des neuen Tests Einfluss genommen haben, werden im Kapitel 5.3 im Zuge der Darstellung des Aufbaus des Tests genauer dargestellt.

bestehen darin, ein passendes Bild mit einem Text zu verbinden oder den korrekten Satz auszuwählen, der zu einem Text passt (Diehl & Hartke 2012a).

4.3.1 Kritische Diskussion

Auf den ersten Blick fällt auf, dass die Durchführungsdauer (90 Minuten) sehr umfangreich ist, was sich primär bei einem Test, welcher aktiv für die erste Klasse konzipiert wurde, als sehr herausfordernd erweist. Studien zeigen auf, dass Schüler:innen der ersten Klasse eine begrenzte Konzentrationsspanne aufweisen und ein systematischer Zusammenhang zwischen der visuellen Aufmerksamkeitsspanne und der Leistung im Bereich der Schriftsprache besteht (Valdois et al. 2024). Eine derart lange Testdauer kann somit zu Ermüdung und Frustrationen führen, was wiederum die Testleistung beeinträchtigen und die Ergebnisse verfälschen kann. Die Dauer des Tests geht offenkundig mit einer sehr hohen Anzahl an Subtests einher, die sich ebenfalls in weitere Aufgaben untergliedern. Daher werden die allgemeinen Aspekte genannt, die kritisch hinterfragt und mit Beispielen aus ausgewählten Subtests unterstützt werden.

Wie bereits in den anderen Tests kann ebenfalls festgehalten werden, dass grundlegend keine Einheitlichkeit in vielen verschiedenen Bereichen besteht. So gibt es in dem ersten Subtest und in der ersten Aufgabe, in dem ein Phonem identifiziert werden soll, eine ungleiche Anzahl an An-, In- und Auslauten (Diehl & Hartke 2012a). Zudem sind nur in ausgewählten Subtests Übungsisems vorhanden. Diese sind jedoch von großer Relevanz, da sie das Verständnis des Aufgabenformats auf Seiten der Kinder gewährleisten, bevor sie mit der Bearbeitung starten. Dies nimmt Unsicherheiten und stellt sicher, dass die Kinder sich auf die inhaltliche Bearbeitung konzentrieren können. Übungsisems tragen somit dazu bei, valide Testergebnisse zu gewährleisten, da sie durch das frühzeitige Auflösen von möglichen Unklarheiten die Widerspiegelung der tatsächlichen Kompetenzen ermöglichen.

Ebenfalls wie in den anderen Tests, wird die Auswahl der Wörter nicht begründet. Was jedoch kenntlich wird, ist der Einsatz niedrigfrequenter Wörter sowie subtestübergreifende Wortdoppelungen. Diese sind problematisch, da sie die Validität der Testergebnisse beeinträchtigen können. Taucht ein Wort in mehreren Subtests auf, besteht die Gefahr, dass die Schüler:innen das Wort wiedererkennen. Solche Wiederholungen können Lerneffekte erzeugen, die die Vergleichbarkeit und Trennschärfe zwischen den Subtests verringern. Es gibt zahlreiche Studien, die belegen, dass wiederholtes Lesen eines Textes starke, anhaltende Auswirkungen auf das Lernen von Wörtern hat (Williams & Horst 2014; Mendes & Kirby 2024). Des Weiteren zeigt beispielsweise eine frühere Studie von Griffiths und Johnson (1995) auf, dass sich bei wiederholtem Kontakt mit Wörtern sich die schriftsprachlichen Leistungen verbessern. Sie untersuchten die Frikativwahrnehmung bei Kleinkindern, indem sie wiederholt Tests durchführten (Griffiths & Johnson 1995). Die Studie ergab, dass sich die Leistung bei der Wortidentifizierungsaufgabe fortlaufend verbesserte (Griffiths & Johnson 1995).

Dass die Aussagekraft durch Faktoren aufgrund des Testdesigns beeinträchtigt wird, zeigt sich des Weiteren an anderer Stelle. In Subtest fünf und in der Aufgabe zwei geht es darum, dass die Proband:innen Bilder sehen und das dazugehörige Wort schreiben sollen (Diehl & Hartke 2012a). Die testleitende Person ist laut den Durchführungshinweisen dazu angehalten, die Bilder mit den richtigen Wörtern zu nennen, um einer Fehlinterpretation der Bilder entgegenzuwirken (Diehl & Hartke 2012b). Obgleich zuzustimmen ist, dass der Einsatz von Bildern stets eine Gefahr einer von der intendierten Antwort abweichenden Antwort birgt, führt die verbale Benennung der Items zu einer Verfälschung der Zielkompetenz des Subtests. Nennt die testleitende Person die Bezeichnung des Items, wird den Proband:innen die Aufgabe abgenommen, das Wort selbstständig zu erkennen, aus dem mentalen Lexikon zu aktivieren und zu verschriftlichen. Stattdessen kann es dazu führen, dass nur die Fähigkeit der Kinder gemessen wird, ein bereits vorgegebenes Wort korrekt zu verschriftlichen. Insbesondere bei Subtests, welche sich schwerpunktmäßig mit der phonologischen Bewusstheit, dem Wortschatzabruf und den orthografischen Kompetenzen beschäftigen, beeinträchtigt eine solche Vorgehensweise die Validität der Tests erheblich.

Darüber hinaus zeichnen sich einige Subtests durch eine hohe Komplexität aus. Ein Beispiel hierfür stellt die erste Aufgabe des vierten Subtests dar, in denen die Kinder die fehlenden Buchstaben in unterschiedliche Worte setzen sollen (Diehl & Hartke 2012a). Es sollen sowohl Vokale und Konsonanten als auch Diphthonge und Doppelkonsonanten eingesetzt werden. Wird die Artikulationsart betrachtet, wird ebenfalls die Komplexität deutlich: Die einzusetzenden Grapheme sind Plosive, Nasale und Frikative.⁵ In diesem Kontext sind im Besonderen komplexe Doppelkonsonanten wie <qu> und schwerer erlernbare Konsonanten wie <z> und <x> zu nennen. Diese sind in der deutschen Sprache zudem seltener vorkommend. Für Schüler:innen der ersten Klasse kann der Einsatz solcher Items sehr herausfordernd sein und zu Überforderungen führen. Die dritte Aufgabe des sechsten Subtests unterstreicht diese Annahme und stellt die Aufgabe an die Proband:innen, ein Bild anhand der aus einem Satz entnommenen Informationen malerisch zu vervollständigen (Diehl & Hartke 2012a). Wenngleich dieser Subtest das Satzverständnis sinnvoll testet, wird für die Items die vorangestellte Dativkonstruktion genutzt. In der ersten Klasse wird allerdings vielfach mit der Grundstruktur deutscher Sätze (Subjekt – Verb – Objekt) gearbeitet. Die vorangestellte Dativkonstruktion ist deutlich komplexer, weil es die gewohnte Reihenfolge der Satzteile verändert (Sassenroth 2000). In der Grundstruktur folgt das Verb direkt dem Subjekt, sodass direkte Referenzen gezogen werden können, was die Verarbeitung erleichtert. Bei vorangestellten Dativkonstruktionen muss die syntaktische Funktion des Dativs im Zusammenhang mit dem Verb verstanden

⁵ Einen umfassenden und differenzierten Einblick in die jeweiligen Definitionen bietet Mehlum (2024, S. 34f.).

werden, bevor das Subjekt gelesen wird. Diese Umstellung erfordert eine höhere kognitive Anstrengung.

Der IEL-1 hebt sich von den anderen Tests in der Art der Entwicklung ab, da diese auf der Basis des Curriculums geschehen ist (Diehl & Hartke 2012b). Die meisten standardisierten Testinstrumente „zielen auf die Diagnose grundlegender Fähigkeiten ab und sind deshalb nicht auf spezifische Unterrichtsinhalte hin konstruiert.“ (Lenhard 2019, S. 72). Dies birgt jedoch den Vorteil, dass die Zielsetzungen des Tests in enger Verbindung zu den schulischen Lernzielen und -anforderungen stehen. Dies kann die Validität des Tests erhöhen, da dies impliziert, dass die tatsächlichen Kompetenzen der Schüler:innen in Bezug auf den spezifischen Lernstoff erfasst werden (Marx 2007).

Des Weiteren ist die genutzte Staffelung in Phonem-, Silben-, Wort-, Satz- und Textebene sinnvoll, da sie die unterschiedlichen Ebenen der Sprachverarbeitung systematisch erfasst und eine gezielte Analyse des Schriftspracherwerbs ermöglicht (del Toro & Thomé 2021). Jede dieser Ebenen spielt eine spezifische Rolle im Prozess des Schriftspracherwerbs, sodass eine differenzierte Betrachtung der unterschiedlichen Prozesse, ermöglicht wird.

4.4 Zwischenfazit

Nach der kritischen Auseinandersetzung mit dem ELFE 1-6 (Lenhard & Schneider 2006a), dem BAKO 1-4 (Stock et al. 2003a) und dem IEL-1 (Diehl & Hartke 2012a), lässt sich festhalten, dass einige Fragen bezüglich der Testkonstruktion offenbleiben. Viele der dargestellten Aspekte beeinträchtigen die Aussagekraft der Tests enorm. Hierbei ist primär die unbegründete Wahl der Items zu nennen, die sich durch alle drei präsentierten Tests zieht. Durch einen inkonsequenten und nicht nachvollziehbaren Einsatz von bestimmten Faktoren wie beispielsweise semantischer Gemeinsamkeiten oder Unterschieden, der Anzahl der Items sowie der Nichtberücksichtigung der Frequenz leidet die Vergleichbarkeit der Ergebnisse. Darüber hinaus werden infolgedessen bestimmte Einflussfaktoren nicht konstant gehalten, sodass sie großen Einfluss auf die zu testenden Variablen haben können und demnach unter Umständen nicht das getestet wird, was getestet werden soll. Dies zieht die Konsequenz nach sich, dass diese bereits bestehenden Tests zur Erfassung des Schriftspracherwerbs nicht ohne Anpassungen eingesetzt werden können.

5. Entwicklung des neuen Lesetests für die erste Klasse

Nachdem die kritische Auseinandersetzung mit den bereits bestehenden Tests zur Erfassung des Schriftspracherwerbs viele Punkte hervorbrachte, die die Tests offenkundig nicht in der Form einsetzbar machen, in der es notwendig ist, wurde an einem neuen Test gearbeitet. Die detaillierte Entwicklung und die Einzelheiten des Tests werden im Folgenden dargestellt. Dafür werden zunächst die Zielsetzung und die Anwendungsbereiche und nachfolgend die Schritte der Testkonstruktion beschrieben. Anschließend wird der Aufbau des Lesetests präsentiert und die Durchführung des Tests dargestellt. Das Kodierungssystem für die Auswertung, welches sich mit der Bepunktung der einzelnen Subtests beschäftigt, rückt abschließend in den Mittelpunkt der Betrachtung.

5.1 Zielsetzung und Anwendungsbereiche

Der Test verfolgt das Ziel, den aktuellen Stand des Schriftspracherwerbs bei Schüler:innen der ersten Klasse zu ermitteln. Er dient zur Überprüfung der phonologischen Bewusstheit, der orthografischen Kompetenz des Leseverständnisses auf Wort-, Satz- und Textebene.

Aufgrund der bereits dargestellten positiven Eigenschaften verfolgt der Test das Ziel, curriculumbasiert vorzugehen und kann daher an diesen Kriterien überprüft werden. Die Items orientieren sich an den Lehrwerken. Da der Test für einen Einsatz in der Schule konstruiert wurde, hat dieses Vorgehen den Vorteil, dass die Aufgabenformate der Subtests den Proband:innen bereits bekannt sind und somit Barrieren abgebaut werden können. An dieser Stelle muss aber angemerkt werden, dass sich der Anwendungsbereich nicht auf die Schule begrenzt. Grundsätzlich können alle Kinder, die im Alter der Kinder der ersten Klasse sind, mit diesem Test getestet werden.

Wie bereits dargestellt, besteht der Schriftspracherwerb aus vielen verschiedenen Teilkompetenzen und -bereichen. Dies spiegelt sich auch in den offiziellen Richtlinien und Empfehlungen für die Schule wider. So wird im Lehrplan für die Grundschule für das Fach Deutsch explizit geschrieben, dass sich der Schriftspracherwerb als ein „Zusammenspiel von Lesen und Schreiben“ (QUA-LiS NRW 2024b) beschreibt. Nachfolgend werden in dem Kontext vier Kompetenzbereiche dargestellt:

- 1) Sprechen und Zuhören
- 2) Schreiben
- 3) Lesen – mit Texten und Medien umgehen
- 4) Sprache und Sprachgebrauch untersuchen (QUA-LiS NRW 2024a)

Mit diesen verschiedenen Bereichen gehen explizit erläuterte Kompetenzerwartungen einher, die jeweils aufgeteilt sind in welche, die für das Ende der Schuleingangsphase entwickelt und

welche, die für das Ende der vierten Klasse definiert wurden. Da für diese Arbeit nur der erste Zeitraum relevant ist, wird sich nachfolgend auch nur auf diesen bezogen.

Die Anforderungen des Tests und die Kompetenzerwartungen des Bereichs *Schreiben* überschneiden sich an einigen Stellen. So wird unter dem Schwerpunkt *Über Schreibfähigkeit verfügen* das Folgende geschrieben: „Die Schülerinnen und Schüler schreiben flüssig und formklar“ (QUA-LiS 2024a). Insbesondere die hier angesprochene Formklarheit lässt sich in dem Test, vor allem bei dem zweiten und fünften Subtest, aber auch bei Subtest vier finden, da die Schüler:innen in diesen Subtests selbstständig schreiben und die Orthografie bei der Bewertung berücksichtigt wird. Dies wird ebenso im Kontext des Schwerpunktes *Richtig schreiben* ersichtlich: „Die Schülerinnen und Schüler [...] wenden beim Schreiben eigener Texte erste Rechtschreibmuster und rechtschriftliche Kenntnisse an (z. B. Einhalten der Wortgrenzen, Großschreibung nach Satzschlusszeichen, Endungen -en und -er sowie Schreibung von Wörtern mit au, ei, eu, ch, sch, st, sp und qu)“ (QUA-LiS NRW 2024a).

Im Anschluss daran wurden wichtige Fähigkeiten und Kenntnisse im Rechtschreiben für Klasse eins bis vier formuliert. Diese gliedern sich in die Laut-Buchstaben-Ebene, die Wortebene und die Satzebene (QUA-LiS NRW 2024a). Bereits diese Art der Gliederung weist Parallelen mit der des entwickelten Tests auf, wobei dort noch die Textebene als zusätzliche Ebene hinzukommt. Speziell der Bereich der Laut-Buchstaben-Ebene erfordert im Kontext der Entwicklung des Tests eine besondere Betrachtung. Hierunter zählen gemäß des Lehrplans folgende Kompetenzen:

- „regelhafte Laut-Buchstaben-Zuordnungen der Schreibung kennen und anwenden
- ähnliche Laute und Lautfolgen unterscheiden und sie den entsprechenden Buchstaben zuordnen
- stimmhafte und stimmlose Konsonanten unterscheiden
- lang und kurz gesprochene Vokale unterscheiden
- Abweichungen von der regelhaften Laut-Buchstaben-Zuordnung erkennen und beachten“ (QUA-LiS NRW 2024a)

All diese erwähnten Fähigkeiten sind für eine erfolgreiche Bearbeitung des Tests ebenfalls relevant. Im Zuge der Entwicklung des Tests ist die Wortebene von großer Bedeutung, vor allem die folgenden Aspekte:

- „Silbentrennung beachten
- Namen und Nomen mit großen Anfangsbuchstaben schreiben (soweit das ohne Bezug zum Satz möglich ist)“ (QUA-LiS 2024a)

Das Erkennen von Silbengrenzen stellt eine zentrale Kompetenz des Schriftspracherwerbs dar, welcher im Test eine große Bedeutung zukommt. Das Beachten der Großschreibregelung

am Satzanfang und von Nomen wird vorrangig in Subtest drei und sechs relevant, da hier die Proband:innen selber Begriffe schreiben sollen und hierbei auch die Orthografie, einschließlich Großschreibung, betrachtet wird.

Die Satzebene beinhaltet im Lehrplan nur zwei Aspekte, die sich für die Testentwicklung als irrelevant herausstellen, da die Proband:innen keine Sätze selber schreiben müssen (QUA-LiS 2024a). Jedoch werden daran anschließend noch das Erlernen von Ausnahmen betont, die orthografische Besonderheiten aufweisen (QUA-LiS 2024a). Diese Irregularitäten wurden auch in dem Test mit aufgegriffen. Hierzu zählen Doppelkonsonanten wie aus Subtest drei das Item *Badewanne* sowie Grapheme, welche Artikulationsähnlichkeiten aufweisen. Bezogen auf den Test sind hierfür beispielsweise die Items *Wolke* und *Feuerwehr* aus Subtest fünf zu nennen, da /w/ und /f/ gemeinsame phonetische Merkmale aufweisen, wie beispielsweise, dass sie beide labiodental gebildet werden und zu der Gruppe der Frikative zählen. Ein weiteres Beispiel für eine Besonderheit innerhalb eines Wortes stellt die Auslautverhärtung dar, wie in Subtest sechs das Item *Mond*. Unter einer Auslautverhärtung versteht man einen phonologischen Prozess, bei dem stimmhafte Konsonanten im Auslaut zu stimmlosen Konsonanten werden (Bredel et al. 2017).

Die Bildungsstandards für das Fach Deutsch für die Primarstufe, entwickelt von der Kultusministerkonferenz (KMK), weisen ebenfalls verschiedene, ausdifferenzierte Kompetenzbereiche auf, wobei die ersten drei prozessbezogen und die letzten beiden domänenspezifisch sind (KMK 2022a). Während die fachbezogenen Kompetenzen sowohl fachspezifisch als auch fachübergreifend von Bedeutung sind, beziehen sich die domänenspezifischen Kompetenzen konkret auf ein Fach. Die Kompetenzen gliedern sich wie folgt:

- 1) Sprechen und Zuhören
- 2) Schreiben
- 3) Lesen
- 4) Sich mit Texten und anderen Medien auseinandersetzen
- 5) Sprache und Sprachgebrauch untersuchen (KMK 2022a)

An dieser Stelle findet sich folglich auch der Bereich des *Schreibens* wieder. Wie auch der Lehrplan orientiert sich die KMK an der Entwicklung des korrekten orthografischen Schreibens: „Die Schülerinnen und Schüler kennen wortbezogene und zum Teil auch satzbezogene Rechtschreibstrategien und wenden diese an“ (KMK 2022a, S. 14). Indes wird das Erschließen von Strukturen von Wörtern gesondert hervorgehoben, insbesondere das Gliedern „in Silben und Wortbestandteile, verlängern, flektieren, ableiten und erweitern“ (KMK 2022a, S. 14). Diese Kompetenzen werden auch in dem Test geprüft.

Sowohl im Lehrplan als auch in den Bildungsstandards für das Fach Deutsch findet sich der Bereich *Lesen* wieder, wobei unterschiedliche Schwerpunkte gesetzt werden. Die KMK rückt das Leseverstehen in den Mittelpunkt der Betrachtung, indem herausgestellt wird, dass die Schüler:innen „explizite oder durch einfaches Schließen zu gewinnende Textinformationen auf lokaler Ebene“ ermitteln und dabei „sprachliche Mittel zur Sicherung des Textzusammenhangs“ (KMK 2022a, S. 17) beachten. Dieser Aspekt wird vor allem bei Subtest acht des Tests relevant, da hier ein kurzer Text gelesen wird. Obwohl die KMK im Kontext des Bereichs *Flüssig lesen* im Besonderen den Lesefluss thematisiert, welcher in dem Test nicht explizit betrachtet wird, werden in dem Zuge Kompetenzen definiert, die in diesem trotz dessen unabdingbar sind:

„Die Schülerinnen und Schüler

- nutzen die schriftsprachlichen Informationen (u. a. Silbe, Morphem, Komma, Punkt, Fragezeichen, Großschreibung) beim Lesen,
- verfügen bei häufig vorkommenden Wörtern über eine weitgehend sichere Worterkennung und korrigieren sich ggf. selbst,
- erlesen unvertraute Wörter anhand größerer Einheiten wie Silbe oder Morphem“ (KMK 2022a, S. 16)

Dies wird im Test spätestens dann relevant, wenn die Proband:innen einen Satz lesen sollen, wie es in Subtest sieben der Fall ist, da an dieser Stelle das semantische Verständnis allgemein in den Mittelpunkt der Betrachtung gestellt wird. Allerdings werden diese Fähigkeiten bereits subtil in vorherigen Subtests geprüft, da die Subtests, die sich mit der Phonem-, Silben- und Wortebene beschäftigen, die Grundlage darstellen.

Im Lehrplan nennt sich dieser Bereich konkret *Lesen – mit Medien und Texten umgehen*. Unter dem Schwerpunkt *Über Lesefähigkeit verfügen* wird Folgendes gefasst: „Die Schülerinnen und Schüler [...] lesen kurze altersgemäße Texte und beantworten Fragen zum Text“ (QUA-LiS NRW 2024a). Wenngleich die Proband:innen während des Tests keine Fragen zu einem Text beantworten müssen, werden die dafür benötigten Kompetenzen während anderer Aufgabenstellungen relevant. So sollen die Proband:innen in Subtest sechs den Inhalt des gelesenen Satzes mit Playmobil nachbauen. Um diesen Subtest erfolgreich absolvieren zu können, müssen die Proband:innen den Satz nicht nur lesen, sondern auch verstehen. Sowohl das Beantworten von Fragen als auch das Nachbauen eines gelesenen Satzes mit Playmobil erfordert, relevante Informationen aus dem Text herauszufiltern.

Schlussendlich kann festgehalten werden, dass der Test darauf abzielt, den aktuellen Stand des Schriftspracherwerbs bei Schüler:innen der ersten Klasse zu erfassen, zentrale

Teilkompetenzen wie phonologische Bewusstheit, orthografische Fertigkeiten und das Leseverständnis zu überprüfen.

5.2 Schritte bei der Testkonstruktion

Als Erstes muss an dieser Stelle herausgestellt werden, dass der gesamte Prozess der Testentwicklung aus der Zusammenarbeit mit dem Team *Empirische und experimentelle Linguistik des Deutschen – Psycholinguistik* unter der Leitung von Prof. Dr. Barbara Mertins an der TU Dortmund entstand, sodass alle in den nachfolgend angeführten Schritten ihren Ursprung in gemeinsamen Gedankengängen finden. Ausgangspunkt für die Entwicklung des Tests bestand in dem bereits in dem dritten Kapitel dargestellten geringen Forschungsstand zu dem Schriftspracherwerb in der ersten Klasse. Daraufhin folgte eine Sichtung der ohnehin wenigen, bereits bestehenden Verfahren zur Erfassung des Schriftspracherwerbs für die erste Klasse, woraufhin die Tests ELFE 1-6 (Lenhard & Schneider 2006a), BAKO 1-4 (Stock et al. 2003a) und IEL-1 (Diehl & Hartke 2012a) näher beleuchtet wurden. Mit der anfänglichen Intention einen dieser Tests einzusetzen, wurden diese kritisch betrachtet, wie in Kapitel vier detailliert dargestellt. Auch wenn auf der Basis dieser Ausführungen erkenntlich wird, dass diese Intention aufgrund der vielfachen methodischen Schwächen nicht umsetzbar ist, wurde der Frage nachgegangen, welche Stärken diese Tests aufweisen, auf denen aufgebaut werden kann und welche Schwächen vorhanden sind, die eine Überarbeitung erfordern.

