

# ENTWICKLUNG UND ERFORSCHUNG INKLUSIVER BILDUNGSPROZESSE

**Masterthesis**

**Geschlechterspezifische Unterschiede in der Entwicklung von Leseflüssigkeit und sinnentnehmendem Lesen hinsichtlich zusätzlicher Fördermaßnahmen im Rahmen einer Längsschnittuntersuchung in drei Grundschulklassen der Jahrgangsstufe zwei in Levumi**

vorgelegt von

**Melina Bacha**

**[melina.bacha@tu-dortmund.de](mailto:melina.bacha@tu-dortmund.de)**

Master Lehramt für sonderpädagogische Förderung nach LABG 2009

**Betreuende: Prof. Dr. Markus Gebhardt**

**Dr. Michael Schurig**

**ausgegeben am: 17.07.2018**

**eingereicht am: 15.10.2018**

**I Inhaltsverzeichnis**

<b>I Inhaltsverzeichnis</b> .....	<b>II</b>
<b>II Abstract</b> .....	<b>IV</b>
<b>1 Einleitung</b> .....	<b>5</b>
1.1 Problemstellung.....	5
1.2 Zielsetzung.....	6
1.3 Vorgehensweise.....	7
<b>2 Theorierahmen</b> .....	<b>8</b>
2.1 Modelle zum Schriftspracherwerb .....	8
2.2 Lesekompetenz.....	11
2.2.1 Leseflüssigkeit.....	12
2.2.2 Sinnentnehmendes Lesen.....	14
2.2.3 Einflussfaktoren.....	15
2.3 Forschungsstand geschlechterspezifischer Unterschiede im Lesen .....	17
2.3.1 Geschlechterspezifische Unterschiede der Lesesozialisation und Leseleistung .....	18
2.3.2 Geschlechterspezifische Unterschiede im Lernverlauf im Rahmen von Levumi .....	20
2.4 Leseförderung .....	21
2.5 Der Kieler Leseaufbau.....	23
<b>3 Formative Diagnostik und Leistungsmessung</b> .....	<b>26</b>
3.1 Anforderungen an Diagnostik im schulischen Kontext .....	26
3.1.1 Gütekriterien.....	26
3.1.2 Bezugsnormen .....	27
3.2 Curriculum-based Measurement .....	28
3.3 Lernverlaufsdagnostik Lesen.....	30
3.4 Einzelfallforschung .....	33
3.5 Die Online-Plattform Levumi.....	34
3.5.1 Lernverlaufsmessung mit Levumi .....	34
3.5.2 Förderansatz von Levumi .....	37
<b>4 Fragestellung</b> .....	<b>39</b>
4.1 Darlegung der Relevanz der Fragestellung im Kontext von Schul- und Unterrichtsentwicklung.....	39
4.2 Darlegung der Fragstellungen .....	40

---

<b>5 Forschungsmethoden .....</b>	<b>43</b>
5.1 Stichprobe .....	43
5.2 Testdurchführung .....	44
5.3 Levumi-Förderung .....	45
5.4 Auswertungsmethoden.....	46
<b>6 Auswertung und Ergebnisse.....</b>	<b>49</b>
<b>7 Diskussion.....</b>	<b>85</b>
<b>8 Zusammenfassung und Ausblick .....</b>	<b>98</b>
<b>III Literaturverzeichnis .....</b>	<b>V</b>
<b>IV Abbildungsverzeichnis .....</b>	<b>XI</b>
<b>V Tabellenverzeichnis .....</b>	<b>XII</b>
<b>VI Anhang .....</b>	<b>XIII</b>
<b>Eidesstattliche Versicherung .....</b>	<b>XXXI</b>

---

## II Abstract

### Zusammenfassung

In vielfältigen Studien konnten geschlechterspezifische Unterschiede in den schriftsprachlichen Kompetenzen zu Gunsten der Mädchen in der Primar- und Sekundarstufe aufgedeckt werden. Im Kontext der vorliegenden Forschungsarbeit mit dem Schwerpunkt der Lesekompetenz wurden die jeweils sechs leseschwächsten Schülerinnen und Schüler aus drei verschiedenen zweiten Klassen über neun Messzeitpunkte hinweg mit der Online-Plattform Levumi getestet. Dabei handelt es sich um ein neu entwickeltes Instrument der Lernverlaufsdagnostik, mit welchem wöchentlich drei Speedtests zur Leseflüssigkeit und ein Test zum sinnentnehmenden Lesen durchgeführt wurden. Ziel dieser Arbeit besteht in der Beantwortung der Fragestellung nach geschlechterspezifischen Unterschieden im Lernverlauf bei zusätzlicher Förderung auf Grundlage vorangegangener Diagnostik. Die aus insgesamt vier einzelnen Forschungsfragen resultierenden Ergebnisse zeigen eine Geschlechtertendenz zu Gunsten der Jungen vor der Förderung sowie zu Gunsten der Mädchen am Ende der Förderung. Damit zusammenhängend zeigen die Mädchen größere Leistungszuwächse in der Leseflüssigkeit als die Jungen. Demzufolge entsprechen die vorliegenden Ergebnisse denen vorangegangener Studien zum Teil.

### Abstract

Various studies have detected gender-related differences in the written language competencies in favour of girls at primary and secondary education schools. In the context of this research paper with its focus on reading skills, six pupils with the lowest reading performances from each of the three second year classes were tested over nine times of measurement with the help of the online platform Levumi. The platform is a recently developed tool for curriculum-based measurement. With its help, three reading fluency speed tests were performed weekly, as well as one reading comprehension test with a focus on extracting meaning. The aim of this paper is to investigate gender-related differences in the learning process while promoting the pupil's competence based on previous diagnostics. The results of the four research questions demonstrate a tendency to the benefit of boys before promoting their competence and to the benefit of girls after the promotion of competence. Linked to these results, girls show a higher increase in their reading skills performance than boys. Thus, the results of this paper correspond to previous studies to some extent.

## 1 Einleitung

Lesekompetenz stellt die Grundlage für schulischen Erfolg dar. Nicht nur im Fach Deutsch, sondern für alle Unterrichtsfächer ist eine ausreichende Lesekompetenz von fundamentaler Bedeutung, um Aufgaben und Inhalte zu verstehen sowie um weiterführende Kompetenzen zu erwerben. Darüber hinaus bietet die Fähigkeit zu lesen zahlreiche Möglichkeiten für auto-didaktische Lernprozesse sowie für berufliche und gesellschaftliche Teilhabe (Heinzl et al., 2015). Der Institution Schule, insbesondere dem Primarbereich, kommt insofern eine besondere Bedeutung zu, als dass sie den primären Ort des Schriftspracherwerbs darstellt (Kuhl, Gebhardt, Bienstein, Käßler, Quinten, Ritterfeld, Tröster, & Wember, 2017).. Damit Schülerinnen und Schüler ihre Lesekompetenz sukzessive erweitern, sind von Lehrkräften unterschiedliche Herausforderungen zu bewältigen und Probleme zu lösen, welche in diesem Kapitel aufgegriffen werden. Anschließend wird das Ziel dieser Forschungsarbeit sowie die Vorgehensweise in kürze dargelegt.

### 1.1 Problemstellung

Eine hohe Bildungsqualität ergibt sich aus einem hohen Leistungsniveau bei geringer Streuung der Leistungen. Die Leseleistungen von Schülerinnen und Schülern der vierten Klasse, die im Rahmen der internationalen Grundschul-Leseuntersuchung 2016 (IGLU) erhoben wurden, zeigen jedoch eine auffällig hohe Streubreite. Die Standardabweichung und die damit verbundene Leistungsheterogenität stieg seit der Erhebung im Jahr 2001 immer weiter an. Insgesamt befinden sich die Leseleistungen der deutschen Viertklässlerinnen und Viertklässlern im internationalen Vergleich im Mittelfeld. Darüber hinaus erreichen ca. 19% der Schülerinnen und Schüler nicht die Kompetenzstufe III, wobei Stufe V die höchste Stufe darstellt. Die niedrigen Leseleistungen dieses Anteils der Schülerinnen und Schüler können fächerübergreifende Probleme in der Sekundarstufe I hervorrufen (Bos et al., 2017). Die zunehmende Leistungsheterogenität innerhalb eines Klassenzimmers auf Grund von Zuwanderung, Inklusion und verschiedenen sprachlichen Vorerfahrungen sowie individuellen Lernfaktoren fordert eine bedürfnisorientierte Förderung. Dabei stellt Diagnosekompetenz eine zentrale Fähigkeit von Lehrkräften dar, die benötigt wird, um leistungsschwächere Schülerinnen und Schüler zu identifizieren. An dieser Stelle beschreibt jedoch die *Wait-to-fail-Problematik* die Schwierigkeit, der späten Identifikation und Einleitung von Fördermaßnahmen für lese-schwache Schülerinnen und Schülern mit oder ohne sonderpädagogischen Förderbedarf (Huber und Grosche, 2012). Werden Schwierigkeiten im Schriftspracherwerb erst erkannt, wenn die Leistungsunterschiede innerhalb einer Klasse auffallend hoch sind, so besteht die Gefahr einer weit auseinandergelassenen Leistungsschere. Diese Leistungsdifferenzen kön-

nen, je größer sie werden, immer schwieriger aufgeholt werden. Überdies stellt die individuelle Förderung ein Recht jeden Kindes dar, welches im § 1 des Schulgesetzes fest verankert ist (Schulgesetz NRW, 2005). Neben den Leistungsunterschieden innerhalb einer Klasse wurden ebenfalls Geschlechterunterschiede in der Lesekompetenz zum Vorteil der Mädchen festgestellt. Bessere Leseleistungen der Mädchen zeigen sich sowohl am Ende der Primarstufe als auch in der Sekundarstufe I (Niklas und Schneider, 2012).

Um auf Schwierigkeiten im Schriftspracherwerb zeitnah zu reagieren, ist eine regelmäßige und leistungssensible Diagnose notwendig. Auf diese Weise kann frühzeitig noch im Lernprozess auf Stagnation oder Sinken der Lesekompetenz mit individueller und bedürfnisorientierter Förderung reagiert werden. Diagnoseinstrumente im Rahmen des Curriculum-based Measurements (CBM) bzw. der Lernverlaufdiagnose (LVD) bieten ebensolche Möglichkeiten der formativen Diagnostik. Mit Hilfe dessen kann der Lernverlauf dokumentiert und anschließend evaluiert werden. Auf Grundlage dessen können pädagogische Entscheidungen getroffen werden und eine passende Förderung kann stattfinden. Neben Forschungen im amerikanischen Raum, welche insbesondere durch (Deno, 2003a) vorangebracht wurden, gibt es in Deutschland unter anderem die noch junge Online-Plattform Levumi (Lern-Verlaufs-Monitoring), welche auf dem Kieler Leseaufbau (KLA) basiert. Sie wurde entwickelt von Markus Gebhardt, Kirsten Diehl und Andreas Mühling, welche ebendiese formative Diagnostik mit zusätzlichen Fördermaterialien anbietet (Gebhardt et al., 2016a). Durch regelmäßige Erhebungen der Lesekompetenz können Schwierigkeiten im Schriftspracherwerb rechtzeitig erkannt und darauf aufbauend die Unterrichtsqualität verbessert werden. An dieser Stelle bietet die Verzahnung von Diagnose und Förderung die Möglichkeit individuelle Lernprozesse bedürfnisorientiert zu gestalten.

## **1.2 Zielsetzung**

Ziel dieser Forschung ist es, die Leseleistung von leseschwachen Zweitklässlerinnen und Zweitklässlern in den Bereichen Leseflüssigkeit und sinnentnehmendem Lesen über einen Zeitraum von 9 bzw. 21 Wochen unter dem Aspekt der geschlechterspezifischen Unterschiede im Kontext der Durchführung zusätzlicher Fördermaßnahmen zu untersuchen. Dabei liegt der Schwerpunkt sowohl auf dem Leistungszuwachs als auch auf der Leseleistung selbst. Im Rahmen dessen soll geklärt werden, inwiefern Geschlechtsunterschiede vorhanden sind und ob diese mit den Ergebnissen von Forschungen anderer Größenordnungen übereinstimmen.

### 1.3 Vorgehensweise

Vorweg ist zu erwähnen, dass nachfolgend für eine geschlechtersensible Schreibweise das Binnen-I neben geschlechtsneutralen Formulierungen verwendet wird. Auf diese Weise wird der Lesefluss nicht unterbrochen und die Schreibweise beinhaltet beide Geschlechter.

In der vorliegenden Forschungsarbeit wird zunächst ein theoretischer Rahmen zur Beantwortung der Forschungsfragen formuliert. Im Rahmen dessen wird der Schriftspracherwerb mitsamt den zentralen Erwerbsmodellen skizziert (Kapitel 2.1). Darauf aufbauend wird die Lesekompetenz in die Schwerpunkte Leseflüssigkeit und sinnentnehmendes Lesen aufgeteilt und mit den jeweiligen Besonderheiten dargelegt (Kapitel 2.2.1 und 2.2.2). Diese Bereiche stellen die Grundlage für das verwendete Forschungsinstrument Levumi dar, bei welchem die zur Verfügung stehenden Tests auf ebendiesen Bereichen aufbauen. Anschließend werden mögliche Einflussfaktoren auf die Lesekompetenz erläutert (Kapitel 2.2.3), insbesondere der Forschungsstand bezüglich geschlechterspezifischer Unterschiede in der Lesekompetenz (Kapitel 2.3). Da im Erhebungszeitraum eine Förderung im Rahmen von Levumi stattfand, werden die Grundlagen von Leseförderungen aufgezeigt (Kapitel 2.4). Daran anknüpfend werden die Stufen des Kieler Leseaufbaus (Kapitel 2.5), welcher den Tests der Online-Plattform Levumi zugrunde liegt, dargelegt, um im weiteren Verlauf der Arbeit die Schwierigkeitsstufen mit den Niveaustufen der Levumi-Tests vergleichen zu können.

Auf die theoretischen Grundlagen folgt ein Kapitel zur formativen Diagnostik und Leistungsmessung. Zu Beginn werden die Anforderungen an Diagnostik im schulischen Kontext mitsamt den erforderlichen Testgütekriterien sowie Bezugsnormen zur Einordnung von Schülerleistungen dargestellt (Kapitel 3.1). Es folgen die Spezifika von curriculum-based Measurement sowie der Lernverlaufsdagnostik (Kapitel 3.2 und 3.3). Da es sich bei der Lernverlaufsdagnostik um Einzelfallforschung handelt, wird dieser Bereich mit dessen möglichen Abläufen ebenfalls dargelegt (Kapitel 3.4). Daraufhin wird die Online-Plattform Levumi sowie der Levumi-Förderansatz vorgestellt (Kapitel 3.5). Kapitel vier beinhaltet die Relevanz der Fragestellung im Kontext von Schul- und Unterrichtsentwicklung sowie die Darlegung und Begründung der vier Fragestellungen auf Grundlage des zuvor skizzierten Forschungsstandes. Im Rahmen des fünften Kapitels zu den Forschungsmethoden wird die Stichprobe hinsichtlich der übergeordneten Studie vorgestellt (Kapitel 5.1). Es folgt die Beschreibung der Testdurchführung (Kapitel 5.2) und der im Zeitraum der Erhebung durchgeführten Förderung auf Grundlage des Levumi-Förderansatzes (Kapitel 5.3). Abschließend werden die Auswertungsmethoden im Kapitel 5.4 erläutert worauf die Auswertung der Daten folgt (Kapitel 6). Die Ergebnisse werden im Kontext des Forschungsstandes diskutiert und eingeordnet (Kapitel 7),

ehe die vorliegende Forschungsarbeit mit einer Zusammenfassung und einem Ausblick abschließt (Kapitel 8).

## **2 Theorierahmen**

Dieses Kapitel dient der Darlegung relevanter theoretischer Grundlagen, auf welchen die vorliegende Arbeit und die damit verbundene Forschung basiert. Nachfolgend werden zuerst drei verbreitete Modelle zum Schriftspracherwerb vorgestellt. Anschließend folgt eine Unterteilung der Lesekompetenz sowie mögliche Einflussfaktoren auf ebendiese. Insbesondere der Forschungsstand bezüglich geschlechterspezifischer Unterschiede im Lesen wird auf Grund der besonderen Relevanz des Themas umfassend skizziert. Dabei wird sowohl Bezug auf die Ergebnisse (inter-)nationaler Studien als auch auf die bisherigen Ergebnisse der übergeordneten Studie genommen. Diese Erkenntnisse beziehen sich auf Daten, die in der ersten Hälfte der Längsschnittuntersuchung mit Levumi im Rahmen einer anderen Masterarbeit erhoben und ausgewertet wurden. Darauf folgt ein Unterkapitel zur Leseförderung, da im Erhebungszeitraum eine spezifische Förderung stattfand, um somit eine Einordnung dieser in den Gesamtkontext von Leseförderung zu gewährleisten. Das Kapitel zwei schließt mit der Darstellung der Spezifika des KLA ab, welcher den Tests des verwendeten Instruments Levumi zur Lernverlaufdiagnose zu Grunde liegt. Aus den in diesem Kapitel dargestellten Informationen sowie aus den Ergebnissen bereits durchgeführter Forschungen, ergeben sich die in Kapitel vier beschriebenen Forschungsfragen.

### **2.1 Modelle zum Schriftspracherwerb**

Der Schriftspracherwerb ist ein viele Jahre andauernder Prozess, welcher sowohl den Erwerb von Lese- als auch von Schreibkompetenz umfasst. Es handelt sich um verschiedene Phasen, die jedes Kind mit einer individuellen Geschwindigkeit durchläuft. Darüber hinaus ist der Schriftspracherwerb als wechselseitiger Prozess zwischen Lesen und Schreiben zu verstehen (Dürscheid, 2012), wobei davon ausgegangen wird, dass sich eine qualitative Veränderung der verwendeten Lesestrategien vollzieht (Schneider, 2017). Grundsätzlich gilt, dass der Erwerb fundamentaler Lesestrategien im Rahmen des Schriftspracherwerbs ca. gegen Ende der zweiten Klasse abgeschlossen ist. Darauf folgt der Erwerb weiterführender Lesekompetenzen (Voss, 2006). Zu den verbreiteten Modellen der Entwicklung des Schriftspracherwerbs zählt das Stufenmodell von Frith (1985), welches aus dem englischsprachigen Raum stammt und drei Phasen beinhaltet sowie das Fünf-Phasen-Modell von Günther (1986) aus dem deutschsprachigen Raum. Beide Modelle gehen von einem Zusammenspiel zwischen Lesen und Schreiben aus (Dürscheid, 2012). Zuletzt wird das Kompetenzentwicklungsmodell des

Lesens nach Klicpera, Schabmann, Gasteiger-Klicpera und Schmidt (2017) dargestellt, welches insbesondere die Forschungsbefunde aus dem deutschen Sprachraum berücksichtigt. Nachfolgend werden die genannten Modelle in ihrem Aufbau und den jeweiligen Annahmen der AutorInnen vorgestellt.

Das Drei-Phasen-Modell von (Frith, 1985) umfasst die logographische, die alphabetische und die orthographische Phase. Alle drei Phasen beziehen sich sowohl auf die Lese- als auch auf die Schreibkompetenz. Dabei verläuft der Erwerb jedoch nicht unabhängig voneinander, sondern vielmehr mit wechselseitigem Einfluss aufeinander (Dürscheid, 2012). Die logographische Phase ist dadurch gekennzeichnet, dass sich die Kinder an optisch charakteristischen Merkmalen eines Wortes orientieren und dabei dessen Bedeutung, jedoch noch nicht die mit dem Wort verbundene Lautfolge kennen (Dürscheid, 2012). Eben diese Art des Lesens wird auch *naiv-ganzheitliche Worterfassung* genannt. Auf der Ebene des Schreibens können Kinder, welche sich in beschriebener Phase befinden, ihnen bekannte Wörter insofern *schreiben*, als dass sie diese als Ganzes wahrnehmen und in gewisser Weise als Bild reproduzieren (Schründer-Lenzen, 2013). Laut (Frith, 1985) beginnt die logographische Phase mit dem Lesen. Anschließend bzw. überlappend dazu fangen Kinder an auf dieser Ebene zu schreiben. Neben der zeichnerischen Darstellung kann es zur Schreibung der prägnanten Konsonanten eines Wortes kommen, wodurch der Beginn der alphabetischen Phase auf der Ebene des Schreibens gekennzeichnet ist. In der alphabetischen Phase erlesen Kinder Wörter Laut für Laut. Dabei werden die Grapheme in Phoneme rekodiert. Bei diesem Vorgang handelt es sich um das phonologische Rekodieren (Dürscheid, 2012). Nachdem das geschriebene Wort im Sinne der Graphem-Phonem-Korrespondenz in die Lautsprache übersetzt wird, folgt das Dekodieren, welches die anschließende Sinnentnahme nach dem Rekodieren bezeichnet (Schneider, 2017). Die Verschriftlichung von Wörtern ist in dieser Phase oftmals von Fehlern geprägt, da die Kinder phonographisch, also rein lautorientiert schreiben. In der orthographischen, der dritten Phase nach Frith (1985), werden bekannte Wörter nicht mehr phonologisch rekodiert, sondern ganzheitlich gelesen. Dabei wird sich an den Buchstabenfolgen und -kombinationen orientiert. Morpheme oder Buchstabengruppen werden als Ganzes wahrgenommen und können somit schneller erkannt werden. Im Anschluss an die Veränderung der Lesestrategie in der orthographischen Phase folgt die Veränderung auf der Ebene des Schreibens, wobei phonographische Schreibweisen mehr und mehr zu Gunsten orthographischer Prinzipien weichen. Darüber hinaus werden auch morphologische Regularitäten zunehmend berücksichtigt. Mit der orthographischen Phase ist der Schriftspracherwerb grundsätzlich abgeschlossen, wobei die alphabetische Strategie stets bei unbekanntem Wörtern zum Tragen

kommt (Dürscheid, 2012). Da es sich um ein Modell aus dem angelsächsischen Raum handelt, wird oftmals angemerkt, dass sowohl das Schriftsystem als auch die institutionelle Steuerung des Schriftspracherwerbs Einfluss auf ebensolchen hat und nicht in gleichem Maße auf den Schriftspracherwerb der deutschen Sprache übertragen werden kann (Dürscheid, 2012).

(Günther, 1986) entwickelte für den deutschsprachigen Raum ein Fünf-Phasen-Modell, welches auf dem dargestellten Modell von Frith (1985) beruht, dieses jedoch um weitere Phasen ergänzt. Um inhaltliche Redundanz durch Wiederholung zu vermeiden, werden nachfolgend lediglich die Neuerungen nach Günther (1986) ergänzt. Die sogenannte präliterale-symbolische Phase stellt er der logographischen Phase voran. Diese Phase wird vollzogen, wenn Kinder bildliche Darstellungen auf zweidimensionaler Ebene auf dreidimensionale Körper beziehen und dies darüber hinaus auch auf zeichnerischer Ebene umsetzen (Dürscheid, 2012). Verwenden Kinder dabei bereits Symbole, denen eine arbiträre kulturelle Bedeutung zugeordnet ist, so wird die präliterale-symbolische Strategie angewendet. In dieser Vorstufe findet sowohl Rezeption als auch Produktion in Form von intensiver Nachahmung von Lese- und Schreibprozessen Erwachsener statt. Günther (1986) beschreibt diese Phase als eine Vorbereitung auf den Schriftspracherwerb, welcher jedoch erst mit der logographischen Phase beginnt. Auch die orthographische Phase ergänzt er um eine weitere, die integrativ-automatisierte Phase. Sie beschreibt den Zustand der automatisierten und sicheren Anwendung verschiedener Lese- und Schreibstrategien (Sassenroth, 2000). Streng genommen handelt es sich um keine zusätzliche Strategie, die integrativ-automatisierte Phase beschreibt vielmehr den schriftlichen Sprachgebrauch kompetenter Leser und Schreiber (Dürscheid, 2012). Es wird deutlich, dass die hinzugefügten Phasen von Günther (1986) weniger den Schriftspracherwerbsprozess selbst beschreiben, sondern ihn gewissermaßen umrahmen.

Das Kompetenzentwicklungsmodell des Lesens nach Klicpera et al. (2017) beinhaltet keine direkte Abfolge von Entwicklungsphasen, wie es in den zuvor dargestellten Modellen der Fall ist, sondern orientiert sich an den wesentlichen Lesekompetenzen, die zu erwerben sind (Klicpera et al., 2017). In diesem Modell wird davon ausgegangen, dass sich das nicht-lexikalische und das lexikalische Lesen weitgehend parallel entwickeln. Bei nicht-lexikalischen Leseprozessen wird das Wort phonologisch rekodiert und bei lexikalischen Leseprozessen kann auf das mentale Lexikon zurückgegriffen werden, in welchem Wörter und die damit verbundenen visuellen, phonologischen und semantischen Eigenschaften abgespeichert sind (Klicpera et al., 2017). Der lexikalische Zugriff bietet zwar einen Geschwindigkeitsvorteil, bei unbekanntem Wörtern wird jedoch nach wie vor das phonologische Rekodieren im Sinne des

nicht-lexikalischen Zugriffs benötigt. Zur Speicherung im mentalen Lexikon ist eine gewisse Sicherheit im Rekodieren notwendig, jedoch baut das lexikalische Lesen laut Klicpera et al. (2017) nicht auf einer alphabetischen Phase auf, wie es im Modell von Frith (1985) dargestellt wird. In diesem Prozess wächst die Größe der verarbeiteten Einheiten kontinuierlich an. Zu Beginn, im Bereich des nicht-lexikalischen Lesens, werden erst kleine Einheiten wie Buchstaben bzw. Grapheme rekodiert, wobei leichte Kombinationen schnell ins Lexikon überführt werden. Darauf folgen größere, sublexikalische Einheiten wie Silben und Morpheme. Es kommt somit zu einer sukzessiven Ausweitung der Lesekompetenz, begonnen mit dem phonologischen Rekodieren bis hin zum primär lexikalischen Lesen ganzer Wörter (Klicpera et al., 2017).