Auf der Basis der bisherigen Ausführungen konnte das größte Entwicklungspotenzial dem IEL-1 (Diehl & Hartke 2012a) zugeschrieben werden, sodass die einzelnen Subtests erneut differenziert und kritisch durchleuchtet wurden, um zu entscheiden, welche Subtests eine geeignete Grundlage für eine Weiterentwicklung sind. Im IEL-1 (Diehl & Hartke 2012a) finden sich aufgrund seiner Komplexität viele Subtests, die eine geeignete Basis für eine Weiterentwicklung darstellen. Es wurde sich jedoch darauf konzentriert eine möglichst große Facette des Schriftspracherwerbs abzubilden bei gleichzeitig möglichst wenigen Subtests. Daher wurden keine Subtests ausgewählt, welche das Gleiche testen. Darüber hinaus wurde ein Subtest des BAKO 1-4 (Stock et al. 2003a) als vielversprechend angesehen, wenn auch herausfordernd. Obwohl diese Subtests einige Schwachstellen aufweisen und demnach auch eine starke Bearbeitung notwendig ist, birgt ein solches Bauen auf bestehenden standardisierten Tests den Vorteil, dass sie bereits auf umfangreiche empirische Untersuchungen und Daten basieren (Lenhard 2019). Es stellt sich als effizienter heraus, gezielt Schwächen zu identifizieren und diese anzupassen, anstatt ein von Grund auf eigenständiges Instrument zu schaffen. Dies trägt zu einer kumulativen Weiterentwicklung des Wissens in einem Forschungsfeld bei (Preddiger et al. 2012). Aufgrund dessen liefen die Subtests durch mehrere Korrekturschleifen.

Die Pilotierung des Tests hat im Zuge dessen zu einer Weiterentwicklung beigetragen, welche aus drei Erwachsenen im Alter zwischen 23 und 28 Jahren, einem Kind einer vierten Klasse

sowie vier Kindern einer ersten Klasse bestand. Die Erstklässler:innen besuchten die Max-Wiethoff-Grundschule in Herne. Das Ziel der Pilotierung besteht darin, die Funktionalität, Verständlichkeit und Praktikabilität des Instruments zu überprüfen, potenzielle Schwäche zu identifizieren und notwendige Anpassungen vorzunehmen (van Teijlingen & Hundley 2001). Schlussendlich entstanden acht Subtests, die im Folgenden umfassend erläutert werden.

5.3 Aufbau des Lesetests

Wie bereits erläutert basiert der Test schwerpunktmäßig auf dem IEL-1 (Diehl & Hartke 2012a), wobei ebenso zu einem geringen Teil auf dem BAKO 1-4 (Stock et al. 2003a). Fortlaufend werden die einzelnen Subtests des entwickelten Testinstrumentes erklärt, indem zunächst kurz auf die ursprüngliche Version dieser bereits bestehenden Tests eingegangen wird. Hierbei wird sich ausschließlich auf die Aspekte konzentriert, die nicht bereits in der kritischen Auseinandersetzung mit den jeweiligen Tests ausführlich dargestellt wurden. Anschließend wird die darauf aufbauende und weiterentwickelte Version beschrieben.⁶ Zunächst wird jedoch auf grundlegende Überlegungen eingegangen, welche für die gesamte Entwicklung der Subtests relevant sind.

Wie bereits erwähnt kommt der Frequenz der Items eine hohe Bedeutung zu, was von den bereits bestehenden Tests nicht beachtet wird. Aufgrund dessen wurde dies zu einem zentralen Ausgangspunkt für die Weiterentwicklung des neuen Tests, sodass ausschließlich hochfrequente Wörter genutzt wurden. Die Worthäufigkeiten wurden mit Hilfe des Frequenzbarometers des digitalen Wörterbuchs der deutschen Sprache (DWDS) erschlossen (Berlin-Brandenburgische Akademie der Wissenschaften o.A.).⁷ Eingesetzt wurden nur Wörter, welche auf der Basis von gegenwartssprachlichen Korpora eine Frequenz von mindestens drei aufweisen, was eine Häufigkeit zwischen 163.577 und 1.635.766 bedeutet (Berlin-Brandenburgische Akademie der Wissenschaften o.A.). Darüber hinaus wurde der Grundwortschatz NRW einbezogen (QUA-LiS NRW 2024c).⁸ Ausnahmen stellen der zweite und vierte Subtest dar, da dort Pseudowörter und aktiv niedrigfrequente Wörter eingesetzt wurden. Des Weiteren wurde Wert daraufgelegt, dass sich keine Items doppeln.

Es wurde sich an der Grundstruktur des IEL-1 orientiert, sodass die Subtests unterteilt werden können in die Phonemebene (Subtest eins und zwei), die Silbenebene (Subtest drei und vier), die Wortebene (Subtest fünf und sechs) und die Satz- und Textebene (Subtest sieben und

⁶ Die Testseiten und Protokollbögen sind im Anhang (siehe Anhang 1 und 2) aufgeführt.

⁷ Das Frequenzbarometer des DWDS beinhaltet eine siebenstufige, logarithmische Skala und basiert auf unterschiedlichen, aktuellen Korpora, welche stetig aktualisiert werden (Berlin-Brandenburgische Akademie der Wissenschaften o.A.). An der Stelle ist jedoch anzumerken, dass sich hierbei ausschließlich auf schriftliche Dokumente bezogen wird und die daraus berechnete Frequenz nicht in gleicher Weise auf den mündlichen Gebrauch übertragen werden kann. Trotz dessen bietet das Frequenzbarometer des DWDS einen guten Anhaltspunkt.

⁸ Der Grundwortschatz NRW enthält 533 Wörter, die grundlegende und damit grundschulrelevante Rechtschreibphänomene abbilden (QUA-LiS NRW 2024c).

acht). Die Phonemebene beschäftigt sich mit der phonologischen Bewusstheit im engeren Sinne, während die Silbenebene die phonologische Bewusstheit im weiten Sinne in den Mittelpunkt der Betrachtung rückt. Die Wortebene stellt die Orthografie in den Fokus und bei der Satz- und Textebene wird die Semantik in den Blick genommen. Die Items sind stets so aufgebaut, dass mit mindestens einem Übungssitem begonnen wird, welches für den Subtest möglichst repräsentativ dargestellt ist. Anschließend staffeln sich die Items nach ihrer Komplexität im Sinne einer progressiven Leistungsanforderung, sodass die individuellen Kompetenzen der Proband:innen genauer erfasst werden können. Dies bietet eine geeignete Grundlage für eine systematische Analyse des Kompetenzaufbaus. Zudem schaffen leichtere Items zu Beginn Erfolgserlebnisse, die motivierend wirken, womit Frustrationen vermieden werden können. In vielen Subtests dienen Bilder als Unterstützung, da sie helfen, den Fokus auf die Zielkompetenz des Tests zu lenken, was die Verarbeitung der Aufgaben erleichtert. Bei den Bildern handelt es sich um einfache, ikonische und schwarz-weiße Motive, um eine Ablenkung zu vermeiden.

Der erste Subtest basiert auf der ersten Aufgabe des ersten Subtests des IEL-1. Die Aufgabe besteht darin, das Kästchen anzukreuzen, an welcher Stelle der jeweilige Vokal gehört wird (Diehl & Hartke 2012a). Es sind pro Item drei Kästchen vorzufinden, welche repräsentativ für den An-, In- und Auslaut stehen. Dieses Grundprinzip des Subtests wurde übernommen, wobei die Darstellungsweise geändert wurde. Während die Items in der Version des IEL-1 pro Vokal spaltenweise präsentiert werden, sind die Items nun in Zeilen vorzufinden. Dies ist mit Blick auf die Leserichtung erheblich intuitiver und erleichtert die Bearbeitung.

Die Items wurden so angepasst, dass pro Vokal eine gleiche Anzahl an Items zu finden ist und zusätzlich eine gleiche Anzahl an An-, In- und Auslauten pro Vokal. Dadurch entsteht eine Vergleichbarkeit zwischen den Items. Zudem wurde pro Vokal jeweils ein Item hinzugefügt, bei welchem aufgrund eines zweifachen Vorkommens des Vokals zwei Kreuze gesetzt werden sollen. Das Erkennen von Vokalen in einem Wort erfordert eine differenzierte Analyse der Wortstruktur. Die Komplexität dieser Anforderung steigt, wenn diese mehrfach in einem Wort vorkommen. Des Weiteren wurden Übungssitems hinzugefügt. Dabei haben die Pilotierungsergebnisse gezeigt, dass es sinnvoll ist, hierbei ein Item einzuschließen, in dem der Vokal zwei Mal vorkommt. Dadurch müssen zwei Kreuze gesetzt werden. Auch wenn den Schüler:innen das in dem Subtest zugrundeliegende Aufgabenformat bekannt ist, ist es für sie meistens neu, dass teilweise zwei Kreuze gefordert werden. Um einer Fehlinterpretation der Bilder entgegenzuwirken und da es in diesem Subtest nicht um das Testen des Wortschatzes geht, wurde entschieden, dass die testleitende Person die Bearbeitung verbal begleitet, indem das Item vorgesagt wird und die Proband:innen anschließend das Kreuz setzen. In Folge des strikten Einsatzes von ausschließlich hochfrequenten Items, wurde sich in diesem Subtest für den

Einbezug von Eigennamen entschieden, die aber nicht eindeutig auf einem Bild darzustellen sind. Da die Betonung des Wortes zusätzlich eine immense Auswirkung auf die Antwort der Proband:innen haben kann, wurden vorab Tonaufnahmen aufgenommen, welche Wort für Wort einzeln abgespielt werden. So kann die Objektivität der Testdurchführung gewährleistet werden, was die Vergleichbarkeit unter den Proband:innen ermöglicht.

Die Items dieses Subtests wurden im Kontrast zu den anderen Subtests nicht nach Schwierigkeit sortiert, wobei Yalcinkaya et al. (2010) in einer Studie aufzeigen konnten, dass der Erwerb des Identifizierens des An-, In- und Auslauts in verschiedenen Entwicklungsphasen stattfindet. Der Anlaut wurde bereits bei 75% der Kinder zwischen dem zweiten und vierten Lebensjahr erworben, während die Entwicklung des Bewusstseins für den Auslaut bis zu dem sechsten Lebensjahr verlief (Yalcinkaya et al. 2010). Die Inlaut-Position konnte als die Komplexeste herausgestellt werden (Yalcinkaya et al. 2010). Eine Staffelung auf der Basis dieser Erwerbsschritte würde allerdings dazu führen, dass die Items pro Vokal jeweils gleich angeordnet werden müssten. Infolge dessen könnten die Proband:innen die Antwortmöglichkeiten prognostizieren, sodass die Fähigkeiten des Schriftspracherwerbs, welche getestet werden sollen, nicht mehr im Vordergrund stehen würden. Daher wurden die Items innerhalb der Zeilen randomisiert.

Dieser Subtest rückt die phonologische Bewusstheit im engeren Sinne in den Fokus. Es wird die Fähigkeit getestet, einzelne Laute zu erkennen und zu unterscheiden sowie die Kompetenz ein Wort in die Lautbestandteile zu segmentieren. Um den Subtest erfolgreich bearbeiten zu können, müssen die Kinder ein Bewusstsein dafür entwickeln, wie sich verschiedene Laute im Wort anhören sowie wie diese Laute durch unterschiedliche Vokale dargestellt werden.

Der zweite Subtest ist der Einzige, der ohne Veränderungen aus dem BAKO 1-4 (Stock et al. 2003a) übernommen wurde. Der ursprünglich sechste Subtest thematisiert die Vokallänge, indem entweder ein Wort mit einem kurzen Vokal aus einer Wörtergruppe mit langem Vokal oder umgekehrt das Wort mit einem langen Vokal aus einer Wörtergruppe mit kurzem Vokal benannt werden soll (Stock et al. 2003a). Dieser Subtest ist verbal-sprachgebunden, wodurch es keinen Testbogen gibt, sondern lediglich einen Protokollbogen für die testleitende Person. Diese spielt die jeweiligen Items ab und die Proband:innen nennen das Wort, welches sich aufgrund der Vokallänge von den anderen unterscheidet.

Die Analyse der Vokallänge gehört ebenfalls zu der phonologischen Bewusstheit im engeren Sinne (Schneider 2017). Um den Subtest erfolgreich bearbeiten zu können, müssen die Kinder über die Kompetenz verfügen, Vokale zu unterscheiden und bewusst zu analysieren. Hierbei bildet das Erkennen der Vokallänge eine komplexe Anforderung, obwohl diese Fähigkeit von besonderer Relevanz für die deutsche Schriftsprache ist. Beispielsweise stellt dies eine Grundlage für die orthografische Kompetenz dar (Bosse 2015), da die Länge des Vokals häufig mit

orthografischen Markierungen wie der Verdopplung des darauffolgenden Konsonanten (beispielsweise in <Hass>) oder des Dehnungs-h (beispielsweise in <Hahn>) einhergeht.

Der dritte Subtest basiert erneut auf dem IEL-1, genauer auf der zweiten Aufgabe des zweiten Subtests. Hierbei sind Bilder abgebildet, mit der jeweils korrekten Anzahl an Silbenbögen zu dem jeweiligen Item, bei denen jedoch nur ein Silbenbogen dick gedruckt wurde (Diehl & Hartke 2012a). Die Aufgabe besteht darin nur die Silbe in dem dicken Silbenbogen zu verschriftlichen. Um Irritationen auf der Seite der Proband:innen zu vermeiden, wurde sich bei der Weiterentwicklung für eine intuitivere Aufgabenstellung entschieden: Die Kinder sollen *alle* Silben in die jeweiligen Silbenbögen schreiben. Die Items bestehen aus einem Übungsitem und anschließend aus drei dreisilbigen Wörtern und drei viersilbigen Wörtern. Um die Items nach Schwierigkeit zu staffeln, wurde sich einerseits auf die Silbenanzahl bezogen und andererseits auf den Erwerb bestimmter Phoneme. Das bedeutet, dass sich die Schwierigkeit mit aufsteigender Silbenanzahl erhöht und zudem die Erwerbsschritte berücksichtigt werden. Marx (2007) stellt heraus, dass „die Komplexität der Silbenstruktur [] in erster Linie durch das Vorkommen von Konsonantenhäufungen bestimmt“ (S. 24) wird. Darüber hinaus ist es von großer Relevanz, ob es sich um offene oder geschlossene Silben handelt, da offene Silben häufig einfacher zu erkennen sind als geschlossene Silben, die vielmals komplexere Konsonantenstrukturen beinhalten (Marx 2007). Dass die Komplexität von Silben den Schriftspracherwerb erheblich beeinflussen, betrachten auch Seymour et al. (2003) in ihrer Studie. Sie betonen, dass Sprachen mit vielen geschlossenen Silben, welche unter anderem oft Konsonantencluster enthalten, sowohl das Erlernen der Phonem-Graphem-Korrespondenz erschweren als auch den Dekodierungsprozess verlangsamen (Seymour et al. 2003). Im Gegensatz dazu fördern Sprachsysteme mit offenen Silben eine schnellere Entwicklung der schriftsprachlichen Kompetenzen (Seymour et al. 2003).

Die Studie von Fitriana & Agustina (2019) setzt sich mit den Erwerbsschritten der Phonologie auseinander und stellt dabei heraus, dass der Erwerb der Vokale zuerst erfolgt. Der Erwerb der Konsonanten erstreckt sich über einen längeren Zeitraum, da insbesondere schwierige Phoneme wie Frikative oder Vibranten eine erhöhte artikulatorische Präzision erfordern und daher später vollständig beherrscht werden. Anschließend werden Konsonantencluster und Diphthonge erworben. Es konnte die Schlussfolgerung getroffen werden, dass Kinder bis zu einem Alter von vier Jahren die meisten Phoneme erworben haben, jedoch einige komplexe Phoneme noch Schwierigkeiten bereiten (Fitriana & Agustina 2019). Auch wenn sich diese Erwerbsprozesse nicht als strikt nacheinander ablaufende Phasen charakterisieren lassen, wird trotzdem eine grundlegende Reihenfolge aufgezeigt, die die Grundlage der folgenden Ausführungen darstellt.

Das Item *Banane* besteht aus drei Silben, die nach dem einfachen Konsonanten-Vokal-Muster aufgebaut sind. Offene Silben und regelmäßige Strukturen sind für Kinder leichter zu segmentieren (Marx 2007). Das Item *Rakete* hat ebenfalls drei Silben, wobei die Silbenstruktur ähnliche wie bei dem Item *Banane* ist. Allerdings stellt es sich als phonologisch etwas anspruchsvoller heraus, da das Phonem [ʁ] herausfordernder ist als das Phonem [b]. Das letzte dreisilbige Item ist *Zitrone*. Die Segmentierung und Verschriftlichung wird an dieser Stelle insbesondere durch den Konsonantencluster [tʁ] erschwert. Das Item *Regenbogen* weist hingegen vier Silben auf, sodass die längere Wortstruktur eine größere phonologische Herausforderung für das Silbensegmentieren darstellt. Des Weiteren erschwert die Wiederholung der Phoneme [gʁ] die Analyse. Bei dem viersilbigen Item *Schokolade* ist der komplexe Anlaut als Konsonantencluster [ʃ] entscheidend, welcher die Schwierigkeit deutlich erhöht. Auch das letzte Item *Badewanne* zeichnet sich aufgrund des Doppelkonsonanten, welcher großen Einfluss auf die Silbensegmentierung hat, durch eine große Komplexität aus.

Dieser Subtest testet die Silbenbewusstheit der Proband:innen. Es wird geprüft, ob Kinder in der Lage sind Wörter in Silben zu segmentieren und darüber hinaus auch die Silbengrenzen zu erkennen. Dies kann häufig herausfordernd sein, denn nach Bredel et al. (2017) ist das Benennen der Silbenanzahl einfacher als die genaue Festlegung von Silbengrenzen. Die Fähigkeit, ein Wort in phonologische Bestandteile wie Silben zu zerlegen, ist Teil der phonologischen Bewusstheit im weiten Sinne. Da die Proband:innen die Silben aufschreiben sollen, spielt ebenfalls das Verständnis der Phonem-Graphem-Korrespondenz eine wichtige Rolle. Auch wenn diese Kompetenzen eindeutig im Fokus dieses Subtests stehen, wird aufgrund der Betrachtung der Orthografie ebenfalls die orthografische Kompetenz getestet. Dies ist bei Silben von großer Bedeutung, da der Frage nachgegangen werden kann, wie beispielsweise mit Doppelkonsonanten umgegangen wird.

Die Basis des vierten Subtests findet sich erneut in dem IEL-1, wobei das Design stark verändert wurde. In der dritten Aufgabe des zweiten Subtests sollen Silbenbögen unter die Wörter gesetzt werden (Diehl & Hartke 2012a). Obwohl es sich hierbei nicht um einen schlecht konstruierten Subtest handelt, stellt der weiterentwickelte Test einen anderen, interessanten Erweiterungspunkt zu dem vorherigen Subtest dar. Die Items wurden durch niedrigfrequente Fremdwörter ersetzt. Diese stammen aus dem PPVT-4⁹ (Lenhard et al. 2015) sowie aus dem OWID Fremdwörterbuch (Leibniz-Institut für Deutsche Sprache o.A.). Fünf Items weisen eine Frequenz von null auf und fünf Items eine Frequenz von eins. Die Staffelung der Items nach Schwierigkeit wurde auf der Basis der Silbenanzahl hergestellt. Ziel war es, ein Design zu kreieren, in dem die Gestalt der Items überrascht, die Silben jedoch nicht kompliziert sind. Ein

⁹ Bei dem PPVT-4 handelt es sich um einen Einzeltest zur Diagnose des rezeptiven (Hör-)Wortschatzes (Lenhard et al. 2015).

Beispiel hierfür stellt das Item *Furage* dar. Auch wenn das Wort den meisten Kindern wahrscheinlich unbekannt ist, lässt es sich aufgrund der offenen Silben und der konsequenten Konsonant-Vokal-Silbenstruktur leicht segmentieren.

Um die Silben dieser niedrigfrequenten Wörter korrekt zu segmentieren, müssen sich die Kinder auf die phonologische Struktur konzentrieren und können sich nicht auf vertraute Wortmuster verlassen. Sie können keine Erwartungen aufbauen, wie es bei bekannten Wörtern der Fall ist (Marx 2007). Es wird die Anforderung an die Proband:innen gestellt, zu erkennen, wie Silben typischerweise in der deutschen Schriftsprache gebildet werden. Schneider (2017) stellt heraus, dass, wenn es sich um die reguläre Orthografie handelt und die Kinder die Phonem-Graphem-Korrespondenz beherrschen, sie die Wörter korrekt rekodieren können. Dieser Subtest liefert folglich Einblicke in das Verständnis von Silbenstrukturen unabhängig von der Bedeutung des Wortes. Demnach testet der Subtest erneut die phonologische Bewusstheit im weiten Sinne und rückt das metalinguistische Bewusstsein in den Fokus, da gezeigt wird, inwieweit Kinder bereits abstrakte Sprachmuster erkannt haben und flexibel auf neue sprachliche Herausforderungen reagieren. Die Fähigkeit mit unbekannten Wörtern umzugehen, gibt Aufschluss über den Schriftspracherwerb, da es ein immer wiederkehrendes Phänomen des Lesens darstellt (Sassenroth 2000).

Die erste Aufgabe des vierten Subtests des IEL-1 zeigt die Basis des fünften Subtests auf. Es wird die Aufgabe gestellt, fehlende Buchstaben in die Lücken von Wörtern zu setzen (Diehl & Hartke 2012a). Der IEL-1 weist von dieser Art des Testdesigns zwei Versionen auf: einmal mit Bildunterstützung und einmal ohne die Verwendung von Bildern (Diehl & Hartke 2012a). Es wurde sich für die Weiterentwicklung für den Einsatz von Bildern entschieden, damit mehrere Optionen von Antwortmöglichkeiten ausgeschlossen werden können. Wie bereits in der kritischen Auseinandersetzung mit dem IEL-1 dargestellt, wird in der Version mit sehr herausfordernden Graphemen gearbeitet. Um diese Komplexität etwas zu reduzieren, wurden bei der Weiterentwicklung die Wörter, die Doppelkonsonanten aufweisen, ausgeschlossen. Des Weiteren sind in der Version des IEL-1 nur Ein- und Zweisilber aufzufinden und es besteht zwischen diesen beiden eine ungleiche Anzahl. Zudem müssen die Grapheme nur als Anlaut und Inlaut eingesetzt werden, sodass die Endposition gänzlich unbeachtet bleibt. Diese Aspekte wurden in der neuen Version grundlegend verändert: Die Items bestehen aus jeweils vier ein-, zwei- und dreisilbigen Worten. Darüber hinaus werden sowohl An- und Inlaute als auch Auslaute gesucht.

Auch in diesem Subtest wurden die Items nach ihrer Komplexität in Bezug auf den Schriftspracherwerb sortiert. Wie in Subtest drei beruhen diese einerseits auf der Anzahl der Silben und andererseits auf den bereits dargestellten Erwerbsstufen, die besagen, dass der Erwerb zunächst mit den Vokalen beginnt, gefolgt von den Konsonanten und schließlich den

Diphthongen sowie Konsonantenclustern (Fitriana & Agustina 2019). Für diesen Subtest ist darüber hinaus noch die von del Toro und Thomé (2021) herausgestellte erhöhte

Schwierigkeit der Doppelkonsonanten relevant. Die Markierung der Vokalkürze durch doppelte Konsonanten muss erlernt werden und stellt in der ersten Klasse häufig noch eine Herausforderung dar (del Toro & Thomé 2021).

Der Fokus dieses Subtests liegt auf der Überprüfung der orthografischen Kompetenz, da die Proband:innen die korrekte Orthografie der Wörter kennen müssen, um die fehlenden Grapheme ergänzen zu können. In diesem Kontext ist besonders hervorzuheben, dass die Phonem-Graphem-Korrespondenz bei diesem Vorgang ausschlaggebend ist. Um die Lücken korrekt zu füllen, müssen die Proband:innen die Beziehung zwischen den Phonemen und den entsprechenden Graphemen verstehen. Die phonologische Bewusstheit spielt hier auch eine Rolle, denn es wird die Anforderung gestellt, die Phoneme der Items in Segmente zu unterteilen, um das fehlende Graphem zu identifizieren.

Die Grundlage für den sechsten Subtest wurde in der zweiten Aufgabe des fünften Subtests gefunden. Die Proband:innen sollen auf der Basis der Bilder die dazugehörigen Wörter schreiben (Diehl & Hartke 2012a). Da das grundlegende Design als sinnvoll für das Testen der orthografischen Kompetenz bewertet wurde, war das vorrangige Ziel der Weiterentwicklung, die Items mit Blick auf die Frequenz anzupassen. Anschließend sollten sie nach ihrer Komplexität auf der Basis ihrer phonologischen, orthografischen sowie silbischen Eigenschaften geordnet werden. Die Items *Nase* und *Rose* haben eine einfache Silbenstruktur mit offenen Silben und enthalten keine orthografischen Besonderheiten, sodass diese beiden Items relativ leicht zu erfassen sind. Anders stellt sich dieser Aspekt bei den Items *Baum* und *Mond* dar, denn diese haben eine geschlossene Silbe. Darüber hinaus findet sich in dem Wort *Baum* ein Diphthong und in dem Wort *Mond* eine Auslautverhärtung, welche zusätzliche Herausforderungen bilden. Bei dem Item *Krone* findet sich am Wortanfang der Konsonantencluster, was die Segmentierung erschwert. Das Item *Löffel* besitzt neben den geschlossenen Silben und des Konsonantenclusters einen Umlaut, welcher für Kinder oft schwerer zu verschriftlichen ist. Das Item *Erdbeere* enthält ebenfalls mehrere geschlossene Silben, den Doppelkonsonanten sowie den Konsonantencluster. Aufgrund dessen erfordert dies die sichere Kenntnis über orthografische Konventionen und daraus entstehende Silbenverbindungen. Das letzte Item *Hubschrauber* enthält einen sehr komplexen Konsonantencluster sowie einen Diphthong, weshalb dieses Item eine hohe Schwierigkeit aufweist.

Dieser Subtest stellt ebenfalls die Testung der orthografischen Kompetenz in den Mittelpunkt der Betrachtung. Getestet wird, inwiefern die Proband:innen die korrekte Orthografie der Wörter beherrschen. Das beinhaltet die Kenntnis von orthografischen Regeln und Besonderheiten. Auch in diesem Subtest spielt die Phonem-Graphem-Korrelation eine bedeutende Rolle.

Das grundlegende Prinzip hinter dem siebten Subtest basiert auf der dritten Aufgabe des sechsten Subtests des IEL-1, in welcher die Proband:innen einen Satz lesen und nachfolgend ein Bild zu sehen ist. Die Aufgabe besteht darin, dem Satz die Information zu entnehmen, welche auf dem Bild noch nicht zu sehen ist und dieses Detail malend zu ergänzen (Diehl & Hartke 2012a). Eine Herausforderung dieses Designs bei einer Anwendung in der ersten Klasse stellt allerdings das Malen dar. Wenngleich der Kerngedanke sinnvoll ist, um das Satzverständnis zu testen, ohne dass dafür geschrieben werden muss, kann es sein, dass das Malen bei vielen Kindern viel Zeit einnimmt. Um den Zeitaspekt zu reduzieren, wurde sich im Kontext der Weiterentwicklung dafür entschieden, dass die Sätze mit Playmobil nachgestellt werden. Kauschke (2012) hat die Vorteile solcher Methoden des Ausagierens herausgearbeitet und indes betont, dass diese „einen Einblick in die Verständnis- und Interpretationsstrategien des Kindes“ (S. 16) gewähren. Darüber hinaus zeichnen sich diese Formate durch eine hohe Motivationsfähigkeit aus (Kauschke 2012). Dies geht damit einher, dass es keinen Testbogen gibt, sondern nur einen Protokollbogen für die testleitende Person. Der Fokus liegt weiterhin auf den Präpositionen. Den Proband:innen wird eine Auswahl an verschiedenen Playmobil-Teilen dargeboten, inklusive verschiedener Varianten einer Sache. Dies ist wichtig, da die Proband:innen auf diese Weise ihre Fähigkeit unter Beweis stellen sollen, gelesene Informationen korrekt zu erfassen und umzusetzen. Indem ihnen verschiedene Varianten einer Sache, wie unterschiedlich gefärbte Hunde oder verschieden große Stühle, zur Verfügung gestellt werden, wird sichergestellt, dass sie den Satz präzise interpretieren und die beschriebenen Szenen entsprechend nachbauen können. Um dies näher zu erläutern, kann das zweite Item herangezogen werden: *Der Schal liegt auf dem kleinen Stuhl*. Zuerst muss das Subjekt identifiziert werden und anschließend das Objekt. In dem Fall gehört zu dem Objekt noch ein Adjektiv, welches beachtet werden muss. Nun müssen anhand der restlichen Informationen das Subjekt und das Objekt in Beziehung gesetzt werden. Dazu gehört das Miteinbeziehen des Verbes sowie die adverbiale Bestimmung des Ortes. Nur wenn all diese Faktoren beachtet wurden, kann die Aufgabe korrekt gelöst werden. Dies ermöglicht eine gezielte Überprüfung des Leseverständnisses und der Übertragung in eine praktische Handlung.

Darüber hinaus wurde bereits in der kritischen Auseinandersetzung mit dem Test herausgearbeitet, dass die im IEL-1 genutzte vorangestellte Dativkonstruktion sehr herausfordernd ist. Aufgrund dessen wurde die Komplexität der Sätze angepasst, sodass diese fortlaufend ansteigt. Der Subtest beinhaltet zu Beginn ein Übungsitem, welches die Grundstruktur (Subjekt-Verb-Objekt) besitzt. Diese Grundstruktur findet sich im nächsten Item wieder, wobei sie um ein Adjektiv ergänzt wird. Das nächste Item zeichnet sich durch die vorangestellte Dativkonstruktion und dem Einsatz eines Adjektivkompositums aus. Auch wenn das letzte Item dieses Subtests ebenfalls eine vorangestellte Dativkonstruktion beinhaltet, enthält dieser Satz mehrere Nomina im Plural mit unterschiedlichen Attributen. Die Pluralität erfordert eine korrekte

semantische Verknüpfung der Mengenangaben mit den Adjektiven und dem Substantiv. Durch die konjunktive Verbindung zwischen den attributiven Mengenangaben, müssen die Proband:innen Verknüpfungen und Referenzen ziehen. Das Prädikat steht im Plural, was die grammatische Aufmerksamkeit besonders erfordert. Die Kongruenz zwischen dem Subjekt und dem Verb ist nicht direkt offensichtlich.

Der Subtest testet das semantische Verständnis auf Satzebene. Die Proband:innen müssen die Bedeutungen der Wörter und des gesamten Satzes erfassen, um die richtigen Figuren auszuwählen und sie in sinnvoller Weise anzuordnen. Bredel et al. (2017) stellen die Herausforderung des Satzlesens heraus, wenn Kinder die einzelnen Wörter nicht zu Wortgruppen zusammenfassen und aufgrund dessen an der korrekten Erfassung des Gesamtsatzes scheitern. In diesem Kontext ist auch das Syntaxverständnis involviert, da die grammatischen Strukturen des Satzes verstanden werden müssen, um die richtige Anordnung der Figuren und Gegenstände nachvollziehen zu können (Kauschke 2012). Das beinhaltet das korrekte Identifizieren von Subjekt, Prädikat und Objekt sowie das anschließende Erkennen von Beziehungen zwischen diesen Elementen.

Der letzte Subtest basiert auf der ersten Aufgabe des siebten Subtests des IEL-1, in dem die Anforderung an die Proband:innen gestellt wird, einen Text zu lesen und anschließend aus drei Bildern das zu dem Text passende zu finden (Diehl & Hartke 2012a). Innerhalb der Items, welche jeweils aus vier Sätzen bestehen, ist stets eine Information die einzig entscheidende, um das richtige Bild auszuwählen. Die Bilder unterscheiden sich folglich nur in einem Aspekt. Das grundlegende Design dieses Subtests wurde übernommen, wobei die Darstellungsweise angepasst wurde. So wurden in der Version des IEL-1 die Sätze untereinander aufgeführt. Die Pilotierungsergebnisse haben jedoch eindeutig gezeigt, dass es sich dabei um eine unnatürliche Präsentation handelt, da die Proband:innen die einzelnen Sätze mit den Bildern verbunden haben, was zur Folge hatte, dass diese Antworten nicht auswertbar gewesen sind. Die Art des Designs hat offensichtlich dazu geführt, dass die Sätze nicht zusammenhängend als Text wahrgenommen wurden. Daraus folgte die Schlussfolgerung, die Sätze aneinanderzureihen, um der üblichen Darstellung eines Textes näher zu kommen. Des Weiteren hat die Pilotierung Ergebnisse dazu geliefert, dass manche Dinge auf den Bildern uneindeutig waren. So steht in dem ersten Item, dass das Spielzeug in der Kiste liegt, wobei auf den Bildern zu sehen ist, dass das Spielzeug auch vor der Kiste liegt. Daraus folgerte das getestete Kind, dass keins der Bilder richtig sei. Aufgrund dessen wurden die Bilder noch einmal überarbeitet, um solchen Uneindeutigkeiten entgegenzuwirken.

Das lokale Textverständnis steht bei diesem Subtest im Vordergrund. Der Test prüft, inwiefern die Proband:innen spezifische Informationen im Text erkennen und diese mit den Bildern verknüpfen können. Da sich die Bilder nur in einem Aspekt unterscheiden, müssen sie genau

lesen. Auch wenn das globale Textverständnis an dieser Stelle nicht vorrangig ist, lässt sich festhalten, dass die Proband:innen dennoch den gesamten Text verstehen müssen, um den Kontext des Details zu erfassen. Das globale Textverständnis ermöglicht es den Kindern, die Bedeutung des Textes als Ganzes zu begreifen und zu erkennen, wie die Informationen in den Gesamtzusammenhang passen. Der Miteinbezug von Vor- und Weltwissen ist an dieser Stelle jedoch nicht gefragt.

Zusammenfassend lässt sich festhalten, dass die Entwicklung des neuen Lesetests auf einer sorgfältigen Analyse und Weiterentwicklung bestehender Testverfahren basiert. Durch die Fokussierung auf hochfrequente Wörter, die Berücksichtigung linguistischer Komplexitäten sowie die Staffelung nach Schwierigkeitsgraden wurde ein Instrument geschaffen, das eine differenzierte Erfassung der Kompetenzen im Schriftspracherwerb ermöglicht.

5.4 Durchführung des Tests

Der Test ist für einen Einsatz in der ersten Klassenstufe konzipiert worden. Darüber hinaus wurde sich von einer festen Zeitvorgabe für die jeweiligen Subtests bewusst distanziert, da dies zu einer Drucksituation führen kann, welche die Leistungen der Proband:innen negativ beeinflussen kann.¹⁰ Dies wird den Proband:innen auch immer wieder in Erinnerung gerufen. Es wird Abstand von der Annahme genommen, dass eine schnelle Bearbeitungszeit einer Aufgabe gleichzusetzen ist mit einer qualitativ guten Bearbeitung.

Es handelt sich um eine Einzeltestung, welche in einem separaten und vorbereiteten Raum stattfinden soll, sodass alle notwendigen Materialien in Einsatznähe verfügbar sind. Bei dem Raum sollte es sich um einen ruhigen Raum mit guten Lichtverhältnissen handeln. Auf der einen Seite sollte gewährleistet werden, dass die Proband:innen nicht durch Geräusche, beispielsweise von Mitschüler:innen, abgelenkt werden. Auf der anderen Seite ist es wichtig, die in dem Test vorkommenden Audios in angemessener Lautstärke abspielen zu können.

Vor der Testdurchführung gilt es folgende Materialien bereitzulegen:

- 1) pro Testung einen Testbogen und einen Protokollbogen, die vorab ausgedruckt werden
- 2) zwei Stifte (davon einen Ersatzstift)
- 3) Testinstruktionen
- 4) Playmobil (für den siebten Subtest zum *Satzverständnis*)
- 5) Endgerät zum Abspielen der Audios

¹⁰ Wenngleich Studien aufzeigen, dass die Lesegeschwindigkeit mit fortschreitender Schulzeit, insbesondere in den ersten Schuljahren, ansteigt (z.B. Alves et al. 2021), messen diese Studien lediglich die Dekodierfähigkeit (Schneider 2017). Studien belegen, dass Zeitdruck die Leistungen negativ beeinflussen kann (z.B. Rojo López et al. 2021). Aufgrund der Basis dieser bestehenden Korrelation wird sich in dieser Arbeit von dem Einsatz von Zeitlimits distanziert.

Die Aussagekraft des Tests kann nur dann gewährleistet werden, wenn dieser unter möglichst vergleichbaren Bedingungen durchgeführt wird (Lenhard 2019). Die folgenden Punkte nehmen dafür eine besonders bedeutsame Rolle ein und sollten zwingend beachtet werden. Zunächst sollten die Instruktionen möglichst wortwörtlich wiedergegeben werden. Es dürfen keine zusätzlichen Hilfestellungen gegeben werden. Dies bedeutet auch, dass die Reihenfolge der Subtests nicht verändert werden darf. Des Weiteren sollte davon abgesehen werden, dass für die Proband:innen ersichtlich wird, ob sie richtige oder falsche Antworten gegeben haben. Dies dient einerseits dazu, mögliche Schlussfolgerungen für die nachfolgenden Items zu vermeiden und andererseits dazu, die Motivation der Proband:innen aufrechtzuerhalten. Werden diese Aspekte nicht beachtet, leidet die Vergleichbarkeit unter den Proband:innen und demnach auch die Aussagekraft der Testergebnisse maßgeblich (Lenhard 2019).

Darüber hinaus ist es von großer Tragweite und bestenfalls auch mit der Klassenlehrkraft und/oder den Erziehungsberechtigten abzusprechen, dass gegenüber den Kindern nicht der Begriff *Test* genutzt wird. Dies kann große Ängste schüren, immensen Druck aufbauen und im Zweifel die Durchführbarkeit des Tests stark negativ beeinflussen. Eine Studie von Abe und Gbenro (2024) zeigte, dass die Wahrnehmung mathematischer Aufgaben innerhalb eines Tests durch die Schüler:innen das Ausmaß ihrer Prüfungsangst beeinflusst. Dies deutet darauf hin, dass die Präsentation eines Tests die Prüfungsangst der Teilnehmenden beeinflussen und somit in direktem Zusammenhang mit deren Leistung stehen kann. Somit gilt es alternative Begriffe wie beispielsweise *Übungen zum Lesen* zu verwenden. Weiterhin ist es bedeutsam den Kindern zu vermitteln, dass nicht die Bewertung der Leistung fokussiert wird, sondern es darum geht einen Einblick darüber zu erlangen, was die Proband:innen bereits gelernt haben.

5.4.1 Durchführungsinstruktionen

Im Folgenden werden die Instruktionen für die Durchführung tabellarisch dargestellt. Mit diesen Durchführungshinweisen sollte sich die testleitende Person frühzeitig vertraut machen, da von einem aktiven und starren Ablesen abzuraten ist, um eine möglichst stressfreie und angenehme Testsituation zu erschaffen. Die Instruktionen sind aufgeteilt in zwei Bereiche: Die linke Spalte besteht aus Hinweisen für die testleitenden Personen, in denen wichtige und weiterführende Erläuterungen vorzufinden sind, welche jedoch nicht vorzulesen sind. In der rechten Spalte sind die Instruktionen aufgeführt, welche den Proband:innen möglichst wörtlich wiedergegeben werden sollen.

Tabelle 1: Durchführungsinstruktionen

Hinweise für die Testleitenden	Wörtliche Instruktion
Einstieg	
Anmerkung vorab für die generelle Durchführung: Es wird immer nur der zu bearbeitende Subtest herausgegeben, um zu verhindern, dass die Kinder blättern Vermeiden des Begriffs „Test“	Wir brauchen deine Hilfe, um zu verstehen, wie Kinder lesen lernen. Du lernst das ja gerade. Kannst du mir dabei helfen? Ich habe ein paar Übungen zum Lesen mitgebracht. Du kannst ganz langsam machen. Nimm dir so viel Zeit wie du willst.
Untertest 1: Bestimmung der Vokallänge	
Bei diesem Subtest wird keine Testseite an die Kinder herausgegeben, sondern es handelt sich um einen verbal-sprachgebundene Subtest Bereitlegen des Protokollbogens 2x abspielen möglich Audio abspielen → Nr. 60: längerer Laut gesucht: [a] Lösung: saan → Nr. 61: kürzerer Laut gesucht: [i] Lösung: nitz Positives Feedback zu jedem Beispiel geben	Wir machen jetzt ein Spiel, bei dem du ganz genau hinhören musst. Es werden dir Wörter vorgespielt, die in der Mitte den gleichen Vokal haben. Weißt du, was ein Vokal ist? Die Vokale sind A, E, I, O, U. Du hörst gleich Quatschwörter, also wunder dich nicht, dass du die Wörter nicht kennst. Du hörst gleich eine Reihe von vier Wörtern. Eins hört sich ein bisschen anders an. Kannst du mir sagen welches? Lass uns das mal probieren: <i>fatt-black-saan-wast</i> Welches Wort ist anders? <i>miel-nitz-spien-rien</i> Welches Wort ist anders? So machen wir jetzt weiter.
Untertest 2: Buchstaben-Laut-Zuordnung	
Austeilen der ersten Testseite sprachliche Hinweise durch Zeigen auf dem Testbogen begleiten	Auf der Seite siehst du verschiedene Reihen. In den Reihen sollst du immer auf einen bestimmten Vokal achten. Du bist ja jetzt schon Vokalprofi! In den Reihen siehst du immer viele Bilder und darunter kleine Kästchen. Dir wird gleich ein Wort vorgespielt, das zu dem Bild passt. Du kreuzt bitte dann das Kästchen an, an welcher Stelle du den gesuchten Vokal hörst. Ganz oben siehst du drei Beispiele. Hier geht es um den Vokal E. Lass uns die einmal gemeinsam durchgehen: Was siehst du auf dem ersten Bild? <i>DELFIN</i> – Das „E“ hört man in der Mitte, deshalb ist im 2. Kästchen ein Kreuz. Was siehst du auf dem zweiten Bild? <i>KIRSCHEN</i> – Das „E“ hört man am Ende, deshalb ist im letzten Kästchen ein Kreuz.

Audios einzeln abspielen und warten bis ein Kreuz gesetzt wurde	Was siehst du auf dem dritten Bild? <i>ERDE</i> – Das „E“ hört man am Anfang und am Ende. Deswegen sind hier zwei Kreuze. Nun löst du die anderen Aufgaben. Lass dir Zeit so viel du willst. Wir starten in der ersten Reihe und fangen ganz links an. An welcher Stelle hörst du ein „O“? OMA AUTO OTTO BOOT Gut gemacht! Jetzt sind wir in der zweiten Reihe und fangen wieder ganz links an. Wo hörst du ein „A“? AFFE ZAHN SOFA ANNA Super! Jetzt sind wir in der dritten Reihe und fangen wieder ganz links an. Wo hörst du ein „I“? KNIE IGEL KIWI BRIEF Gut gemacht! Jetzt sind wir in der vierten Reihe und fangen wieder ganz links an. Wo hörst du ein „E“? EMIL STEG SCHNEE ESEL Super! Jetzt sind wir in der letzten Reihe und fangen wieder ganz links an. Wo hörst du ein „U“? BUCH UHU UHR AKKU Gut gemacht!
Untertest 3: Wörter in Silben segmentieren <i>Silbe in die Silbenbögen schreiben</i>	
Austeilen der zweiten Testseite Methodik „Silbenbögen“ sollte bekannt sein, aber zur Reaktivierung des Wissens nachfragen Silben klatschen zulassen Am Ende der Übung positives Feedback geben	Bei dieser Übung siehst du wieder verschiedene Bilder und Silbenbögen. Silbenbögen kennst du sicherlich aus der Schule. Kannst du mir das einmal an dem ersten Beispiel erklären? <i>Optional und abhängig von der Antwort des Kindes:</i> <i>Auf dem Bild ist eine HOSE zu sehen. Du siehst daneben zwei Silbenbögen, weil das Wort HO SE aus zwei Silben besteht. Die erste Silbe ist „Ho“, deswegen steht sie im ersten Silbenbogen. Die zweite Silbe ist „se“, deswegen steht sie im zweiten Silbenbogen.</i> Auch bei dieser Übung kannst du dir Zeit lassen. Du kannst jetzt starten.
Untertest 3: Wörter in Silben segmentieren <i>Silbenbögen unter die Wörter zeichnen</i>	
Austeilen der dritten Testseite Silben klatschen zulassen (laut) Vorlesen zulassen Am Ende der Übung kurz auf Eindeutigkeit der Silbengrenzen prüfen und ggf. nachfragen Am Ende der Übung positives Feedback geben	Bei dieser Übung siehst du wieder viele Quatschwörter. Es ist nicht schlimm, wenn du sie nicht kennst. Du sollst die Wörter lesen und dann die Silbenbögen unter die Silben setzen. Lass uns einmal das Beispiel anschauen. Kannst du mir einmal vorlesen was da steht? <i>PANTALON</i>. PAN TA LON hat drei Silben. Deswegen sind unter dem Wort drei Silbenbögen. Achte darauf, dass du die Bögen genau unter die Silben setzt. Lass dir so viel Zeit wie du brauchst. Du kannst jetzt starten.

Untertest 4: Buchstaben zu Wörtern verbinden	
Austeilen der vierten Testseite	In dieser Übung siehst du wieder viele Bilder und darunter ein passendes Wort, bei dem aber ein paar Buchstaben verloren gegangen sind. Du sollst die fehlenden Buchstaben wieder einsetzen, damit das Wort wieder ganz ist. Wie hier in den Beispielaufgaben: Was siehst du auf dem ersten Bild? <i>MANN</i> - Darunter siehst du „Ma“ und einen Strich daneben. Was fehlte da? Genauso ist es bei dem anderen Bild. Was siehst du da? <i>ELEFANT</i> - Darunter steht „Ele“, dann kommt ein Strich und „ant“. Was fehlte hier, um das Wort wieder ganz zu machen?
Am Ende der Übung positives Feedback geben	Jetzt bist du dran. Nimm dir die Zeit, die du brauchst zum Überlegen.
Untertest 5: Wörter	
Austeilen der fünften Testseite	Auf dieser Seite siehst du wieder viele verschiedene Bilder. Du sollst zu den Bildern die passenden Wörter schreiben. Kannst du mir das an dem Beispiel einmal erklären?
Am Ende der Übung positives Feedback geben	Das machst du jetzt mit den anderen Bildern genauso. Lass dir Zeit.
Untertest 6: Satzverstehen	
Austeilen der sechsten Testseite	Bei dieser Aufgabe darfst du etwas mit Playmobil bauen. Lies zuerst den Satz laut vor. Danach sollst du das mit Playmobil nachbauen. Baue nur das nach, was in dem Satz gesagt wird.
Playmobil bereitlegen	Wir machen den ersten Satz einmal gemeinsam. Kannst du einmal bitte vorlesen, was da steht? <i>DER MANN STEHT NEBEN DEM BAUM</i> . Was brauchen wir also an Playmobil dafür? In dem Satz steht, dass der Mann <i>NEBEN</i> dem Baum steht. Wie müssen wir das jetzt hinstellen?
Auf den jeweiligen Satz einmal zeigen	Jetzt bist du dran! Lies bitte den nächsten Satz laut vor. Kannst du das mit Playmobil nachbauen? Lies jetzt den nächsten Satz laut vor. Bau es jetzt nach.
Am Ende jeder Übung positives Feedback geben.	Jetzt kannst du den letzten Satz vorlesen. Bau es bitte mit Playmobil nach.
Untertest 7: Textverstehen	
Austeilen der siebten Testseite	Auf dieser Seite siehst du Kästchen mit drei Bildern und einem Text darin. Du sollst das richtige Bild zum Text suchen und die beiden verbinden. Lies den Text für dich. Schau dir die Bilder dazu genau an. Nur ein Bild passt.
Laut vorlesen zulassen	
Am Ende jeder Übung positives Feedback geben	Du kannst jetzt starten.
Abschluss	
	Wir sind fertig! Du hast super gearbeitet. Danke, dass du uns geholfen hast!