Die Einordnung der Schülerleistungen bezüglich des Stufenmodells ist für Lehrerinnen und Lehrer insofern von Bedeutung, als dass sie Rückschlüsse auf den Entwicklungsstand und die individuell verwendeten Strategien gibt. Auf diese Weise können angemessene Fördermaßnahmen eingeleitet werden, um das Erreichen qualitativ höherer Phasen anzubahnen (Sassenroth, 2000). Von besonderer diagnostischer Relevanz ist hierbei die Analyse von Fehlern, insbesondere bei Spontanschreibungen, welche Aufschluss über die derzeitigen verwendeten Strategien geben (Sassenroth, 2000).

## **2.2 Lesekompetenz**

In der Betrachtung der Kompetenzmodelle, die den Erwerb von Lesekompetenz mit unterschiedlicher Schwerpunktsetzung erklären, wird deutlich, dass zahlreiche Elemente für eben diesen Erwerb von Bedeutung sind. Das Lesen ist eine multiple Tätigkeit, die insbesondere für Leseanfänger einige Herausforderungen mit sich bringt. Auf der einen Seite umfasst Lesekompetenz die Fähigkeit auf Buchstaben- und Wortebene zu rekodieren und zu dekodieren (Bertschi-Kaufmann, 2011), auf der anderen Seite das Leseverstehen, womit die Konstruktion von Sinn in einer größeren Einheit wie einem Satz oder einem Text gemeint ist. Es kann unterschieden werden zwischen hierarchieniedrigen und hierarchiehöheren Leseprozessen, wobei es sich nicht um eine Wertung handelt, sondern vielmehr um die Komplexität und den mentalen Aufwand beim Leseprozess (Philipp, 2017a). Bei den hierarchieniedrigen Prozessen steht das visuelle Erkennen geschriebener Wörter im Fokus, welches, sofern bereits eine Automatisierung stattfindet, für den oder die LeserIn mit einem verhältnismäßig geringen Aufwand verbunden ist. Bei den hierarchiehöheren Prozessen ist hingegen das Verstehen von Sinnzusammenhängen kontinuierlicher schriftsprachlicher Texte zentral, welches mit einer höheren Denkleistung verbunden ist (Hurrelmann, 2011). Es wird deutlich, dass die Fähigkeit

auf hierarchieniedriger Ebene zu lesen, grundlegend für komplexe Leseleistungen hinsichtlich des Textverständnisses ist (Philipp, 2017a). In verschiedenen Kompetenzmodellen und Forschungen zur Lesekompetenz ist oftmals das Lesen von Texten und die Verständnisleistung im Fokus des Interesses. Aufgrund der Erhebung der Leseflüssigkeit und des sinnentnehmenden Lesens auf Wort-Satzebene in der vorliegenden Arbeit, wurde für dieses Kapitel ebendieser Schwerpunkt gewählt. Leseflüssigkeit und das sinnentnehmende Lesen sowie die zugrunde liegenden Prozesse werden nachfolgend dargelegt.

### 2.2.1 Leseflüssigkeit

Der Begriff der Leseflüssigkeit beinhaltet das genaue, schnelle, automatisierte (Müller et al., 2013), prosodisch angemessene sowie das relativ mühelose Lesen von Wörtern und Sätzen (Walter, 2014). Die Leseflüssigkeit wird durch die verwendeten Strategien bedingt und zählt zu den hierarchieniedrigen Leseprozessen, wozu ebenso die Buchstaben- und Worterkennung sowie die syntaktische und semantische Analyse gehört (Lindauer und Schneider, 2011). Zwei verschiedene Verarbeitungsprozesse, *bottom-up* und *top-down*, können herangezogen werden, um die Leseflüssigkeit näher zu beschreiben. Bei diesem Erklärungsansatz handelt es sich um ein Zwei-Wege-Modelle. *Bottom-up*-Prozesse beschreiben den Leseprozess auf niedriger Stufe, bei der Wörter durch die indirekt-phonologische Strategie erlesen werden (Schründer-Lenzen, 2013). Buchstaben bzw. Grapheme werden im Sinne der Graphem-Phonem-Korrespondenz rekodiert und innerhalb eines Wortes miteinander verbunden bis es zur indirekten Worterkennung durch die phonologische Repräsentation des Wortes im mentalen Lexikon kommt (vgl. Kapitel 2.1). Den erlesenen Graphemen werden demzufolge Phoneme, die lautliche Repräsentation der Grapheme, zugeordnet. Es ist jedoch zu betonen, dass durch das reine Rekodieren auch ohne das Erkennen des Wortes durch das mentale Lexikon, phonologisch richtig gelesen werden kann (Richter und Müller, 2017). Bereits Kinder, die primär diese Strategie anwenden, können mehr oder weniger stark auf Vorinformationen zurückgreifen, wodurch der oder die LeserIn Vermutungen bezüglich der nachfolgenden Phoneme anstellen kann. Dies kann die weitere Wahrnehmung beeinflussen und wird bereits als *Top-Down*-Prozess verstanden. Grundsätzlich handelt es sich bei dem indirekten Weg der Worterkennung um eine eher fehleranfällige Strategie (Richter und Müller, 2017).

*Top-Down*-Prozesse stehen für die Nutzung des lexikalischen Wortspeichers. Im mentalen Lexikon sind alle relevanten Informationen des jeweiligen Wortes bzw. des Graphems, der Silbe oder beispielsweise einer Endung im Langzeitgedächtnis gespeichert und die Elemente können schneller erkannt werden. Bei häufig gelesenen Wörtern wird das mentale Lexikon besonders schnell aktiviert. Neben der orthographischen Repräsentation von Wörtern ist

auch die phonologische Repräsentation vorhanden, die es ermöglicht, das bekannte Wort korrekt auszusprechen (Richter und Müller, 2017). Es handelt sich um einen direkt-lexikalischen Zugriff auf den Wortspeicher, welcher in gewisse Teilbereiche untergliedert werden kann. Das Dual-Route Cascaded Model (DRC-Modell) zählt zu den Zwei-Wege-Modellen und beschreibt verschiedene Elemente des mentalen Lexikons. (Coltheart, 2005) benennt das semantische, orthographische und das phonologische Lexikon, welche die jeweiligen Informationen beinhalten.

In früheren Veröffentlichungen wurden die genannten Zugangswege als unabhängig voneinander betrachtet, was heute jedoch kritisch diskutiert wird. Darüber hinaus ist zu betonen, dass auch geübte LeserInnen den indirekten Weg des phonologischen Rekodierens bei unbekanntem Wörtern verwenden. Ausnahmen stellen sogenannte *exception words* dar, wobei es sich um Wörter handelt, die trotz der phonologischen Rekodierung nicht korrekt gelesen werden können. Zentral für das Rekodieren ist eine gewisse Regelmäßigkeit der gelesenen Sprache hinsichtlich der Graphem-Phonem-Korrespondenz. Bei der deutschen Sprache handelt es sich um eine eher regelmäßige Sprache im Vergleich zum Englischen (Klicpera et al., 2017). Während Leseanfänger mit dem Rekodieren einzelner Buchstaben und Grapheme beginnen, schreitet der Kompetenzerwerb mit dem Lesen größerer sublexikalischer Einheiten fort, bis es zum Erlesen ganzer Wörter durch den direkten Zugang kommt. Zum Ende der Grundschulzeit beherrscht die Mehrzahl der Kinder die beschriebenen hierarchieniedrigen Leseprozesse (Richter und Müller, 2017). Demzufolge hängt die Leseflüssigkeit stark von den verwendeten Strategien der Kinder ab. Sind die zu lesenden Wörter im mentalen Lexikon gespeichert, so kann ein automatisiertes Lesen erfolgen. Daraus ergibt sich unweigerlich eine höhere Lesegeschwindigkeit sowie eine höhere Genauigkeit als bei der Verwendung der indirekt-phonologischen Strategie. Im schulischen Kontext ist der Schereneffekt zwischen schwachen und starken LeserInnen zu beachten. Anfänglich gute Leseleistungen gehen mit einer erhöhten Lesemotivation einher, welche sich wiederum positiv auf die Festigung der Leseprozesse auswirkt. Umgekehrt führen eher schwächere Leistungen zu einer verminderten Lesemotivation, woraus sich ggf. ein Verbleiben auf einem niedrigen Kompetenzniveau ergibt (Richter und Müller, 2017).

### 2.2.2 Sinnentnehmendes Lesen

Grundsätzlich kann man die Begriffe sinnentnehmendes Lesen und Leseverständnis synonym verwenden. Es handelt sich um hierarchiehöhere Prozesse, die der Entnahme von Informationen auf Satz- und Textebene dienen (Philipp, 2017a). Bei den zu entnehmenden Informationen handelt es sich nicht immer nur um solche, die genau das beinhalten, was sie rein faktisch bedeuten, sondern ggf. auch um schwer zu erkennende Informationen, die im Kontext einer gewissen Deutung bedürfen. Das sinnentnehmende Lesen ist als eigener Teilbereich der Lesekompetenz zu betrachten, da die Gruppe leseschwacher Kinder sowohl solche beinhaltet, die Schwierigkeiten im Rekodieren haben, als auch solche, die in dieser Hinsicht hohe Leistungen zeigen. Demzufolge ist die Leseflüssigkeit nur eine Einflussgröße auf das Leseverständnis, welches als mehrdimensionales Konstrukt zu verstehen ist (Klicpera et al., 2017).

Der Erwerb von Lesekompetenz zielt auf den eigenständigen Umgang mit Texten hinsichtlich des Entnehmens von relevanten Informationen ab. Grundlegend dafür sind die zuvor erläuterten Fähigkeiten hinsichtlich des Rekodierens und Dekodierens von Wörtern. Liegt der Großteil der Aufmerksamkeit beim Lesen auf dem korrekten Rekodieren, so kann das Erfassen größerer Sinnzusammenhänge erschwert werden, da es zu einer höheren Belastung des Arbeitsgedächtnisses kommt (Heinzl et al., 2015). Nicht nur eine mangelnde Dekodierfähigkeit, sondern auch ein geringer Wortschatz kann das Verständnis maßgeblich beeinträchtigen, sofern eine Kompensation durch den Kontext nicht ausreichend ist. Ist der Leseprozess automatisiert, so kann der Inhalt von Sätzen und Texten kontinuierlich und nahezu simultan erschlossen werden (Klicpera et al., 2017). Nicht nur unbekannte Inhaltswörter, insbesondere Funktionswörter wie Präpositionen, Partikel oder Adverbien können das Verständnis eines ganzen Satzes maßgeblich beeinflussen (Heinzl et al., 2015).

Da der Fokus der vorliegenden Arbeit beim sinnentnehmenden Lesen auf Wort-Satzebene liegt, werden komplexere Zusammenhänge nachfolgend nur in Kürze dargestellt. Neben den basalen Lesefähigkeiten, die für jeden Satz oder Text erforderlich sind, kann der Text selbst gewisse Anforderungen an die Fähigkeiten des oder der LeserIn stellen. Neben dem Vorwissen, dem semantischen Wissen und kontextgebundener Wortbedeutung ist das Verständnis von Textstrukturen und Diskursformen auf der lokalen Anforderungsebene eines Textes von Bedeutung (Klicpera et al., 2010). Die Komplexität, hinsichtlich der Fähigkeitsanforderungen des Textes, ist von dessen objektiver Schwierigkeit und den individuellen Voraussetzungen des Lesers oder der Leserin abhängig. Bei längeren Texten ist darüber hinaus eine Bearbeitung auf der Makroebene erforderlich (Klicpera et al., 2010). Im Rahmen der internationalen

Grundschul-Lese-Untersuchung 2006 (IGLU) werden verschiedene Aspekte für die Verstehensleistung formuliert. Dabei bildet das Erkennen und das Wiedergeben von Textinformationen die Grundlage für komplexere Leistungen. Daran schließt das Ziehen einfacher Schlussfolgerungen an, mit späterer Erweiterung der Fähigkeit zum Ziehen und Begründen komplexer Schlussfolgerungen sowie das Interpretieren. Die höchste Verstehensleistung ist das „Prüfen und Bewerten von Inhalt und Sprache“ (Bos et al., 2008). Der Lesekompetenzbegriff der Untersuchung von Programme for international Student Assessment (PISA) beschränkt sich hingegen auf die drei Dimensionen: Informationen ermitteln, textbezogenes Interpretieren sowie Reflektieren und Bewerten (Hurrelmann, 2002).

### **2.2.3 Einflussfaktoren**

Nachfolgend werden verschiedene Faktoren dargestellt, die den individuellen Schriftspracherwerb bzw. insbesondere den Erwerb der Lesekompetenz beeinflussen können. Zu den individuellen Faktoren eines jeden Kindes zählt unter anderem die kognitive Leistungsfähigkeit, welche die schulischen Leistungen im Allgemeinen prägt. Die Leistungsfähigkeit im Sinne unterdurchschnittlicher Kognition kann verschiedenen Schweregraden zugeordnet werden. Dabei handelt es sich um Lernbehinderungen und Lernstörungen, welche Lernschwierigkeiten im Schriftspracherwerb erklären können (Zöllner und Roos, 2009). Obwohl eine hohe allgemeine Intelligenz nicht zwingend Schulerfolg mit sich bringt, sind die Lese- und Rechtschreibleistungen bedeutend von der kognitiven Leistungsfähigkeit beeinflusst (Zöllner und Roos, 2009). Sowohl IGLU als auch PISA ermittelten moderate bis mittelhohe Korrelationen zwischen der Intelligenz und der Deutsch- bzw. der Rechtschreibnote (Gebhardt, Sälzer, Mang, Müller, & Prenzel, 2015). Inwiefern der Schriftspracherwerb und die Lesekompetenz im Einzelnen von der Intelligenz beeinflusst werden, ist nicht ausreichend bekannt (Schneider, 2017). Ebenso wird in diesem Kontext der phonologischen Bewusstheit, dem Arbeitsgedächtnis und der sprachbezogenen Informationsverarbeitung Bedeutung beigemessen (Schneider, 2017). Darüber hinaus sind weitere körperliche Faktoren für den Schriftspracherwerb von Bedeutung, denn die optischen Reize geschriebener Sprache müssen für das Lesen sowohl visuell richtig wahrgenommen werden als auch korrekt im Gehirn verarbeitet werden. Ebenso die auditive Verarbeitung prägt den Erwerbsprozess (Fischer, 2015).

Das Geschlecht als Einflussfaktor stellt ein den Schriftspracherwerb beeinflussendes individuelles Merkmal dar. In verschiedenen Studien wurde herausgefunden, dass es einen Geschlechterunterschied im Schriftspracherwerb gibt. Dabei sind die Mädchen den Jungen sowohl im Lesen als auch im Rechtschreiben überlegen (Schneider, 2017). Da der Einfluss des

Geschlechts auf die Leseleistung in der vorliegenden Arbeit von besonderer Bedeutung ist, wird der aktuelle Forschungsstand diesbezüglich in Kapitel 2.3 ausführlich dargelegt.

Der Motivationspsychologie sind unterschiedliche Konstrukte von Motivation mit verschiedenen Schwerpunkten zuzuordnen. Motivation, als weiterer individueller Einflussfaktor, beschreibt den inneren Antrieb etwas zu tun, unabhängig ob dieser Antrieb von inneren oder externen Faktoren beeinflusst wird. Im Kontext der Lesemotivation handelt es sich um die Tatsache, dass mit der Aktivität des Lesens oder den damit verbundenen Folgen gewisse Anreize verbunden sind, welche dazu führen etwas zu lesen (Philipp, 2017b). Motivation ist ein komplexes Konstrukt und kann aus unterschiedlichen Zielsetzungen resultieren. Es wird unterschieden zwischen intrinsischen und extrinsischen Anreizen für die jeweilige Aktivität. Intrinsische Motivation beinhaltet die Ausübung einer „Aktivität wegen der ihr inhärenten Befriedigung [...]“, wie beispielsweise Vergnügen oder Neugier (Philipp, 2013b). Bei extrinsischer Motivation handelt es sich hingegen um die Ausübung einer Aktivität aus Gründen, welche außerhalb der Aktivität selbst liegen, wie beispielsweise Anerkennung, gute Noten oder Wettbewerb. Die Aktivität dient demzufolge dem Erreichen eines Ziels, welches aus dieser selbst resultiert und wird als instrumentelle Aktivität verstanden. Die intrinsische Lesemotivation gilt als zielführender für die Entwicklung der Lesekompetenz verglichen mit der extrinsischen Motivation (Förster und Souvignier, 2014). Neben dem Anreiz selbst ist auch die Wahrnehmung der eigenen Fähigkeiten von Bedeutung, welche das Maß der Motivation beeinflussen kann (Philipp, 2013a). Besonders im Kontext der Lesekompetenz können als unzureichend bewertete Fähigkeiten, den Antrieb die Aktivität auszuüben negativ beeinflussen. Kommt es zu Beginn des Erwerbsprozesses zu Schwierigkeiten, so sinkt die Lern- bzw. Lesemotivation, was zu einem verringerten Selbstbewusstsein führen kann. Kommt es aufgrund der geringen Lesemotivation dazu, dass der oder die SchülerIn weniger liest, so mangelt es an Übung und die Leistungsschere zwischen starken und schwachen LeserInnen vergrößert sich (Dummer-Smoch und Hackethal, 2016). Passend dazu wurde im Rahmen der IGLU 2016 herausgefunden, dass insgesamt jedes sechste Kind einer vierten Klasse eine niedrige Lesemotivation aufweist. Die Lesemotivation stellt einen zentralen Einflussfaktor auf das Leseverhalten dar (Bos et al., 2017).

Nachdem zuvor die individuellen Einflussfaktoren dargestellt wurden, folgen nun die äußeren, sozialen Faktoren, die sich auf die Lesekompetenz auswirken können. Beispielsweise in den Studien von Zöllner und Roos (2009) sowie von Dummert et al. (2014) konnte bereits in der ersten Klassenstufe ein Einfluss des sozioökonomischen Status auf die Lesekompetenz verzeichnet werden (Schneider, 2017). Ein hoher sozioökonomischer Status geht mit einer eher

positiven Leistungsentwicklung und ein niedriger sozioökonomischer Status mit einer geringeren Leistungsentwicklung einher. Ursächlich für den Einfluss des sozioökonomischen Status und der elterlichen Stellung in der Gesellschaft ist der damit verbundene Zugang zu Einkommen und Bildung (Zöllner und Roos, 2009). Nicht nur die ungleichen Lebensbedingungen gehen mit Unterschieden des Lernmilieus einher, auch die verfügbaren Ressourcen können die Gestaltung des Entwicklungsumfeldes des Kindes hinsichtlich einer lernförderlichen Atmosphäre prägen. Im Allgemeinen haben somit der Bildungshintergrund der Eltern und die oftmals damit verbundenen finanziellen Möglichkeiten einen Einfluss auf außerschulische Bildungsangebote, die Vielfältigkeit von Aktivitäten und kulturellen Erfahrungen, welche wiederum die Entwicklung von Kindern beeinflussen. Auf den Erwerb der Lesekompetenz bezogen sind die Erweiterung des Wortschatzes, der Zugang zu Medien und das außerschulische Lesen von besonderer Bedeutung. Das Vorhandensein dieser Aspekte im Alltag von Kindern kann von einem niedrigen sozioökonomischen Status der Familie eingeschränkt werden (Zöllner und Roos, 2009).

Für den schulischen Erfolg ist eine ausreichende Sprachkompetenz in der Unterrichtssprache elementar. Demzufolge stellt auch die Zuwanderungsgeschichte der Familie insofern ein Einflussmerkmal dar, als dass diese oftmals mit einer anderen Muttersprache einhergeht (Zöllner und Roos, 2009). Daraus ergibt sich der Erwerb der deutschen Sprache als Zweitsprache. Dieser externe Bedingungsfaktor beinhaltet das Potential den Schriftspracherwerb zu erschweren. Sowohl Qualität und Quantität des sprachlichen Inputs prägen den vorschulischen Spracherwerb. Ist dieser Input nicht ausreichend im Sinne eines umfangreichen und abwechslungsreichen Sprachangebots, so können weder die grundlegenden Prinzipien der deutschen Sprache noch ein ausreichender Grundwortschatz entwickelt werden. Darüber hinaus ist das Merkmal Deutsch als Zweitsprache oftmals mit einem niedrigen sozioökonomischen Status verbunden, dessen mögliche Auswirkungen bereits dargestellt wurden. Besonders eine Kombination dieser beiden Faktoren können den Schriftspracherwerb und demzufolge auch die Entwicklung der Lesekompetenz maßgeblich beeinflussen (Marx, 2017).

### **2.3 Forschungsstand geschlechterspezifischer Unterschiede im Lesen**

Das Geschlecht gilt als bedeutender Einflussfaktor auf die Schulleistungen im Allgemeinen und auch auf spezifische Teilbereiche wie die Lesekompetenz (Schneider, 2017). Nachfolgend wird der Forschungsstand hinsichtlich geschlechterspezifischer Unterschiede in der Lesekompetenz und dem Leseinteresse bzw. der Häufigkeit des Lesens im Alltag dargelegt. Anschließend werden verschiedene Erklärungsansätze für die tendenziell besseren Leseleistungen der Mädchen erläutert. Im zweiten Teil dieses Kapitels wird Bezug zu den Ergebnis-

sen einer Arbeit genommen, welche im Rahmen desselben Forschungsprojekts herausgearbeitet wurden. Die Darstellung dieser Ergebnisse von Wenner (2018) ist insofern von besonderem Interesse, als dass es sich um dieselben SchülerInnen handelt, die auch in dieser Arbeit mit der formativen Lernverlaufdiagnose-Plattform Levumi getestet wurden. Die Daten der vorliegenden Arbeit knüpfen an diese an, wodurch die Betrachtung der Entwicklung geschlechterspezifischer Unterschiede im gesamten Lernverlauf ein besonderes Forschungsinteresse darstellt.

### **2.3.1 Geschlechterspezifische Unterschiede der Lesesozialisation und Leseleistung**

Geschlechterunterschiede in den schulischen Leistungen äußern sich in unterschiedlichem Ausmaß in verschiedenen Teilbereichen. Da im Rahmen dieses Forschungsprojektes die Lesekompetenz von GrundschülerInnen im Fokus steht, werden hier die Ergebnisse der IGLU Studie, welche sich auf den Primarbereich spezialisiert, beschrieben. Es ist jedoch zu betonen, dass die im Rahmen von IGLU durgeführten Tests hierarchiehöhere Prozesse hinsichtlich der Verständnisleistung beim Lesen von Texten in den Vordergrund stellen und nicht weiter auf hierarchieniedrige Leseprozesse eingehen (Bremerich-Vos et al., 2012), was womöglich an der untersuchten Jahrgangsstufe vier und einer damit zu erwartenden höheren Lesekompetenz liegt. Nachfolgend werden unterschiedliche Ergebnisse dieser Untersuchungen aus den Jahren 2011 und 2016 dargelegt.

Die Ergebnisse der IGLU 2011 ergeben, dass in allen teilnehmenden Staaten, in denen ein Geschlechterunterschied signifikant ist, stets die Mädchen das leistungsstärkere Geschlecht sind. In Deutschland liegen die Mädchen mit acht Punkten (insg. 545P) auf der Gesamtskala Lesen vor den Jungen (insg. 537P) (Bos et al., 2012). Die Ergebnisse der IGLU 2016 zeigen ebenfalls signifikante Unterschiede zu Gunsten der Mädchen. Diese erreichen auf der Gesamtskala Lesen einen Vorsprung von 11 Punkten gegenüber den Jungen. Darüber hinaus zeigt sich ein Vorsprung der Mädchen bezüglich ihres Leseselbstkonzepts, ihrer Lesemotivation und dem Leseverhalten (Bos et al., 2017).

Auch die Verortung in die fünf von IGLU benannten Lesekompetenzstufen zeigt gewisse Unterschiede hinsichtlich des Geschlechts. Kompetenzstufe eins ist die niedrigste Stufe und beinhaltet ein rudimentäres Leseverständnis, Kompetenzstufe fünf als höchstmögliche Stufe umfasst selbstständiges Begründen und Interpretieren auf Grundlage des gesamten Textes bzw. Textpassagen (Bremerich-Vos et al., 2012). Auf der ersten Kompetenzstufe befinden sich  $\approx 2\%$  der Mädchen und  $\approx 3\%$  der Jungen. 11.9% der Mädchen und 14.2% der Jungen gehören Stufe zwei an. Der Kompetenzstufe drei gehören 37.3% der Mädchen und 39.2% der Jungen an. Es zeigt sich, dass in den unteren drei Kompetenzstufen insgesamt mehr

Jungen mit  $\approx 56\%$  als Mädchen mit  $\approx 52\%$  vertreten sind. Der Stufe vier sind 38% der Mädchen und 35% der Jungen zuzuordnen. Der höchsten Stufe (fünf) gehören 10.4% der Mädchen und 8.8% der Jungen an. Daraus ergibt sich, dass mit 48.4% mehr Mädchen als Jungen mit 43.8% in den beiden höchsten Kompetenzstufen vertreten sind (Bos et al., 2012). Hinsichtlich des außerschulischen Leseverhaltens geben 16% der Jungen an nie oder fast nie zum Vergnügen zu lesen, wohingegen es bei den Mädchen nur 6% sind (Bos et al., 2012). Eine Verteilung der nach Geschlecht getrennten Zugehörigkeit zu den Kompetenzstufen wurde im Rahmen von IGLU 2016 nicht in dieser Form dargestellt. Dennoch zeigt ein Anteil von 18.9% GrundschülerInnen der Klasse vier, die den Kompetenzstufen eins und zwei angehören, dass ein Großteil leseschwacher SchülerInnen vorhanden ist (Bremerich-Vos et al., 2017). Darüber hinaus ist zu betonen, dass auf Grundlage der Klassifikationen durch den ICD-10 zwei bis dreimal so viele Jungen von Sprachentwicklungsstörungen betroffen sind wie Mädchen (Rißling et al., 2014).

Die KIM Studie 2016 (Kindheit, Internet und Medien) erfasst den Medienkonsum von Kindern im Alter zwischen 6 und 13 Jahren und erhebt diesen sowohl durch die Kinder selbst als auch mit Hilfe der Eltern. Neben verschiedenen digitalen Medien werden auch die Nutzung und Präferenzen von Büchern der Kinder erhoben. 65% der Mädchen geben an, gerne oder sehr gerne zu lesen, wohingegen nur 38% der Jungen diese beiden Antwortmöglichkeiten wählen. Nicht so gerne lesen 22% der Mädchen und 36% der Jungen. Die beiden Antwortmöglichkeiten *lese gar nicht gerne* und *lese nie* beantworten fast doppelt so viele Jungen (25%) wie Mädchen (13%). Es lässt sich deutlich erkennen, dass Mädchen grundsätzlich lieber lesen. Diese Tendenz spiegelt sich auch bei der Nutzungsfrequenz von Büchern wieder. Auch hier geben 59% der Mädchen an, jeden/fast jeden Tag oder ein/mehrmals in der Woche zu lesen, wohingegen dies nur 39% der Jungen angeben. 41% der Mädchen geben an selten oder nie zu lesen, bei den Jungen sind es 62% (Feierabend et al., 2017). Das Interesse am und die Regelmäßigkeit des Lesens ist wegen der außerschulischen Übung, welche die Lesekompetenz positiv beeinflusst, von Bedeutung.

Neben diesen, bereits als signifikant geltenden geschlechterspezifischen Unterschieden in der Lesekompetenz zu Gunsten der Mädchen, sei erwähnt, dass dieser Unterschied bei den 15-Jährigen der PISA Studie 2016 wesentlich höher war. Daraus lässt sich schließen, dass die Unterschiede mit wachsendem Alter weiter ansteigen (McElvany et al., 2017).

Im Rahmen der Forschung von (Niklas und Schneider, 2012) wurden geschlechterspezifische Unterschiede in mathematischen und schriftsprachlichen Kompetenzen sowie dem Selbstkonzept von über 900 Kindern untersucht. Der Fokus dieser Forschung lag, anders als bei der IGLU, auf dem Ende der Kindergarten- und dem Anfang der Schulzeit. Die Studie ergab

signifikante Unterschiede zwischen den Jungen und Mädchen bezüglich der schriftsprachlichen Kompetenzen, also dem Lesen und Schreiben, am Ende der ersten Klasse zu Gunsten der Mädchen. Darüber hinaus zeigten die Mädchen höhere Ausprägungen des Selbstkonzepts bezüglich des Lesens und Rechtschreibens als die Jungen (Niklas und Schneider, 2012).

### **2.3.2 Geschlechterspezifische Unterschiede im Lernverlauf im Rahmen von Levumi**

Im Rahmen der Masterarbeit von Wenner (2018) wurde die Lesekompetenz von 18 Schülerinnen, je sechs aus drei verschiedenen Klassen, über acht Messzeitpunkte (MZP) hinweg mit der Onlineplattform Levumi erhoben. Dabei wurden drei Tests zur Leseflüssigkeit, Silben-, Wörter- und Pseudowörtertests, sowie ein Test zum sinnentnehmenden Lesen durchgeführt. Bei dem ersten MZP handelt es sich um eine Statusdiagnose der gesamten drei Klassen mit einer Grundgesamtheit von N=66. Bei den 18 regelmäßig getesteten SchülerInnen handelt es sich jeweils um die Leseschwächsten aus den drei Klassen. Der Schwerpunkt dieser Arbeit lag, ähnlich wie in der hier vorliegenden, auf den geschlechterspezifischen Unterschieden der Leseflüssigkeit. Jedoch wurde im Messzeitraum keine zusätzliche Förderung durchgeführt und das sinnentnehmende Lesen wurde nicht näher betrachtet. Die Auswertung der Ergebnisse zeigt, dass die Mädchen über die acht MZP hinweg im Mittelwert nur beim Silbenlesetest den Jungen überlegen waren. Im Wörter- und Pseudowörterlesetest waren die Jungen den Mädchen sowohl bei der Statusdiagnostik als auch im Mittelwert der acht MZP überlegen (Wenner, 2018). Insofern zeigen sich an dieser Stelle andere Ergebnisse als diese auf Grundlage des Forschungsstandes zu erwarten wären, welcher besagt, dass die Mädchen in der Regel höhere Leseleistungen als die Jungen zeigen (vgl. Kapitel 2.3.1). Ein zusätzlicher Lehrerfragebogen bezüglich des außerschulischen Leseverhaltens bei zwei der drei Klassen bestätigte hingegen die Tendenz der Forschung, dass eher die Mädchen in ihrer Freizeit lesen. Die Jungen lasen im Durchschnitt ein bis zwei Silben oder Wörter mehr falsch als die Mädchen. Ebenfalls bei der Statusdiagnose lag die Fehleranzahl bei den Jungen über dem Wert der Mädchen. Zusammenfassend kann formuliert werden, dass die Jungen zwar insgesamt in der Itembearbeitung bessere Leistungen zeigen, jedoch gleichzeitig auch mehr Fehler als die Mädchen machen. Als Erklärungsansatz für die hohe Fehlerzahl der Jungen wurde herangezogen, dass insbesondere zwei der Schüler ihren Fokus bei der Itembearbeitung eher auf die Quantität, also die Schnelligkeit, statt die Qualität gerichtet haben.

Über die acht MZP hinweg wurde deutlich, dass bei den Mädchen durchschnittlich ein höherer Lernzuwachs als bei den Jungen zu verzeichnen war. Beide Geschlechter verringerten die Fehleranzahl über acht MZP geringfügig. Bei der Reduktion lagen die Jungen jedoch vor

den Mädchen. Die geringere Fehleranzahl kann nach Wenner (2018) ggf. mit der höheren Leseintensität der Mädchen zusammenhängen, woraus eine höhere Konzentration beim Lesen und somit weniger Fehler resultieren. Grundsätzlich verbesserten sowohl die Mädchen als auch die Jungen ihre Leistungen in der Leseflüssigkeit und konnten ihre Fehleranzahl im Durchschnitt verringern.

## 2.4 Leseförderung

In diesem Kapitel wird vorerst auf die Relevanz von Leseförderung und anschließend auf die Arten und Rahmenbedingungen von Fördermaßnahmen eingegangen. Leseförderung in der Grundschule ist insofern von besonderer Relevanz, als dass Lehrkräfte bei 17% der SchülerInnen der vierten Klasse in Deutschland einen Förderbedarf sehen, womit eine nicht ausreichende Lesekompetenz verbunden ist (Bremerich-Vos et al., 2017). Daraus können sich fächerübergreifend erhebliche Schwierigkeiten in der zukünftigen Schullaufbahn ergeben. Auch wenn Deutschland hinsichtlich der internationalen Rangfolge der IGLU 2011 im oberen Drittel liegt, zeigt sich ein eher geringer Anteil an SchülerInnen auf der höchsten Kompetenzstufe. Der Anteil der SchülerInnen, welche noch in der Sekundarstufe eins eine basale Leseförderung benötigen wird als hoch eingeschätzt (Tarelli et al., 2012). In der IGLU 2016 befinden sich die Leseleistungen der deutschen Viertklässler lediglich im mittleren Bereich hinsichtlich des internationalen Vergleichs. Darüber hinaus geht die Leistungsschere zwischen den schwächsten und stärksten Kindern weit auseinander. Bezüglich dieser hohen Streuung ist in den 15 Jahren seit der IGLU 2001 eine Zunahme zu verzeichnen (Bos et al., 2017). Diese hohe Streuung kommt unter anderem dadurch zustande, dass einige Kinder, sog. Responder, im Verlauf des normalen Unterrichts Lernfortschritte erzielen und andere, sog. Non-Responder, nicht mit Fortschritten auf die im Unterricht verwendeten Methoden reagieren (Fuchs et al., 2004). Auf diese Weise profitieren nicht alle SchülerInnen vom selben Unterricht gleichermaßen, wodurch klasseninterne Leistungsunterschiede mit der Zeit ansteigen können. Auf Grundlage von drei unterschiedlichen Messverfahren zur Lesekompetenz und den Ergebnissen von drei verschiedenen Stichproben benennt Walter (2011) eine Gruppe von 25% der GrundschulernInnen als Non-Responder, welche im Zeitraum von einem Schuljahr keine Fortschritte hinsichtlich ihrer Lesekompetenz erreichen können (Walter, 2011). Im Unterricht, ohne regelmäßige Erhebungen der Leistungsentwicklung, fallen diese SchülerInnen den Lehrkräften erst auf, wenn der Abstand zu den Mitschülern bereits kritisch zu bewerten ist. Diese Problematik wird als *Wait-to-fail-Problematik* bezeichnet (Huber und Grosche, 2012). Um dem entgegen zu wirken, sind frühzeitige Fördermaßnahmen auf Grundlage diagnostischer Instrumente unerlässlich. Inwiefern diese diagnostischen Instrumente verwendet werden können, wird in Kapitel drei ausführlich dargestellt.

Eine Leseförderung kann an unterschiedlichen Orten erfolgen. Dabei handelt es sich um schulische, familiäre oder externe Förderorte (McElvany und Schneider, 2009). Neben dem Ort und den damit verbundenen Personen, die die Förderung durchführen, wie LehrerInnen, Eltern oder externe Personengruppen, kann ebenso in der Art der Förderung unterschieden werden. Eine mögliche Art von Förderung zielt auf die Lesemotivation ab. Dabei ist das Schaffen von Lesevorbildern, das zur Verfügung stellen von Ressourcen sowie das Einbringen von Zeiten und Orten zum eigenständigen oder gemeinsamen Lesen von Bedeutung. Die Förderung der Motivation erhöht die Lust am Lesen, wodurch es vermehrt zur Übung der Lesefähigkeiten kommt. Sie kann prinzipiell an allen drei Orten stattfinden (Lenhard, 2013).

Gezielte Fördermaßnahmen hingegen finden primär in der Institution Schule statt. Sie können sowohl präventiv als auch intervenierend erfolgen und finden auf individueller, Klassen- oder Kleingruppenebene statt. Bei einer Kleingruppenförderung können beispielsweise leistungsschwache SchülerInnen, SchülerInnen mit Migrationshintergrund und ggf. mit Deutsch als Zweitsprache oder SchülerInnen mit sonderpädagogischem Förderbedarf im inklusiven Setting im Vordergrund stehen. Die Förderung muss demzufolge an die Bedürfnisse der jeweiligen Zielgruppe angepasst werden (McElvany und Schneider, 2009). Um diese Passung von Leistungsstand und Förderung zu ermöglichen, ist die Diagnose der Lesekompetenz erforderlich. Erst wenn der individuelle Leistungsstand von den Lehrkräften erfasst wurde, kann, ausgehend von den bereits erworbenen Kompetenzen und den auftretenden Schwierigkeiten, eine Förderung geplant und durchgeführt werden (Stecker et al., 2005). Die Durchführungshäufigkeit von Fördermaßnahmen stellt ein zentrales Merkmal für den Erfolg der Intervention dar. Treten bei SchülerInnen schwere Lese-Rechtschreib-Schwierigkeiten auf, wird davon ausgegangen, dass eine intensive mehrmonatige Einzel- oder Kleingruppenförderung nötig ist, um einen sichtbaren Leistungsanstieg zu erreichen (Torgesen, 2005).

Gefördert werden können sowohl die Lesekompetenz im Allgemeinen als auch unterschiedliche Teilkompetenzen dieser, die individuellen Voraussetzungen für erfolgreiches Lesen sowie unterschiedliche Vorläuferfähigkeiten (McElvany und Schneider, 2009). Darüber hinaus werden zwei Arten von Leseschwierigkeiten voneinander unterschieden. Auf der einen Seite können Probleme hinsichtlich des Dekodierens oder der Worterkennung auftreten, auf der anderen Seite kann es sich um Schwierigkeiten beim Leseverstehen handeln (Nation, 2005). Beide Möglichkeiten erfordern eine spezifische Förderung. Probleme des Dekodierens und der Worterkennung verweisen auf Schwierigkeiten der basalen Lesefähigkeiten. Anknüpfungspunkte stellen hier die phonologische Bewusstheit, die auditive Wahrnehmung sowie die auditive Merkfähigkeit dar. Schwierigkeiten beim Leseverständnis beinhalten hierarchiehöhere Prozesse und bedürfen einer Förderung hinsichtlich des Wortschatzes, des Wissens

über Semantik, Grammatik, Morphologie und Syntax (Klicpera et al., 2017) sowie bezüglich des Arbeitsgedächtnisses.

Neben motivational oder rezeptionsästhetisch orientierten Modellen zur Leseförderung gibt es Förderprogramme, welche den Fokus auf die Förderung expliziter Teilkompetenzen des Lesens legen (Kruse, 2011). Der KLA beispielsweise stellt ein Programm dar, welches systematisch aufgebaute Übungen anbietet. Dabei wird die Verinnerlichung der Graphem-Phonem-Korrespondenz mit Hilfe von Lautgebärden unterstützt (Dummer-Smoch und Hackethal, 2016). Der KLA wird aufgrund seiner Relevanz in dieser Arbeit ausführlich im nachfolgenden Kapitel dargestellt.

Das Unterrichtskonzept *Leseflüssigkeit fördern* von Rosebrock et al. (2011) greift beispielsweise auf das Prinzip des Lesetandems zurück. Dabei lesen sich zwei Partner, ein leistungsstarker sowie ein leistungsschwächerer, simultan wiederholt die gleichen Texte vor. Ziel ist die Steigerung des Lesetempos und der Gewinn von Sicherheit beim Leseprozess (Rosebrock et al., 2011). Es handelt sich nicht um ein geschlossenes Programm. Es ist vielmehr als eine in den Unterricht zu integrierende Förderroutine gedacht (Lenhard, 2013). Auch das Förderprogramm *Wir werden Textdetektive* von Gold et al. (2004), welches im Rahmen einer formativen Evaluation entwickelt wurde, sowie *Wir werden Lesedetektive* von Rühl et al. (2006) als Anpassung für leseschwache SchülerInnen sind im deutschsprachigen Raum umfassend evaluiert. *Textdetektive* zielt darauf ab, den SchülerInnen sieben verschiedene kognitive bzw. metakognitive Lesestrategien zu vermitteln. *Lesedetektive* legt den Fokus ebenfalls auf das Textverstehen, jedoch in einer quantitativ reduzierten Form (Lenhard, 2013).

## 2.5 Der Kieler Leseaufbau

Der Kieler Leseaufbau ist ein Förderprogramm, welches aus dem Bereich der Lese-Recht-schreib-Schwäche (LRS-Therapie) stammt und ist insbesondere für Kinder mit Lernschwierigkeiten wirksam (Jungjohann et al., 2017). Der KLA dient als Grundlage für die Konstruktion der Tests im Rahmen der Online Plattform Levumi. Gemeinsamkeiten und Unterschiede in der Struktur werden in Kapitel 3.5 dargelegt.

Beim KLA handelt es sich um ein Phonem-basiertes Lesetraining, welches auf die Festigung des automatisierten Lesens abzielt. Im Fokus des Trainings steht das Lesen, welches mit vielfältigen spielerischen Methoden geübt wird (Groth et al., 2013). Der KLA beinhaltet 14 verschiedene Schwierigkeitsstufen und orientiert sich an fünf Prinzipien sowie an fünf Strategien. Bei den Prinzipien, welche sich auf die ersten vier Stufen beziehen, handelt es sich um Konsonant-Vokal-(Konsonant)-Verbindungen, die als einfache Wortstrukturen gelten. Dabei

gibt es die Struktur Konsonant-langer Vokal (KV) und Konsonant-kurzer Vokal (Kv). In Stufe eins bis vier wird sich auf Wörter mit einer 1:1-Zuordnung von Buchstabe und Phonem sowie auf Wörter mit dehnbaren Konsonanten beschränkt. Darüber hinaus werden Konsonanten Verbindungen (KK) wie <Bl> oder <Kr> weggelassen. Auf kurze Vokale in Endungen außer <e> wird ebenfalls verzichtet (Dummer-Smoch und Hackethal, 2016). Die fünf Strategien gehen über die vierte Stufe hinaus und beinhalten das Lautieren anstelle des Buchstabierens sowie die Strategien „Verschleifen zweier Laute zur Silbe, Silbenschwingen und Dehnsprechen beim Schreiben, die Strategie der nachträglichen Selbstkontrolle mit Lautgebärden beim Schreiben (und) die Strategie zum Unterscheiden ähnlicher Laute nach KOSSOW“ (Hervorhebungen im Original) (Dummer-Smoch und Hackethal, 2016). Die Verwendung von Lautgebärden dient der Kompensation von Beeinträchtigungen in der Lautwahrnehmung. Dabei ist das Erlesen eindeutiger Graphem-Phonem-Verbindungen zentral, weswegen auf Konsonantenverbindungen wie <pf>, <st>, <sp>, <str>, die Lautverbindungen <ng> und <nk> sowie auf Dopplungs- und Dehnungsschreibung verzichtet wird (Dummer-Smoch und Hackethal, 2016).

Nachfolgend werden die 14 Stufen des KLA nach Schwierigkeitsstufen zusammengefasst dargestellt. Stufe eins bis vier führt lange Vokale sowie die Diphthonge <au> und <ei>, dehnbare Konsonanten und das kurze <e> in den Endungen <en> und <er> ein. Den gewählten Wörtern liegt die Struktur KV-Kv und KV-KvK zugrunde. Die Stufen fünf und sechs übernehmen die erwähnten Aspekte der vorherigen Stufen und ergänzt werden die Plosivlaute und die nicht dehnbaren Konsonanten, wie <b>, <d>, <g>, <p>, <t>, <k> (Dummer-Smoch und Hackethal, 2016). Den Stufen 7 bis 10 wird das <eu>, das kurze <e> in <el> und alle noch fehlenden Konsonanten hinzugefügt. Dabei werden pro Stufe jeweils drei neue Konsonanten eingeführt. In Stufe 11 kommt es zur Einführung einer neuen Wortstruktur KKV-KvK. Die folgende Stufe 12 beinhaltet nun auch kurze Vokale in der ersten Silbe des Wortes KKv-Kv(K). Stufe 13 bezieht sich auf die Inhalte der beiden vorangegangenen Stufen und beinhaltet eine Gegenüberstellung der bereits eingeführten Wörter. Die abschließende Stufe 14 sieht eine Übung mit Wörtern vor, welche aus mehr als drei Silben bestehen und nimmt Bezug zu allen durchlaufenen Stufen (Dummer-Smoch und Hackethal, 2016).

Das Training des KLAs beinhaltet vielfältiges Material, welches nach und nach weiter ergänzt wurde. Das ursprüngliche Material umfasst ein Handbuch mitsamt verschiedenen Spielanweisungen und Lautgebärden, Spielkarten für jede Stufe, Arbeitsblätter, Lese- und Übungstexte. Ein Übungsheft, Lesehefte und weitere Lesetexte wurden dem Material-Pool nachträglich hinzugefügt. Auch andere Autoren entwickelten zusätzliches Material (Dummer-Smoch

und Hackethal, 2016). Auf Grundlage dieser Stufenfolge findet sowohl Diagnostik als auch Förderung statt. Diese Förderung richtet sich an SchülerInnen, welche in der ersten oder zweiten Klassenstufe kein ausreichendes Fundament für das automatisierte Lesen entwickeln konnten. Die Automatisierung stellt anschließend die Grundlage zur Sinnerfassung dar. Die Förderung kann sowohl einzeln als auch in Kleingruppen stattfinden (Dummer-Smoch und Hackethal, 2016).

### 3 Formative Diagnostik und Leistungsmessung

Diagnostik und Leistungsmessung stellen eine wichtige Grundlage für die Gestaltung von Unterricht und die Planung von gezielten Fördermaßnahmen dar. Formative Diagnostik dient nicht, wie klassische standardisierte Instrumente der Statusdiagnostik, dem punktuellen Erfassen des Leistungsstandes zu einem oder zwei Messzeitpunkten im Schuljahr, sondern der engmaschigen und veränderungssensiblen Abbildung von Leistungsveränderungen mit dem Ziel das pädagogische Handeln während des Lernprozesses an den Leistungsstand der SchülerInnen anzupassen (Walter, 2014). Da ein großer Teil der SchülerInnen keine signifikanten Fortschritte in der Lesekompetenz macht und die Lehrkräfte mit Hilfe von formativer Diagnostik diese Stagnation zeitnah erkennen können, zeigt sich die Verwendung solcher Instrumente als lernförderlich (Walter, 2011). Auf Grundlage dessen werden nachfolgend die Anforderungen an ein solches Instrument im schulischen Kontext, die Spezifika des Curriculum-based Measurements (CBM) und der Lernverlaufdiagnostik (LVD) sowie der Einzelfallforschung näher erläutert. Das Kapitel schließt mit der Darstellung der Online-Plattform Levumi, welche in vorliegender Forschung zur Datenerhebung hinsichtlich der Leistungen der Lesekompetenz von Zweitklässlern verwendet wurde.

#### 3.1 Anforderungen an Diagnostik im schulischen Kontext

Für Instrumente zur Diagnostik im schulischen Kontext sind sowohl grundlegende Qualitätskriterien der Testkonstruktion als auch bereichsspezifische Kriterien bedeutsam. Im Folgenden werden sowohl verschiedene Gütekriterien als auch mögliche Bezugsnormen zur Einordnung von Schülerleistungen dargelegt.

##### 3.1.1 Gütekriterien

Gütekriterien dienen der Qualitätsbeurteilung psychologischer Tests. Neben den Hauptgütekriterien werden nachfolgend insbesondere die Nebengütekriterien mit dem Fokus auf die besonderen Anforderungen formativer Diagnostik im schulischen Kontext dargestellt.

Moosbrugger und Kelava (2012) nennen zehn verschiedene Gütekriterien. Von besonderer Bedeutung sind hohe Werte hinsichtlich der Objektivität, Reliabilität und Validität. Die Objektivität ist für die Vergleichbarkeit von Messergebnissen von Bedeutung. Wird das gemessene Merkmal unabhängig vom Testleiter und Testauswerter gemessen, so gilt ein Test als objektiv. Eindeutige Regeln hinsichtlich der Durchführung und Auswertung sind dafür wichtig. Eine Standardisierung ist gegeben, sofern die Testperson während der Messung die einzige Variationsquelle darstellt (Moosbrugger und Kelava, 2012). Die Reliabilität beschreibt die Messgenauigkeit eines Tests. Diese beinhaltet, inwiefern ein Test das zu untersuchende Merkmal frei von Messfehlern misst (Moosbrugger und Kelava, 2012). Das Gütekriterium der Validität

bezieht sich auf den inhaltlichen Kern eines Tests. Es wird gemessen, ob ein Instrument tatsächlich das Merkmal, das gemessen werden soll, misst (Moosbrugger und Kelava, 2012). Als Nebengütekriterien gelten unter anderem Normierung, Nützlichkeit, Zumutbarkeit und Unverfälschbarkeit. Ist ein Test normiert, so wurde ein Bezugssystem erstellt, welches es ermöglicht die Ergebnisse einer Person beispielsweise mit denen von Gleichaltrigen zu vergleichen. Die Nützlichkeit beinhaltet die Relevanz der Untersuchung eines Merkmals und ein positives Verhältnis von Nutzen und Schaden bzw. Kosten. Mit Zumutbarkeit ist gemeint, inwiefern die Durchführung eines Tests für die Testperson hinsichtlich ihrer individuellen zeitlichen, psychischen und physischen Voraussetzungen zumutbar ist (Moosbrugger und Kelava, 2012). Die Unverfälschbarkeit ergibt sich daraus, dass die zu testende Person ihre Testwerte nicht gezielt durch ihr Verhalten steuern oder verzerren kann (Moosbrugger und Kelava, 2012).