5.5 Kodierungssystem für die Auswertung

Ein eindeutiges Kodierungssystem reduziert subjektive Interpretationen, sodass Fehler bei der Auswertung minimiert werden können. Dies allein ermöglicht die nötige Vergleichbarkeit zwischen den Proband:innen (Marx 2007). Des Weiteren gewährleistet es Transparenz, da die Bewertungsentscheidungen klar nachvollziehbar sind. Ohne ein klar definiertes Kodierungssystem kann nicht sichergestellt werden, dass die Ergebnisse tatsächlich die beabsichtigten Konstrukte messen und nicht durch uneinheitliche Bewertung verzerrt werden. Die Art und Weise wie die Punkte verteilt werden hat einen starken Einfluss auf die Ergebnisse, weshalb das Treffen dieser Entscheidungen nicht unreflektiert und aus mehreren Perspektiven erfolgen sollte. Aufgrund dessen ist es sinnvoll, stets alternative Antwortmöglichkeiten der Proband:innen zu antizipieren, die ebenfalls als richtig gewertet werden sollten. Es ist ein bedeutsamer, aber komplexer Teil der Testentwicklung, ein Kodierungssystem zu entwerfen aus dem unmissverständlich hervorgeht, was als richtig und was als falsch bewertet wird und welche Konsequenzen daraus für die Bepunktung gezogen werden. Was dies konkret für den dargestellten Test bedeutet, wird im Folgenden anhand der einzelnen Subtests näher erläutert.¹¹

Die Bepunktung bei Subtest eins ergibt sich aus der Anzahl der richtig gesetzten Kreuze. Hierbei gilt, dass jedes richtig gesetzte Kreuz einen Punkt und jedes falsch gesetzte Kreuz null Punkte bedeuten. Demnach ergeben Items, wie beispielsweise Qma, eine Höchstpunktzahl eines Punktes, während Items wie beispielsweise Anna eine Höchstpunktzahl von zwei Punkten aufweisen. Da dieser Subtest auf eindeutig richtig oder falsch gesetzten Kreuzen und daraus folgend auf einem eindeutigen richtigen oder falschen Identifizieren der Phoneme baut, gibt es keine abweichenden Antworten der Proband:innen, welche als richtig bewertet werden können. Die Pilotierung hat jedoch aufgezeigt, dass dieses Aufgabenformat den Proband:innen aus der Schule bekannt ist, sie jedoch häufig den Buchstaben statt Kreuze in das Kästchen schreiben. Aufgrund dessen wird jeder richtig gesetzter Buchstabe ebenfalls mit einem Punkt bewertet.

Das System in Subtest zwei ergibt sich aus einer ähnlichen Systematik, wobei der Unterschied sich daraus ergibt, dass dieser verbal-sprachgebunden funktioniert, sodass die Proband:innen selber keine schriftlichen Antworten geben. Stattdessen geben sie mündliche Antworten und die testleitende Person setzt das Kreuz bei dem Wort, welches von den Teilnehmenden genannt wird. Daraus ergibt sich erneut die Regel für die Auswertung, dass ein richtiges Kreuz ein Punkt und jedes falsche Kreuz null Punkte bedeutet. An dieser Stelle kommt jedoch die Frage auf, welche Antworten der Proband:innen als richtig bewertet werden, sodass das Kreuz an die Stelle gesetzt werden kann. Dies wird an einer beispielhaften Erläuterung an dem ersten Item des Subtests deutlich: *maar, raas, dack, laaf*. Die Musterlösung für dieses Item ist die

¹¹ Die tabellarische Darstellung des Kodierungssystems ist im Anhang (siehe Anhang 3) zu finden.

verbale Äußerung des Begriffs *dack*. Geben die Teilnehmenden jedoch an, dass sich der gesuchte Begriff beispielsweise an dritter oder vorletzter Position befindet, ist dies ebenfalls als korrekt zu werten. Dies bedeutet folglich, dass die Angabe der Position des richtigen Begriffs ebenfalls benannt werden darf. Da es sich bei den Begriffen um Pseudowörter handelt, ist dies von besonders großer Relevanz, da die Teilnehmenden aufgrund der höheren Belastung des Arbeitsgedächtnisses größere Schwierigkeiten haben, sich die Wörter exakt zu merken. Bei dem Subtest liegt jedoch der Schwerpunkt nicht auf dem Merken von Wörtern, sondern auf den Kompetenzen, die benötigt werden, um die unterschiedlichen Vokallängen zu identifizieren.

Für das Kodierungssystem von dem dritten Subtest hat sich die Pilotierung als besonders gewinnbringend erwiesen. Bei diesem Subtest liegt die Aufmerksamkeit auf der Silbenstruktur, sodass zu Beginn festgehalten worden ist, dass jede richtig geschriebene Silbe einen Punkt ergibt. Während der Auswertung der Pilotierungsphase hat sich jedoch die Frage ergeben, wie Antworten von Proband:innen gehandhabt werden, die die Silbe orthografisch nicht richtig geschrieben haben, jedoch die Silben richtig getrennt haben. Auf der einen Seite zeigt sich an der Stelle, dass der Fokus des Subtests, nämlich das Verständnis von Silbenstrukturen, verstanden wurde. Auf der anderen Seite ist es aufgrund des Aspektes der Orthografie nicht als gleichwertig zu bezeichnen im Vergleich mit Antworten von Teilnehmenden, die neben der richtigen Silbenstruktur auch diese Silben orthografisch korrekt geschrieben haben. Wie bereits dargestellt kann die Orthografie wertvolle Hinweise auf den kindlichen Zugriff auf die Schriftsprache liefern (Schröder-Lenzen 2004). Die Lösung für dieses Dilemma lässt sich in einer Zweiteilung der Punkte finden. Daraus folgt eine eigenständige Bepunktung mit Blick lediglich auf die Silbenstruktur, wobei jede richtig identifizierte Silbe einen Punkt und jede falsch identifizierte Silbe null Punkte ergibt. In einem zweiten Schritt wird die Orthografie in den Blick genommen, sodass jede orthografisch richtig geschriebene Silbe einen Punkt und jede orthografisch falsch geschriebene Silbe null Punkte bedeutet. Da die Basis dieses Subtests das Erkennen der Bilder bildet, ist es von Bedeutung etwaige alternative Begriffe für das Bild zu antizipieren. So wird bei dem zweiten Item *Rakete* ebenfalls die Mehrzahl *Raketen* akzeptiert sowie der Begriff *Raumschiff*. Da es für die Gleichwertigkeit relevant ist, dass diese Begriffe eine gleiche maximale Höchstpunktzahl ergeben, es sich hierbei jedoch unterschiedliche Silbenanzahlen ergeben, gilt es bei dem Begriff *Raumschiff* für jede Silbe 1,5 Punkte anstelle von einem Punkt zu geben. Ähnlich ist es bei dem sechsten Item *Badewanne*. Begriffe, welche ebenfalls als richtig zu werten sind, sind *Dusche* sowie *Wanne*. An dieser Stelle ergibt sich erneut die Herausforderung der abweichenden Silbenzahlen im Vergleich zu der Musterlösung, sodass bei den beiden Begriffen zwei statt einem Punkt pro Silbe gegeben wird.

Da die Proband:innen bei dem vierten Subtest eingeständig Silbenbögen setzen, dienen diese als Grundlage der Bepunktung. Daraus folgt, dass jeder richtig gesetzte Silbenbogen mit einem Punkt gleichzusetzen ist und im Umkehrschluss jeder falsch gesetzte Silbenbogen mit null Punkten bewertet wird. Eine Besonderheit ergibt sich bei dem dritten Item *Epigone*, da hier sowohl die Silbentrennung *Epi-go-ne* als auch *E-pi-go-ne* als richtig bewertet wird, da beide Varianten linguistisch nachvollziehbar sind, wenngleich sie auf unterschiedlichen sprachlichen Prinzipien basieren. Die Trennung *E-pi-go-ne* orientiert sich an der phonotaktischen Regel, dass Silben so getrennt werden, dass jede Silbe einen Vokal enthält. Diese Segmentierung stellt eine intuitive Aufteilung basierend auf der gesprochenen Sprache dar und spiegelt die natürliche Artikulation wider. Die Silbensegmentierung *Epi-go-ne* berücksichtigt die morphologische Struktur des Wortes, indem das Präfix *Epi-* als Einheit belassen wird. Da der Test die Kompetenzen zur Silbenerkennung in den Vordergrund rückt und nicht die morphologische Analyse und das Kennen von orthografischen Konventionen, werden beide Varianten als richtig bewertet.

Im fünften Subtest stellen die von den Proband:innen eingesetzten Buchstaben das Fundament für die Bepunktung dar. Dabei ergibt jeder richtig gesetzte Buchstabe einen Punkt. Da in diesem Test die orthografischen Kompetenzen im Fokus sind, werden Grapheme, die sich an der Lautsprache orientieren, als falsch und demnach mit null Punkten bewertet. Dies bedeutet, dass das Einsetzen des Graphems <eu> in *Flugz__g* mit zwei Punkten bewertet wird, während das Einsetzen der Grapheme <oi> mit null Punkten bewertet wird. Darüber hinaus bedeutet jeder falsche zusätzliche Buchstabe einen Minuspunkt. Dieses Vorgehen wird bei der Betrachtung des Beispiels *L__ter* erkennbar, welches korrekterweise mit dem Graphem <ei> vervollständigt werden müsste. Tragen Proband:innen stattdessen die Grapheme <ai> ein, welches sich offensichtlich an der Lautsprache orientiert, gibt es null Punkte für das Graphem <a> und einen Punkt für das Graphem <i>. Eine Ausnahme besteht bei dem Item *Heft*, bei dem das Graphem <t> korrekterweise eingesetzt werden soll. Wird stattdessen <te> eingesetzt, sodass das Item *Hefte* ergibt, wird das ebenfalls mit einem Punkt und folglich als gleichermaßen richtig bewertet. Ein ähnlicher Fall resultiert aus dem Item *Fernsehe__*, wo intendiert wurde das Graphem <r> einzusetzen, allerdings wird ebenfalls das Einsetzen des Graphems <n> als richtig bewertet.

Da im sechsten Subtest der Schwerpunkt der Betrachtung auf der orthografischen Kompetenz liegt, werden lautsprachliche Schreibweisen nicht als richtig bewertet. Um die Auswertung zu erleichtern, wird sich an der Silbenstruktur der Begriffe orientiert, sodass sich die Kodierungsregel ergibt, dass jede orthografisch richtig geschriebene Silbe einen Punkt und jede orthografisch falsch geschriebene Silbe null Punkte bedeutet. Da in diesem Subtest erneut die Interpretation der Bilder als Grundlage der Antworten der Proband:innen dienen, kommt die Frage

nach alternativ richtigen Antwortmöglichkeiten auf. So ist bei dem zweiten Item die Musterlösung *Rose*, wobei die Begriffe *Tulpe* und *Blume* ebenfalls als richtig bewertet werden. Bei dem letzten Item des Subtests *Hubschrauber* gelten die Antworten *Helikopter*, *Flugzeug* und *Flugschrauber* ebenfalls. Durch die abweichende Silbenanzahl werden die Punkte für die einzelnen Silben angepasst.

Subtest sieben basiert, ähnlich wie der zweite Subtest, nicht auf einer schriftlichen Antwort der Proband:innen. Während diese den Inhalt der Sätze mit Playmobil nachbauen, füllt die testleitende Person den Protokollbogen aus. Der Protokollbogen ist so aufgebaut, dass der Satz in seine Einzelteile segmentiert wurde und die Items jeweils eine maximale Höchstpunktzahl von vier Punkten aufweisen. Die Proband:innen können für jedes Item einen Punkt für die Wahl des richtigen Subjekts, die Wahl des richtigen Objekts, die Beachtung des Adjektivs und die Beachtung der adverbialen Bestimmung des Ortes erlangen. Auf dem Bogen wird ein Haken für die Aspekte gegeben, die erfüllt worden sind und ein Kreuz für welche, die nicht beachtet worden sind. Daraus ergibt sich die Kodierungsregel, dass jeder Haken mit einem Punkt zu bewerten ist und jedes Kreuz mit null Punkten gleichzusetzen ist.

Bei dem achten Subtest sollen die Proband:innen den Text mit dem dazugehörigen Bild verbinden. Hierbei gibt es für jedes richtige Verbinden einen Punkt.

Abschließend ist anzumerken, dass solche Kodierungssysteme nicht überinterpretiert werden dürfen. Dies bedeutet, dass schnell Gefahr gelaufen wird, Antworten von Proband:innen unreflektiert um jeden Preis in die Kategorien *richtig* oder *falsch* zu einzuordnen, ohne dabei ausreichend auf Zwischennuancen oder Kontextfaktoren einzugehen.

6. Methode

Im Anschluss an die Analyse und Entwicklung des neuen Testinstrumentes, wird der Frage nachgegangen, inwiefern dieser Test die Kompetenzen im Schriftspracherwerb bei Kindern in der ersten Klasse misst. Hierfür wurde eine Studie entworfen, deren Zielsetzung im Anschluss ausgeführt wird. Nachfolgend wird die Stichprobe und das eingesetzte Material beschrieben, woraufhin das Vorgehen erläutert wird.

6.1 Zielsetzung der Studie

Übergeordnet setzt sich diese Studie zum Ziel, den Test als Instrument zu evaluieren. Es sollen Antworten auf die Frage gefunden werden, ob der Test tatsächlich das misst, was er vorgibt zu messen: die schriftsprachlichen Kompetenzen in den Bereichen der phonologischen Bewusstheit, der Orthografie und der Semantik. Daraus lässt sich ableiten, dass unter anderem die Testdurchführung in den Mittelpunkt der Betrachtung gerückt werden muss. Hierbei geht es um die Frage nach der praktischen Anwendbarkeit des Tests. Die Ergebnisse sollen

Hinweise und Anhaltspunkte dazu liefern, den Test weiterzuentwickeln und fundiert anzupassen. Prediger et al. (2012) stellen an die fachdidaktische Entwicklungsforschung das Ziel gegenstands-, prozessorientiert, iterativ sowie vernetzt vorzugehen. Dies schließt ein zirkuläres Vorgehen ein. Zu diesem Prozess gehört das Strukturieren und Spezifizieren von Lerngegenständen, indem der fachliche Hintergrund durchleuchtet wird sowie gegenstandsspezifische Lernziele festgelegt werden (Prediger et al. 2012). Anschließend wird das Design entwickelt, in dessen Zuge empirische und theoretische Lernverläufe gesichtet und sich mit typischen Hürden auseinandergesetzt werden (Prediger et al. 2012). Nachdem das Experiment auf dem daraus resultierenden Design durchgeführt und ausgewertet wurde, können lokale Theorien sowohl zu gegenstandsspezifischen als auch -übergreifenden Prozessen entwickelt werden (Prediger et al. 2012). Diese bilden „die neue und erweiterte Grundlage für den nächsten Durchlauf im Entwicklungsforschungszyklus“ (S. 7). Somit kann das Ziel verfolgt werden, immer genauere und differenzierte Aussagen über den Schriftspracherwerb zu ermöglichen und die Lücke zwischen Theorie und Praxis zu schließen (Prediger et al. 2012).

Darüber hinaus sollen Entwicklungsunterschiede im Bereich des Schriftspracherwerbs aufgedeckt werden und erste Erkenntnisse zu diesen Kompetenzen bei Schüler:innen der ersten Klasse liefern, die aufgrund der bereits dargestellten mangelnden Studienlage oft unberücksichtigt bleiben.

6.2 Stichprobe

Die Stichprobe setzte sich aus 47 Schüler:innen der ersten Klasse der städtischen katholischen Grundschule an der Bergstraße in Herne zusammen. Die Rekrutierung der Teilnehmenden erfolgte durch den bereits bestehenden Kontakt zu der Grundschule, da es sich hierbei um einen Kooperationspartner des *BeeMehr*-Projekts¹² handelt. Die Proband:innen sind verteilt auf zwei Parallelklassen, welche unter anderem verschiedene Lehrkräfte im Deutschunterricht aufweisen.¹³ 23 Teilnehmende besuchen die Klasse 1a und 24 Schüler:innen die Klasse 1b, sodass ein ausgeglichenes Verhältnis entstanden ist. Auf der Basis der Informationen der Lehrkräfte kann festgehalten werden, dass 33 Mädchen und 14 Jungen teilnahmen. Da nur eine geringe Anzahl an Elternfragebögen zurückkamen, fehlen einige weiterführende und relevante Informationen zu den einzelnen Schüler:innen. Dadurch können keine Aussagen beispielsweise zu dem sozioökonomischen Status oder zu der Mehrsprachigkeit getroffen werden. Mit Blick auf die explizite Schule wurde ein Schulsozialindex von eins festgelegt (MSB

¹² Das Projekt *BeeMehr – Warum Mehrsprachigkeit einfach mehr ist* wird an der TU Dortmund von Prof. Dr. Barbara Mertins geleitet. Weitere Informationen zu dem Projekt können der Website (Technische Universität Dortmund o.A.) entnommen werden.

¹³ Es gibt zahlreiche Studien, welche darlegen, dass die Lehrkraft einen immensen Einfluss auf die Leistungen der Schüler:innen haben kann (z.B. Chaminzo-Nieto et al. 2021; Wang 2022; Sáez-Delgado et al. 2022). Auch wenn dieser Aspekt in dieser Arbeit nicht im Fokus ist, ist es aufgrund dieser Korrelation trotzdem relevant zu erwähnen.

2024a). Bei dem Schulsozialindex handelt es sich um ein Instrument, „mit dessen Hilfe es möglich ist, die soziale Zusammensetzung der Schülerschaft einzelner Schule mit einem Wert abzubilden“ (MSB 2024b). Für die Berechnung werden vier Indikatoren hinzugezogen: Kinder- und Jugendarmut, Anteil der Schüler:innen mit vorwiegend nichtdeutscher Familiensprache, Anteil der Schüler:innen mit eigenem Zuzug aus dem Ausland sowie der Anteil der Schüler:innen mit dem Förderschwerpunkt Lernen, emotionale und soziale Entwicklung und Sprache (MSB 2024b). Hierbei gibt es neun Sozialindexstufen, wobei Stufe eins als niedrigste Belastung und Stufe neun als höchste Belastung gekennzeichnet wird.¹⁴ Da die Schule der Proband:innen einen Sozialindex von eins aufweist, lässt sich folglich festhalten, dass die soziale Belastung an dieser Schule als äußerst gering einzustufen ist. Diese Rahmenbedingungen können sich positiv auf den Schriftspracherwerb auswirken, da empirisch bereits häufig belegt werden konnte, dass eine Korrelation zwischen dem sozioökonomischen Status und den Kompetenzen der Kinder besteht. Ein hoher sozioökonomischer Status geht hierbei mit höheren Leistungen inklusiver besserer Lernvoraussetzungen einher (z.B. Linberg & Wenz 2017; Wawire et al. 2024; Walper & Grgic 2013).

Alle Proband:innen nahmen freiwillig an der Studie teil. Darüber hinaus wurden die Erziehungsberechtigten vor Beginn der Studie umfassend über das Forschungsvorhaben informiert und sie gaben ihre Zustimmung dazu, dass ihre Kinder an der Studie teilnehmen durften. Darüber hinaus haben sie der anonymen Verwendung der gesammelten Daten für wissenschaftlichen Zwecke zugestimmt.

6.3 Vorgehen

Die Datenerhebung erfolgte in einem Zeitraum von zwei Wochen kurz vor den Sommerferien, sodass festgehalten werden kann, dass das Ende der ersten Klasse Gegenstand der Testung ist. Die Erziehungsberechtigten erhielten vor der Erhebung einen Informationsbrief, der sowohl eine Einverständniserklärung zur Nutzung der Daten als auch Informationen zum Datenschutz enthielt.

Grundlegend sollte die Erhebung auf der Basis von zwei Methoden erfolgen: einem Online-Fragebogen für die Erziehungsberechtigten und der Testdurchführung mit den Erstklässler:innen. Bei dem Fragebogen handelt es sich um den Elternfragebogen der *psycholinguistics laboratories* der TU Dortmund. Er besteht aus zwei Varianten: Ein Fragebogen ist für Bilinguale und ein Fragebogen für Monolinguale konstruiert. Dieser wird stetig weiterentwickelt und beschäftigt sich mit Aspekten wie beispielsweise dem Grad der Mehrsprachigkeit, dem

¹⁴ Das Schulministerium gibt hier den Begriff der *Belastung* vor, weshalb dieser auch in diesem Kontext genutzt wird. Die Verwendung dieses Begriffs ist jedoch stark zu kritisieren, da er negativ konnotiert ist und beispielsweise impliziert, dass Schüler:innen mit einem bestimmten sozioökonomischen Status eine Last darstellen. Dies kann stigmatisierend wirken und die wertneutrale Beschreibung von Diversität und sozialen Herausforderungen erschweren.

sozioökonomischen Status und dem Input innerhalb des lesenden Umfeldes des Kindes. Standardisierte Fragebögen gelten in der Spracherwerbsforschung als eine bewährte Methode, die insbesondere durch ihre Effizienz überzeugt (Kauschke 2012). Obwohl stets berücksichtigt werden muss, dass den Proband:innen die Forschungssituation bewusst ist und ihre Antworten dadurch potenziell beeinflusst werden können, belegen frühere Studien, dass Eltern verlässliche Angaben machen können (Kauschke 2012). Folglich sind Informationen über diese Kriterien in Kombination mit den Punkten der Tests von großem Interesse für die Auswertung.¹⁵ Aufgrund einer zu geringen Anzahl an beantworteten Fragebögen, können diese allerdings nicht ausgewertet werden, weshalb der Fragebogen im Folgenden nicht weitergehend erläutert wird.