Im Kontext der schulischen formativen Diagnostik sind darüber hinaus die Gütekriterien Testökonomie, Fairness, Änderungssensibilität und Eindimensionalität von besonderer Bedeutung (Casale et al., 2015). Ein Test gilt als ökonomisch, sofern dieser, gemessen an dessen Erkenntnisgewinn, keine über die nötigen Maße hinausgehenden finanziellen und zeitlichen Ressourcen beansprucht. Die Fairness eines Tests ergibt sich, wenn die Ergebnisse zu keiner systematischen Benachteiligung von Personengruppen hinsichtlich ihrer ethnischen, soziokulturellen und geschlechterspezifischen Zugehörigkeit führen (Moosbrugger und Kelava, 2012). Besonders im inklusiven schulischen Kontext ist die Fairness hinsichtlich des Migrationshintergrundes oder einem ggf. vorhandenen sonderpädagogischen Förderbedarf von Bedeutung (Gebhardt et al., 2016b). Im Rahmen formativer Diagnostik in heterogenen Gruppen ist auch die Änderungssensibilität von Bedeutung. Sie beschreibt inwiefern ein Test auch geringe Veränderungen erfassen kann (Klauer, 2011). Die Eindimensionalität eines solchen Tests gewährleistet, dass das Antwortverhalten der Testperson nur auf eine Kompetenz zurückzuführen ist (Casale et al., 2015).

### **3.1.2 Bezugsnormen**

Normen sind Bezugssysteme, die im Kontext dieser Arbeit der Einordnung von Leistung dienen (Moosbrugger und Kelava, 2012). Die Leistung einer einzelnen Testperson kann auf Grund von unterschiedlichen Bezugsnormen interpretiert werden, woraus sich teils hohe Differenzen hinsichtlich der Leistungsbewertung ergeben können. Es wird von drei verschiedenen Bezugsnormen gesprochen: Die kollektive, die individuelle und die kriteriumsorientierte Bezugsnorm.

Die *kollektive Bezugsnorm* gilt als die am häufigsten verwendete Bezugsnorm und orientiert sich am Durchschnitt einer gewissen Gruppe, welche unterschiedlich groß sein kann. Im schulischen Kontext wird sich primär an den durchschnittlichen Leistungen der Klasse und eher seltener an einer anderen Gruppe gleichaltriger SchülerInnen, wie beispielsweise aus Nordrhein Westfalen, orientiert (Kostorz, 2016). Die Leistungsbewertung kann je nach Niveau der kollektiven Bezugsnorm anders ausfallen, was wiederum die Einordnung der Leistung für Lehrkräfte erschweren kann. Die *individuelle Bezugsnorm* betrachtet Leistung hinsichtlich des individuellen Lernfortschritts. Dafür wird die Entwicklung des individuellen Lerntempos und der vergangenen Lernzuwächse herangezogen. Auf diese Weise kann die Leistung eines Kindes nach Verwendung einer kollektiven Bezugsnorm weit unterdurchschnittlich sein, nach Verwendung der individuellen Bezugsnorm dennoch positiv (Kostorz, 2016). Die dritte Bezugsnorm ist die *kriteriumsorientierte Norm*. Sie zeichnet sich durch die Leistungsbewertung hinsichtlich des Erreichens eines zuvor operationalisierten Ziels aus. Inwiefern das Ziel erfüllt wurde, ist zentral für die Leistungsbewertung (Kostorz, 2016). Dabei kann es sich um das Formulieren von einheitlichen Zielen für die ganze Klasse oder um individuelle Leistungsziele im inklusiven Kontext handeln. Es wird deutlich, dass die gewählte Bezugsnorm einen maßgeblichen Einfluss auf die Einordnung und Bewertung von Schülerleistungen hat sowie auf die individuelle Bewertung der eigenen Leistung mit Einfluss auf die Motivation von Seiten der SchülerInnen.

### 3.2 Curriculum-based Measurement

Der Begriff Curriculum-based Measurement (CBM) stammt aus dem sonderpädagogischen Kontext im amerikanischen Raum. Der Ansatz wurde 1972 entwickelt und ist geprägt von Stanley Deno (2003). Bei Curriculum-based Assessment (CBA) handelt es sich um die Curriculum-basierte Einschätzung von Schülerleistungen. CBM hingegen ist ein Teilbereich dessen, welcher auf die Messung des Lernfortschritts abzielt und aus dem Modell data-based-programm modification (DBPM) entstanden ist (Diehl und Hartke, 2007). Es handelt sich um eine Konzeption, die der formativen Diagnostik zuzuordnen ist. Formative Diagnostik beschreibt die häufige Evaluation des Lernprozesses, indem der Lernverlauf über einen gewissen Zeitraum hinweg dargestellt wird. Summative Evaluation findet im Gegensatz dazu erst am Ende eines Lernprozesses statt und erhebt, ob ein Lehrziel erreicht wurde (Klauer, 2006). Der Begriff CBM beinhaltet die Orientierung am Curriculum und das Verwenden von aus dem Unterricht stammenden Materialien, was jedoch einen gewissen Aufwand hinsichtlich der Erstellung des Aufgabenmaterials für Lehrkräfte impliziert (Diehl und Hartke, 2007). Die Aufgaben müssen repräsentativ für die geforderte Leistung sein. Darüber hinaus ist eine gewisse Größe des Itempools unabdingbar, damit zwar die Aufgabenstellung bei jeder Messung

gleich ist, jedoch die Items immer zufällig gewählt werden können. Dabei beinhaltet jede Messung den gleichen Schwierigkeitsgrad. Die Leistungen an einem bestimmten Messzeitpunkt können sowohl im intra- als auch im interindividuellen Vergleich betrachtet werden (Diehl und Hartke, 2007).

Tests im Rahmen von CBM messen Schülerleistungen in regelmäßigen Abständen jeweils zum selben Lehrziel (Klauer, 2006) über einen längeren Zeitraum hinweg (Klauer, 2014). Die Leistungsveränderungen von SchülerInnen werden demzufolge im Verlauf der Zeit dokumentiert. Die Leistungen zu den Messzeitpunkten können graphisch dargestellt werden, wodurch der Lernverlauf visuell sichtbar wird und leichter zu interpretieren ist (Walter, 2009). Die Dauer einer Messung sollte pro Kind und Test drei Minuten nicht überschreiten (Deno, 2003b). CBM dient weder der Selektion leistungsschwächerer SchülerInnen, noch profitieren nur diese vom CBM (Klauer, 2006). Auch bei SchülerInnen, die den Lehrkräften eher weniger auffallen, können stagnierende oder sinkende Leistungen frühzeitig erkannt werden (Diehl und Hartke, 2007).

Aus dem reinen Erfassen des Lernverlaufs resultiert jedoch keine Steigerung der Leistungen. Vielmehr ist die Interpretation der erhobenen Daten durch Lehrkräfte von besonderem Mehrwert hinsichtlich der Unterrichtsentwicklung sowie des Lernzuwachses. Die Analyse der Verlaufsgraphen ermöglicht die Evaluation der Effektivität des eigenen Unterrichts und der verwendeten Methoden (Deno, 2003a). Die Lehrkraft erlangt durch regelmäßige Messungen die Möglichkeit ihre zukünftigen pädagogischen Entscheidungen an die individuellen Bedürfnisse der SchülerInnen anzupassen (Stecker et al., 2005). Auf diese Weise können Lernschwierigkeiten bereits während und nicht erst am Ende einer Unterrichtseinheit erfasst werden, wodurch gravierendere Schwierigkeiten durch die Anpassung des Unterrichts oder einer Fördermaßnahme abgewendet werden können (Diehl und Hartke, 2007). In der systematischen Untersuchung von Fuchs et al. (1993) bezüglich der Lernzuwächse der Lesekompetenz, konnten Normwerte je Klassenstufe für die Messtechniken CBM-LL und CBM-Maze in Zuwachs pro Woche erhoben werden. Die Messtechnik CBM-LL steht für das laute Lesen eines Textabschnittes mit dem Kompetenzmaß Wörter pro Minute (Walter, 2009). Beim CBM-Maze lesen die SchülerInnen einen Text, bei welchem jedes siebte Wort fehlt und das richtige Wort aus drei Optionen mit zwei Distraktoren gewählt werden muss (Walter, 2010). In der zweiten Jahrgangsstufe ist für CBM-LL, mit dem Indikator Anzahl der richtig gelesenen Wörter in einer Minute, ein Zuwachs von 1.46 Wörtern pro Woche zu erwarten. Für die Messtechnik CBM-Maze mit dem Indikator Anzahl der in 2.5 Minuten richtig ausgewählten Alternativen ist in der zweiten Jahrgangsstufe ein Zuwachs von 0.39 pro Woche zu erwarten (Fuchs et al., 1993).

Hinsichtlich eines Leistungsziels im Rahmen von zusätzlicher Förderung, wird auf den mittleren Wert eine Standardabweichung addiert. Daraus resultiert beim CBM-LL ein Wert von ca. zwei Wörtern sowie einer von 0.74 beim CBM-Maze (Walter, 2010). Anhand dieser Werte können Ziele über mehrere Wochen durch Addition des mittleren Zuwachses formuliert werden.

### 3.3 Lernverlaufsdiagnostik Lesen

Lernverlaufsdiagnostiken sind kontrollierte Einzelfallstudien, bei welchen die Leistungsentwicklung über einen gewissen Zeitraum hinweg betrachtet wird (Wilbert, 2014). Aufgrund des Schwerpunkts der vorliegenden Arbeit auf dem Bereich Lesen wird der Fokus nach einer allgemeinen Erläuterung von LVD auf ebendiesen Bereich gelegt. Im Gegensatz zum CBM werden Tests der LVD nicht nach der klassischen Testtheorie (KTT) konstruiert (Deno, 2003b). Da es sich hierbei aufgrund der nötigen Überprüfung von Dimensionalität, Messinvarianz und Änderungssensibilität nicht mehr um den aktuellen wissenschaftlichen Stand handelt, werden bei der LVD Konzepte der Item-Response-Theory (IRT) verwendet (Gebhardt, Heine, Zeuch, & Förster, 2015); Klauer, 2014; Wilbert und Linnemann, 2011). Die IRT ist als Erweiterung der klassischen Testtheorie zu verstehen (Moosbrugger, 2012) und wird aufgrund der expliziten Prüfung der Eigenschaften einer Messung im engeren Sinn als Grundlage für Tests der LVD empfohlen (Wilbert, 2014). Im Gegensatz zur KTT, bei welcher der Fokus auf den Test an sich gelegt wird, setzt die IRT ihren Schwerpunkt auf die einzelnen Items. Dabei wird untersucht inwieweit Rückschlüsse auf Einstellungs-, Persönlichkeits- oder Fähigkeitsmerkmale allein auf Grundlage der gegebenen Antworten gezogen werden können (Moosbrugger, 2012). Bei diesem stochastischen Modell wird die Lösungswahrscheinlichkeit eines Items in Abhängigkeit von Fähigkeitsvariablen angegeben. Bedingung dafür ist, dass die Schwierigkeit der Items weitgehend unabhängig von individuellen Merkmalen außerhalb der Personenfähigkeit ist, was für eine hohe Testfairness notwendig ist (Mühling et al., 2017).

Beim Generieren der Aufgaben stellt die Nutzung eines robusten Indikators eine Möglichkeit dar, eine komplexe Kompetenz ökonomisch zu erfassen. Es handelt sich um einen solchen, wenn er mit unterschiedlichen Teilkompetenzen einer größeren Kompetenz korreliert (Walter, 2008). Dieser Indikator ist ein Maß für die Leistung, woraufhin auf eine übergeordnete Kompetenz geschlossen werden kann. Das flüssige Lesen mit dem Maß richtig gelesener Wörter pro Minute wurde beispielsweise als robuster Indikator für allgemeine Lesekompetenz identifiziert (Walter, 2009). Außerdem kann LVD nach dem Ansatz des Curriculum Samplings

erfolgen. Dabei wird die am Schuljahresende geforderte Kompetenz in Teilkompetenzen untergliedert. Anschließend werden Items mit der gleichen Schwierigkeit erstellt. Das beinhaltet, dass jeder Paralleltest dieselbe Anzahl an zuvor festgelegten Aufgabentypen beinhaltet (Voß, 2014).

Der Begriff Lernverlaufsdiagnostik wurde insofern durch Klauer (2006) geprägt, als dass er seine eigene Formulierung von Lernfortschritts- zu Lernverlaufsdiagnose änderte. Diese Abwandlung ist von Bedeutung, da es sich bei der Lernentwicklung in der Realität nicht immer um einen Fortschritt handelt. Der Lernverlauf ist ebenso geprägt von Lernstillständen oder gar Verlusten (Klauer, 2011). Die LVD kann in zwei Formen unterschieden werden. Zum einen können Tests die Verbesserung einer Kompetenz messen, wie beispielsweise die Leseflüssigkeit, zum anderen kann die substanzielle Erweiterung einer Kompetenz gemessen werden, wie es im Bereich der Mathematik eher der Fall ist, in dem stets neue Komponenten erworben werden (Klauer, 2011). Zentral bei der LVD im Bereich Lesen ist die Messung derselben Kompetenz im gesamten Messzeitraum durch für diese Kompetenz repräsentative Aufgaben, woraus sich die Validität des Tests ergibt (Klauer, 2011). Daraus resultiert, dass es sich bei der zu messenden Kompetenz um die Zielkompetenz zum Ende des Messzeitraums handelt. Der Test ist demzufolge so konstruiert, dass die SchülerInnen zu Beginn nur sehr wenige und mit der Zeit immer mehr Aufgaben lösen können (Klauer, 2014). Um dem Gütekriterium der Eindimensionalität zu entsprechen, muss das Ergebnis allein auf die gemessene Kompetenz zurückzuführen sein (Wilbert und Linnemann, 2011). Die einzelnen Tests werden pro SchülerIn und Messzeitpunkt (MZP) zufällig auf Grundlage des Itempools zusammengestellt. Dieses Vorgehen nennt sich Itemsampling, was beinhaltet, dass sich beliebig viele unterschiedliche Tests generieren lassen (Strathmann und Klauer, 2010), wodurch ein Leistungszuwachs durch einen Übungseffekt weitgehend vermieden wird. Folglich bekommt jede(r) SchülerIn eine eigene Zufallsstichprobe (Klauer, 2011). Neben den zufällig gezogenen Items ist der Aspekt der homogenen Schwierigkeit von besonderer Relevanz (Strathmann und Klauer, 2010; Klauer, 2011). Um Lernzuwachs tatsächlich mit den gewachsenen Fähigkeiten der SchülerInnen begründen zu können, was für positive Werte hinsichtlich der Reliabilität unabdingbar ist (Wilbert und Linnemann, 2011), ist von hoher Relevanz, dass der Test zu jedem MZP von gleicher Schwierigkeit ist. Auch die Änderungssensibilität ist für die LVD von besonderer Bedeutung. Sie beschreibt, inwiefern auch kleine Veränderungen in den Leistungen der SchülerInnen durch das Instrument sichtbar gemacht werden können. Besonders bei leistungsschwächeren SchülerInnen oder im sonderpädagogischen Kontext müssen auch geringe Veränderungen erkennbar sein (Klauer, 2011). Je näher die MZP beieinander liegen, desto kleiner sind in der Regel auch die Leistungsveränderungen.

Die Lesekompetenz kann aus zwei unterschiedlichen Perspektiven betrachtet werden. Die Lernfortschrittsdiagnostik Lesen (LDL) bezieht sich auf elementare Prozesse des Wortlesens (*Bottom-Up*-Perspektive vgl. Kapitel 2.2.1), wohingegen die Verlaufsdiagnostik sinnerfassendes Lesen (VSL) die Lesekompetenz aus der *Top-Down*-Perspektive betrachtet (Walter, 2014). Bereits evaluierte lernverlaufsdiagnostische Verfahren im Bereich der Lesekompetenz können in drei verschiedene Arten von Messverfahren unterteilt werden, welche als robuste Indikatoren für die Lesefähigkeit gelten (Walter, 2008). Bei der ersten Form handelt es sich um das *laute Lesen* (oral reading fluency), wobei den SchülerInnen ein Textabschnitt vorgelegt wird, welcher so richtig und flüssig wie möglich vorgelesen werden soll. Die Dauer beträgt ein bis drei Minuten. Kommt es beim Lesen zur Auslassungen oder Einfügungen von Graphemen, zur Ersetzung dieser sowie zu langem Zögern oder einer falschen Aussprache des Wortes, so wird dieses Item als Fehler gewertet (Walter, 2014). Das Kompetenz-Maß ergibt sich aus den richtig gelesenen Wörtern pro Minute. Die Bezeichnung *maze selection* stammt aus dem englischsprachigen Raum und ist die zweite Form der drei möglichen Messverfahren. Die SchülerInnen lesen ebenfalls einen Textabschnitt. Der Unterschied zur ersten Form ist, dass der Abschnitt leise gelesen wird und anstelle jedes siebten Wortes stehen drei verschiedene Wörter zur Auswahl in Klammern. Die Aufgabe ist es, das inhaltlich passende Wort aus den zwei Distraktoren zu erkennen und auszuwählen. Das Kompetenz-Maß resultiert aus der Anzahl der richtig markierten Wörter in den Klammern (Walter, 2014). Die dritte Form ist die *Wortidentifikation*. Ein bis drei Minuten lang werden einzelne, hoch frequente Wörter laut vorgelesen. Das Kompetenz-Maß sind die richtig gelesenen Wörter, dessen Bewertung nach denselben Kriterien wie denen der oral *reading fluency* erfolgt (Walter, 2011).

Walter (2010) veröffentlichte im deutschsprachigen Raum durch den Hogrefe Verlag den Test zur *Lernfortschrittsdiagnostik Lesen*. Etwas später folgte das Testverfahren *Verlaufsdiagnostik sinnerfassendes Lesen*, ebenfalls entwickelt von Walter (2013) (Klauer, 2014). Aufgrund der hohen Ökonomie hinsichtlich der Durchführung, Auswertung und Rückmeldung bei LVD entwickelte Souvignier (2014) das Quop-System, ein internetbasierter Ansatz der Lernverlaufsdiagnostik für Mathematik und Lesen, welcher seit mehreren Jahren erfolgreich in Grundschulen eingesetzt wird. Auch bei diesem Ansatz wird zwischen basalen Lesefertigkeiten und Leseverständnisleistungen unterschieden (Souvignier et al., 2014).

### 3.4 Einzelfallforschung

Kontrollierte Einzelfallstudien basieren auf regelmäßigen Messungen eines Zielverhaltens bzw. einer Zielkompetenz von Individuen oder wenigen Personen. Die Einzelfallforschung ist auf Grund von eher kleinen und heterogenen Stichproben, die neben der Psychologie auch oft im Bereich der Sonderpädagogik vorzufinden sind, auch in vorliegender Forschung von Bedeutung (Grünke, 2012). Wiederholte Messungen sind aufgrund von Verhaltens- bzw. Leistungsschwankungen oder vorhandenen Messfehlern bedeutsam. Einzelfallversuchspläne können auf dem *AB-Plan* beruhen. Dieser beschreibt verschiedene Phasen im Zeitraum der Datenerhebung. Phase A steht für die Baseline bzw. die Grundrate, welche die Leistung der abhängigen Variable vor Beginn einer Intervention darstellt. Sie dient als Grundlage für eine Leistungsprognose und um den Interventionseffekt am Ende der Forschung zu bestimmen. Eine möglichst stabile Baseline, die nur wenigen Schwankungen unterliegt, ist von Vorteil. In Phase B folgt die Intervention selbst, in welcher weiterhin das Verhalten oder die Leistung, nun aber mit Einfluss einer bestimmten Maßnahme, dokumentiert wird (Jain und Spieß, 2012). Der *AB-Plan* ist das einfachste Modell, bei welchem die Länge der Phasen variieren können. Dieser Versuchsplan kann jedoch beliebig erweitert und verändert werden, indem eine erneute A-Phase oder weitere Interventionsphasen (B-Phasen) hintereinandergeschaltet werden. Im Rahmen der Lernverlaufdiagnostik ist ein Multiple-Baseline-Design, bei welchem mehrere Testpersonen die gleiche Förderung ggf. auch mit unterschiedlichem Beginn der Intervention erhalten, besonders sinnvoll, da Zusammenhänge zwischen Intervention und Leistungsanstieg nachgewiesen werden können. Bei Multiple-Baseline-Designs wird zwar keine Korrelation bewiesen, jedoch besitzen diese eine höhere interne Validität (Wilbert, 2014). Der Fokus bei diesem Design kann auf einer Testperson liegen, bei welcher verschiedene Interventionen durchgeführt werden, oder auf unterschiedlichen Personen, bei welchen jedoch dieselbe Intervention durchgeführt wird (Kazdin, 1982).

Eine stabile Grundrate ist vorhanden, sofern keine Trends, weder nach unten noch nach oben, zu erkennen sind. Auf diese Weise wird eine solide Grundlage für die Bewertung der Intervention gelegt. Die Stabilität der Baseline wird durch die Anzahl der Datenpunkte bestimmt, welche nach Kern (1997) in einem bestimmten Bereich liegen sollten. Es gilt, dass 80 bis 90% der Datenpunkte in einem Bereich von 15% um den Mittelwert liegen sollten. Liegen mehr Werte außerhalb, so handelt es sich um eine variable Grundrate. Darüber hinaus kann man einen Trend, die Steilheit der Daten, bestimmen, wobei dieser entweder aufsteigend oder absteigend ist (Kern, 1997). Es ist zu erwähnen, dass sich die Angaben auf eine Grundratenerhebung ohne Förderung beziehen und dem Bereich der systematischen Verhaltensbeobachtung zugeordnet werden (Kern, 1997). Dennoch ist die Orientierung an

diesen Werten von Nutzen, um die SchülerInnen bezüglich der Stabilität ihres Lernverlaufs einzuordnen.

### 3.5 Die Online-Plattform Levumi

Die Online-Plattform Levumi ist ein computergestütztes Instrument der Lernverlaufsdagnostik. Es handelt sich um ein gemeinsames Forschungsprojekt von Dr. Markus Gebhardt, Kirsten Diehl und Andreas Mühling (Gebhardt et al., 2016a; Jungjohann, DeVries, Gebhardt & Mühling, 2018). Der Name der Plattform stellt gleichzeitig den des Maskottchens dar, bei welchem es sich um einen kleinen Drachen handelt.

Mit Levumi werden drei unterschiedliche Ziele verfolgt. Einerseits soll der Schulpraxis ein praktikables und kostenlos zur Verfügung stehendes Instrument zur Lernverlaufsmessung für die Fächer Deutsch und Mathematik angeboten werden. Andererseits dient es der Erweiterung zur Forschung im Rahmen von Lernverlaufsmessung und der Verbreitung hinsichtlich der Akzeptanz bei Lehrkräften. Das dritte Ziel stellt die Nutzung der gesammelten Daten zur Verbesserung diagnostischer Maßnahmen und damit verbundenen gezielten Fördermaßnahmen dar. Dazu sollen den Lehrkräften verschiedene Maßnahmen zur Verfügung gestellt werden (Mühling et al., 2017). Im Sinne eines *whole-in-one-Pakets* können Lehrkräfte mit Levumi sowohl diagnostizieren als auch darauf abgestimmte Förderungen durchführen (Gebhardt et al., 2016b). Die Durchführung und Konstruktion der Lesetests werden in den nun folgenden Unterkapiteln dargelegt.

#### 3.5.1 Lernverlaufsmessung mit Levumi

Im Bereich Lesen werden vier verschiedene Tests angeboten. Dabei handelt es sich um drei Tests zur Erhebung der Leseflüssigkeit, den Silben- Wörter- und Pseudowörtertest und darüber hinaus um einen Test zum sinnentnehmenden Lesen. Die Silben- und Wörtertests wurden von Kirsten Diehl im Rahmen des Rügener Inklusionsmodells (RIM) entwickelt und für Levumi adaptiert. RIM beruht auf dem *Response-to-Intervention-Ansatz* (RTI), wobei die Mehrebenenprävention, die Evidenzbasierung und das Monitoring-System als dessen Kernelemente gelten (Gebhardt et al., 2016b). Die Pilotstudie von Levumi wurde mit Diagnose- und Förderklassen des Arbeitskreises Schwaben im Schuljahr 2015/2016 durchgeführt. Im Rahmen dessen wurde ein Lehrerhandbuch verfasst und Fehler wurden behoben. Für den Silbentest N3 wurden Raschmodelle zur Veränderungsmessung verwendet. Insgesamt wurden für diesen Test reliable Ergebnisse bei der Pilotstudie aufgezeigt. Es handelt sich jedoch nur um vorläufige Ergebnisse, da es sich um eine geringe Stichprobe von 90 SchülerInnen mit sonderpädagogischem Förderbedarf im Bereich Lernen handelt. Obgleich die Tests aus

bereits als reliabel geltenden Tests entwickelt wurden, steht ein endgültiger Nachweis hinsichtlich der Reliabilität und Validität noch aus (Gebhardt et al., 2016b).

Für die Nutzung von Levumi ist keine Installation oder das Durchführen von Updates notwendig, sondern lediglich ein aktueller Browser. Das Format ist insofern von Vorteil, als dass zeitliche Ressourcen der Lehrkräfte hinsichtlich der Organisation und Auswertung größtenteils geschont werden. Durch die Plattform werden den Lehrkräften Lernverlaufstests inklusive einer qualitativen und quantitativen Auswertung angeboten. Zuerst legen die Lehrkräfte alle SchülerInnen der Klasse in Levumi an. Es gibt eine Klassen- und eine Schüleransicht, wobei die Schüleransicht jeweils einen Graphen zur Anzahl und Quote der gelösten und ungelösten Items sowie die Aufgaben anzeigt. Bei der Klassenansicht können die Lernverläufe aller SchülerInnen miteinander verglichen werden (Gebhardt et al., 2016b). Anschließend können Tests verschiedener Teilbereiche und Niveaustufen ausgewählt und neue Messzeitpunkte erstellt werden. Da in dieser Arbeit die Erhebung der Lesekompetenz in Niveaustufe vier im Vordergrund steht, wird nachfolgend ausführlicher auf diese Kompetenzstufe im Kontext des KLA eingegangen.

Die Lesekompetenz wird in Tests zur Leseflüssigkeit, also der Lesegeschwindigkeit sowie Lesegenauigkeit, und in Tests zum sinnentnehmendem Lesen eingeteilt (Gebhardt et al., 2016a). Die Leseflüssigkeitstests sind lehrerzentriert und als zufällig generierte Speedtests konstruiert (Dauer von einer Minute), wobei die Lehrkraft am Computer nach dem Vorlesen jedes Items entweder die Taste 1 für richtig oder die Taste 0 für falsch betätigt. Es handelt sich um das *Oral-Reading*, jedoch lesen die SchülerInnen keinen Text, sondern einzelne Silben bzw. Wörter und Pseudowörter laut vor. In der Konstruktion orientieren sich die drei Leseflüssigkeitstests sowohl am Lehrplan der Grundschule als auch am KLA (Gebhardt et al., 2016b). Levumi unterscheidet sich hinsichtlich der Unterteilungen der Niveaustufen und teilweise bei der Reihenfolge der Einführung neuer Buchstaben vom KLA. Nachfolgend werden zuerst die drei Tests zur Leseflüssigkeit und anschließend der Test zum sinnentnehmenden Lesen vorgestellt.

Die drei Tests zur Leseflüssigkeit, den Silben-, Wörter- und Pseudowörterlesetest, gibt es in jeweils vier bzw. fünf Niveaustufen. Die Tests basieren auf dem Aufbau und den Schwierigkeitsstufen des KLA. Die Niveaustufen N0 bis N4 unterscheiden sich hinsichtlich der verwendeten Buchstaben und der Komplexität der Struktur bei den verwendeten Silben und Wörtern (Gebhardt et al., 2016a). Die Niveaustufe 0 entspricht der Vorstufe des KLAs (vgl. Kapitel 2.4). Stufe 0 gibt es nur für die Tests Silben- und Pseudowörterlesen und verwendet die Buchstaben <m>, <l>, <a>, <e>, <i>, <o> und <u>. Niveaustufe 1 entspricht den Stufen 1 bis

2 des KLA und ergänzt die Buchstaben <m>, <r>, <s>, <n>, <f>, <l>. Niveaustufe 2 wird aufgeteilt in N2a und N2b. N2a entspricht den Stufen 4 bis 5 des KLA und ergänzt die Buchstaben <h>, <w>, <p>, <t> und <d>. Niveaustufe N2b entspricht den Stufen 6 bis 7 des KLA und führt die Buchstaben bzw. Graphemeinheiten <ch>, <k>, <b>, <sch>, und <g> ein. Die Stufen 8 bis 10 des KLA finden sich in der Niveaustufe 3 wieder. Hier werden ergänzend zu den bereits eingeführten Buchstaben nun <j>, <v>, <ß>, <sp>, <st> (ohne <tz> und <ck>), <z>, <qu>, <x>, <y> sowie <eu> und die Umlaute <ä>, <ö> und <ü> verwendet. Tests der Niveaustufe 4 gehören der höchsten Schwierigkeitsstufe an. Sie prüfen letztendlich alle Buchstaben ab und können den Stufen 11 bis 14 des KLA zugeordnet werden (Gebhardt et al., 2016a). Demzufolge beinhaltet der Test Silbenlesen (N4) Reibelaute, dehnbare Konsonanten und Plosive sowie die Vokale unter Verwendung offener Silben aus den vorherigen Stufen. Der Test Wörterlesen (N4) beinhaltet ergänzend zum Silbentest auch Wörter mit Kurzvokalen sowie alle drei Schwierigkeitsstufen hinsichtlich der Wortstruktur. Die komplexeste Wortstruktur ist demzufolge KVK-KVK. Der Silbentest gilt bei gleicher Niveaustufe verglichen mit dem Wörter- und dem Pseudowörtertest allein aufgrund der Itemlänge als einfacher (Jungjohann et al., 2017). Die Items des Pseudowörtertests ergeben sich aus denen des Silbentests. Sie sind aufgrund der offenen Silbenstruktur der mehrsilbigen Wörter ohne semantischen Inhalt, weswegen die SchülerInnen nicht auf das semantische Lexikon zurückgreifen können. Bei diesem Test wird die Rekodierfähigkeit abgefragt.

Der Test zum sinnentnehmenden Lesen ist schülerzentriert und erfordert die Anmeldung mit einem Schüleraccount. Jede(r) SchülerIn benötigt demzufolge einen eigenen Computer bzw. ein eigenes Tablet. Der Test zum sinnentnehmenden Lesen ist vergleichbar mit dem Prinzip *maze-selection* (vgl. Kapitel 4.3), jedoch handelt es sich ebenfalls nicht um Textabschnitte mit Auslassungen, sondern um einzelne Sätze, bei welchen ein Wort fehlt. Jeweils unter dem Satz mit der Lücke stehen vier Wörter zur Auswahl, wobei es sich bei drei von vier Wörtern um Distraktoren und bei dem anderen um das semantisch passende Wort handelt. Die Dauer des Tests beträgt ca. acht Minuten. Bei Niveaustufe 4 werden alle Buchstaben sowie der Sichtwortschatz verwendet, wobei die Wörter aus maximal zwei Silben bestehen. Die Sätze können hinsichtlich der Position, Satzanfang, -mitte oder -ende, der Wortart und der Funktion des ausgelassenen Wortes unterschieden werden. Bei dem fehlenden Wort kann es sich um Auslassungen der Argumentstruktur, Subjekt oder Objekt, der Prädikatstruktur, Verb oder Adjektiv, sowie um Auslassungen eines Konjunktors oder einer Präposition handeln. Verschiedene Verständnisseleistungen werden auf diese Weise getestet. Der Umfang des Wortschatzes ist insofern von Bedeutung, als dass dieser für das Erkennen des passenden Wor-

tes auf Grundlage des Kontextes notwendig ist. Darüber hinaus ist das Wissen um die Bedeutung eines einzelnen Wortes im Satz nicht ausreichend, um dessen Inhalt umfassend zu verstehen. Dafür müssen die gelesenen Wörter im Kontext aufeinander bezogen werden (Klicpera et al., 2017).

### 3.5.2 Förderansatz von Levumi

*Förderansätze im Lesen mit Levumi* ermöglichen die Verknüpfung von Diagnose und Förderung. Es ist zu betonen, dass der Förderansatz mitsamt dem Material kein umfassendes Förderkonzept darstellt, sondern dazu dient, Lehrkräfte hinsichtlich der Verwendung und Gestaltung von Fördermaterialien zu unterstützen (Jungjohann et al., 2017; Jungjohann & Gebhardt, 2018). Der Förderansatz beruht auf der Annahme, dass alle Kinder dieselben Phasen beim Schriftspracherwerb durchlaufen, jedoch in einem unterschiedlichen Tempo. Folgende fünf Förderbausteine liegen dem Ansatz zugrunde.



Abbildung 1 Förderbausteine und Entwicklungsbereiche LEVUMI (Jungjohann et al., 2017, S.9)

Der erste Förderbaustein beinhaltet die Förderung der Vorläuferfähigkeiten und bildet die Grundlage für den Erwerb nachfolgender Lesekompetenzen. Zentraler Aspekt dabei ist die *phonologische Bewusstheit*, welche unter anderem die Wahrnehmung der einzelnen Sequenzen der Lautsprache beinhaltet (Jungjohann et al., 2017). Beim zweiten Baustein steht die Buchstabe-Lautbeziehung im Fokus der Förderung. Bezüglich der *Graphem-Phonem-Korrespondenz* ist es zentral, dass Kinder erkennen, inwiefern einzelne Buchstaben(-kombinationen) (Grapheme) bestimmte Laute (Phoneme) repräsentieren. Ziel der intensiven Übungen dieses Bausteins ist die sichere und automatisierte Buchstabenidentifikation, -benennung und -kodierung. Aufgaben hinsichtlich des Findens von Anlauten oder Buchstabengitter

werden als sinnvoll betrachtet (Jungjohann et al., 2017). Der dritte Förderbaustein *richtiges Lesen* stellt das phonologische Rekodieren als indirekte Lesestrategie in den Fokus (vgl. Kapitel 2.2). Ziel dessen ist das Verbinden einzelner Phoneme zu ganzen Wörtern sowie die artikulatorische Verschmelzung dieser. In diesem Zusammenhang ist ebenso die Silbensegmentierung für ein flüssiges Lesen von Bedeutung. Buchstaben oder Silben zu einem Wort zu verbinden, das Finden der Anfangssilbe sowie das Malen von Silbenbögen unter verschiedene Wörter wird als sinnvolle Übung betrachtet (Jungjohann et al., 2017). Bei dem vierten Förderbaustein *Wörter kennen und erkennen* ist die automatisierte Worterkennung im Fokus. Das beinhaltet, dass ganze Wörter oder kleinere Einheiten im mentalen Lexikon vorhanden sind (vgl. Kapitel 2.2). Auf diese Weise können bekannte Wörter im Ganzen erkannt werden und ein phonologisches Rekodieren ist nicht mehr notwendig. Das automatisierte Lesen ermöglicht, dass der Fokus beim Lesen mehr auf den Inhalt selbst gelegt werden kann, was für das sinnentnehmende Lesen von besonderer Bedeutung ist. Baustein fünf *Inhalte Verstehen* zielt auf die Förderung des semantisch-lexikalischen Lesens ab, bei welchem der Sinn von Wörtern, Sätzen und Texten erfasst wird. Der Einsatz von Bildern in Übungsaufgaben zur Aktivierung des semantischen Wissens erweist sich als sinnvoll. Der sechste Förderbaustein *mit anderen Lesen* hat die Förderung der Leseflüssigkeit zum Ziel, welche als Brückenfunktion zwischen hierarchiehöheren und hierarchieniedrigen Prozessen fungiert (Jungjohann et al., 2017).

Durch Levumi wird zweierlei Material zur Verfügung gestellt. Einerseits werden fertige Arbeitsblätter, andererseits Vorlagen für Arbeitsblätter im Word-Format angeboten, wodurch sie an die Bedürfnisse der SchülerInnen angepasst werden können. Die Aufgabentypen sind den Niveaustufen der Lernverlaufsdagnostik von Levumi zugeordnet und bieten den Lehrkräften Orientierung (Jungjohann et al., 2017).

## 4 Fragestellung

Nachfolgend wird die Relevanz dieser Forschung im Kontext der Schul- und Unterrichtsentwicklung herausgearbeitet. Darüber hinaus werden die sich aus dem theoretischen Teil dieser Arbeit ergebenden Forschungsfragen dargelegt und das Forschungsinteresse wird begründet.

### 4.1 Darlegung der Relevanz der Fragestellung im Kontext von Schul- und Unterrichtsentwicklung

Die Bereiche Lernverlaufdiagnose, Förderung und das Geschlecht stellen allesamt sowohl Möglichkeiten als auch Herausforderungen für den schulischen Alltag und darüber hinaus für die Schul- und Unterrichtsentwicklung dar. Um sowohl den im Curriculum festgelegten Rahmenbedingungen als auch der heterogenen Schülerschaft gerecht zu werden, bietet die Lernverlaufdiagnose die Möglichkeit der zeitnahen Evaluation des eigenen Unterrichts in Form von regelmäßiger Diagnostik. Auf diese Weise können Schülerleistungen dokumentiert, beobachtet und analysiert werden. Dies dient der ggf. nötigen Anpassung der Didaktik und Methodik im Unterricht, ermöglicht jedoch auch eine frühzeitige Intervention durch zusätzliche und auf den oder die SchülerIn abgestimmte individuelle Förderung. Neben der Identifikation von leistungsschwachen SchülerInnen, möglicherweise auch von Non-Respondern, können Diagnose und Förderung miteinander verzahnt werden um eine höhere Qualität des Unterrichts zu ermöglichen. Für den Erwerb der Lesekompetenz ist das frühzeitige Erfassen dieser SchülerInnen von besonderer Bedeutung, da die wachsenden Fähigkeiten im Rekodieren, Dekodieren und ebenso im sinnentnehmenden Lesen, eine Basiskompetenz für Erfolge in allen Schulfächern sowie der gesamten Schullaufbahn darstellen. Im Rahmen dieser Forschung werden neben dem Leistungszuwachs auch geschlechterspezifische Unterschiede ausgewertet. Unterschiedliche Studien, welche Leistungsunterschiede zu Gunsten der Mädchen ergaben, machen Lehrkräfte auf ebendiese aufmerksam. Inwiefern und in welchem Ausmaß diese jedoch auch in der eigenen Klasse vorkommen, kann auf Grundlage dessen nicht vorhergesagt werden. Die Sensibilität für Geschlechterunterschiede im Lernen kann durch lernverlaufdiagnostische Verfahren spezifischer betrachtet werden. Im Rahmen dessen wurden in der vorliegenden Forschung Daten zur Lesekompetenz erhoben, um unter anderem zu untersuchen, in welchem Maß ein Leistungszuwachs während einer fünfwöchigen Förderung zu verzeichnen ist. Ebenso kann die Auswertung ggf. Aufschluss darüber geben, ob und in welcher Form auch bei dieser Stichprobe Unterschiede zwischen den Geschlechtern auftreten.

## 4.2 Darlegung der Fragstellungen

1. Inwiefern gibt es Unterschiede zwischen der ersten und letzten Messung hinsichtlich des Leistungszuwachses der SchülerInnen unter Berücksichtigung des Geschlechts?

Durch die Darstellung unterschiedlicher Forschungsergebnisse wurde deutlich, dass sich die Jungen in ihren Leseleistungen grundsätzlich unter denen der Mädchen befinden. Jungen gehören beispielsweise auch eher niedrigeren Kompetenzstufen an und lesen in ihrer Freizeit weniger gerne, was die Studien IGLU 2011 und KIM 2016 ergaben. Neben den Unterschieden in der Leistung ist jedoch auch ein Vergleich des Zuwachses von besonderem Interesse, welcher, falls vorhanden, auch die Leseleistung selbst im zeitlichen Verlauf prägt. Insofern ist im Rahmen dieser Forschung von Bedeutung, inwiefern sich die Ergebnisse von den im Kapitel 2.3.1 dargelegten unterscheiden. Darüber hinaus ergibt sich aus der, die Längsschnittstudie umrahmenden Diagnostik<sup>1</sup>, die Möglichkeit eine größere Grundgesamtheit über fast ein ganzes Schulhalbjahr hinsichtlich geschlechterspezifischer Unterschiede und ihres Leistungszuwachses betrachten zu können und mit den Ergebnissen der 18 SchülerInnen zu vergleichen

2. Inwieweit sind zu zwei Messzeitpunkten geschlechterspezifische Unterschiede zu erkennen?
  - 2a) Geschlechterspezifische Unterschiede zum Messzeitpunkt 11 vor Beginn der Förderung mit dem Unterricht als primären Einflussfaktor
  - 2b) Geschlechterspezifische Unterschiede zum Messzeitpunkt 17 nach der Förderung mit dieser als zusätzlichen Einflussfaktor

Bei dieser Forschungsfrage ist es zentral zu klären, ob es im Verlauf der Intervention zu Veränderungen der geschlechterspezifischen Leistungen kommt. Daraus ergibt sich der erste Teil der Forschungsfrage, welcher klären soll, inwiefern sich die Leseleistung zu einem Zeitpunkt vor der Förderung, mit dem Unterricht als einzigen Einflussfaktor, zwischen den Geschlechtern voneinander unterscheiden. Dieser Leistungsstand in den vier Tests dient als Grundlage für den zweiten Teil der Forschungsfrage. Durch die punktuelle Darstellung der Leistung nach bzw. am Ende der Förderung können auf Grundlage der Leistung vor der Förderung Veränderungen hinsichtlich der geschlechterspezifischen Leseleistung ausgemacht werden. Es stellt sich die Frage, inwiefern die Leistungen den Forschungsergebnissen zu

---

<sup>1</sup> Da die Levumi-Tests sowie die Erhebung der Lesekompetenz mit ELFE-II mehr als einmal durchgeführt wurden, ist der Begriff *Statusdiagnostik*, da eine Entwicklung erkennbar ist, nicht zutreffend.

Gunsten der Mädchen entsprechen oder ob sich die Ergebnisse zu den zwei Messzeitpunkten voneinander unterscheiden.

Darüber hinaus ist ebenfalls von Interesse, ob die ggf. vorhandenen Unterschiede denen der vorausgegangenen Forschung (Wenner, 2018) mit denselben SuS ähneln oder, ob es durch die Förderung zu anderen Tendenzen kommt. Es soll geklärt werden, ob Jungen hinsichtlich ihres Leistungszuwachses anders auf die Förderung reagieren als die Mädchen.

3. Inwieweit zeigen sich geschlechterspezifische Unterschiede im Lernverlauf während der Förderung über fünf Messzeitpunkte hinweg?

Forschungsfrage drei ist als Erweiterung der vorangegangenen zweigeteilten Forschungsfrage zu betrachten. Neben der Auswertung der punktuellen Leseleistung vor und nach der Förderung ist an dieser Stelle von Interesse, inwiefern auch im Lernverlauf über fünf Messzeitpunkte der B-Phase geschlechterspezifische Unterschiede vorhanden sind. Darüber hinaus findet eine Einordnung der SchülerInnen statt, indem deren Entwicklung unter dem Aspekt des Leistungsfortschritts bewertet wird. Damit verbunden ist die Benennung von Respondern und Non-Respondern in Anlehnung an die errechneten Mittelwerte bezüglich eines erwartungsgemäßen Zuwachses nach Fuchs et al. (1993) (vgl. Kapitel 3.2). Auch die intraindividuelle Heterogenität der Leistungen ist von Interesse, weswegen die Stabilität der Lernverläufe nach den Richtwerten von Kern untersucht werden (vgl. Kapitel 3.4). Auf diese Weise werden die SchülerInnen nicht mehr im arithmetischen Mittel, sondern auf individueller Ebene betrachtet. Die Ergebnisse sollen ebenfalls Aufschluss darüber geben, inwieweit eine auf Grundlage von Lernverlaufsdagnostik durchgeführte Förderung zu Leistungszuwachs führt.

4. Gibt es einen Zusammenhang zwischen den Leistungen der SchülerInnen bei den Leseflüssigkeitstests und dem Test zum sinnentnehmenden Lesen in Abhängigkeit zur Testschwierigkeit unter Berücksichtigung geschlechterspezifischer Unterschiede?

Auf Grundlage der in Kapitel 2.2 dargelegten Spezifika der Entwicklung von Leseflüssigkeit und dem sinnentnehmenden Lesen, ist davon auszugehen, dass es sich beim Silben-, Wörter- und Pseudowörterlesen um unterschiedlich geforderte Lesestrategien mit ansteigender Schwierigkeit handelt. Es ist von Interesse, inwieweit die steigende Schwierigkeit der drei Tests zur Leseflüssigkeit sich ebenfalls in den Ergebnissen der SchülerInnen widerspiegelt. Sofern es ebensolche Zusammenhänge gibt, ist von Relevanz, in welchem Ausmaß auch hier Geschlechterunterschiede zu verzeichnen sind. Das sinnentnehmende Lesen hingegen gilt als hierarchiehöherer Prozess verglichen mit den geforderten Fähigkeiten im Rahmen der Leseflüssigkeit. Auf Grundlage dessen kann erwartet werden, dass geringe Leistungen in der

Leseflüssigkeit auch mit ebensolchen beim sinnentnehmenden Lesen einhergehen. Geht man von einem engen Zusammenhang aus, so müssten umgekehrt SchülerInnen mit hohen Leistungen in den Tests zur Leseflüssigkeit auch dementsprechende Ergebnisse im sinnentnehmenden Lesen erzielen. Insgesamt sind demzufolge niedrigere Leistungen im sinnentnehmenden Test als in der Leseflüssigkeit auf Grund der Zuordnung zu Hierarchie höheren Prozessen zu erwarten.

## 5 Forschungsmethoden

Nachfolgend wird die dieser Forschung zugrunde liegende Stichprobe sowie die Methoden zur Auswahl dieser dargelegt. Anschließend wird die Durchführung der Tests mitsamt den Rahmenbedingungen und des zeitlichen Umfangs erläutert.

Die vorliegende Forschung wurde als Teil einer Längsschnittuntersuchung mit insgesamt sechs zweiten Klassen an drei verschiedenen Grundschulen durchgeführt. In der gesamten Studie wurde wöchentlich die Lesekompetenz von 32 SchülerInnen, je sechs aus jeder Klasse, mit drei Tests zur Leseflüssigkeit und einem zum sinnentnehmenden Lesen erhoben. Über den gesamten Zeitraum hinweg wurde die Lesekompetenz an 18 Messzeitpunkten über 21 Wochen gemessen. Dieser Zeitraum kommt durch Feiertage und Ferien zustande, währenddessen die Erhebungen nicht stattfinden konnten. Am ersten und letzten Termin wurden jeweils alle 126 SchülerInnen der sechs Klassen sowohl mit den Tests von Levumi als auch mit dem ELFE II, einem normierten Leseverständnistest, getestet. Gegen Ende des gesamten Erhebungszeitraums wurde mit den insgesamt 32 SchülerInnen zweimal pro Woche über fünf Wochen hinweg zusätzlich zur wöchentlichen Erhebung eine Leseförderung durchgeführt. Am letzten MZP fand erneut, wie schon zu Beginn, eine Erhebung mit Levumi und ELFE II mit allen Kindern der sechs Klassen statt.

### 5.1 Stichprobe

Um die 32 SchülerInnen auszuwählen, wurden die ersten beiden Untertests des ELFE II zum Wort- und Satzverständnis durchgeführt. Dabei handelt es sich um den Wort- und Satzverständnistest im *paper-pencil* Format. Dabei müssen die SchülerInnen zum einen Wörter den passenden Bildern zuordnen und zum anderen Sätze sinnvoll vervollständigen. Die Tests wurden mit Hilfe der Prozenträge ausgewertet, um die leseschwächsten SchülerInnen zu identifizieren. Für die regelmäßige Datenerhebung kamen SchülerInnen mit Leseleistungen in den unteren beiden Prozenträgen ( $PR \leq 25$ ) in Frage. Auch die durch Levumi erhobenen Daten sowie das Lehrerurteil wurde für die Auswahl der 32 SchülerInnen herangezogen. Besonders die Meinung der KlassenlehrerInnen wurde berücksichtigt, da diese die Leistungen ihrer SchülerInnen über einen langen Zeitraum beobachten konnten. Die schwächsten SchülerInnen der vier Levumi-Tests wurden ermittelt, indem geprüft wurde wie häufig ein Kind zu den sechs leistungsschwächsten Kindern der Klasse zählte. Auf Grundlage des ELFE II Tests, den Levumi-Daten sowie dem Lehrerurteil wurden letztendlich die 32 SchülerInnen ausgewählt. Auf eine ausführlichere Darstellung des Auswahlprozesses wird an dieser Stelle

verzichtet, da dieser nicht von der Studierenden selbst durchgeführt wurde. Eine solche Darstellung des Auswahlprozesses ist Wenner (2018) zu entnehmen. Die gesamte Studie wurde folgendermaßen auf vier Studierende aufgeteilt:

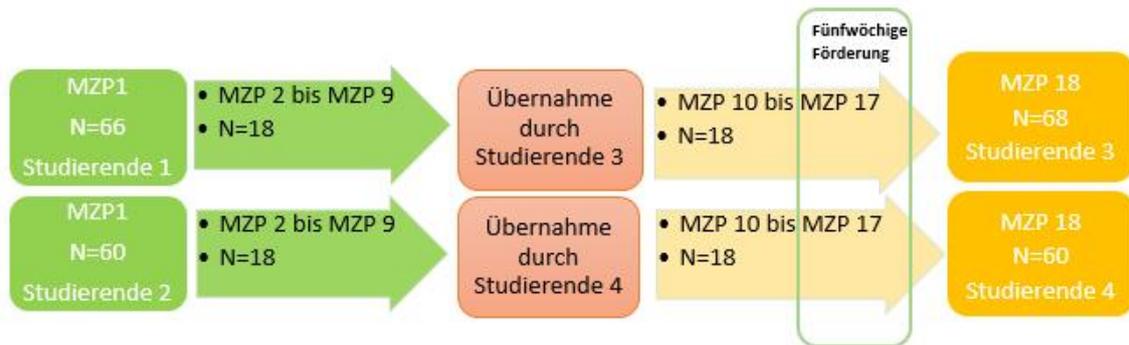


Abbildung 2 Ablauf der übergeordneten Studie (eigene Darstellung)

Insofern wurden im Rahmen der vorliegenden Forschung N=18 SchülerInnen, je sechs aus drei verschiedenen Klassen, getestet. Es handelt sich demnach um dieselben SchülerInnen, die auch von Studierender 1 im ersten Messzeitraum der Untersuchung getestet wurden. Von den 66 Schülerinnen besitzt die Mehrheit mit N=57 einen Migrationshintergrund. Die Stichprobe von N=66 setzt sich aus 40 Jungen und 26 Mädchen zusammen. Von den 18 SchülerInnen der wöchentlichen Erhebung sind 8 männlich und 10 weiblich. Da es zu unterschiedlichen Messzeitpunkten zu verschiedenen Schülerzahlen kommt, dient Tabelle 2 für die MZP 1 und 18 der Orientierung.