Die Haupterhebung fand in den Räumlichkeiten der Grundschule während der Schulzeit statt. Dies hat den Vorteil mit sich gebracht, dass der Anfahrtsweg als potenzielle Hürde für die Testdurchführung entfiel, da die Kinder ohnehin bereits in der Schule anwesend waren. Um die Ablenkung so gering wie möglich zu halten, wurde die Erhebung in einem separaten Raum durchgeführt, welcher ruhig und gut beleuchtet war. Da die Durchführungen in immer demselben Raum stattfanden, wurde dieser stets gleich vorbereitet. In dem Zuge lagen alle Materialien griffbereit. Hierzu gehörte der Testbogen, welcher aus einzelnen Blättern bestand und umgedreht auf den Tisch lag. Somit wurde sichergestellt, dass die Proband:innen den Subtest immer erst dann sahen, wenn dieser bearbeitet werden sollte. Wenn ein Subtest fertig bearbeitet wurde, wurde das Blatt zur Seite gelegt und der nächste Subtest herausgegeben. Auch das Playmobil wurde zunächst abgedeckt, sodass es die Kinder von der Bearbeitung der anderen Subtests nicht ablenkte. Darüber hinaus lagen Stifte bereit, aus denen die Teilnehmenden sich einen aussuchen durften. Die Audioaufnahmen wurden mit dem Smartphone abgespielt. Die Proband:innen wurden nacheinander mit Absprache der Klassenlehrkräfte aus den Klassen abgeholt und einzeln getestet.

Während der Durchführung wurde sich möglichst genau an den Durchführungsinstruktionen orientiert, um vergleichbare Bedingungen zu kreieren. Aufgrund dessen war es von großer Bedeutung, sich vor den Erhebungen mit den Instruktionen vertraut zu machen, um sich von dem Vorlesecharakter zu lösen. Das Ziel bestand darin, möglichst nah an dem Wortlaut der Instruktionen zu bleiben, bei gleichzeitiger interaktiver Freiheit und Flexibilität mit den Teilnehmenden. Von inhaltlichen Antworten auf Fragen der Proband:innen wurde strikt abgesehen und indes versucht, die Schüler:innen zu motivieren, sich selbstständig der Antwort anzunähern. Inwiefern diese schlussendliche Antwort richtig oder falsch ist, wurde in dem Kontext

¹⁵ Dass der sozioökonomische Status einen erheblichen Einfluss haben kann, wurde bereits dargestellt. Viele Studien haben sich darüber hinaus mit der Bedeutung der Mehrsprachigkeit (z.B. Berghoff & Bylund 2023; Thomas-Sunesson et al. 2018) und dem Input (z.B. Schulze & Saalbach 2022; McElvany et al. 2009; Pfost et al. 2010; Lehl et al. 2012) für den Schriftspracherwerb auseinandergesetzt.

nicht thematisiert, sodass die Proband:innen keine Rückschlüsse aus diesen Aussagen ziehen konnten.

Die Auswertung erfolgte zunächst anhand des bereits dargestellten Kodierungssystems des Tests, wobei die Tests ein zweites Mal von einer weiteren, unabhängigen Person ausgewertet wurden. Dies dient dem Zweck, Beurteilungsfehler aufzudecken und somit die Genauigkeit und Zuverlässigkeit der Ergebnisse zu erhöhen. Des Weiteren kann auf der Basis dessen die Konsistenz der verschiedenen Bewertungen geprüft werden, sodass Aussagen zu der Eindeutigkeit des Kodierungssystems getroffen werden können. Nach der Bepunktung erfolgte die statistische Auswertung.

7. Darstellung der Ergebnisse

Da die Informationen aus dem Fragebogen fehlen, wird auf der Basis der deskriptiven Statistik ausgewertet. Somit stehen bei der Auswertung die Punktzahlen im Vordergrund, welche auf der Basis des Kodierungssystems vergeben wurden. Um die Reliabilität dieser Punkte zu gewährleisten, wurde die Interrater-Reliabilität berechnet. Dies stellt ein Maß dafür dar, wie stark die Bewertungen mehrerer Beurteilenden übereinstimmen, wenn sie dasselbe Phänomen bewerten. Sie zeigt somit an, wie konsistent und zuverlässig das Kodierungssystem und die daraus entstehenden Urteile sind. Dies bedeutet, dass eine hohe Interrater-Reliabilität auf eine hohe Objektivität des Bewertungsprozesses hinweist, während eine niedrige Interrater-Reliabilität Hinweise auf Inkonsistenzen liefert. Der Ursprung solcher Ergebnisse kann häufig in subjektiven Urteilen oder unklaren Bewertungsrichtlinien gefunden werden. Die Basis für die Berechnung des prozentualen Übereinstimmungswertes stellen die Anzahlen an übereinstimmenden und abweichenden Urteilen dar, bezogen auf die Gesamtzahl der Beurteilungen. Die Formel für die Berechnung der prozentualen Übereinstimmung lautet:

$$\text{Prozentuale Übereinstimmung} = \frac{\text{Anzahl der Übereinstimmungen}}{\text{Gesamtzahl der Beurteilungen}} \times 100$$

Die Tests der Proband:innen wurden von zwei Personen unabhängig voneinander bewertet.¹⁶ Daraus ergab sich folgende Berechnung:

$$\text{Prozentuale Übereinstimmung} = \frac{368}{376} \times 100 \approx 97,87 \text{ (\%)}$$

Folglich lässt sich festhalten, dass die Beurteilungen zwischen den Bewertenden zu 97,87% deckungsgleich sind. Dieser Wert zeigt eine sehr hohe Übereinstimmung an, was darauf hindeutet, dass die Bewertungen äußerst konsistent sind. Somit ist eine Auswertung auf Basis des Kodierungssystems zulässig.

¹⁶ Die genaue Aufstellung der Bepunktung kann im Anhang (siehe Anhang 4 und 5) eingesehen werden.

In diesem Zuge muss jedoch zunächst der Fokus auf den zweiten Subtest, in welchem die Vokallänge bestimmt werden soll, gelegt werden. Werden die Mittelwerte der einzelnen Subtests verglichen, sticht der zweite Subtest im besonderen Maße heraus.

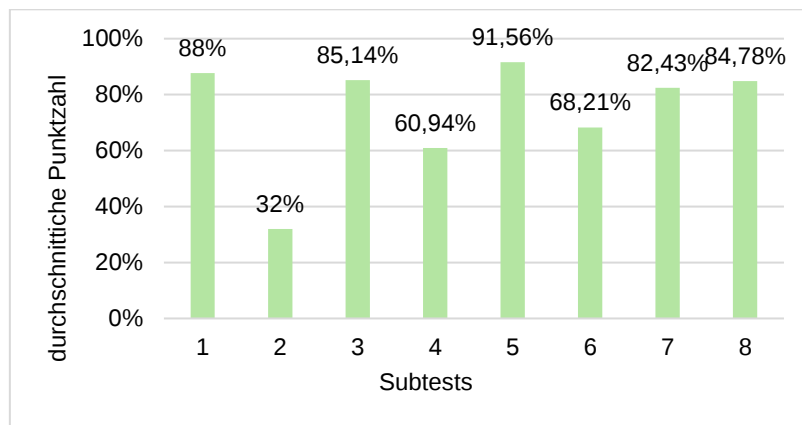


Abbildung 8: Mittelwerte der einzelnen Subtests im Vergleich

Während sich bei Subtest zwei ein Mittelwert von 32% ergibt, weisen die Mittelwerte der anderen Subtests einen durchschnittlichen Wert von 80,15% auf. Es ist ein immenser Unterschied zwischen diesen Werten festzustellen. Hieraus lässt sich schlussfolgern, dass Subtest zwei die Daten verzerrt. Ein Hauptgrund dafür, dass dieser Bias für die Testauswertung problematisch ist, besteht darin, dass verzerrte Daten zu falschen oder irreführenden Ergebnissen führen können. Auf der einen Seite scheint der zweite Subtest zu einer immensen Überforderung auf Seiten der Proband:innen zu führen. Dies wird allein daran ersichtlich, dass die Durchführung dieses Subtests bei einer Testung (P47) erst gar nicht möglich war. Das Kind zeigte eine derart starke Überforderung und daraus resultierende Verunsicherung, welche die Durchführung der nächsten Subtests unmöglich gemacht hätte, sodass der zweite Subtest abgebrochen wurde. Auf der anderen Seite haben viele Proband:innen die Aufgabe nicht verstanden, was daran kenntlich wurde, dass sie dies verbal äußerten. Diese Aussage kann insofern unterstützt werden, dass zwölf Proband:innen bei allen Items dieses Subtests konsequent das letzte Wort genannt haben. Einerseits kann das auf die Kapazität des Arbeitsgedächtnisses zurückgeführt werden, da der Einsatz von Pseudowörtern und zudem das verbal-sprachgebundene Format dieses stark beanspruchen. Andererseits bestätigt dies zudem die These, dass einige Proband:innen wahrscheinlich häufig geraten haben. Der Subtest zwei scheint folglich nicht das zu testen, was er zu testen vorgibt. Es lässt sich festhalten, dass sich Subtest zwei als nicht einsetzbar herausstellt und demnach aus der nachfolgenden Auswertung ausgeschlossen werden muss.

Um Unterschiede sowie Gemeinsamkeiten in der Verteilung der Daten aufzeigen und Abweichungen sichtbar zu machen, wurde ein Boxplot erstellt. Dieser ermöglicht einen schnellen Vergleich und liefert eine Grundlage für tiefergehende Analysen (Meindl 2011).

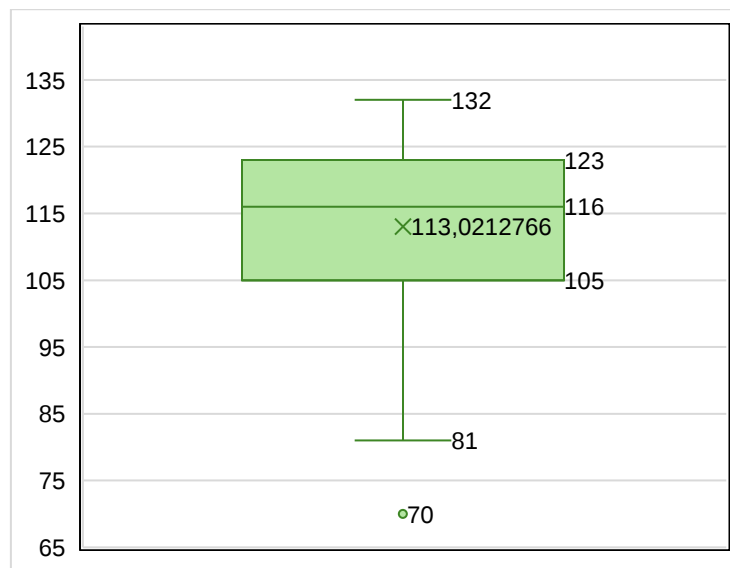


Abbildung 9: Boxplot (Gesamtpunktzahl ohne Subtest 2)

An diesem Boxplot wird deutlich, dass sich die mittleren 50% der Verteilung zwischen 105 Punkten (73,43%) und 123 Punkten (86,01%) befinden. Die Werte liegen demnach relativ nah beieinander. Die Spannweite ergibt sich aus einem Maximum von 132 Punkten (92,31%) und einem Minimum von 81 Punkten (56,64%), was eine leichte Streuung aufzeigt, wobei ein einzelner Ausreißer mit 70 Punkten (48,95%) nach unten erkennbar ist. Dennoch bleibt die Streuung insgesamt in einem gewissen Rahmen, sodass dies darauf hinweist, dass die Leistungen der Proband:innen relativ konsistent sind. Unterstützt wird diese Aussage mit Hilfe der Berechnung der Standardabweichung. Je größer die Standardabweichung ist, desto weiter streuen die Daten um das arithmetische Mittel (Meindl 2011). Für die Stichprobe kann eine Standardabweichung von 13,22 Punkten errechnet werden, was 9,24% ergibt. Da ein gewisses Maß an Leistungsfluktuation gängig ist, insbesondere in einer ersten Klasse, kann die vorherige Aussage einer stabilen Leistung in der Stichprobe bestätigt werden. Da der Median mit 116 Punkten (81,12%) leicht über den Mittelwert von 113,02 Punkten (79%) liegt, deutet dies auf eine leicht linksschiefe Verteilung hin. Das bedeutet, dass einige niedrige Werte den Mittelwert etwas nach unten ziehen, während der Median als robusterer Lageparameter weniger beeinflusst wird (Meindl 2011). Der einzelne Ausreißer beeinflusst den Mittelwert leicht, hat aber keinen großen Einfluss auf den Median, was die Aussage unterstützt, dass die Leistungen der Proband:innen insgesamt konsistent mit nur kleineren Schwankungen sind.

Weitere tiefgreifendere Erkenntnisse können auf der Basis einer differenzierteren Betrachtung der einzelnen Subtests beziehungsweise der einzelnen Bereiche des Tests gewonnen werden.

Aufgrund der unterschiedlichen Höchstpunktzahlen der Subtests, basieren diese auf ungleichen Skalen. Während Subtest drei, aufgrund der Zweiteilung der Punkte, eine Höchstpunktzahl von 42 Punkten aufweist, kann in dem achten Subtest eine maximale Punktzahl von zwei Punkten erreicht werden. Dies führt dazu, dass die Ergebnisse in absoluten Zahlen nicht ohne Weiteres miteinander verglichen werden können. Daher müssen die Skalen in einem ersten Schritt vereinheitlicht werden. Eine gängige Lösung für dieses Problem besteht in der Weiterarbeit mit relativen Zahlen. Wenngleich dieses Vorgehen die Vergleichbarkeit der Ergebnisse erhöht und klare Tendenzen aufgezeigt werden können, ist weiterhin nicht außer Acht zu lassen, dass die Subtests auf stark abweichenden Skalen basieren. Für die Zusammenfassung in Bereiche hat dies zur Folge, dass niedrige Punktzahlen stärker ins Gewicht fallen. Dies gilt es bei der Auseinandersetzung mit den nachfolgenden Daten stets mitzudenken.

Auf der Grundlage der bereits dargestellten Abbildung der durchschnittlichen Gesamtpunktzahlen (Abbildung 8), zeigen sich stabile, relativ hohe Ergebnisse, was erneut auf eine leistungsstarke Stichprobe hinweist. An diesen Ergebnissen wird deutlich, dass das Verständnis über die Phonem-Graphem-Korrespondenz, welche insbesondere in dem ersten Subtest getestet wird, bereits vorhanden ist. Des Weiteren kann aufgezeigt werden, dass die Proband:innen bereits eine hohe Punktzahl in den Subtests erreichen, die die orthografische Kompetenz in den Vordergrund rücken. Dies spricht für einen fortschreitenden Weg von einer lautgetreuen Schreibweise zu der orthografisch korrekten Schreibweise. Die Proband:innen weisen ebenfalls im Bereich der Semantik, welche sich schwerpunktmäßig mit dem globalen Verständnis beschäftigt, eine hohe Kompetenz auf. Die Berechnung der Standardabweichung wird bei einer Betrachtung der einzelnen Bereiche des Tests im besonderen Maße interessant.

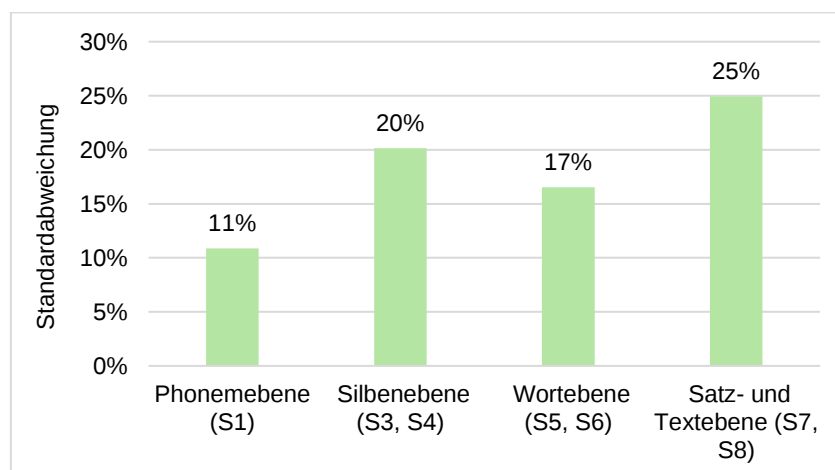


Abbildung 10: Standardabweichungen der linguistischen Bereiche der Subtests im Vergleich

Anhand dieser Grafik wird der aufbauende Schwierigkeitsgrad deutlich, welcher bei der Testkonstruktion intendiert wurde. Sie zeigt auf, dass die Unterschiede zwischen den Leistungen der Proband:innen im Verlauf des Tests fortlaufend größer werden, sodass die Daten stärker

streuen. Daraus kann schlussgefolgert werden, dass die Subtests immer anspruchsvoller werden. Auffallend ist jedoch, dass es mit Blick auf die Wortebene, in welcher der Fokus auf die Orthografie gelegt wurde, eine Unterbrechung dieser Kontinuität entsteht, da die Streuung an dieser Stelle etwas zurückgeht. Es kann die Annahme getroffen werden, dass ein Grund für diese kurzfristige Abweichung des allgemeinen Trends des ansteigenden Schwierigkeitsgrades darin gefunden werden kann, dass es sich bei den Subtests zu der phonologischen Bewusstheit im weiten Sinne auf der Silbenebene um herausfordernde Subtests handelt. Bereits im Vergleich des Mittelwerts der einzelnen Subtests (Abbildung 8), fällt der vierte Subtest mit einer deutlich niedrigeren durchschnittlichen Gesamtpunktzahl (60,94%) auf.

Darüber hinaus kann anhand der Streuung der Bereiche festgestellt werden, dass die Phonebene, in der die phonologische Bewusstheit im engeren Sinne in den Mittelpunkt der Betrachtung gerückt wird, die geringste Streuung aufweist. Daraus lässt sich die Annahme ableiten, dass dieser Subtest aufzeigt, dass die Kompetenzen in der phonologischen Bewusstheit eine Grundlage für die weiteren Kompetenzen des Schriftspracherwerbs darstellen. Um diese Annahmen weiterhin zu bestärken, ist ein vergleichender Blick auf die Proband:innen mit hohen und niedrigen Testergebnissen hilfreich. Es ist jedoch an dieser Stelle festzuhalten, dass es in solchen Tests nicht primär darum geht, einzelne Schüler:innen miteinander zu vergleichen, sondern vielmehr darum, ein gesamtbildliches Verständnis der Entwicklung der schriftsprachlichen Kompetenzen zu gewinnen. Trotz dessen kann eine solche Betrachtungsweise an geeigneter Stelle die Aussagekraft bestimmter Ergebnisse erhöhen und exemplarisch aufzeigen.

Tabelle 2: Vergleich der erreichten Punktzahlen in den Subtests einzelner Proband:innen

	Phonemebene (phonologische Bewusstheit im engeren Sinne)	Silbenebene (phonologische Bewusstheit im weiten Sinne)		Wortebene (Orthografie)		Satz-/ Textebene (Semantik)		
Subtest Proband:in	S1	S3	S4	S5	S6	S7	S8	Gesamt
P14	96%	95,24%	96,55%	100%	81,25%	83,33%	0%	92,31%
P37	100%	97,62%	65,52%	100%	87,5%	100%	100%	90,91%
P29	60%	71,43%	55,17%	58,82%	25%	50%	0%	56,64%
P6	56%	61,90%	13,79%	70,59%	50%	50%	0%	48,95%

Zunächst bestätigt der Blick auf diese Tabelle die Aussage, dass mit einer hohen Punktzahl in dem ersten Subtest, welcher die Phonemebene in den Blick nimmt, hohe Ergebnisse in den anderen Bereichen einhergehen. Erreichen Proband:innen niedrige Punktzahlen in dem

ersten Subtest, erreichen sie auch niedrigere Ergebnisse in den weiteren Subtests. Wenn- gleich in der Tabelle nur exemplarisch einzelne Proband:innen dargestellt werden, kann dieser Trend in den gesamten Ergebnissen festgestellt werden. Aufgrund dessen kann geschlussfol- gert werden, dass die phonologische Bewusstheit im engeren Sinne eine immense Vorhersa- gekraft für die weiteren schriftsprachlichen Kompetenzen hat. Zusätzlich scheinen Schüler:in- nen der ersten Klasse Aufgaben in diesem Bereich sicher absolvieren zu können. Darüber hinaus stellt sich der vierte Subtest auch bei dieser Betrachtung als sehr herausfordernd dar. Die große Streuung in der gesamten Stichprobe innerhalb dieses Subtests (16,44%), spiegelt sich wider.

Bei Subtest eins bedarf es einer genaueren Betrachtung, da sich hier Auffälligkeiten bezüglich der Items gezeigt haben. Erwartungsgemäß waren Items, die das Phonem zweimal aufweisen und demnach zwei Kreuze gesetzt werden müssen, für die Proband:innen besonders heraus- fordernd. Allerdings ist im Prozess schnell aufgefallen, dass es Items gibt, bei denen die Pho- neme eher korrekt identifiziert wurden als bei anderen Items. Dies wird beispielsweise bei den Items *Uhu* und *Ente* augenscheinlich: Die Anzahl der Proband:innen, die bei *Uhu* zwei Kreuze gesetzt haben beträgt 38. Im Gegensatz dazu haben nur 16 Proband:innen bei *Ente* zwei Kreuze gesetzt. Die Herausforderung dieses Items bestätigt sich ebenfalls bei einer Betrach- tung der Proband:innen, welche alle Items außer *Ente* richtig bearbeitet haben: Fünf Schü- ler:innen haben somit gezeigt, dass sie Phoneme korrekt und sicher identifizieren können und ein Grundverständnis für die Phonem-Graphem-Korrespondenz entwickelt haben. Es kommt die Frage auf, wieso das Item *Ente* dennoch so häufig zu Schwierigkeiten führt. Die Antwort findet sich in der Repräsentation der Phoneme, welche offensichtlich einen starken Einfluss auf die Antworten der Proband:innen haben. Was hiermit konkret gemeint ist, wird bei der Be- trachtung der einzelnen Phoneme der Items kenntlich:

Ente ['ɛntə]

Anna [a'na:]

Otto ['ɔto:]

Kiwi ['ki:vi]

Uhu ['u:hu]

Es wird an diesen Beispielen deutlich, dass mit Blick auf die Vokale kein Phonem zweimal auf identische Art und Weise vorhanden ist. Während sich die Unterschiede beispielsweise bei dem Item *Anna* auf einem kurzen und einem längeren ausgesprochenen [a] begrenzen, grün- den die Unterschiede in anderen Items auf komplett unterschiedlichen Phonemen. Insbeson- dere bei *Ente* finden sich ein [ɛ] sowie ein [ə]. Bei dem e-Schwa handelt es sich um einen Reduktionslaut, welcher akustisch schwächer und kürzer als betonte Vokale ist und in unbe- tonten Silben vorkommt. Dadurch kann es zudem anderen reduzierten Phonemen ähneln, was

die Identifikation erschwert (del Toro & Thomé 2021). Im Gegensatz dazu wird das [ɛ] artikuliert und damit spezifischer und eindeutiger wahrnehmbar. Es kann folglich festgehalten werden, dass die Repräsentation der Phoneme deutliche Konsequenzen für die Art der schriftsprachlichen Umsetzung hat (Schründer-Lenzen 2004). Das Item *Uhu* ist demnach einfacher, da beide Vokale voll artikuliert und betont sind, wodurch sie akustisch klar und eindeutig wahrnehmbar bleiben. Im Gegensatz dazu enthält das Item *Ente* ein Schwa am Wortende, das durch seine reduzierte und unbetonte Natur schwerer zu identifizieren ist.

Während der Testdurchführung ist darüber hinaus aufgefallen, dass es zu starken Unterschieden in der für den Test benötigten Zeit kam. Grund hierfür lag in dem Aspekt der Testkonstruktion, sich von einem Test unter Zeitdruck zu distanzieren. Dies führte dazu, dass eine Zeitspanne von 16 Minuten bis 35 Minuten entstanden ist, wobei eine durchschnittliche Dauer von 24 Minuten berechnet werden kann. Dies stellt auf der bereits erläuterten theoretischen Grundlage eine zu lange Testdauer dar, weshalb in dem Zusammenhang der Faktor der Motivation der Schüler:innen eine große Rolle spielt. Diese war bei den Proband:innen zu nahezu jedem Zeitpunkt sehr hoch. Dies kann einerseits an den bekannten Aufgabenformaten liegen, wodurch Sicherheit und Vertrauen in das eigene Können entstanden ist, aber andererseits ebenso an der Testkonstruktion und -durchführung selbst. Wenn die Motivation über die Dauer des Tests etwas geringer wurde, ist der siebte Subtest auf große Begeisterung gestoßen. Mit dem Einsatz des Playmobils geht ein spielerischer Charakter einher. Darüber hinaus begünstigt es eine Abwechslung im Testablauf.

8. Diskussion

In dieser Arbeit wurde sich bewusst von dem Begriff der *Diagnostik* distanziert, wenngleich dieser Begriff vielfach im wissenschaftlichen sowie didaktischen Kontext Verwendung findet. Insbesondere bei der Erfassung von Kompetenzen im schulischen Bereich fällt häufig der Begriff der diagnostischen Kompetenz von Lehrkräften (KMK 2022b; Lenhard 2019; Helmke 2015). Dieser ist jedoch als problematisch zu betrachten, wobei die Gründe hierfür ihren Ursprung in mehreren Faktoren finden, die aufgrund ihrer hohen Relevanz für diese Arbeit in diesem Kapitel direkt zu Beginn benannt werden.

Zunächst stammt der Begriff ursprünglich aus der Medizin und ist eng mit dem Erkennen und Klassifizieren von Krankheiten verbunden (Becker-Lenz & Lüscher 2012). In nicht-medizinischen Kontexten, wie dem Schriftspracherwerb oder der Bildungswissenschaft, geht der Begriff daher mit einer pathologisierenden oder defizitorientierten Konnotation einher. Vielmehr geht es in diesen Bereichen jedoch darum, individuelle Kompetenzen, Stärken und Entwicklungsbedarfe zu erkennen, ohne dabei Kinder durch Etikettierung zu stigmatisieren (Heiner 2018). Kinder als Proband:innen sollen primär als diejenigen gesehen werden, die bereits

vielfältige Kompetenzen und Potenziale mitbringen, deren spezifische Ausprägungen es zu erfassen gilt. Die durch den Begriff der Diagnostik implizierte Defizitorientierung suggeriert hingegen fälschlicherweise, dass etwas *falsch* sei und fördert damit eine schwächenzentrierte Perspektive auf Kinder. Darüber hinaus wird dadurch impliziert, dass es sich um etwas Endgültiges handelt, wobei sich gerade der Schriftspracherwerb durch eine starke Dynamik und Kontinuität charakterisieren lässt (Schröder-Lenzen 2004).

In der Literatur wird diese Thematik häufig als selbstverständlich angesehen und nicht kritisch hinterfragt, wobei die kritische Auseinandersetzung mit dem Begriff zunimmt. So werden mittlerweile häufiger alternative Begrifflichkeiten verwendet. Helmke (2015) empfiehlt beispielsweise im Kontext der benötigten Kompetenzen der Lehrkräfte den Begriff der Expertise, da dieser umfassender sei und methodisches sowie prozedurales Wissen beinhalte. Es lässt sich demnach feststellen, dass der Begriff Diagnostik zwar in der Wissenschaft etabliert ist, aber aufgrund der Assoziationen für die Verwendung im Zuge des Schriftspracherwerbs kritisch bewertet wird.

Um die schriftsprachlichen Kompetenzen erfassen zu können, stehen im wissenschaftlichen Diskurs standardisierte Testverfahren im Fokus. Diese bieten laut Lenhard (2019) einige Vorteile, da sie auf Normierungen basieren und sich den Anforderungen an den Kriterien der Objektivität, Reliabilität und Validität orientieren müssen. Das oberste Ziel ist es demnach, das zu messen, was sie zu messen vorgeben (Bredel et al. 2017). Nach einer detaillierten Auseinandersetzung mit dem ELFE 1-6 (Lenhard & Schneider 2006a), dem BAKO 1-4 (Stock et al. 2003a) und dem IEL-1 (Diehl & Hartke 2012a) sind einige Aspekte aufgefallen, die das Erreichen dieses Zieles infrage stellen. Lenhard (2019) hat bereits auf die niedrige Reliabilität vieler Instrumente verwiesen, allerdings wird bei der kritischen Betrachtung der Einfluss verschiedener Faktoren auf die Testvalidität kenntlich. Eine Sache, die am deutlichsten bei allen Tests herauszustellen war, ist die nicht vorhandene Einheitlichkeit bei der Auswahl der Items. So bestehen meist ungleiche Anzahlen an Items, auch innerhalb eines Subtests. Zudem besteht kein begründeter Zusammenhang zwischen semantischen, phonemischen und graphemischen Unterschieden. Das hat zur Folge, dass die Vergleichbarkeit leidet. Des Weiteren bleiben die Frequenzen unbeachtet, obwohl Harden (2023) herausstellt, dass hochfrequente Wörter einen schnelleren Zugriff im mentalen Lexikon bedeuten. Marx (2007) hat bemängelt, dass die Möglichkeit von Bodeneffekten besteht, was sich auch in der Analyse der ausgewählten Testverfahren bestätigt hat. Lenhard (2019) spricht von einer zu langen Testdauer mancher Tests. Dies spiegelte sich ebenfalls in dem IEL-1 wider, da dieser 90 Minuten umfasst (Diehl & Hartke 2012b). All diese Aspekte haben einen starken Einfluss auf die Validität des Tests. Auf Basis dieser Ausführungen kann die erste Hypothese, dass methodische Schwächen dazu

führen, dass bestehende Lesetests nicht die intendierten Aspekte der Lesekompetenz erfassen, verifiziert werden.

Die Tragweite dieser Erkenntnis wird noch deutlicher, wenn die Ergebnisse der Studie miteinbezogen, in welcher der weiterentwickelte Test eingesetzt wurde. Im Zuge der Entwicklung des neuen Tests wurde lediglich ein Subtest ohne anschließende Veränderung aus dem BAKO 1-4 (Stock et al. 2003a) übernommen. Dieser Subtest erwies sich als nicht durchführbar, da er mit einer Überforderung der Schüler:innen, mangelndem Verständnis sowie übermäßigen Anforderungen an das Arbeitsgedächtnis einherging, insbesondere durch den Einsatz von Pseudowörtern. Aufgrund dieser Problematiken führte der Subtest zu verzerrten Ergebnissen und musste aus der Datenauswertung ausgeschlossen werden. Hierbei handelt es sich jedoch um einen Test, welcher weiterhin in Schulen eingesetzt wird und als Grundlage für weitreichende Entscheidungen, wie beispielsweise Schullaufbahnentscheidungen, herangezogen wird (Helmke 2015). Dies birgt das Risiko, unreflektierte und ungerechtfertigte Entscheidungen zu treffen, die nachhaltige Auswirkungen haben können.

Die zweite Forschungsfrage beschäftigt sich mit der Eignung des weiterentwickelten Tests, um den Schriftspracherwerb zu erfassen. Grundlegend hat sich die Durchführbarkeit inklusive der Durchführungsinstruktionen und des Kodierungssystems als sinnvoll einsetzbar herausgestellt. Ein zentraler Ausgangspunkt für die Auswertung des entwickelten Tests bestand in einer Abkehr von einer starren, ergebnisorientierten Bewertungsgrundlage, die ausschließlich auf das Erzielen von 100% der möglichen Punkte abzielt. Stattdessen liegt der Fokus auf dem Gewinnen eines differenzierten Einblicks in den aktuellen Stand des Schriftspracherwerbs. Dies mag zunächst paradox erscheinen, da die Testergebnisse natürlich in Form von Punkten ausgewertet werden, jedoch dient diese Bepunktung primär der Analyse und nicht der reinen Leistungskategorisierung. Vielmehr sollen die Kinder als Expert:innen ihres eigenen Lernprozesses wahrgenommen werden. Diese Perspektive nimmt Abstand von einer Defizitorientierung, sodass Fehler „nicht mehr als Ausdruck des Scheiterns [gelten], sondern als Zeichen phasentypischer Zugriffsweise auf Schrift“ (Schröder-Lenzen 2004, S. 37). Fehler haben bei dem Verfolgen dieses Kerngedankens sowohl einen subjektiven Sinn als auch eine sinnvolle Notwendigkeit (Schröder-Lenzen 2004). Dieser Ansatz ermöglicht eine wertschätzende und konstruktive Betrachtung des Schriftspracherwerbs, die den individuellen Lernprozess in den Vordergrund rückt und den Kindern signalisiert, dass ihre Bemühungen und Herangehensweisen anerkannt und geschätzt werden.

Aufgrund dieser Ausführungen kann gesagt werden, dass sich das entwickelte Kodierungssystem als möglichst fair mit Blick auf die Schüler:innen erwiesen hat. Auch die Interrater-Reliabilität hat bestätigt, dass dieses zuverlässig funktioniert, allerdings ist an einigen Stellen ersichtlich geworden, dass trotz dessen Verbesserungspotenzial besteht. Um die zweite

Forschungsfrage nach der Einsetzbarkeit des Tests beantworten zu können, muss auch dies diskutiert werden. Zunächst haben die unterschiedlichen Skalen der Subtests die statistische Auswertung der Testergebnisse erschwert. Die Subtests können deshalb nicht ohne Weiteres miteinander verglichen werden. Darüber hinaus spiegeln die Punkte nicht die Komplexität der Subtests wider. Dies wird beispielsweise an dem achten Subtest deutlich, welcher eine Höchstpunktzahl von zwei Punkte aufweist. Damit ist dieser Subtest der Test mit der geringsten Maximalpunktzahl, wobei es sich um den mit den größten Anforderungen handelt, da die Semantik auf Textebene in den Mittelpunkt der Betrachtung gerückt wird. Eine Lösung dieses Problems kann auf zwei Wegen erfolgen: Entweder alle Subtests erhalten dieselbe Punktzahl, sodass eine einheitliche und vergleichbare Skala entsteht oder die Punkte steigen mit der Komplexität der Subtests. Für diese Arbeit wurde aufgrund dessen viel mit relativen Zahlen gearbeitet, was die Ergebnisse trotzdem bis zu einem gewissen Maße vergleichbar macht.

Darüber hinaus hat die Kodierungsregel für den Subtest drei zu Schwierigkeiten geführt. Diese besagt, dass für jede richtig getrennte Silbe und zusätzlich für jede orthografisch richtig geschriebene Silbe ein Punkt vergeben wird. Um die Herausforderung verständlich darstellen zu können, werden zwei Beispiele herangezogen:



Abbildung 11: Beispiel einer Proband:innenantwort in dem dritten Subtest

In diesem Beispiel kann nur ein Punkt für die korrekte Silbe <de> gegeben werden. Zusätzlich wird ein Punkt für diese Silbe gegeben, da diese orthografisch korrekt geschrieben wurde. Daraus ergeben sich für dieses Item insgesamt zwei Punkte. Obwohl das Wort bei einer Nichtbeachtung der Silbenstruktur vollständig orthografisch richtig geschrieben ist, bekommt das Kind nur einen von vier möglichen Punkten für die Orthografie. Ursache dafür ist, dass die Bepunktung an die Silbenstruktur gekoppelt ist. Dieses Dilemma lässt sich allerdings schwer lösen und insofern legitimieren, dass der Subtest die phonologische Bewusstheit auf Silbenebene in den Blick nimmt und dadurch der Schwerpunkt eindeutig auf der Silbenstruktur liegt, was sich demnach auch in der Bepunktung zeigt.

Da die Proband:innen bei der Bearbeitung des vierten Subtests die Silbenbögen selbst setzen sollen, hat sich dies an vielen Stellen als schwierig auswertbar herausgestellt. Es war oft unklar an welcher Stelle der Silbenbogen anfängt und an welcher dieser aufhört. In diesem Zuge ist erneut auf die hohe Interrater-Reliabilität zu verweisen, die aufzeigt, dass die Bepunktung trotz dessen objektiv verlief, allerdings ist für eine Weiterarbeit mit diesem Test über Alternativen nachzudenken. Eine Möglichkeit bestünde darin, die Silbenbögen durch beispielsweise Striche zwischen den Buchstaben zu ersetzen, welche eindeutige Ergebnisse liefern sollten.

Allerdings ist der Einsatz von Silbenbögen ein fester Bestandteil der ersten Klasse, sodass die Schüler:innen die Anwendung dieser bereits verinnerlicht haben.

Die Kodierungsregel des fünften Subtests hat sich im praktischen Einsatz als schwierig erwiesen. In dem Subtest sollen die Teilnehmenden die in dem Wort fehlenden Grapheme ergänzen. Die Regel für die Vergabe der Punkte besteht darin, dass jeder richtige Buchstabe einen Punkt bedeutet und jeder falsche, zusätzliche Buchstabe mit einem Minuspunkt einhergeht. Dies führt dazu, dass lautgetreue Schreibweisen wie beispielsweise *Flugzoig* als vollständig falsch gelten und null Punkte vergeben werden. Das Vorgehen ist jedoch in dem Aspekt zu legitimieren, dass der Schwerpunkt dieses Subtests auf der orthografischen Kompetenz liegt. Es soll getestet werden, inwieweit die Schüler:innen bereits ein Verständnis über die Phonem-Graphem-Korrespondenz aufweisen und orthografische Regeln anwenden können. Ebenfalls hat die Kodierungsregel zur Folge, dass beispielsweise die Schüler:innenantwort *Broot* ebenfalls mit null Punkten bewertet wird. Dies ergibt sich daraus, dass für ein <o> ein Punkt vergeben wird und für das zweite, falsche <o> wiederum ein Punkt abgezogen wird. Dies impliziert, dass auf Basis der Punkte fälschlicherweise angenommen werden könnte, die Antworten *Broot* und *Brat* seien auf einer vergleichbaren Ebene orthografischer Kompetenz einzuordnen. Für eine Weiterarbeit mit dem Test ist an der Stelle eine Überarbeitung dieser Kodierungsregel notwendig.

Bei der Auswertung des sechsten Subtests ergibt sich die Schwierigkeit daraus, dass Items mit unterschiedlich hohen linguistischen Anforderungen zu einer gleichen Punktzahl gelangen und demnach das Geschriebene durch die Punkte nicht hundertprozentig wiedergeben. Die Regel für die Bepunktung dieses Subtests besteht daraus, dass jede orthografisch richtig geschriebene Silbe mit einem Punkt bewertet wird. Daraus geht hervor, dass bei dem Item *Rose* die Antworten <Rone> und <Pane> gleichermaßen mit einem Punkt bewertet werden, obwohl alle Grapheme, die im Wort <Rone> vorkommen, auch in <Krone> zu finden sind im Gegensatz zu dem Wort <Pane>. Ein ähnliches Phänomen ergibt sich aus der Betrachtung der alternativen, ebenfalls als richtig bewerteten Wörter. Diese sind häufig schwerer oder leichter als die intendierten Items. Ein Beispiel stellt das Item *Hubschrauber* dar, an welcher Stelle viele Proband:innen *Helikopter* geschrieben haben. Die Herausforderung bei dem Item *Hubschrauber* ergibt sich aus den mehrfachen Konsonantenclustern und der daraus entstehenden komplexen Silbenstruktur. Aufgrund dieser Schwierigkeiten werden bei einer Bewertung der orthografischen Kompetenz häufig qualitative Fehleranalysen für sinnvoll erachtet, wobei gesagt werden muss, dass in diesem Zusammenhang keine hohe Reliabilität zu erwarten ist (Marx 2007). Des Weiteren kann trotz dessen nicht „auf die Prozesse geschlossen werden [], die die Fehler verursacht haben“ (Marx 2007, S. 89). Daher wurde sich im Kontext dieser Arbeit gegen einen Einsatz qualitativer Verfahren entschieden.

Insgesamt kann aus diesen Ausführungen abgeleitet werden, dass die Ergebnisse der Proband:innen nicht überinterpretiert werden dürfen. Ein niedriges Testergebnis sagt nicht zwangsläufig aus, dass es sich um ein schlecht lesendes Kind handelt. Insbesondere bei einem Einsatz in der ersten Klasse muss stets bedacht werden, dass Kinder Entwicklungsschritte unterschiedlich schnell durchlaufen und in diesem Kontext häufig ihre individuellen Lösungswege suchen (Sassenroth 2000). Dies schmälert die Aussagekraft dieser Tests allerdings nicht, denn die Ergebnisse liefern trotzdem klare Hinweise und Tendenzen und ermöglichen eine anschließend tiefergehende Betrachtung.

Nach dem expliziten Einsatz des Testinstruments in der ersten Klasse kann festgehalten werden, dass die Teilnehmenden den ersten Subtest, welcher die phonologische Bewusstheit im engeren Sinne in den Blick genommen hat, sicher absolviert haben. Dies deutet auf einen bereits länger da gewesen Lernverlauf hin. Diese Feststellung deckt sich mit der von Mehlem (2024) vertretenen Position, dass der Erwerb der phonologischen Bewusstheit bereits im Vorschulalter einsetzt. Es bestätigt auch die Darstellungen von Schröder-Lenzen (2004) sowie von Marx (2007), welche die phonologische Bewusstheit als Vorläuferfertigkeit betiteln. Eine zentrale Aussage der Studie von Schneider und Näslund (1993) bestand darin, dass, obwohl vielfache Faktoren Einfluss auf den Schriftspracherwerb haben, die phonologische Bewusstheit besonders hervorzuheben gilt. Hogan et al. (2005) haben ebenfalls die phonologische Bewusstheit als Prädiktor für spätere schriftsprachliche Leistungen herausgestellt. An dieses Erkenntnis schließen Milankov et al. (2021) an, die in ihrer Studie niedrige Kompetenzen im Bereich der phonologischen Bewusstheit als Hauptursache für spätere Leseschwierigkeiten herausstellten. Diese Erkenntnisse gehen mit den Ergebnissen dieser Arbeit einher, indem herausgestellt werden konnte, dass Schüler:innen, die eine hohe Punktzahl im Bereich der phonologischen Bewusstheit erreichten, ebenfalls hohe Punkten in anderen Bereichen erzielten.

Dessen ungeachtet ist jedoch aufgefallen, dass die Phonemrepräsentation einen erheblichen Einfluss auf die Antworten der Proband:innen hatte. Diese zeichnet sich durch eine hohe Komplexität aus, da ein Graphem durch unterschiedliche Phoneme artikuliert werden kann (Schneider 2017). Es existieren demnach mehr Phoneme als Grapheme, sodass eine korrekte Verwendung ein Bewusstsein über diese Tatsache voraussetzt. Die Teilnehmenden der Studie hatten hierbei an manchen Stellen Schwierigkeiten, was die von Schröder-Lenzen (2004) herausgestellte Herausforderung unterstützt, dass Grapheme in verschiedenen Wortkontexten anders artikuliert werden. An der Stelle lässt sich festhalten, dass es für eine Weiterarbeit mit dem Testinstrument sinnvoll wäre, die Items des ersten Subtests zu überarbeiten mit Blick auf die Repräsentation der Phoneme.

Es konnte herausgearbeitet werden, dass der Test, der in verschiedene linguistische Bereiche unterteilt wurde, so konzipiert ist, dass diese Bereiche systematisch aufeinander aufbauen. Obwohl die Literatur die phonologische Bewusstheit im weiten Sinne als Vorläufer der phonologischen Bewusstheit im engeren Sinne kennzeichnet (Bredel et al. 2017; Reber 2017), deutet der Aufbau des Tests, einschließlich der Auswertung, darauf hin, dass das Verhältnis umgekehrt ist. Auf der einen Seite kann ein Grund in dem dritten Subtest gefunden werden, da sich hier die Bereiche der phonologischen Bewusstheit und der orthografischen Kompetenz überschneiden. Das liegt an der zusätzlichen Vergabe der Punkte für korrekte Orthografie. Dies lässt sich allerdings auf der theoretischen Grundlage legitimieren, denn diese stellt klar heraus, dass die Silbenstruktur und die Orthografie korrelieren. So stellen beispielsweise Hogan et al. (2005) in ihrer Studie heraus, dass die phonologische Bewusstheit und das Lesen von Worten eng zusammenhängen. Die Silbenstruktur ist für das Lesen von Wörtern von großer Relevanz, da sie als grundlegende phonologische Einheiten die Segmentierung, Dekodierung, orthografische Verarbeitung, Leseflüssigkeit und Leseverständnis erleichtern (Schründer-Lenzen 2004). Auf der anderen Seite kann der vierte Subtest ausschlaggebend gewesen sein, da er sich als sehr herausfordernd herausgestellt hat. Seine Komplexität wird primär durch den Einsatz von niedrigfrequenten Wörtern erzeugt. Um diese korrekt zu segmentieren, ist ein Einblick in die phonologische Struktur notwendig, ohne auf vertraute Muster zurückgreifen zu können (Marx 2007). Um die Aufgabe korrekt zu lösen, müssen die Teilnehmenden demnach ihr wörterbezogenes Verständnis über Silben generalisieren und anwenden. Wenn gleich der Subtest sich durch eine hohe Komplexität auszeichnet, ist die Fähigkeit, auf neue sprachliche Herausforderungen zu reagieren, ein zentrales Element des Schriftspracherwerbs (Sassenroth 2000). Darüber hinaus kann es den Standpunkt von Marx (2007) unterstreichen, dass gerade in der ersten Klasse die intensive Auseinandersetzung mit der Schrift zu einem großen Kompetenzzuwachs führt. Zudem wird die Aussage von Schneider (2017) bestätigt, dass sich die Entwicklung dieser Bereiche nicht isoliert vollzieht, sondern ineinander übergeht.

Im wissenschaftlichen Diskurs ist die Reihenfolge des Einflusses der orthografischen Kompetenz und des semantischen Bewusstseins häufig divergent. In diesem Zuge bezeichnen Retelsdorf und Köller (2024) den Einfluss des Leseverstehens auf die orthografische Kompetenz größer als andersherum, während Vettori et al. (2023) die orthografische Kompetenz als Schlüsselkompetenz für weitere Bereiche des Schriftspracherwerbs einordnet. Obwohl die Ergebnisse dieser Arbeit eher die zweite These unterstützen, kann die Frage nach dem Einfluss nicht hinreichend beantwortet werden, sodass dies weiter überprüft werden muss.

Wie Bredel et al. (2017) anführen, ist die Schrift vielmehr eine Abbildung von Grammatik als eine Abbildung von Lauten. Die Teilnehmenden der Studie dieser Arbeit zeigten, dass sie an der Stelle des Erwerbs sind, an der sich die orthografische Kompetenz fortschreitend

entwickelt. Dies schließt eine stetige Verinnerlichung der Phonem-Graphem-Korrespondenz sowie die fortlaufende Erweiterung des mentalen Lexikons im Bereich der Orthografie mit ein.