## 5.2 Testdurchführung

Jeden Montagvormittag wurden an zwei unterschiedlichen Bochumer Schulen nacheinander insgesamt 18 SchülerInnen mit vier verschiedenen Tests der Online-Plattform Levumi getestet. Dabei handelt es sich um die drei Tests zur Leseflüssigkeit, Silben-, Wörter- und Pseudowörterlesen, sowie um einen Test zum sinnentnehmenden Lesen, welche alle der Niveaustufe vier angehören. Die Erhebung der Lesekompetenz der sechs SchülerInnen der Klasse 2c wurde mit Hilfe von sechs Tablets, einem Laptop und einem Hotspot durchgeführt. Die insgesamt 12 SchülerInnen der Klassen 2e und 2d der anderen Schule konnten im schuleigenen Computerraum getestet werden. Pro Klasse dauerte die wöchentliche Erhebung ca. 45 Minuten. Bei der Erhebung der Klassen 2c und 2e waren jeweils gleichzeitig sechs Kinder im Raum. Ein Kind wurde am Laptop bzw. am Computer durch die Studierende getestet, während die anderen Aufgaben der Klassenleitung bearbeiteten oder malten. Die sechs SchülerInnen der Klasse 2d waren jeweils allein mit der Studierenden im Raum und wurden nacheinander getestet. Bei allen drei Klassen wurden erst die drei Tests zur Leseflüssigkeit,

in der Reihenfolge Silben-, Wörter- und Pseudowörterlesetest, durchgeführt und anschließend bearbeiteten alle SchülerInnen am eigenen Tablet bzw. Computer den Test zum sinnentnehmenden Lesen. Am letzten MZP 18 den 09.07. bzw. 11.07. wurden alle SchülerInnen der drei Klassen (N=68) sowohl mit den vier Levumi-Tests als auch mit ELFE II getestet (vgl. Abb. 3 Studierende 3). Insgesamt wurde über einen Zeitraum von zehn Wochen erhoben, wobei es auf Grund eines Feiertages zwischen MZP 11 und 12, an dem keine Daten erhoben wurden, zu neun Messzeitpunkten kommt.

### 5.3 Levumi-Förderung

Die fünf- bzw. sechswöchige<sup>2</sup> Förderung mit insgesamt zehn Fördersitzungen wurde nicht von der Studierenden selbst durchgeführt, weshalb dessen Effektivität nicht im Fokus dieser Arbeit steht. Aufgrund des Themas der vorliegenden Arbeit, hinsichtlich geschlechterspezifischer Unterschiede mit Bezug zur Förderung, wird diese nachfolgend der Vollständigkeit halber in Kürze dargestellt.

*Tabelle 1 Übersicht der Förderzeiträume nach Klassen*

Klasse	Zeiträume der Förderung		
	MZP vor der Förderung	Zeitraum Förderung	MZP nach der Förderung
2c	MZP 13 04.06.	05.06. bis 06.07.	MZP 18 11.07.
2d	MZP 11 14.05.	15.05. bis 29.06	MZP 17 02.07.
2e	MZP 11 14.05.	15.05. bis 29.06	MZP 17 02.07.

Das Leseflüssigkeitstraining *Levumi und Fredo auf Schatzsuche* basiert auf den in Kapitel 3.5.2 dargestellten Förderbausteinen und den in diesem Kontext empfohlenen Übungsformaten. Es umfasst derzeit zehn Sitzungen à 45 Minuten. Zentral bei diesem Training ist, dass die Übungen individualisierbar sind, wofür die Orientierung an den qualitativen und quantitativen Ergebnissen der Levumi-Lernverlaufdiagnostiktests von besonderer Bedeutung ist. Auf diese Weise können Fehlerschwerpunkte gezielt in der Förderung aufgegriffen und aufgearbeitet werden. Das Training beinhaltet Materialien wie Lesegeschichten, Aufgaben in Silbenschrift oder Bild-Wort-Zuordnungen und ist für Niveaustufe vier konzipiert. Die Lesegeschich-

<sup>2</sup> Alle Klassen wurden an derselben Anzahl von Sitzungen gefördert, jedoch kam es auf Grund von Feiertagen bei den Klassen 2d und 2e zu einem insgesamt längeren Zeitraum.

ten sind zentrale Elemente einer jeden Fördereinheit und sind inhaltlich miteinander verbunden. Das Lösen von Leserätseln im Kontext einer Schatzsuche umrahmt die Förderung. Die Geschichten sind optisch stets gleich strukturiert. Die Silben der Wörter sind farblich hervorgehoben, die Zeilen nummeriert und es befindet sich durchweg nur ein Satz in jeder Zeile. Bei den Namen der Figuren und Orte handelt es sich immer um zwei- bis dreisilbige Pseudowörter mit KKV-Silbenstruktur. Neben den Leserätseln werden Bodenbilder zur Geschichte gelegt, wobei unterschiedliches Material wie beispielsweise Seile, Figuren und Bilder aber auch Wörter verwendet werden, die aus einzelnen Silben zusammengesetzt werden müssen. Eine Fördereinheit beinhaltet sowohl konzentrierte Lese- und Arbeitsphasen als auch kooperatives Lautlesen und handlungsorientierte Phasen. Insgesamt wird durch die Förderung direkt an die vorausgegangene Lernverlaufdiagnose angeknüpft.

#### **5.4 Auswertungsmethoden**

Die bereits dargelegten Forschungsfragen werden auf unterschiedliche Art und Weise beantwortet. Grundsätzlich handelt es sich um eine deskriptive Darstellung der erhobenen Daten ohne die Durchführung statistischer Analysen, insbesondere da es sich primär um eine eher geringe Grundgesamtheit von  $N=18$  handelt. Lediglich im Rahmen der ersten Forschungsfrage wird ein t-Test durchgeführt, welcher sich wegen der größeren Stichprobe von  $N=66$  durch die rahmende Diagnostik anbietet. Alle Daten, bis auf die Messungen am MZP1, wurden von der Studierenden selbst erhoben. An dieser Stelle wurde auf die Daten der Studierenden 1 (vgl. Abbildung 2) im ersten Erhebungszeitraum der übergeordneten Forschung zurückgegriffen, wodurch erst der Vergleich der gesamten Stichprobe über 21 Wochen hinweg ermöglicht wird. Durch einen unabhängigen t-Test soll die Signifikanz des Unterschieds der Mittelwerte beider Geschlechter untersucht werden.

Im Rahmen der Forschungsfrage drei werden die Lernverläufe der 18 SchülerInnen geschlechtergetrennt betrachtet. Um Aussagen über die unterschiedlichen Verläufe treffen zu können, werden nachfolgend die Rahmenbedingungen festgelegt. Um die Lernverläufe hinsichtlich ihrer Stabilität auszuwerten, wird sich an den prozentualen Werten nach Kern (1997) orientiert. Damit ein Lernverlauf als stabil gilt sollten 80 bis 90% der Datenpunkte um nicht mehr als  $\pm 15\%$  vom Mittelwert abweichen (vgl. Kapitel 3.4). Es wurde sich dafür entschieden, die Einordnung auf Grundlage der prozentualen Werte durchzuführen, anstelle mit Hilfe der Anzahl richtig gelesener Items in Form von absoluten Zahlen. Diese Entscheidung gründet darauf, dass die Mittelwerte der SchülerInnen bezüglich der richtig gelesenen Items teilweise geringer als zehn sind und sich das daraus ergebene Intervall, im Vergleich zu Mittelwerten von über 30, stark verringern würde. Es käme zu Intervallspannen von bis zu  $\pm$  sechs

um den Mittelwert sowie zu ebensolchen von lediglich +/- eins. Die prozentualen Werte ermöglichen weitaus fairere Vergleiche hinsichtlich der Stabilität der Verlaufskurven.

Um die SchülerInnen dahingehend zu beurteilen, ob ihr Leistungszuwachs den Erwartungen entspricht und ob es sich damit verbunden um Responder oder Non-Responder handelt, müssen die Erwartungswerte vorweg festgelegt werden. Die im Rahmen der Forschung von Fuchs et al. (1993) erhobenen zu erwartenden Zuwachsraten beziehen sich auf die Messtechniken CBM-LL mit dem Indikator richtige Wörter pro Minute und CBM-Maze, welche in Kapitel 2.4 bereits dargestellt wurden. Bei den vorliegenden Daten handelt es sich jedoch nicht nur um den Wörterlesetest und den Test zum sinnentnehmenden Lesen, sondern ebenso um den Silben- und Pseudowörtertest. Um die Lernverläufe dennoch dahingehend bewerten zu können, ist eine Anpassung auf Grundlage dieser Werte an die vorhandenen Rahmenbedingungen unabdingbar. Obwohl eine direkte Übertragung der Werte nicht möglich ist, handelt es sich bezüglich der geringen Forschungslage mit zur Verfügung stehenden Vergleichswerten um eine passende Orientierung für die folgende Auswertung. Über einen Zeitraum von vier Wochen, ohne eine zusätzlich zum Unterricht stattfindende Förderung, wäre für den Wörtest ein Zuwachs von 5.84 Wörtern in der Jahrgangsstufe zwei zu erwarten. Mit Förderung läge dieser bei 8.6 Wörtern mehr pro Minute als vier Wochen zuvor. Für die Messtechnik CBM-Maze wäre über denselben Zeitraum ohne Förderung ein Zuwachs von 1.56 und mit Förderung von ca. 2.52 zu erwarten. Da es sich bei der zu Grunde liegenden Schülerschaft um die jeweils sechs leseschwächsten einer Klasse handelt und sich die CBM-Werte auf einen Klassendurchschnitt beziehen, ist ein deutlich geringerer wöchentlicher Zuwachs zu erwarten. Die Werte werden insofern an die Schülerschaft angepasst, als dass sie halbiert werden. Auch der SD-Zuwachs, welcher bei Vorhandensein einer Förderung pro Woche hinzuaddiert wird, wird halbiert. Da keine Vergleichswerte für den Silben- und Pseudowörtertest vorliegen, werden diese entsprechend ihrer Schwierigkeit ausgehend von den Werten des CBM-LL definiert. Im Silbentest sind demzufolge höhere Zuwachsraten zu erwarten als im Wörtest, da dieser als leichter gilt. Beim Pseudowörtertest ist ein geringerer Zuwachs als beim Wörtest zu erwarten, da dieser als schwieriger gilt. Wie vorab bereits erwähnt wurde, wird nach der Auswertung Bezug zu den Ergebnissen der Masterarbeit von Wenner (2018) genommen, bei welcher die Lesekompetenz derselben SchülerInnen erhoben wurde. Die Anpassung der Werte ist nachvollziehbar und wird aus Gründen der Vergleichbarkeit der Ergebnisse als Orientierung verwendet. Da im Zeitraum der Erhebung bei vorliegender Arbeit jedoch eine Förderung stattfand, wird in Anlehnung an Fuchs et al. (1993) nach Anpassung an die Schülerschaft eine halbe Standardabweichung zum erwarteten Zuwachs hinzuaddiert (Walter, 2010). Auf Grundlage dessen ergibt sich über einen Messzeitraum von

vier Wochen mit stattfindender Förderung ein erwarteter Zuwachs von 7 (6.6) Silben, 4 (4.3) Wörtern und 3 (3.12) Pseudowörtern. Liegen die Werte der SchülerInnen unter den genannten, so gelten sie als Non-Responder. Die Anpassung der CBM-Maze Werte von Fuchs et al. (1993) für den erwartungsgemäßen Zuwachs im sinnentnehmenden Lesen zeigt sich komplexer, da sich der durchschnittliche wöchentliche Zuwachs von Zweitklässlern auf einen Test mit einer Dauer von 2.5 Minuten bezieht. Der im Rahmen dieser Arbeit durchgeführte Test der Online-Plattform Levumi dauert jedoch acht Minuten. Auf Grundlage dessen wird der Erwartungswert halbiert und eine halbe Standardabweichung hinzuaddiert. Dieser Wert wird anschließend auf eine Dauer von 8 Minuten angepasst und mit der Anzahl der Wochen multipliziert. Daraus ergibt sich ein erwarteter Zuwachs von 4 (4.032) mehr richtig gelösten Items nach vier Wochen (vgl. Anhang F).

## 6 Auswertung und Ergebnisse

Nachfolgend werden die Ergebnisse der vorliegenden Studie dargestellt. Verschiedene Graphiken veranschaulichen die Ergebnisse. Teilweise werden weiterführende Daten im Anhang aufgeführt, welche im Text nur kurz beschrieben werden.

1. Inwiefern gibt es Unterschiede zwischen der ersten und letzten Messung hinsichtlich des Leistungszuwachses der SchülerInnen unter Berücksichtigung des Geschlechts?

Zur Beantwortung dieser Forschungsfrage werden die Daten des ersten und letzten MZP (MZP18) verwendet. Es wurde sich für die Nutzung dieser Daten entschieden, obwohl die des ersten MZP nicht von der Studierenden selbst erhoben wurden, da die Betrachtung dieser Daten eine Auswertung über fast ein ganzes Schulhalbjahr ermöglicht und mit N=66 eine größere Grundgesamtheit beinhaltet als die wöchentlichen Messungen mit N=18. Für die Auswertung des Leistungsfortschritts wird sich an der Lösungswahrscheinlichkeit (LW)<sup>3</sup>, den insgesamt bearbeiteten Items, den richtig und den falsch gelösten Items orientiert. Da insbesondere geschlechterspezifische Unterschiede hinsichtlich des Kompetenzzuwachses im Zeitraum von MZP1 bis MZP18 im Fokus stehen, wird der Zuwachs nach dem Geschlecht getrennt in Form von Balkendiagrammen dargestellt. Es folgen unterschiedliche Darstellungen, die nach den vier Tests aufgeteilt werden. Nach der rein deskriptiven Auswertung folgt ein t-Test zur Prüfung signifikanter Unterschiede zwischen den Geschlechtern zum jeweiligen Zeitpunkt mitsamt einer kurzen Darstellung der tatsächlichen Leistungen.

Tabelle 2 Übersicht der Stichprobengröße von MZP1 und MZP18

	Stichprobengröße pro Test von MZP1 und MZP18							
	Silben		Wörter		Pseudowörter		Sinnentnehmendes Lesen	
	J	M	J	M	J	M	J	M
Ge- schlecht								
MZP1	40	26	40	25	40	26	40	26
N gesamt	66		65		66		66	
MZP18	39	29	39	29	39	29	29	27

<sup>3</sup> Die Lösungswahrscheinlichkeit zeigt den prozentualen Anteil richtig gelesener Items der insgesamt bearbeiteten Items an. Auf diese Weise kann die Leseleistung von SchülerInnen mit unterschiedlichem Tempo und einer damit verbundenen höheren oder niedrigeren Anzahl an bearbeiteten Items miteinander verglichen werden.

N gesamt

68

68

68

56

Am MZP1 wurde die Datenerhebung mit N=66 begonnen und am MZP18 mit N=68 beendet (vgl. Tabelle 2). Die Erhöhung der Grundgesamtheit kommt durch Abwesenheit von Kindern zum ersten MZP oder einzelne Datenverluste durch Verbindungsschwierigkeiten mit dem Internet zustande. Der große Unterschied der Grundgesamtheit hinsichtlich des sinnentnehmenden Lesens am MZP18 (N=56 statt N=68) entstand durch einen unzureichenden Austausch mit einer Klassenleitung, woraus zeitliche Engpässe am Erhebungstag resultierten. Die unterschiedlichen Grundgesamtheiten je Test sind in Tabelle 2 dargestellt.

In Abbildung 3 ist der Leistungszuwachs, welcher sowohl in positiver als auch in negativer Form vorliegen kann, für den Silbentest dargestellt. Die gesamten Ergebnisse beider MZP können den Anhängen A und B entnommen werden, auf eine punktuelle Darstellung der einzelnen MZP wird auf Grund des Schwerpunktes der Forschungsfrage verzichtet.

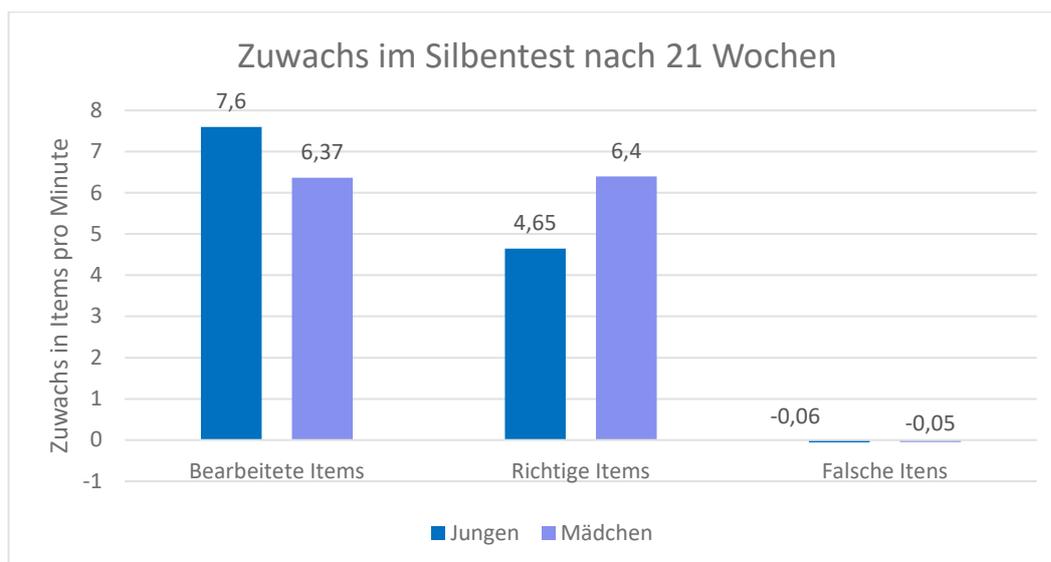


Abbildung 3 Leistungszuwachs der Jungen und Mädchen im Silbentest nach 21 Wochen

Die X-Achse zeigt jeweils einen Aspekt der Leseleistung. Auf der Y-Achse wird der Zuwachs an Items dargestellt, der sich auf alle Teile der Leseleistung bezieht. Beim Silbentest zeigt sich über einen Zeitraum von 18 MZP (21 Wochen) hinweg ein positiver Leistungszuwachs für beide Geschlechter, womit ein negativer Zuwachs bezüglich der Fehlerzahl verbunden ist. Da die Lösungswahrscheinlichkeit in Prozent ausgewertet wird und dementsprechend nicht dieselbe Einheit wie die anderen Bereiche beinhaltet, werden die Zuwächse in der Lösungswahrscheinlichkeit aller vier Tests geschlechtergetrennt in Abbildung 4 einheitlich dargestellt.

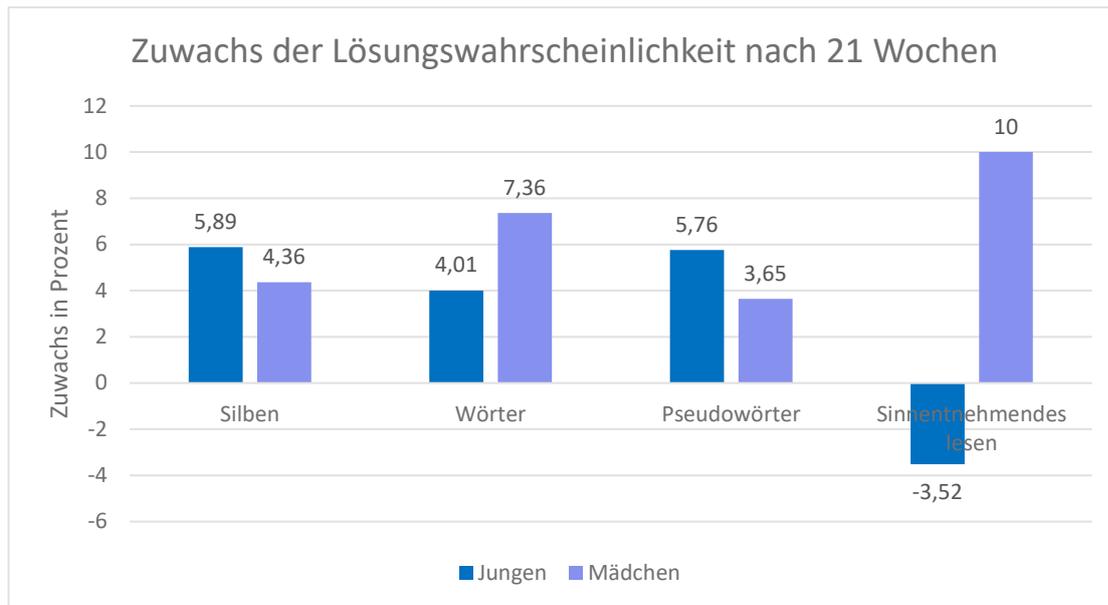


Abbildung 4 Zuwachs der Lösungswahrscheinlichkeit nach 21 Wochen

Die Jungen erhöhen ihre Lösungswahrscheinlichkeit im Mittel ( $M = +5.89\%$ ,  $SD^4 = 17.02$ ) um 1.58% mehr als die Mädchen ( $M = +4.31\%$ ,  $SD = 13.40$ ). Auch bei der Anzahl der bearbeiteten Items verzeichnen die Jungen mit 7.6 ( $SD = 9.15$ ) mehr gelesener Silben einen höheren Zuwachs als die Mädchen ( $M = 6.37$ ,  $SD = 9.52$ ). Hinsichtlich des Zuwachses der richtig gelösten Items lesen die Mädchen am MZP18 6.4 ( $SD = 10.95$ ) Silben mehr richtig als am MZP1 und die Jungen ( $M = 4.65$ ,  $SD = 11.37$ ) im Mittel 1.75 Silben weniger. Beide Geschlechter können ihre Fehlerzahl geringfügig um  $-0.06$  ( $SD(j^5) = 3.90$ ) bzw.  $-0.05$  ( $SD(m) = 3.33$ ) reduzieren. Die Jungen liegen demzufolge in ihrem Zuwachs bezüglich der Lösungswahrscheinlichkeit und den bearbeiteten Items vor den Mädchen. Die Mädchen zeigen einen größeren Zuwachs bei den richtigen Items.

<sup>4</sup> Die Standardabweichungen in Klammern beziehen sich auf den MZP18.

<sup>5</sup> Um eine einheitliche Schreibweise der Abkürzungen für die Geschlechter zu verwenden, steht  $j$  nachfolgen für Jungen und  $m$  für Mädchen.