Die Subtests, die sich mit dem Bereich der Semantik auf Satz- und Textebene befassen, weisen trotz der hohen Anforderung (Bredel et al. 2017) eine hohe durchschnittliche Punktzahl auf. Dies zeigt, dass die Teilnehmenden sowohl die Wörter erkennen und verbinden als auch diese in Beziehung zueinander setzen. Diese Kompetenzen sind laut Marx (2007) unabdingbar, um Sätze und daran anschließend Texte verstehen zu können. Auf dieser Basis können die semantischen Verbindungen im mentalen Lexikon stetig erweitert werden. Wenngleich in der Theorie die Bedeutung des globalen Verständnisses hervorgehoben wird (Schneider 2017), wird in dem entwickelten Test lediglich das lokale Verständnis getestet. Dies schränkt einerseits die Reichweite der Aussagen zu den Kompetenzen im Bereich der Semantik ein, ist jedoch andererseits eine realistische Abbildung, da das globale Verständnis aufgrund der Komplexität des Miteinbezug von Vor- und Weltwissen in der ersten Klasse noch keinen Einhalt findet (Rosebrock & Nix 2015).

Insgesamt weisen die Daten keine weitreichende Streuung auf. Ein Grund hierfür könnte beispielsweise der sozioökonomische Status der Teilnehmenden sein. Wenngleich der fehlende Elternfragebogen keine Informationen konkret zu diesem Faktor zulässt, kann auf der Basis des der Schule zugeschriebenen Sozialindex die Annahme eines hohen sozioökonomischen Status getätigt werden. Linberg & Wenz (2017) sowie Wawire et al. (2024) konnten jedoch in ihren Studien belegen, dass ein Zusammenhang zwischen dem sozioökonomischen Status und den Leistungen der Kinder besteht. Detaillierte und belegte Aussagen können aufgrund der fehlenden Informationen konkret zu der Stichprobe jedoch nicht getroffen werden.

Es wurde bereits angeführt, dass Schüler:innen einer ersten Klasse eine geringe Konzentrationsspanne aufweisen und ein direkter Zusammenhang zwischen der Aufmerksamkeit und der schriftsprachlichen Leistung besteht (Valdois et al. 2024). Die Durchführung des Tests hat teilweise eine sehr lange Testdauer (maximal 35 Minuten) umfasst. Nach Marx (2007) handelt es sich hierbei um eine zu lange Zeit. Im Zuge der Frage, wie in Zukunft damit umgegangen wird, lassen sich nur drei Lösungswege aufzeigen: Erstens können für die Subtests jeweils Zeitlimits bestimmt werden. Zweitens kann die Anzahl an Subtests verringert werden oder drittens kann es bei dem bisherigen Vorgehen belassen werden. Auf der Basis der in dieser Arbeit gewonnenen Erkenntnisse haben durch den nicht vorhandenen Zeitdruck entstehenden Vorteile erheblichen Vorrang. Wenn im Zuge der Erfassung des Schriftspracherwerbs die Bereiche der phonologischen Bewusstheit, der orthografischen Kompetenz sowie der Semantik erhoben werden sollen, können keine Subtests gekürzt oder exkludiert werden. Demnach erweist sich die dritte Möglichkeit als am sinnvollsten. An der Stelle können jedoch auch für die Konzentration unterstützende Faktoren wie die Motivation der Teilnehmenden in den Blick genommen

werden. Diese war bei den Teilnehmenden stets als hoch zu bewerten, sodass die teils lange Testdauer kein Ausschlusskriterium für eine präzise Erfassung der schriftsprachlichen Kompetenzen wurde.

Insgesamt zeigt die Studie auf, dass die verschiedenen Bereiche, die sich aus der phonologischen Bewusstheit, der Orthografie und der Semantik ergeben, ineinander übergehen und nicht klar abgrenzbar sind. Das Ziel des Testinstruments besteht darin, den aktuellen Stand des Schriftspracherwerbs in genau diesen Bereichen zu erfassen. Die vorherigen Ausführungen belegen, dass dies erfolgreich erreicht wurde. Somit lässt sich die zweite Hypothese, dass der entwickelte Lesetest methodisch dazu geeignet ist, um die Kompetenzen von Schüler:innen in den Bereichen der phonologischen Bewusstheit, der orthografischen Kompetenz sowie der Semantik zu erfassen, ebenfalls verifizieren. Es kann festgehalten werden, dass der Test in einigen Bereichen noch Verbesserungspotential aufweist, jedoch im Vergleich zu den bereits etablierten Verfahren zur Erfassung des Schriftspracherwerbs einen großen Fortschritt darstellt. Ein zentraler Ausgangspunkt bei der Konstruktion war der explizite Blick auf die erste Klasse, da es zu diesem Zeitpunkt der Schullaufbahn wenige empirische Befunde gibt. Dabei stellt Marx (2007) beispielsweise heraus, dass eine frühe Feststellung von Schwierigkeiten im Schriftspracherwerb von großem Vorteil ist. Der Einsatz des Tests kann in diesem Zusammenhang dazu beitragen, potenzielle individuelle Problematiken frühzeitig zu erkennen und entsprechende Fördermaßnahmen gezielt einzuleiten.

9. Fazit

Das Ziel der vorliegenden Arbeit zum Thema *Kritische Auseinandersetzung mit bestehenden Verfahren zur Erfassung des Schriftspracherwerbs und Entwicklung eines neuen Lesetests: Eine empirische Untersuchung mit Schüler:innen der ersten Klasse* bestand darin, ein Testinstrument zur differenzierten Erfassung des Schriftspracherwerbs zu entwickeln und dessen Einsatzfähigkeit zu evaluieren. Es sollten die verschiedenen Dimensionen des Schriftspracherwerbs, insbesondere die phonologischen, orthografischen und semantischen Kompetenzen integriert und die Subtests speziell auf die Bedürfnisse von Erstklässler:innen zugeschnitten werden.

In diesem Zuge wurden zunächst bereits bestehende Testinstrumente zur Erfassung von schriftsprachlichen Kompetenzen, genauer der ELFE 1-6 (Lenhard & Schneider 2006a), der BAKO 1-4 (Stock et al. 2003a) und der IEL-1 (Diehl & Hartke 2012a) kritisch analysiert, um deren Stärken und Schwächen zu identifizieren. Diese Erkenntnisse bildeten die Grundlage für die Entwicklung eines neuen Lesetests, der die besonderen Anforderungen der ersten Klasse berücksichtigt. In einer empirischen Studie wurde dieser Test an einer Stichprobe von 47 Erklässler:innen evaluiert.

Die kritische Auseinandersetzung mit den bestehenden, standardisierten Testverfahren ergab eine Offenlegung vielfacher methodischer Schwächen. Hierbei konnten vor allem die unbegründete Auswahl an Items inklusive des Einsatzes von niedrigfrequenten Worten und subtestübergreifenden Dopplungen von Items sowie die lange Testdauer und zu hohe Komplexität für eine erste Klasse hervorgehoben werden. Es wurde die Schlussfolgerung getroffen, dass diese Tests nicht ohne Weiteres für den Einsatz in der Praxis geeignet sind. Diese Aussage konnte weiterhin bestärkt werden, da der einzige unverändert eingesetzte Subtest aus dem BAKO 1-4 aus (Stock et al. 2003a) der Wertung ausgeschlossen werden musste, obwohl dieser eigentlich aufgrund seiner Standardisierung in der Praxis überzeugen müsste.

Der vorrangig auf der Basis des IEL-1 (Diehl & Hartke 2012a) weiterentwickelte Lesetest zeigte in diesen Aspekten eine deutliche Verbesserung. Hervorzuheben ist hierbei der konsequente Einsatz hochfrequenter Wörter, das Schaffen eindeutiger Items und die Abkehr von Zeitlimits. Dennoch wurden Schwächen in der Gestaltung einzelner Items und in der Auswertung aufgezeigt, die in künftigen Überarbeitungen berücksichtigt werden sollten. Trotz dessen konnte der Einsatz des entwickelten Tests einen Beitrag zu der Einsicht in die schriftsprachlichen Kompetenzen am Ende der ersten Klasse liefern. Diese konnten bereits stabile Ergebnisse in den Bereichen der phonologischen Bewusstheit, der orthografischen Kompetenz sowie der Semantik zeigen. In diesem Zuge konnte die phonologische Bewusstheit als eine essenzielle Grundlage für den Schriftspracherwerb definiert werden.

Die Arbeit verdeutlicht, dass Verfahren zur Erfassung des Schriftspracherwerbs nur dann wirksam und valide sein können, wenn sie altersgerecht gestaltet sind, eine klare Zielsetzung verfolgen, die individuellen Kompetenzen der Schüler:innen differenziert erfassen, gleichzeitig praktikabel in der Anwendung sind und auf wissenschaftlich fundierten Standards basieren. Indem diese Arbeit die Schwächen bestehender Tests aufzeigt und ein neues Instrument entwickelt, das an aktuelle wissenschaftliche und pädagogische Anforderungen angepasst ist, trägt sie zur Weiterentwicklung der Erfassung des frühen Schriftspracherwerbs bei. Sie bietet somit eine Grundlage für weiterführende Forschung und praxisnahe Anwendungen.

Um die zweite Hypothese weiter zu stützen, wäre eine erneute Evaluierung der Reliabilität interessant, in der in den Fokus gerückt wird, ob der Test konsistente Ergebnisse liefert, wenn er zu unterschiedlichen Zeitpunkten oder bei verschiedenen Gruppen von Schüler:innen angewendet wird. Für weitere Aussagen zu der Entwicklung des Schriftspracherwerbs dieser Stichprobe könnte eine Längsschnittstudie die Möglichkeit eröffnen, Rückschlüsse zu den Verläufen zu erlangen. Ebenfalls kann eine anschließende Interventionsstudie darauf abzielen, auf Basis der erhobenen Leistungsstände gezielte Fördermaßnahmen zu entwickeln und deren Wirksamkeit empirisch zu überprüfen. Dies trägt zu einem von Lenhard (2019) beschriebenen Kreislauf der systematischen Evaluation bei: Nach der Erhebung des aktuellen Leistungsstandes folgt das Erkennen von Ansatzpunkten für gezielte Förderung und anschließend die Überprüfung der Interventionsmaßnahmen. Wie bereits in vielen, in dieser Arbeit nur kurz genannten, Studien aufgezeigt werden konnte, gibt es im Kontext des Schriftspracherwerbs zahlreiche Einflussfaktoren: der sozioökonomische Status (Linberg & Wenz 2017; Wawire et al. 2024; Walper & Grgic 2013), Mehrsprachigkeit (Berghoff & Bylund 2023; Thomas-Sunesson et al. 2018) und der Input (Schulze & Saalbach 2022; Pfost et al. 2010; Lehl et al. 2012) stellen nur Beispiele dar. Diese konnten in dieser Arbeit aufgrund des nicht auswertbaren Fragebogens nicht näher beleuchtet werden. Der Einbezug dieser Faktoren kann Erkenntnisse zu möglichen Zusammenhängen mit den Testergebnissen und dem daraus resultierenden Stand des Schriftspracherwerbs liefern. Auch der Faktor der neuen Medien ist ein in dem wissenschaftlichen Diskurs stetig aufkommendes Thema (Bredel et al. 2017; Schneider 2017), welches in Kombination mit dem Einsatz des Tests interessante Ergebnisse liefern kann.

Abschließend verdeutlicht diese Arbeit die Notwendigkeit einer wissenschaftlich fundierten Erfassung des Schriftspracherwerbs in der ersten Klasse. Das entwickelte Testinstrument hat sich in diesem Zuge als eine deutliche Verbesserung gegenüber den bestehenden Verfahren erwiesen.

IV. Literaturverzeichnis

- Abe, T. & Gbenro, S. (2024). Effects of Test Anxiety On Students' Performance in Mathematics in Senior Secondary Schools in Ekiti State. *British Journal of Multidisciplinary and Advanced Studies Education, Learning, Training & Development*, 5(3), 25-35.
- Alves, L., dos Santos, L., Miranda, I., Carvalho, I., Ribeiro, G., Freire, L., Martins-Reis, V. & Celeste, C. (2021). Reading speed in Elementary School and Junior High. *CoDAS*, 33(5).
- Becker-Lenz, R. & Lüscher, D. (2012). Der Nutzen der objektiv hermeneutischen Diagnostik in der Sozialen Arbeit. *Neue Praxis*, 5, 475-492.
- Berghoff, R. & Bylund, E. (2024). L2 activation during L1 processing is increased by exposure but decreased by proficiency. *International Journal of Bilingualism*, 28(3), 555-569.
- Berlin-Brandenburgische Akademie der Wissenschaften (o.A.). *Erschließung: Worthäufigkeit*. Verfügbar unter: <https://www.dwds.de/d/worthaeufigkeit> [30.12.2024].
- Bosse, M. L. (2015). Learning to Read and Spell: How Children Acquire Word Orthographic Knowledge. *Child Development Perspectives*, 9(4), 222-226.
- Bredel, U., Fuhrhopp, N. & Noack, C. (2017). *Wie Kinder lesen und schreiben lernen* (2. Aufl.). Narr Francke Attempto.
- Browne Rego, L., & Bryant, P. (1993). The connection between phonological, syntactic and semantic skills and children's reading and spelling. *European Journal of Psychology of Education*, 8, 235-246.
- Byrd, A., Oppenheimer, K., Silvermann, R. & Edwards, J. (2024). Filling in the Blank: The Development of a Writing Screener for Elementary School Students Who Speak African American English. *Language, Speech, and Hearing Services in Schools*, 55(2), 598-606.
- Chamizo-Nieto, M., Arrivillaga, C., Rey, L., & Extremera, N. (2021). The role of emotional intelligence, the teacher-student relationship, and flourishing on academic performance in adolescents: a moderated mediation study. *Frontiers in psychology*, 12, 695067.
- Del Toro, I. & Thomé, G. (2021). Meilensteine im Erwerb der deutschen Orthografie. *Lernen und Lernstörungen*, 10(3), 169-180.
- Diehl, K. & Hartke, B. (2012a). *IEL-1. Inventar zur Erfassung der Lesekompetenz im 1. Schuljahr. Testheft*. Hogreve Verlag.
- Diehl, K. & Hartke, B. (2012b). *IEL-1. Inventar zur Erfassung der Lesekompetenz im 1. Schuljahr. Manual*. Hogreve Verlag.
- Dornheim, D. (2008). *Prädiktion von Rechenleistung und Rechenschwäche: Der Beitrag von Zahlen-Vorwissen und allgemein-kognitiven Fähigkeiten*. Logos Verlag Berlin.
- Fitriana, R. & Agustina, A. (2019). Phonological Acquisition (Case Study on Indonesian Child). *Seventh International Conference on Languages and Arts (ICLA 2018)*, 301, 485-490.
- Frith, U. (1986). Psychologische Aspekte des orthographischen Wissens: Entwicklung und Entwicklungsstörung. In Augst, G. (Hrsg.). *New Trends in Graphemics and Orthography*. De Gruyter, 218-233.

- Frost, J., Madsbjerg, S., Niedersøe, J., Olofsson, Å. & Sørensen, P. (2005). Semantic and Phonological Skills in Predicting Reading Development: From 3–16 Years of Age. *Dyslexia*, 11(2), 79-92.
- Griffiths, S. & Johnson, C. (1995). Effects of training on fricative identification in toddlers. *Applied Psycholinguistics*, 16(4), 443-462.
- Gutzmann, D. (2020). *Semantik. Eine Einführung*. J.B. Metzler Verlag.
- Harden, T. (2023). Wo sind die Wörter im Kopf und wie greift man auf sie zu? Mentales Lexikon. In Hentschel, E. & Harden, T. (Hrsg.). *Einführung in die germanistische Linguistik*. Peter Lang Classics, 15-27.
- Helmke, A. (2015). *Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts*. Klett & Kallmeyer.
- Heiner, M. (2018). Diagnostik in der Sozialen Arbeit. In Otto, H., Thiersch, H., Treptow, R. & Ziegler, H. (Hrsg.). *Handbuch Soziale Arbeit. Grundlagen der Sozialarbeit und Sozialpädagogik*. utb, 242-255.
- Hogan, T., Catts, H. & Little, T. (2005). The relationship between phonological awareness and reading. *Lang Speech Hearv Serv Sch*, 36(4), S. 285-293.
- Huber, C., Grosche, M. & Schütterle, P. (2013). Inklusive Schulentwicklung durch response-to-intervention (RTI)–Realisierungsmöglichkeiten des RTI-Konzepts im Förderbereich Lesen. *Gemeinsam leben*, 2, 79-91.
- Kauschke, C. (2012). *Kindlicher Spracherwerb im Deutschen: Verläufe, Forschungsmethoden, Erklärungsansätze* (Vol. 45). Walter de Gruyter.
- Klauer, K. (2014). Formative Leistungsdiagnostik: Historischer Hintergrund und Weiterentwicklung zur Lernverlaufsdiagnostik. In Hasselhorn, M., Schneider, W. & Trautwein, U. (Hrsg.). *Lernverlaufsdiagnostik*. Hogrefe, S. 1-17.
- Leibniz-Institut für Deutsche Sprache (o.A.). *OWID*. Verfügbar unter: <https://www.owid.de/> [31.12.2024].
- Lehrl, S., Ebert, S., Roßbach, H.-G., & Weinert, S. (2012). Die Bedeutung der familiären Lernumwelt für Vorläufer schriftsprachlicher Kompetenzen im Vorschulalter. *Zeitschrift für Familienforschung*, 24(2), 115-133.
- Lenhard, W. & Schneider, W. (2006a). *ELFE 1-6. Ein Leseverständnistest für Erst- bis Sechstklässler. Testheft*. Hogreve Verlag.
- Lenhard, W. & Schneider, W. (2006b). *ELFE 1-6. Manual*. Hogreve Verlag.
- Lenhard, A., Lenhard, W., Segerer, R. & Suggate S. (2015). *PPVT-4. Peabody Picture Vocabulary Test-4. Ausgabe. Auswertungsbogen*. NCS Pearson, Inc.
- Lenhard, W. (2019). *Leseverständnis und Lesekompetenz: Grundlagen – Diagnostik – Förderung* (2. Aufl.). Kohlhammer.
- Linberg, T., & Wenz, S. (2017). Ausmaß und Verteilung sozioökonomischer und migrationsspezifischer Ungleichheiten im Sprachstand fünfjähriger Kindergartenkinder. *Journal for educational research online*, 9(1), 77-98.
- Marx, P. (2007). *Lese- und Rechtschreiberwerb*. utb.
- McElvany, N., Becker, M., Lüdtke, O. (2009). Die Bedeutung familiärer Merkmale für Lesekompetenz, Wortschatz, Lesemotivation und Leseverhalten. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 41(2), 121-131.

- McElvany, N., Lorenz, R., Frey, A., Goldhammer, F., Schilcher, A. & Stubbe, T. C. (2023). *IGLU 2021. Lesekompetenz von Grundschulkindern im internationalen Vergleich und im Trend über 20 Jahre*. Waxmann.
- Mehlem, U. (2024). *Schriftspracherwerb. Theorie und Praxis für den Anfangsunterricht in der Grundschule*. Kohlhammer.
- Meindl, C. (2011). *Methodik für Linguisten. Eine Einführung in Statistik und Versuchsplanung*. narr Verlag.
- Mendes, B. & Kirby, J. (2024). The Effects of a Morphological Awareness Intervention on Reading and Spelling Ability of Children With Dyslexia. *Learning Disability Quarterly*, 47(4), 222-233.
- Milankov, V., Golubović, S., Krstić, T. & Golubović, Š. (2021). Phonological awareness as the foundation of reading acquisition in students reading in transparent orthography. *International journal of environmental research and public health*, 18(10), 5440.
- Ministerium für Schule und Bild des Landes Nordrhein-Westfalen [MSB] (2024a). *Übersicht über die Sozialindexstufen der Schulen nach Bezirksregierung, Kreis / kreisfreier Stadt und Schulform ab dem Schuljahr 2024/25*. Verfügbar unter: https://www.schulministerium.nrw/system/files/media/document/file/schulsozialindex_schulliste_2024_2025.pdf [07.12.2024].
- Ministerium für Schule und Bild des Landes Nordrhein-Westfalen [MSB] (2024b). *Schulsozialindex*. Verfügbar unter: <https://www.schulministerium.nrw/schulsozialindex> [07.12.2024].
- Organisation for Economic Co-operation and Development [OECD] (2024). *Programme for International Student Assessment (PISA)*. Verfügbar unter: <https://www.oecd.org/en/about/programmes/pisa.html> [11.01.2025].
- Pfost, M., Dörfler, T. & Artelt, C. (2010). Der Zusammenhang zwischen außerschulischem Lesen und Lesekompetenz. Ergebnisse einer Längsschnittstudie am Übergang von der Grund- in die weiterführende Schule. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 42(3), 167-176.
- Prediger, S., Link, M., Hinz, R., Hußmann, S., Thiele, J. & Ralle, B. (2012). Lehr-Lernprozesse initiieren und erforschen – Fachdidaktische Entwicklungsforschung im Dortmunder Modell. In: *MNU*, 65(8), 452–457.
- Qualitäts- und UnterstützungsAgentur - Landesinstitut für Schule [QUA-LiS NRW] (2024a). *Kompetenzerwartungen*. Verfügbar unter: <https://www.schulentwicklung.nrw.de/lehrplaene/lehrplannavigator-grundschule/deutsch/lehrplan-deutsch/kompetenzen/index.html> [03.12.2024].
- Qualitäts- und UnterstützungsAgentur - Landesinstitut für Schule [QUA-LiS NRW] (2024b). *Aufgaben und Ziele*. Verfügbar unter: <https://www.schulentwicklung.nrw.de/lehrplaene/lehrplannavigator-grundschule/deutsch/lehrplan-deutsch/aufgaben-ziele/index.html> [03.12.2024].
- Qualitäts- und UnterstützungsAgentur - Landesinstitut für Schule [QUA-LiS NRW] (2024c). *Grundwortschatz. Wortfilter*. <https://www.schulentwicklung.nrw.de/cms/grundwortschatz-nrw/grundwortschatz/wortfilter/index.html> [30.12.2024].
- Reber, K. (2017). *Lesen und Schreiben lernen: Wie erobern Kinder die Schriftsprache?* Springer.

- Retelsdorf, J. & Köller, O. (2014). Reciprocal effects between reading comprehension and spelling. *Learning and Individual Differences*, 30, 77-83.
- Rojo López, A., Cifuentes Férez, P. & Espín López, L. (2021). The influence of time pressure on translation trainees' performance: Testing the relationship between self-esteem, salivary cortisol and subjective stress response. *PloS one*, 16(9), e0257727.
- Rosebrock, C. & Nix, D. (2015). *Grundlagen der Lesedidaktik und der systematischen schulischen Leseförderung* (7. Aufl.). Schneider.
- Rosebrock, C. & Scherf, D. (2022). *Lesedidaktik? Literaturdidaktik? Ein Dutzend Antworten für Einsteiger* (2. Aufl.). Schneider Verlag Hohengehren.
- Sáez-Delgado, F., López-Agulu, Y., Mella-Norambuena, J., Baeza-Sepúlveda, C., Contreras-Saavedra, C. & Lozano-Peña, G. (2022). Teacher Self-Regulation and Its Relationship with Student Self-Regulation in Secondary Education. *Sustainability*, 14(24), 16863.
- Sassenroth, M. (2000). *Schriftspracherwerb. Entwicklungsverlauf, Diagnostik und Förderung*. Haupt.
- Scheerer-Neumann, G. (2023). *Lese-Rechtschreib-Schwäche und Legasthenie: Grundlagen, Diagnostik und Förderung*. Kohlhammer.
- Schneider, W. & Näslund, J. (1993). The impact of early metalinguistic competencies and memory capacity on reading and spelling in elementary school: Results of the Munich Longitudinal Study on the Genesis of Individual Competencies (LOGIC). *European Journal of Psychology of Education*, 8, 273–287.
- Schneider, W. (2017). *Lesen und Schreiben lernen. Wie erobern Kinder die Schriftsprache?* Springer.
- Schründer-Lenzen, A. (2004). *Schriftspracherwerb und Unterricht. Bausteine professionellen Handlungswissens*. Leske + Budrich.
- Sekretariat der Ständigen Konferenz der Kulturlminister der Länder in der Bundesrepublik Deutsch [KMK] (2022a). *Bildungsstandards für das Fach Deutsch. Primarbereich*. Verfügbar unter: https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2022/2022_06_23-Bista-Primarbereich-Deutsch.pdf [23.11.2024].
- Sekretariat der Ständigen Konferenz der Kulturlminister der Länder in der Bundesrepublik Deutsch [KMK] (2022b). *Standards für die Lehrerbildung: Bildungswissenschaften*. Verfügbar unter: https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_12_16-Standards-Lehrerbildung.pdf [27.12.2024].
- Schulze, C. & Saalbach, H. (2022). Socio-cognitive engagement (but not socioeconomic status) predicts preschool children's language and pragmatic abilities. *Journal of Child Language*, 49(4), 839-849.
- Seymour, P., Aro, M., Erskine, J. & Collaboration with COST Action A8 Network. (2003). Foundation literacy acquisition in European orthographies. *British Journal of psychology*, 94(2), 143-174.
- Stock, C., Marx, P. & Schneider, W. (2003a). *BAKO 1-4. Protokollbogen*. Beltz.
- Stock, C., Marx, P. & Schneider, W. (2003b). *BAKO 1-4. Basiskompetenzen für Lese-Rechtschreibleistungen. Manual*. Beltz.
- Technische Universität Dortmund (o.A.). *Das Projekt*. Verfügbar unter: <https://beemehr.kuwi.tu-dortmund.de/> [26.01.2025].

- Thomas-Sunesson, D., Hakuta, K. & Bialystok, E. (2018). Degree of bilingualism modifies executive control in Hispanic children in the USA. *International journal of bilingual education and bilingualism*, 21(2), 197-206.
- Valdois, S., Zaher, A., Meyer, S., Diard, J., Mandin, S. & Bosse, M. (2024). Effectiveness of Visual Attention Span Training on Learning to Read and Spell: A Digital-game-based Intervention in Classrooms. *Reading Research Quarterly*.
- Van der Straten Waillet, P., Crowe, K., Charlier, B. & Colin, C. (2023). Assessing the speech production of multilingual children: A survey of speech-language therapists in French-speaking Belgium. *Int J Lang Commun Discord* 58(5), 1496-1509.
- Van Teijlingen, E. & Hundley, V. (2001). The importance of pilot studies. *Social research update*, 35, 1-4.
- Vettori, G., Incognito, O. & Bigozzi, L. & Pinto, G. (2023). Relationship between lexical, reading and spelling skills in bilingual language minority children and their monolingual peers. *Frontiers in Psychology*, 14.
- Walper, S. & Grgic, M. (2013). Verhaltens- und Kompetenzentwicklung im Kontext Familie. Zur relativen Bedeutung von sozialer Herkunft, elterlicher Erziehung und Aktivitäten in der Familie. *Zeitschrift für Erziehungswissenschaft*, 16, 503-531.
- Wang, L. (2022). Exploring the relationship among teacher emotional intelligence, work engagement, teacher self-efficacy, and student academic achievement: A moderated mediation model. *Frontiers in psychology*, 12, 810559.
- Wawire, B., Wawire, G., & Kiroro, F. (2024). The Structural Relations of Component Reading Comprehension Skills in Kiswahili: The Influence of Socio-Economic Status and Home Literacy Environment. *Reading Research Quarterly*, 60(1).
- Williams, S. & Horst, J. (2014). Goodnight book: sleep consolidation improves word learning via storybooks. *Frontiers in psychology*, 5, 184.
- Yalcinkaya, F., Muluk, N. & Budak, B. (2010). Speech Sounds Acquisition Evaluated by Speech Sound Development Test (SSDT) in Turkish-Speaking Children. *The Journal of International Advanced Otolaryngology*, 6(1), 60-66.

V. Anhangsverzeichnis

Anhang 1: Testseiten.....	vi
Anhang 2: Protokollbögen.....	xiii
Anhang 3: tabellarisches Kodierungssystem.....	v
Anhang 4: Excel-Auswertung Punktevergabe.....	ix
Anhang 5: Wiederholungsbewertung: zweite Excel-Auswertung Punktevergabe.....	x

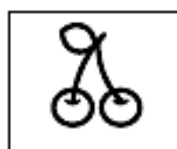
Anhang 1: Testseiten

Buchstaben-Laut-Zuordnung

Wo hörst du den Laut?

Kreuze an.

E


☐ ☒ ☐
☐ ☐ ☒
☒ ☐ ☒

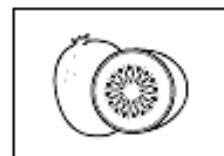
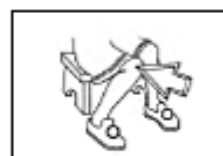
O


☐ ☐ ☐
☐ ☐ ☐
☐ ☐ ☐
☐ ☐ ☐

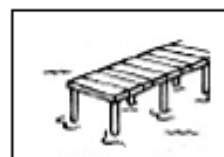
A


☐ ☐ ☐
☐ ☐ ☐
☐ ☐ ☐
☐ ☐ ☐

I


☐ ☐ ☐
☐ ☐ ☐
☐ ☐ ☐
☐ ☐ ☐

E

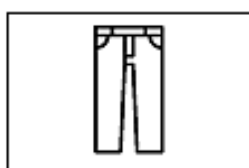

☐ ☐ ☐
☐ ☐ ☐
☐ ☐ ☐
☐ ☐ ☐

U


☐ ☐ ☐
☐ ☐ ☐
☐ ☐ ☐
☐ ☐ ☐

Wörter in Silben segmentieren

Schreib die Silben in die Silbenbögen.



Ho se













Setz Silbenbögen unter die Wörter.

P a n t a l o n

Q u i n t e t t

D e p u t a t

E p i g o n e

F i l t r i e r u n g

K e n o t a p h

B i g a m i e

F u r a g e

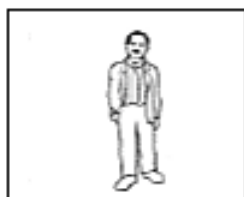
A l k o v e n

O s k u l a t i o n

D i v i n a t i o n

Buchstaben zu Wörtern verbinden

Setz die fehlenden Buchstaben ein.



Ma_nn



Ele_f_ant



Br__t



__auch



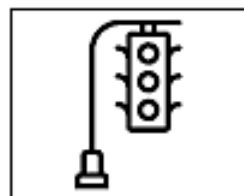
H__s



Hef__



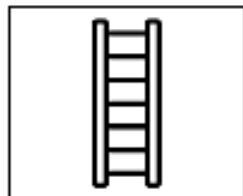
__olke



Ampe__



Flugz__g



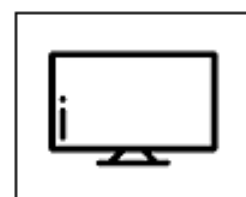
L__ter



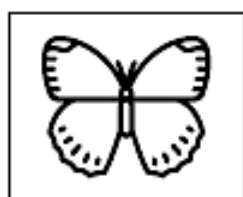
__euerwehr



To__ate



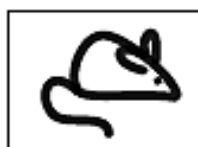
Fernsehe__



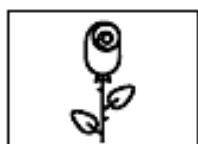
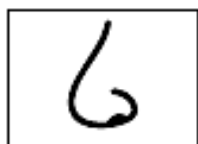
__metterling

Wörter

Schreib die Wörter.



Maus



Satzverstehen

Lies den Satz. Baue mit Playmobil nach.

1. Der Mann steht neben dem Baum.

2. Der Schal liegt auf dem kleinen Stuhl.

3. Vor dem Pferd sitzt ein schwarz-weißer Hund.

4. Neben dem Stein blühen eine kleine und zwei große Blumen.

Textverstehen

Verbinde den Text mit dem passenden Bild.



Die Kinder spielen mit der Puppe. Sie spielen Vater, Mutter und Kind. Die Katze sitzt auf dem Regal. Das Spielzeug ist in der Kiste.



Wir wollen einen Salat aus Obst machen. Mama geht einkaufen. Im Regal liegt viel Obst. Im Wagen liegen Bananen, Äpfel, Birnen und Orangen.

Anhang 2: Protokollbögen

Protokollbogen Vokallänge (S2)

Es soll das Wort benannt werden, welches einen kurzen Vokal in einer Reihe von Wörtern mit langem Vokal oder welches einen langen Vokal in einer Reihe von Wörtern mit kurzem Vokal aufweist.

1	maar ☐	raas ☐	dack ☐	laat ☐
2	deek ☐	kett ☐	pesch ☐	zenn ☐
3	roof ☐	soof ☐	pfoor ☐	jomm ☐
4	reem ☐	feer ☐	nell ☐	beef ☐
5	wuul ☐	duck ☐	tupp ☐	pumm ☐
6	liem ☐	sief ☐	minn ☐	nier ☐
7	bamm ☐	jaal ☐	rack ☐	batz ☐
8	muur ☐	suuk ☐	ruul ☐	luch ☐
9	lotz ☐	boos ☐	konn ☐	goll ☐
10	tiez ☐	lick ☐	gipp ☐	piff ☐

Protokollbogen Satzverstehen (S7)

Die Proband:innen sollen die Sätze lesen und diese mit Playmobil nachbauen.

Das Kind...	☒ / ☐
...wählt den SCHAL aus.	
... wählt die richtige Größe (KLEIN) des Stuhls aus.	
...wählt den STUHL aus.	
... legt den Schal AUF den Stuhl.	
... wählt einen HUND aus.	
... wählt die richtige Farbe (SCHWARZ-WEIß) des Hundes aus.	
... wählt das PFERD aus.	
... setzt den Hund VOR das Pferd.	
... wählt BLUMEN aus.	
... wählt die richtige Anzahl an Blumen aus (EINE KLEINE UND ZWEI GROßE).	
... wählt den STEIN aus.	
... setzt die Blumen NEBEN dem Stein.	

Anhang 3: tabellarisches Kodierungssystem

Auswertung					
Untertest	Kodierungssystem	Mögliche Punkte	Items	Musterlösung	Alternative richtige Lösungen
1	1 richtiges Kreuz = 1 Punkt Alternativ: Richtig gesetzter Buchstabe = 1 Punkt	1	Oma	Anlaut	
		1	Auto	Auslaut	
		2	Otto	An- und Auslaut	
		1	Boot	Inlaut	
		1	Affe	Anlaut	
		1	Zahn	Inlaut	
		1	Sofa	Auslaut	
		2	Anna	An- und Auslaut	
		1	Knie	Auslaut	
		1	Igel	Anlaut	
		2	Kiwi	In- und Auslaut	
		1	Brief	Inlaut	
2	1 richtiges Kreuz = 1 Punkt	1	Emil	Anlaut	
		1	Steg	Inlaut	
		1	Schnee	Auslaut	
		2	Ente	An- und Auslaut	
		1	Buch	Inlaut	
		2	Uhu	An- und Auslaut	
		1	Uhr	Anlaut	
		1	Akku	Auslaut	
		1	maar, raas, dack, laat	dack	„das dritte Wort“ „das vorletzte Wort“
		1	deek, kett, pesch, zenn	deek	„das erste Wort“

		1	roof, soof, pfoor, jomm	jomm	„das vierte Wort“ „das letzte Wort“
		1	reem, feer, nell, beef	nell	„das dritte Wort“ „das vorletzte Wort“
		1	wuul, duck, tupp, pumm	wuul	„das erste Wort“
		1	liem, sief, minn, nier	minn	„das dritte Wort“ „das vorletzte Wort“
		1	bamm, jaal, rack, batz	jaal	„das zweite Wort“
		1	muur, suuk, ruul, luch	luch	„das vierte Wort“ „das letzte Wort“
		1	lotz, boos, konn, goll	boos	„das zweite Wort“
		1	tiez, lick, gipp, piff	tiez	„das erste Wort“
3	1 richtige Silbe = 1 Punkt 1 orthographisch richtig geschriebene Silbe = 1 Punkt	6	Banane	Ba na ne	
		6	Rakete	Ra ke te	Raum schiff (statt 1 Punkte jeweils 1,5), Raketen (6 Punkte)
		6	Zitrone	Zi tro ne	
		8	Regenbogen	Re gen bo gen	
		8	Schokolade	Scho ko la de	
		8	Badewanne	Ba de wan ne	Du sche (statt 1 Punkt, jeweils 2 Punkte) Wan ne (statt 1 Punkt jeweils 2 Punkte)
4	1 richtig gesetzter Silbenbogen = 1 Punkt	-	<i>Quintett</i>	<i>aus der Wertung ausgeschlossen</i>	
		3	Deputat	De pu tat	
		3	Epigone	Epi go ne	E pi go ne (E pi jeweils 0,5 Punkte)
		3	Filtrierung	Fil trie rung	
		3	Kenotaph	Ke no taph	
		3	Bigamie	Bi ga mie	
		3	Furage	Fu ra ge	
		3	Alkoven	Al ko ven	
		4	Oskulation	Os ku la tion	
		4	Divination	Di vi na tion	

5	1 richtig gesetzter Buchstabe = 1 Punkt Pro falschen zusätzlichen Buchstabe ein Minuspunkt	1	Br _ t	Brot	
		1	_ auch	Bauch	
		2	H _ s	Haus	
		1	Hef _	Heft	Hefte
		1	_ olke	Wolke	
		1	Ampe _	Ampel	
		2	Flugz _ g	Flugzeug	
		2	L _ ter	Leiter	Laiter (1 Punkt)
		1	_ euerwehr	Feuerwehr	
		1	To _ ate	Tomate	
		1	Fernsehe _	Fernseher	Fernsehen (1 Punkt)
		3	_ metterling	Schmetterling	Chmetterling (2 Punkte)
6	1 orthographisch richtig geschriebene Silbe = 1 Punkt	2	Nase		Schwanz (2 Punkte)
		2	Rose		Tulpe, Blume
		1	Baum		
		1	Mond		
		2	Krone		
		2	Löffel		
		3	Erdbeere		
		3	Hubschrauber		Heli kop ter (3 Punkte) Flug zeug (statt 1 Punkt 1,5 Punkte = 3 Punkte) Flug schrau ber (3 Punkte)
7	1 Haken = 1 Punkt 1 Kreuz = 0 Punkte	4	Der Schal liegt auf dem kleinen Stuhl.	Auswahl des Schals Auswahl des kleinen Stuhls Auswahl des Stuhls Legen des Schals auf den Stuhl	
		4	Vor dem Pferd sitzt ein schwarz-weißer Hund.	Auswahl des Hundes Auswahl des schwarz-weißen Hundes Auswahl des Pferdes Setzen des Hundes vor das Pferd	
		4	Neben dem Stein blühen eine kleine und zwei große Blumen.	Auswahl von Blumen Auswahl von einer kleinen und zwei großen Blumen	

				Auswahl des Steines Setzen der Blumen neben den Stein	
8	1 richtige Antwort = 1 Punkt	1	Die Kinder spielen mit der Puppe. Sie spielen Vater, Mutter und Kind. Die Katze sitzt auf dem Regal. Das Spielzeug ist in der Kiste.	Bild 3	
		1	Wir wollen einen Salat aus Obst ma- chen. Mama geht einkaufen. Im Re- gal liegt viel Obst. Im Wagen liegen Bananen, Äpfel, Birnen und Oran- gen.	Bild 2	

Anhang 4: Excel-Auswertung Punktevergabe

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Proband:in	Klasse	Subtest 1	Subtest 2		Subtest 3		Subtest 4	Subtest 5	Subtest 6	Subtest 7	Subtest 8	Gesamt	Gesamt ohne S2
2	P01	1a	20	3	12	11	23	21	15	10	7	2	101	98
3	P02	1a	16	2	17	12	29	15	14	9	7	0	92	90
4	P03	1b	23	2	19	17	36	16	14	12	9	2	114	112
5	P04	1b	19	1	16	12	28	14	17	12	11	2	104	103
6	P05	1b	19	1	16	15	31	18	17	9	11	2	108	107
7	P06	1b	14	3	15	11	26	4	12	8	6	0	73	70
8	P07	1b	23	4	21	20	41	20	16	13	9	2	128	124
9	P08	1a	22	4	21	21	42	20	17	12	10	2	129	125
10	P09	1a	22	5	13	13	26	18	15	10	9	2	107	102
11	P10	1b	20	3	17	13	30	19	16	7	11	2	108	105
12	P11	1a	19	3	21	20	41	15	16	11	9	2	116	113
13	P12	1a	21	3	21	16	37	24	13	10	9	1	118	115
14	P13	1a	25	5	21	20	41	22	17	8	11	2	131	126
15	P14	1b	24	4	21	19	40	28	17	13	10	0	136	132
16	P15	1b	23	6	21	19	40	25	17	13	9	2	135	129
17	P16	1b	23	2	21	19	40	9	16	10	10	2	112	110
18	P17	1a	24	2	16	16	32	14	16	10	12	2	112	110
19	P18	1a	23	2	21	20	41	7	17	13	11	2	116	114
20	P19	1a	23	5	21	20	41	19	17	11	11	1	128	123
21	P20	1a	19	3	19	17	36	10	17	11	10	2	108	105
22	P21	1a	24	3	21	21	42	17	14	13	11	2	126	123
23	P22	1a	22	6	21	20	41	17	15	11	12	2	126	120
24	P23	1a	20	3	21	20	41	22	17	11	10	2	126	123
25	P24	1b	18	2	12	10	22	15	15	11	11	1	95	93
26	P25	1b	24	6	19	18	37	20	17	10	8	2	124	118
27	P26	1b	22	3	8	8	16	16	17	10	10	2	96	93
28	P27	1b	22	2	21	20	41	17	17	11	9	2	121	119
29	P28	1b	23	3	21	19	40	22	15	11	8	2	124	121
30	P29	1b	15	2	18	12	30	16	10	4	6	0	83	81
31	P30	1b	24	2	13	12	25	21	17	12	9	2	112	110
32	P31	1b	24	4	17	17	34	18	15	12	11	2	120	116
33	P32	1a	24	4	21	20	41	17	17	13	12	2	130	126
34	P33	1a	25	4	19	18	37	22	17	11	11	1	128	124
35	P34	1a	21	2	21	21	42	20	17	10,5	9	2	123,5	121,5
36	P35	1b	22	3	12	11	23	16	15	10	11	2	102	99
37	P36	1a	18	3	21	21	42	19	15	14	10	2	123	120
38	P37	1b	25	6	21	20	41	19	17	14	12	2	136	130
39	P38	1b	25	1	21	18	39	11	14	8	11	2	111	110
40	P39	1b	23	2	16	14	30	17	13	10	12	1	108	106
41	P40	1b	25	3	21	20	41	24	17	12,5	10	2	134,5	131,5
42	P41	1b	20	4	21	13	34	21	10	8	10	0	107	103
43	P42	1a	22	5	21	20	41	17	17	11	9	1	123	118
44	P43	1a	25	4	21	20	41	15	14	11	10	2	122	118
45	P44	1a	25	2	21	18	39	19	17	12	11	2	127	125
46	P45	1a	20	3	21	20	41	21	14	11	9	2	121	118
47	P46	1a	25	2	16	15	31	9	17	12	10	2	108	106
48	P47	1b	23	nicht ausgefüllt	21	19	40	25	16	10	10	2	126	126
49					Silbenstruktur	Orthographie	Gesamt							

Anhang 5: Wiederholungsbewertung: zweite Excel-Auswertung Punktevergabe

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Proband:in	Klasse	Subtest 1	Subtest 2		Subtest 3		Subtest 4	Subtest 5	Subtest 6	Subtest 7	Subtest 8	Gesamt	Gesamt ohne S2
2	P01	1a	20	3	12	11	23	23	15	10	7	2	103	100
3	P02	1a	16	2	17	12	29	15	14	9	7	0	92	90
4	P03	1b	23	2	19	17	36	16	14	12	9	2	114	112
5	P04	1b	19	1	16	12	28	14	17	12	11	2	104	103
6	P05	1b	19	1	16	15	31	18	17	9	11	2	108	107
7	P06	1b	14	3	15	11	26	4	12	8	6	0	73	70
8	P07	1b	23	4	21	20	41	20	16	13	9	2	128	124
9	P08	1a	22	4	21	21	42	20	17	12	10	2	129	125
10	P09	1a	22	5	13	13	26	18	15	10	9	2	107	102
11	P10	1b	20	3	17	13	30	19	16	7	11	2	108	105
12	P11	1a	19	3	21	20	41	15	16	11	9	2	116	113
13	P12	1a	21	3	21	16	37	24	13	10	9	1	118	115
14	P13	1a	25	5	21	20	41	22	17	8	11	2	131	126
15	P14	1b	24	4	21	19	40	28	17	13	10	0	136	132
16	P15	1b	23	6	21	19	40	25	17	13	9	2	135	129
17	P16	1b	23	2	21	19	40	9	16	10	10	2	112	110
18	P17	1a	24	2	16	16	32	14	16	10	12	2	112	110
19	P18	1a	23	2	21	20	41	7	17	13	11	2	116	114
20	P19	1a	23	5	21	20	41	19	17	11	11	1	128	123
21	P20	1a	19	3	19	17	36	10	17	11	10	2	108	105
22	P21	1a	24	3	21	21	42	17	14	13	11	2	126	123
23	P22	1a	22	6	21	20	41	17	15	11	12	2	126	120
24	P23	1a	20	3	21	20	41	22	17	11	10	2	126	123
25	P24	1b	18	2	12	10	22	15	16	11	11	1	96	94
26	P25	1b	24	6	19	18	37	20	17	10	8	2	124	118
27	P26	1b	22	3	8	8	16	16	17	10	10	2	96	93
28	P27	1b	22	2	21	20	41	17	17	11	9	2	121	119
29	P28	1b	23	3	21	19	40	22	17	11	8	2	126	123
30	P29	1b	15	2	18	12	30	16	10	4	6	0	83	81
31	P30	1b	24	2	13	12	25	21	17	12	9	2	112	110
32	P31	1b	24	4	17	17	34	18	15	12	11	2	120	116
33	P32	1a	24	4	21	20	41	17	17	13	12	2	130	126
34	P33	1a	25	4	19	18	37	22	17	11	11	1	128	124
35	P34	1a	21	2	21	21	42	20	17	10,5	9	2	123,5	121,5
36	P35	1b	22	3	12	10	22	16	15	10	11	2	101	98
37	P36	1a	18	3	21	21	42	19	15	14	10	2	123	120
38	P37	1b	25	6	21	20	41	19	17	14	12	2	136	130
39	P38	1b	25	1	21	18	39	11	14	8	11	2	111	110
40	P39	1b	23	2	16	14	30	18	14	10	12	1	110	108
41	P40	1b	25	3	21	20	41	24	17	12,5	10	2	134,5	131,5
42	P41	1b	20	4	21	13	34	21	10	8	10	0	107	103
43	P42	1a	22	5	21	20	41	17	17	11	9	1	123	118
44	P43	1a	25	4	21	20	41	15	14	11	10	2	122	118
45	P44	1a	25	2	21	19	40	19	17	12	11	2	128	126
46	P45	1a	20	3	21	20	41	21	14	11	9	2	121	118
47	P46	1a	25	2	16	15	31	9	17	12	10	2	108	106
48	P47	1b	23	nicht bearbeitet	21	19	40	25	16	11	10	2	127	127
49					Silbenstruktur	Orthographie	Gesamt							