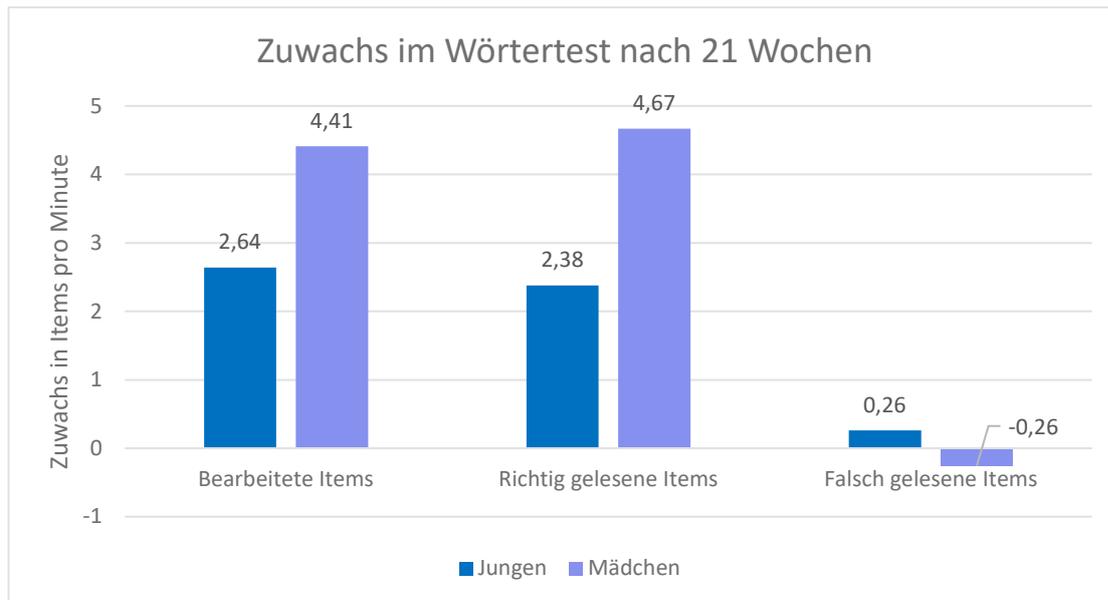


Abbildung 5 Leistungszuwachs der Jungen und Mädchen im Wörtertest nach 21 Wochen

Im Wörterlesetest zeigen sich die Ergebnisse, wie in Abbildung 5 deutlich wird, durchweg zum Vorteil der Mädchen. Auch an dieser Stelle ist zu betonen, dass bei beiden Geschlechtern in allen Bereichen, ausgenommen der falsch gelesenen Items, ein positiver Zuwachs vorhanden ist. In der Lösungswahrscheinlichkeit steigt die Leistung der Mädchen um 7.36% ( $SD = 22.70\%$ ), demzufolge um 3.35% mehr als bei den Jungen mit 4.01% ( $SD = 19.08\%$ ) (vgl. Abb. 4). Die Mädchen lasen am MZP18 im Mittel 4.41 ( $SD = 7.63$ ) Wörter mehr als am MZP1, die Jungen 2.64 ( $SD = 8.37$ ) Wörter mehr. Auch bezüglich der richtig gelesenen Wörter verbesserten sich die Mädchen um 4.67 ( $SD = 9.12$ ) und die Jungen um 2.38 ( $SD = 9.67$ ) Wörter. Bei den falschen Items kommt es zu gegensätzlichen Ergebnissen. Die Jungen erhöhten ihre Anzahl falscher Wörter geringfügig um 0.26 ( $SD = 4.18$ ) und die Mädchen reduzierten ihr Fehleranzahl im Mittel um -0.26 ( $SD = 3.72$ ). Im Wörterlesetest zeigt sich bei den Mädchen ein größerer, teils sogar ein fast doppelt so großer Leistungszuwachs in allen dargestellten Bereichen verglichen mit den Jungen.

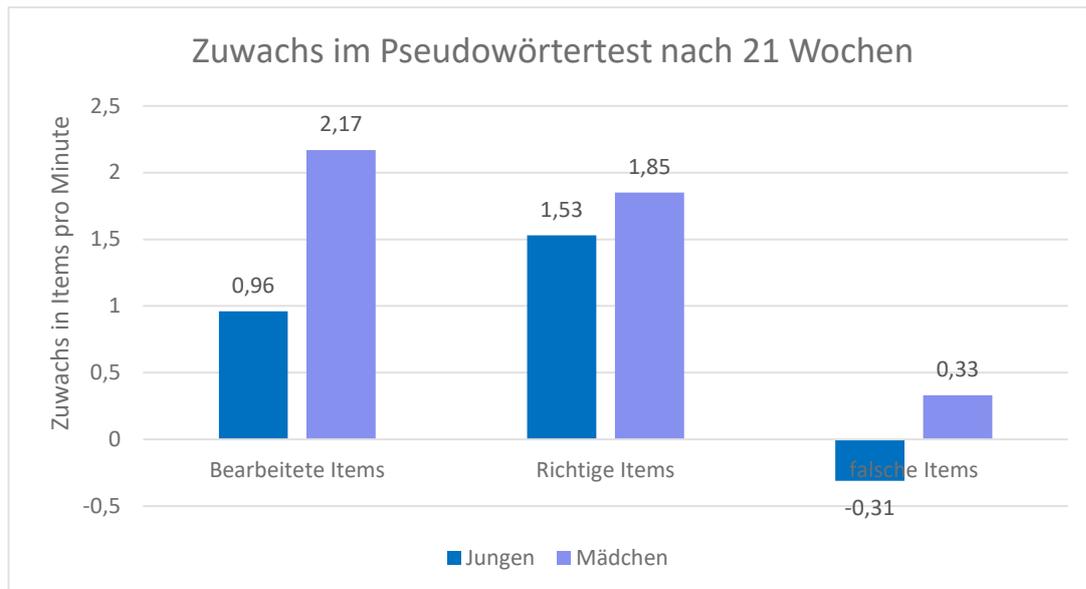


Abbildung 6 Leistungszuwachs der Jungen und Mädchen im Pseudowörtertest nach 21 Wochen

Beim Zuwachs im Pseudowörterlesetest zeigen sich deutlich andere geschlechterspezifische Ausprägungen (vgl. Abb. 6). Bei den Jungen ist ( $M = 5.76\%$ ,  $SD = 24.98\%$ ) ein größerer Zuwachs hinsichtlich der Lösungswahrscheinlichkeit als bei den Mädchen ( $M = 3.65\%$ ,  $SD = 24.44\%$ ) zu verzeichnen (vgl. Abb. 4). Die Mädchen bearbeiteten jedoch 2.17 ( $SD = 5.56$ ) Items mehr als am MZP1, wohingegen die Jungen im Mittel nur 0.96 ( $SD = 5.41$ ) mehr Items bearbeiteten. Bei den richtigen Items liegen die Mädchen mit ihrem Zuwachs von 1.85 ( $SD = 7.19$ ) geringfügig über dem der Jungen ( $M = 1.53$ ,  $SD = 6.43$ ). Bei den falschen Items zeigen sich gegensätzliche Ergebnisse. Die Jungen konnten die Anzahl um -0.31 ( $SD = 4.04$ ) reduzieren, wohingegen sie sich bei den Mädchen um 0.33 Items ( $SD = 3.47$ ) erhöhte. Beim Zuwachs der Leistungen im Pseudowörtertest zeigen sich bei der Betrachtung aller vier Bereiche keine eindeutigen Tendenzen zu Gunsten eines Geschlechts.

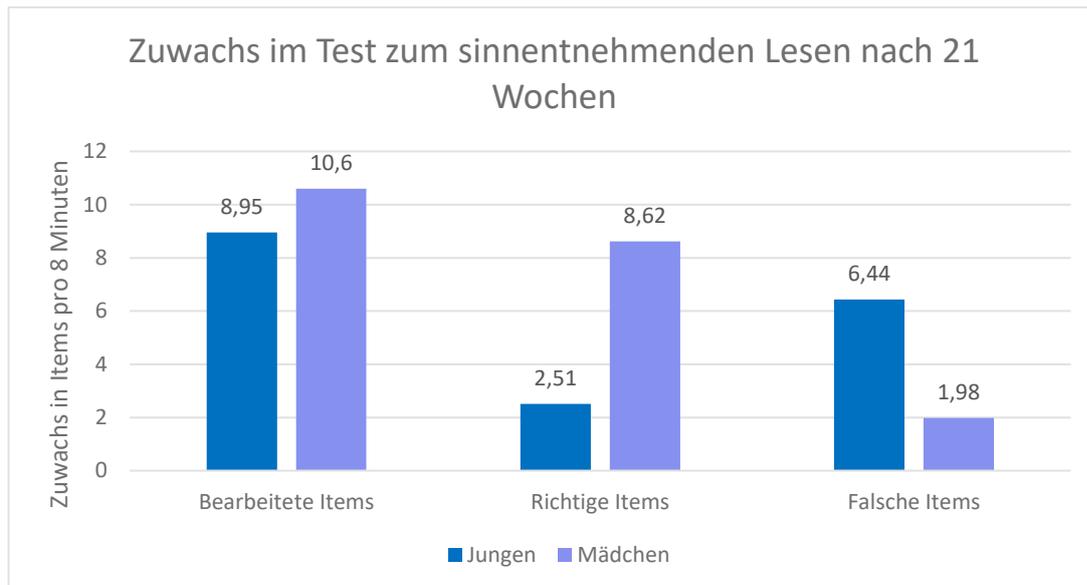


Abbildung 7 Leistungszuwachs der Jungen und Mädchen im Test zum sinnentnehmenden Lesen nach 21 Wochen

Im sinnentnehmenden Lesen zeigt sich grundsätzlich ein höherer Zuwachs für die Mädchen (vgl. Abb. 7). Die Jungen verringern ihre Lösungswahrscheinlichkeit im Mittel um -3.52% ( $SD = 23.12\%$ ) wohingegen sich die Mädchen um ganze 10% ( $SD = 26.3\%$ ) verbessern (vgl. Abb. 4). Beide Geschlechter erhöhen die Anzahl der gelesenen Items, die Jungen mit 8.95 ( $SD = 14.8$ ) und die Mädchen mit 10.6 mehr bearbeiteten Items ( $SD = 12.4$ ). Hinsichtlich der richtig gelösten Items ist bei den Mädchen mit 8.62 ( $SD = 10.19$ ) ein mehr als dreimal so großer Zuwachs als bei den Jungen mit 2.51 ( $SD = 9.48$ ) zu beobachten. Beide Geschlechter erhöhten die Anzahl der falsch gelösten Items. Dabei lösten die Mädchen im Mittel 1.98 ( $SD = 11.80$ ) Items mehr falsch als am MZP1, die Jungen bearbeiteten jedoch 6.44 ( $SD = 13.13$ ) Items mehr falsch. Die Mädchen zeigen in allen Bereichen einen deutlich höheren Leistungszuwachs im sinnentnehmenden Lesen als die Jungen. Sie zeigen dennoch beide einen positiven Zuwachs in der Fehleranzahl, welcher bei den Mädchen jedoch geringer ist.

Um neben der deskriptiven Darstellung des Leistungszuwachses auch die Signifikanz der geschlechterspezifischen Unterschiede bezüglich der Leistungsunterschiede, nicht des Zuwachses, zu prüfen wurde ein t-Test für beide Messzeitpunkte durchgeführt. Da der Wert der Lösungswahrscheinlichkeit die Leistungen der SchülerInnen, mitsamt richtig und falsch gelöster Items, am treffendsten beschreibt, werden nachfolgend ebendiese Werte beider MZP sowie die zentralen Ergebnisse der t-Tests betrachtet. Aufgrund der jedoch teils sehr ähnlichen Ergebnisse werden diese im Fließtext nur exemplarisch dargelegt und zusammengefasst ausgewertet. Die ausführlichen Ergebnisse der t-Tests können Anhang C entnommen werden.

Am MZP1 zeigen die Mädchen ( $M = 84.48\%$ ,  $SD = 16.70\%$ ) im Silbentest hinsichtlich der Lösungswahrscheinlichkeit höhere Werte als die Jungen ( $M = 80.57\%$ ,  $SD = 28.88\%$ ). Der Unterschied zwischen den Geschlechtern zeigt sich als nicht signifikant ( $t(64) = .618$ ,  $p = .539$ ), da der p-Wert weitaus über dem Grenzwert von 5% bzw. 0.05 liegt. Auch am MZP18 zeigen die Mädchen eine höhere Lösungswahrscheinlichkeit ( $M = 88.79\%$ ,  $SD = 13.40\%$ ) als die Jungen ( $M = 86.56\%$ ,  $SD = 17.01\%$ ). Auch hier ist der Unterschied nicht signifikant ( $t(66) = .611$ ,  $p = .543$ ). Im Mittel zeigen die Mädchen verglichen mit den Jungen zu MZP1 lediglich im Silbentest bessere Leistungen. Am MZP18 zeigen sich sowohl beim Silben- und Wörtertest als auch beim sinnentnehmenden Lesen höhere Lösungswahrscheinlichkeiten der Mädchen. Die Größe des Mittelwertunterschiedes ist den Werten der Spalte T im Anhang C zu entnehmen. Grundsätzlich liegen alle p-Werte deutlich über der Signifikanzgrenze von 0.05 bzw. 5%. Auch wenn der p-Wert am MZP18 beim sinnentnehmenden Lesen über dieser Grenze liegt, ist die Signifikanz der Differenz der Mittelwerte mit  $p = .106$  weitaus höher als bei den übrigen Ergebnissen, bei welchen der p-Wert im Mittel bei  $p = .638$  liegt. Da die Ergebnisse des t-Tests von der Höhe der Standardabweichung sowie von der Stichprobengröße abhängen, zeigt sich im Rahmen dieser Forschung eine nicht ausreichende Datenlage, um tatsächlich signifikante Unterschiede zwischen den Geschlechtern zu identifizieren.

2. Inwieweit sind zu zwei Messzeitpunkten geschlechterspezifische Unterschiede zu erkennen?

Die zweite Forschungsfrage zielt darauf ab zu prüfen, inwiefern sich die ggf. vorhandenen geschlechterspezifischen Unterschiede durch die fünfjährige Förderung verändert haben. Dafür werden sowohl die Leistungen des Messzeitpunktes vor Beginn der Förderung und am Ende der Förderung betrachtet. Bei Klasse 2d und 2e wären dies die MZP 11 und 17, bei Klasse 2c auf Grund des späteren Beginns der Förderung MZP 13 und 18. Da jedoch die geschlechterspezifische Auswertung im Durchschnitt aller drei Klassen von Bedeutung ist, wurde sich dafür entschieden auch bei Klasse 2c die MZP 11 und 17 zu betrachten. MZP11 erweist sich als sinnvolle Wahl, da der anschließende MZP auf Grund der Pfingstferien erst zwei Wochen später stattfand und dies ggf. Einfluss auf die Leistungen zu MZP12 hätte (vgl. Tabelle in 1). MZP17 liegt bei Klasse 2c zwar noch im Zeitraum der Förderung, dennoch wurden bereits acht Fördereinheiten durchgeführt. Darüber hinaus wurden zu MZP18 nicht alle der 18 SchülerInnen in allen Bereichen getestet (vgl. Tabelle 2). Die Anpassung hin zu einheitlichen MZP erweist sich für die Auswertung als zielführend. Bei beiden Teilfragen wird nicht wie in Forschungsfrage eins der Zuwachs, sondern die tatsächliche Leistung der beiden

Geschlechter miteinander verglichen. Dafür wird jeweils das arithmetische Mittel der insgesamt bearbeiteten Items, der richtig und der falsch gelösten Items graphisch dargestellt. Da die Lösungswahrscheinlichkeit in Prozent angegeben wird und somit nicht mit den anderen Werten innerhalb einer Graphik verglichen werden kann, wird diese nicht in den Abbildungen dargestellt. Diese und weitere Werte können dem Anhang entnommen werden (vgl. Anhang D und E). Die 18 wöchentlich getesteten SchülerInnen teilen sich in zehn Mädchen und acht Jungen auf. An beiden MZP weicht die Anzahl jedoch auf Grund von Abwesenheit weniger SchülerInnen ab.

2a) Geschlechterspezifische Unterschiede zum Messzeitpunkt elf vor Beginn der Förderung mit dem Unterricht als primären Einflussfaktor

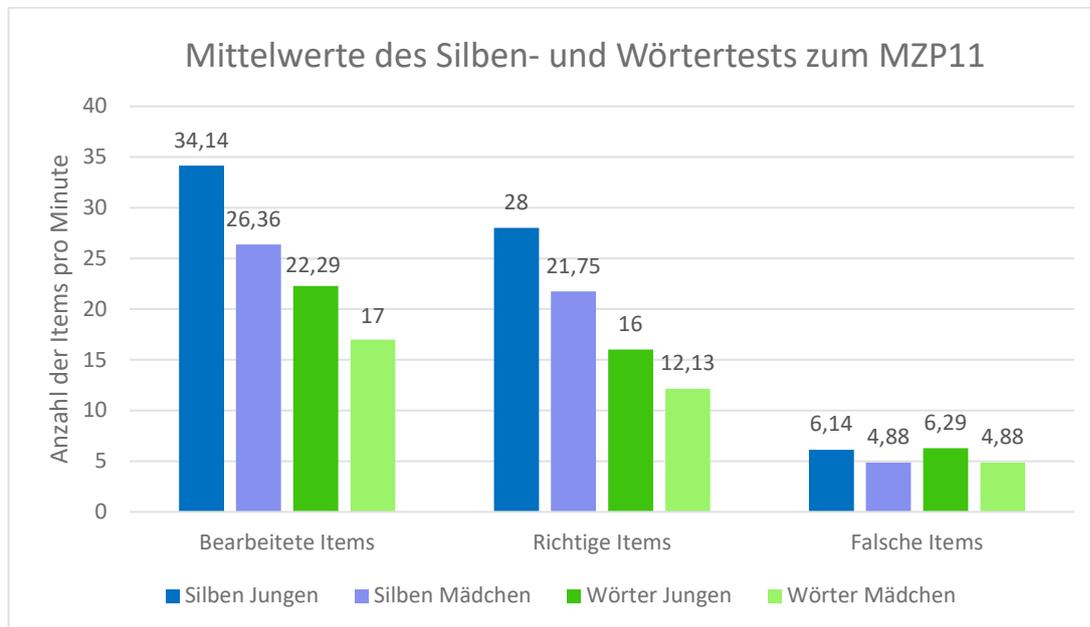


Abbildung 8 Mittelwerte der Jungen und Mädchen im Silben- und Wörterttest zum MZP11

In Abbildung 8 ist die Leistung der SchülerInnen geschlechtergetrennt im arithmetischen Mittel zusammen dargestellt. Die Zahlenwerte der bearbeiteten, richtigen und falschen Items sind 1:1 als Anzahl der Silben bzw. Wörter zu verstehen. Auf Grund der Sensibilität des arithmetischen Mittels für Ausreißer werden Median und Standardabweichung zusätzlich in tabellarischer Form ergänzend angegeben, um den Lesefluss nicht zu unterbrechen.

Tabelle 3 Übersicht zum Mittelwert, der Standardabweichung und des Medians des Silbentests am MZP11

	Silbentest							
	LW		Bearbeitete Items		Richtige Items		Falsche Items	
	J	M	J	M	J	M	J	M
M	82.27%	76.04%	34.14	26.36	28	21.75	6.14	4.88
SD	31.57%	20.07%	5.7	10.68	12.34	11.46	11.45	2.27
Md	92.86%	84.24%	36	28.5	26	23.5	2	4.5

Die Jungen zeigen am MZP11 im Silbentest bessere Leistungen als die Mädchen. Hinsichtlich der Lösungswahrscheinlichkeit liegen die Jungen mit 82.27% vor den Mädchen mit 76.04%. Ebenso bearbeiteten die Jungen mit 34.14 gelesenen Silben im Mittel 7.78 Silben mehr als die Mädchen mit 26.36. Davon lasen die Jungen mit 28 richtig gelesenen Silben 6.25 mehr richtig als die Mädchen mit 21.75. Lediglich bezüglich der Anzahl falsch gelesener Silben zeigen die Jungen geringfügig schwächere Leistungen. Die Jungen lasen im Mittel 6.14 Silben falsch und die Mädchen 4.88. An dieser Stelle ist jedoch die hohe Standardabweichung von 11.45 der Jungen zu erwähnen. Der Median fällt deutlich geringer aus als der Mittelwert der falsch gelesenen Wörter. Im Mittel zeigen die Jungen in den Bereichen Lösungswahrscheinlichkeit, Anzahl der bearbeiteten Items und Anzahl der richtig gelesenen Items bessere Leistungen als die Mädchen. Lediglich hinsichtlich der Fehleranzahl liegen die Mädchen vor den Jungen.

Tabelle 4 Übersicht zum Mittelwert, der Standardabweichung und des Medians des Wörkertests am MZP11

	Wörkertest							
	LW		Bearbeitete Items		Richtige Items		Falsche Items	
	J	M	J	M	J	M	J	M
M	73.84%	66.78%	22.3	17	16	12.1	6.3	4.88
SD	28.62%	22.29%	3.82	6.55	6.58	6.94	8.1	2.95
Md	82.6%	70.1%	23	17.5	17	12.5	4	4

Im Wörterlesetest zeigt sich eine ähnliche geschlechterspezifische Tendenz. Vorab ist jedoch zu erwähnen, dass sich die Leistungen insgesamt verglichen mit denen des Silbentests sichtbar verringern. Die Lösungswahrscheinlichkeit der Jungen mit 73.84% und der Mädchen mit

66.78% beinhaltet einen sehr ähnlichen Geschlechterunterschied mit 7.06% wie beim Silbentest. Die Jungen bearbeiteten im arithmetischen Mittel 22.3 Items und damit 5.3 mehr Items als die Mädchen mit 17. Darüber hinaus lösten die Jungen 16 Items richtig und damit 3.9 mehr als die Mädchen mit 12.1. Hinsichtlich der Fehleranzahl zeigen erneut die Mädchen, wie auch beim Silbentest, eine höhere Leistung mit 4.88, im Gegensatz zu den Jungen mit 6.3 falsch gelesenen Wörtern. Auch bezüglich der Fehlerzahl im Wörtertest zeigt sich erneut eine deutlich höhere Standardabweichung als bei den Mädchen. Im Mittel sind die Jungen den Mädchen im Wörterlesetest ebenfalls in allen Bereichen, außer der Fehleranzahl, überlegen.

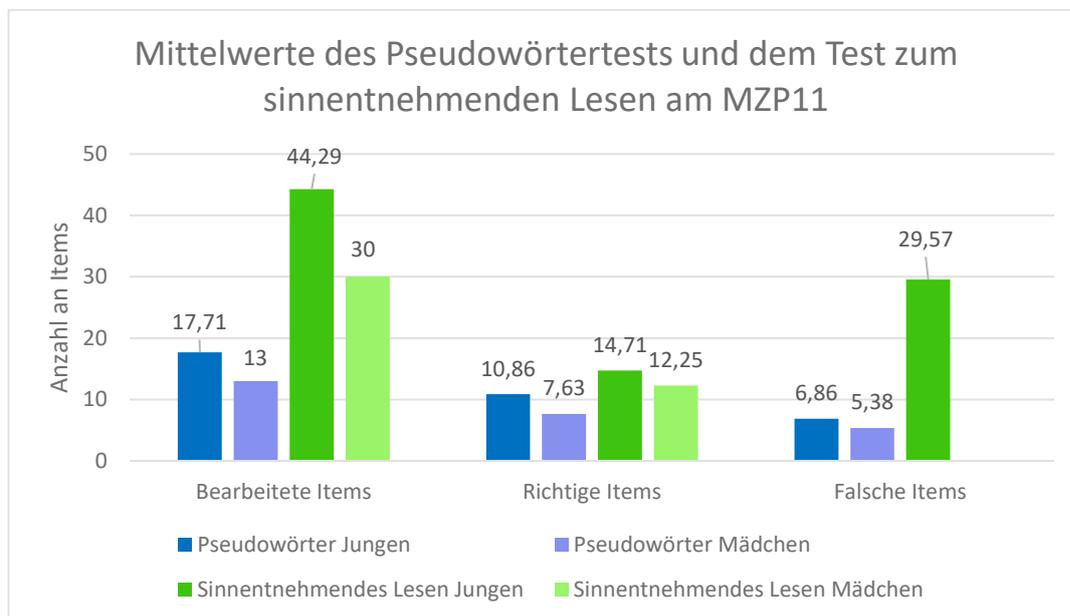


Abbildung 9 Mittelwerte der Jungen und Mädchen des Pseudowörtertests und des Tests zum sinnentnehmenden Lesen am MZP11

Bei dem Pseudowörtertest, dargestellt in Abbildung 9, zeigen sich insgesamt geringere Leistungen als beim Wörtertest. Darüber hinaus zeigen sich erneut Leistungsunterschiede zu Gunsten der Jungen. In Hinblick auf die Lösungswahrscheinlichkeit liegen die Jungen mit 66.04% vor den Mädchen mit 53.29%. Ebenso bearbeiteten die Jungen mit einer Anzahl von 17.71 Items 4.71 Items mehr als die Mädchen mit 13. Die Jungen lasen mit einer Anzahl von 10.86 3.23 mehr Items richtig als die Mädchen mit 7.63 richtigen Items. Wie auch beim Silben- und Wörtertest wurden von den Jungen mit einem Wert von 6.9 mehr Items falsch gelöst als von den Mädchen mit 5.4. Auch im Pseudowörtertest sind die Jungen im arithmetischen Mittel in allen Bereichen, außer der Fehleranzahl, den Mädchen überlegen.

Tabella 5 Übersicht zum Mittelwert, der Standardabweichung und des Medians des Pseudowörtertests am MZP11

	Pseudowörtertest							
	LW		Bearbeitete Items		Richtige Items		Falsche Items	
	J	M	J	M	J	M	J	M
M	66.04%	53.29%	17.71	13	10.86	7.63	6.9	5.4
SD	30.92%	32.41%	4.75	5.24	5.67	5.48	8.59	3.82
Md	71.4%	61.7%	17	12	14	8	4	5

Beim sinnentnehmenden Lesen zeigt sich bei den Jungen im Mittel eine Lösungswahrscheinlichkeit von 38.5% und bei den Mädchen von 44.9%. Die Jungen bearbeiteten 44.3 Items und die Mädchen mit 30 deutlich weniger. Die Jungen konnten 14.7 Items richtig bearbeiten, die Mädchen 12.3. Trotz der im Mittel besseren Leistungen der Jungen bezüglich der richtig gelösten Items, zeigen sich deutliche Unterschiede bei der Fehlerzahl. Die Jungen bearbeiteten 29.6 Items falsch, die Mädchen hingegen mit 17.8 deutlich weniger. Die grundsätzlich hohen Werte der Standardabweichungen zeigen eine große Streuung der Werte um das arithmetische Mittel, woraus sich sowohl bei den Jungen als auch bei den Mädchen eine gewisse Heterogenität zwischen den SchülerInnen ergibt (vgl. Tab. 6).

Tabella 6 Übersicht zum Mittelwert, der Standardabweichung und des Medians des Tests zum sinnentnehmenden Lesen am MZP11

	Sinnentnehmendes Lesen							
	LW		Bearbeitete Items		Richtige Items		Falsche Items	
	J	M	J	M	J	M	J	M
M	38.5%	44.9%	44.3	30	14.7	12.3	29.6	17.8
SD	16.2%	29.98%	20.2	14.72	6.2	9.2	18.37	15.08
Md	42.1%	28.9%	57	29	13	10	33	13.5

Bezüglich der Leistungen am MZP11 kann zusammengefasst werden, dass die Jungen im Mittel im Silben-, Wörter- und Pseudowörtertest stets bessere Leistungen in den Bereichen Lösungswahrscheinlichkeit, bearbeitete und richtig gelöste Items zeigen. Auf der anderen Seite sind jedoch durchweg schwächere Leistungen in der Fehleranzahl zu verzeichnen. Im Test zum sinnentnehmenden Lesen zeigen die Jungen in den Bereichen bearbeitete und richtig gelöste Items bessere Leistungen als die Mädchen, welche sich jedoch bei der Lösungswahrscheinlichkeit und der Fehleranzahl als überlegen herausstellen.

2b) Geschlechterunterschiede in der Leistung zum Messzeitpunkt 17 nach der Förderung mit dieser als zusätzlichen Einflussfaktor

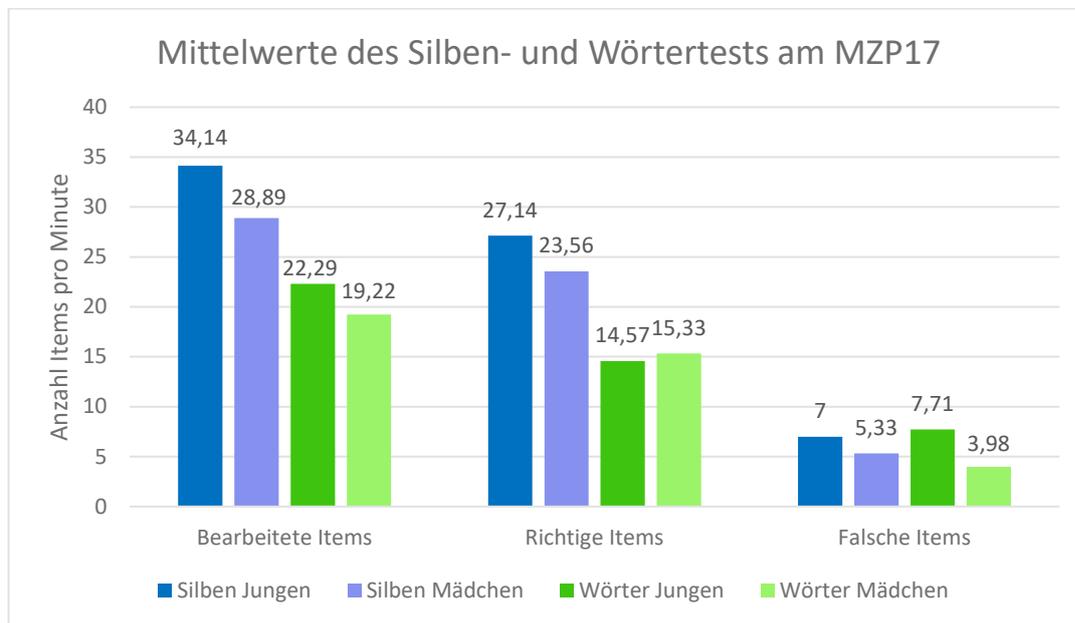


Abbildung 10 Mittelwerte der Jungen und Mädchen im Silben- und Wörterttest am MZP17

Die Leistungen der SchülerInnen zeigen hinsichtlich der geschlechterspezifischen Tendenzen verglichen mit dem MZP11 vor der Förderung sowohl ähnliche als auch unterschiedliche Ausprägungen (vgl. Abbildung 10). Bei der Lösungswahrscheinlichkeit des Silbentests zeigen die Mädchen mit 81% eine durchschnittlich bessere Leistung als die Jungen mit 75%. Die Jungen bearbeiteten jedoch insgesamt mehr Items, wovon sie ebenfalls mehr korrekt lösten als die Mädchen. Die Jungen bearbeiteten 34.1 und die Mädchen 28.9 Items. Davon wurden von Seiten der Jungen im arithmetischen Mittel 27.1 und von Seiten der Mädchen 23.6 Items richtig gelöst. Entgegen dieser Tendenz zu Gunsten der Jungen zeigen diese bezüglich der Fehleranzahl von 7 Items schwächere Leistungen als die Mädchen mit 5.3 falsch gelesenen Wörtern. Somit liegen die Jungen in den Bereichen bearbeitete und richtig gelöste Items vorne, die Mädchen jedoch in den Bereichen Lösungswahrscheinlichkeit und Fehlerzahl. Bei den Jungen sind erneut in allen Bereichen deutlich höhere Standardabweichungen zu verzeichnen (vgl. Tab.7).

Tabelle 7 Übersicht zum Mittelwert, der Standardabweichung und des Medians des Silbentests am MZP17

	Silbentest							
	LW		Bearbeitete Items		Richtige Items		Falsche Items	
	J	M	J	M	J	M	J	M
M	75%	81%	34.1	28.9	27.1	23.6	7	5.3
SD	28.22%	11.79%	9.63	3.55	12.89	5.22	7.59	2.92
Md	89.2%	84.9%	37	27	33	23	4	5

Auch im Wörterlesetest sind teilweise ähnliche Geschlechterunterschiede zu erkennen. Betreffend der Lösungswahrscheinlichkeit liegen erneut die Mädchen mit 79.1% vor den Jungen mit 60%. Der Unterschied zwischen den Mädchen und Jungen bezüglich der Anzahl bearbeiteter Items verringert sich verglichen mit den Ergebnissen des Silbentests. Die Jungen bearbeiteten 22.3 und die Mädchen 19.2 Items. Mit Bezug zu den richtig gelösten Items konnten die Mädchen die Jungen im Mittel geringfügig überholen. Die Mädchen lösten 15.3 und die Jungen 14.6 Items richtig.

Tabelle 8 Übersicht zum Mittelwert, der Standardabweichung und des Medians des Wörkertests am MZP17

	Wörkertest							
	LW		Bearbeitete Items		Richtige Items		Falsche Items	
	J	M	J	M	J	M	J	M
M	60%	79.1%	22.3	19.2	14.6	15.3	7.7	4
SD	34.38%	15.3%	6.63	4.92	8.28	5.57	6.16	2.92
Md	76.2%	84.6%	24	21	19	16	5	2

Die Standardabweichung der Jungen verweist jedoch auf heterogene Leistungen, die das arithmetische Mittel beeinflussen (vgl.Tab.8). Vor diesem Hintergrund ist der Vorsprung der Mädchen nicht als signifikant zu betrachten. Bezüglich der Fehleranzahl zeigen die Jungen nach wie vor schwächere Leistungen als die Mädchen. Die Jungen bearbeiteten 7.7 und die Mädchen 4 Items falsch. Demzufolge zeigen die Jungen beim Wörterlesetest bezüglich der Anzahl der bearbeiteten Items bessere Leistungen, die Mädchen jedoch in den Bereichen Lösungswahrscheinlichkeit, richtig gelöste und falsch gelöste Items.

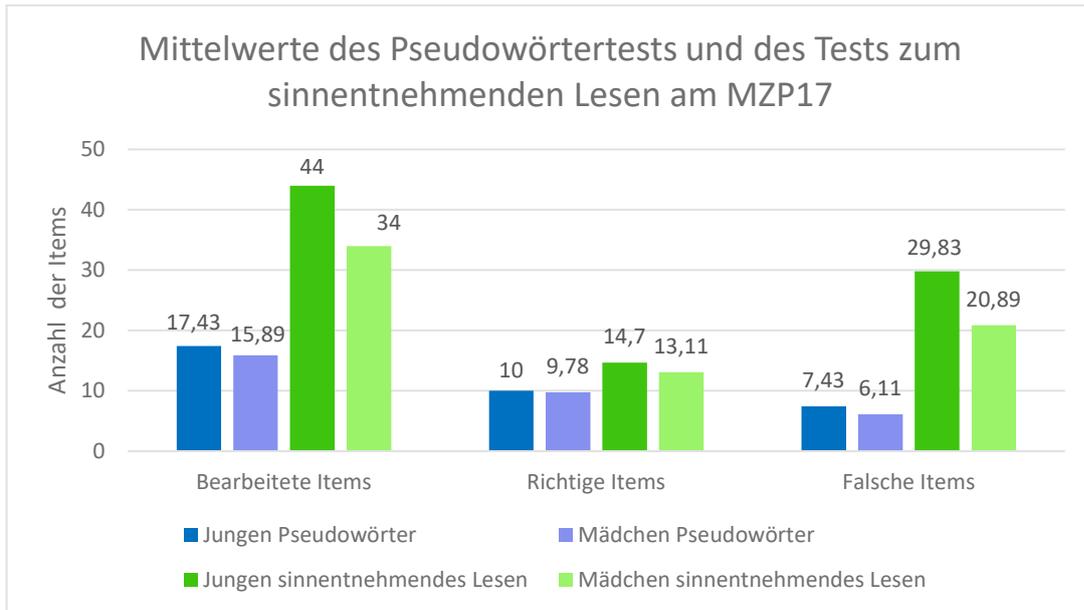


Abbildung 11 Mittelwerte der Jungen und Mädchen des Pseudowörtertests und des Tests zum sinnentnehmenden Lesen am MZP17

Im Pseudowörterlesetest zeigen sich im arithmetischen Mittel nur geringfügige Unterschiede zwischen den Geschlechtern (vgl. Abb. 11). Bei einer Lösungswahrscheinlichkeit der Jungen von 57.2% und bei den Mädchen von 59.2% ist kein großer Unterschied zu erkennen. Lediglich die höhere Standardabweichung der Jungen ist auffällig (vgl. Tab.9). Die Jungen bearbeiteten mit 17.4 geringfügig mehr Items als die Mädchen mit 15.9. Bezüglich der Anzahl richtig gelesener Wörter zeigt sich im Mittel nur ein Unterschied von weniger als einem Wort. Die Jungen lasen 10 und die Mädchen 9.8 Wörter richtig. Bezüglich der falsch gelesenen Wörter zeigen erneut die Mädchen mit einer Anzahl von 6.1 etwas bessere Leistungen als die Jungen mit 7.4 falsch gelesenen Wörtern. Insgesamt zeigen sich im Mittel zwischen den Jungen und Mädchen nur geringe Unterschiede. Die Jungen zeigen einen Vorsprung bezüglich der insgesamt bearbeiteten Items, die Mädchen hinsichtlich der Lösungswahrscheinlichkeit und der Fehleranzahl. Der Unterschied in der Anzahl richtig gelesener Items beträgt weniger als ein Wort, demzufolge ist kein signifikanter Unterschied zu erkennen.

Tabelle 9 Übersicht zum Mittelwert, der Standardabweichung und des Medians des Pseudowörtertests am MZP17

	Pseudowörtertest							
	LW		Bearbeitete Items		Richtige Items		Falsche Items	
	J	M	J	M	J	M	J	M
M	57.2%	59.2%	17.4	15.9	10	9.8	7.4	6.1
SD	40.51%	25.35%	6.45	3.41	6.86	5.22	8.94	3.22
Md	70%	71.4%	19	14	14	11	6	4

Im sinnentnehmenden Lesen sind erneut größere Geschlechterunterschiede zu verzeichnen. Die Jungen erreichen eine Lösungswahrscheinlichkeit von 39.6% und die Mädchen von 35.7%. Bezüglich der bearbeiteten Items liegen die Jungen mit 44 Items 10 vor den Mädchen mit 34. Trotz der höheren Zahl bearbeiteter Items lösten die Jungen mit 14.7 nur geringfügig mehr Items richtig als die Mädchen mit 13.11. Dieser eher geringe Unterschied ist mit einer höheren Fehlerzahl der Jungen zu erklären. Sie lösten 29.8 Items falsch, die Mädchen hingegen 20.9. Auch hier zeigt sich bei beiden Geschlechtern durch die hohen Standardabweichungen eine besondere Heterogenität zwischen den Leistungen (vgl. Tab. 10).

Tabelle 10 Übersicht zum Mittelwert, der Standardabweichung und des Medians des Tests zum sinnentnehmenden Lesen am MZP17

	Sinnentnehmendes Lesen							
	LW		Bearbeitete Items		Richtige Items		Falsche Items	
	J	M	J	M	J	M	J	M
M	39.6%	35.7%	44	34	14.7	13.11	29.8	20.9
SD	26.57%	18.54%	15.75	14.93	4.62	10.6	18.93	9.03
Md	33.7%	34.9%	43.5	27	16	11	29.5	21

Für den MZP17 nach der Förderung lässt sich zusammenfassen, dass die Jungen weiterhin im Silben-, Wörter- und Pseudowörtertest durchweg bessere Leistungen hinsichtlich der Anzahl bearbeiteter und richtig gelöster Items zeigen. In Bezug zur Lösungswahrscheinlichkeit und der Fehlerzahl zeigten sich jedoch die Mädchen als überlegen. Aus den Ergebnissen des Tests zum sinnentnehmenden Lesen ergibt sich ein geringer Vorsprung der Jungen in den Bereichen Lösungswahrscheinlichkeit und der richtig gelösten Items. Ein deutlicher Vorsprung zeigte sich in der Anzahl der bearbeiteten Items. Lediglich hinsichtlich der Fehleranzahl liegen die Mädchen vorne.

Vergleicht man beide MZP miteinander, so erlangt man zwar kein vollständig anderes Ergebnis am MZP17, dennoch zeigen sich deutliche Veränderungen in der Leistung zu Gunsten der Mädchen nach sechs Wochen. In den Leseflüssigkeitstests führen sie zwar nicht in jeweils allen vier jedoch in mindestens zwei oder gar drei Bereichen.

### 3. Inwieweit zeigen sich geschlechterspezifische Unterschiede im Lernverlauf während der Förderung über fünf Messzeitpunkte hinweg?

Zur Beantwortung der dargelegten Forschungsfrage wird die Leistung der SchülerInnen von MZP13 bis MZP17 dargestellt. Dieser Messzeitraum wurde gewählt, um die Leistungen der SchülerInnen der drei unterschiedlichen Klassen miteinander vergleichen zu können, auch wenn die Förderung bei Klasse 2c erst kurz nach der Messung und bei den Klassen 2e und 2d bereits vor diesem Zeitpunkt begann. Diese Verschiebung ist in gleichem Ausmaß bezüglich des MZP17 vorhanden (vgl. Tabelle 1). Im Sinne einer erleichterten Auswertung und der besseren Vergleichbarkeit wurde dieser Zeitraum dennoch als aussagekräftig bewertet.

Als Maßstab für den Vergleich des Lernverlaufs werden nachfolgend die Anzahl der richtig gelesenen Items sowie die Lösungswahrscheinlichkeit verwendet. Die Stabilität des Lernverlaufs wird anhand der prozentualen Angaben der Lösungswahrscheinlichkeit nach Kern (1993) bewertet, indem geprüft wird, ob mindestens 80% der Werte, der fünf dargestellten MZP, in dem jeweiligen Intervall  $\pm 15\%$  um den Mittelwert liegen. Bei fünf vorhandenen Werten darf sich demzufolge nur ein Wert (20% der Werte) außerhalb dieses Intervalls befinden. Sind diese Bedingungen erfüllt, gilt der Lernverlauf der SchülerInnen als stabil (vgl. Kapitel 5.4). Liegt mehr als ein Wert außerhalb des individuell festgelegten Bereichs, so gilt dieser als variabel.

Bezüglich des Zuwachses über die fünf MZP hinweg wird die Anzahl richtig gelöster Items verwendet. Diese werden nach Geschlecht und Art des Tests getrennt graphisch dargestellt. Auf diese Weise wird ein visuell schnell wahrnehmbarer Vergleich der Lernverläufe hinsichtlich eines Zuwachses an tatsächlich mehr richtig gelesenen Items ermöglicht. Auf eine zusätzliche graphische Darstellung der Lösungswahrscheinlichkeit wird aus Gründen der Leseflüssigkeit verzichtet. Darüber hinaus wäre ein Vergleich der tatsächlichen Leistung auf Grundlage des errechneten prozentualen Wertes ohne weitere Angaben nicht möglich. Durch den Zuwachs an richtig gelesenen Items kann durch die angepassten Vergleichswerte nach Fuchs et al. (1993) geprüft werden, inwiefern bei den 18 SchülerInnen ein erwartungsgemäßer Zuwachs zu verzeichnen ist (vgl. Anhang F). Ist dies der Fall, so gelten die SchülerInnen als Responder. Wird dieser Zuwachs nicht erreicht, handelt es sich um Non-Responder. Die

genauen Werte der richtig gelesenen Items sowie die Angaben zur Lösungswahrscheinlichkeit können dem Anhang entnommen werden (vgl. Anhänge G-N). Auf Grundlage dieser Einordnung können anschließend geschlechterspezifische Unterschiede im Lernverlauf dargestellt werden. Die zwei folgenden Abbildungen (Abb. 12 & 13) zeigen pro MZP und Test, geschlechtergetrennt das arithmetische Mittel der richtig gelösten Items. In Abbildung 12 wird der mittlere Lernverlauf des Silben- und des Wörterlesetests über fünf MZP hinweg dargestellt.

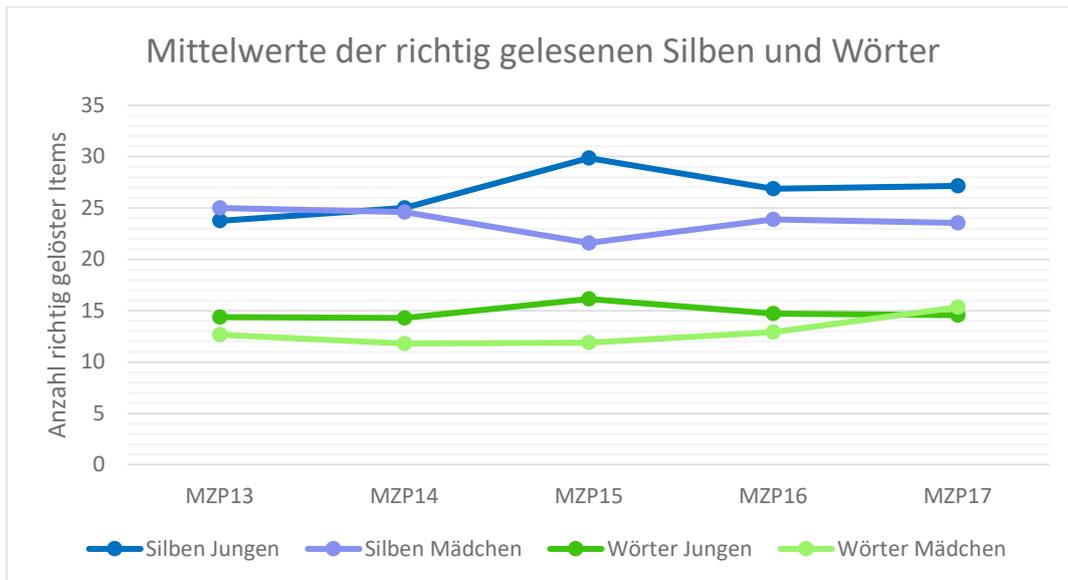


Abbildung 12 Mittelwerte der richtig gelesenen Silben und Wörter über fünf MZP hinweg

Im Mittel zeigen die Jungen beim Silbentest am MZP13 minimal schwächere Leistungen mit einem Unterschied von 1.25 weniger richtig gelesenen Silben als die Mädchen. Anschließend zeigen sich die Jungen am MZP14 minimal stärker mit einem Vorsprung von 0.4 mehr richtig gelesenen Silben. Der mittlere Lernverlauf der Jungen verläuft entgegengesetzt zu dem der Mädchen. Während die Jungen sich um fast fünf Silben verbessern, verringert sich die Anzahl bei den Mädchen um drei Silben. Die Verläufe nähern sich bis auf einen Unterschied von drei Silben am MZP16 an. An MZP17 erhöht sich der Unterschied leicht auf 3.5 Silben zu Gunsten der Mädchen. Im Mittel ist demzufolge über fünf MZP für die Jungen ein positiver Zuwachs von 3.39 und für die Mädchen ein negativer Zuwachs von -1.44 mehr bzw. weniger richtig gelesenen Silben zu verzeichnen. Die Lernverläufe der SchülerInnen bezüglich des Wörterlesetests zeigen ähnliche Ausprägungen wie im Silbentest, beinhalten jedoch abweichende Tendenzen und Unterschiede zwischen den Geschlechtern. Zu Beginn des Zeitraums zeigen sich die Jungen als minimal stärker mit einem Vorsprung von 1.71 mehr richtig gelesenen Wörtern. Die Werte beider Geschlechter sinken geringfügig am MZP14, die der Jungen steigen jedoch zum MZP15 mit einem Zuwachs von fast zwei Wörtern an, wobei sich der mittlere

Wert der Mädchen nur geringfügig verändert. Zu diesem MZP ist der größte Unterschied mit 4.24 mehr richtig gelesenen Wörtern bei den Jungen zu verzeichnen. Daran anschließend nähern sich die Graphen erneut an, wobei es zu einem Unterschied von 1.81 Wörtern kommt. Zum letzten MZP zeigen jedoch die Mädchen bessere Leistungen, auch wenn es sich nur um einen Vorsprung von 0.58 mehr richtig gelesenen Wörtern im Mittel handelt. Insgesamt können sich die Jungen im Mittel nicht verbessern, sondern lediglich verschlechtern, auch wenn es sich nur um einen geringen negativen Zuwachs von -0.19 handelt. Die Mädchen konnten ihre Leistungen über die fünf MZP um 2.66 Items verbessern.

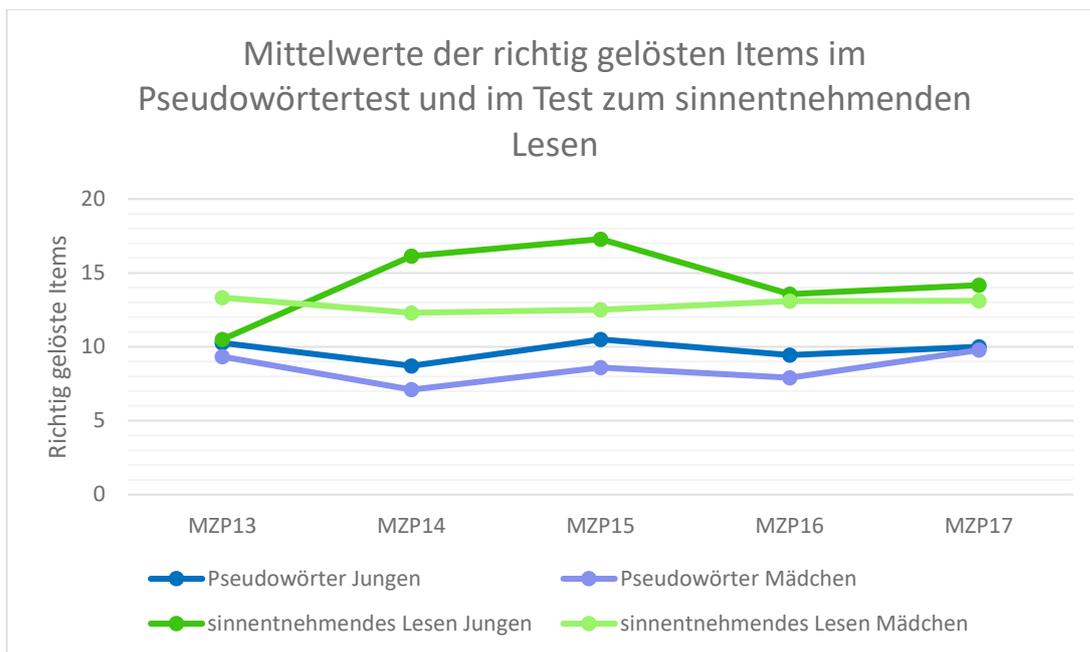


Abbildung 13 Mittelwerte der Jungen und Mädchen der richtig gelösten Items im Pseudowörtertest und im Test zum sinnentnehmenden Lesen über fünf MZP

Im Pseudowörterlesetest zeigen sich geringe, jedoch kontinuierliche Unterschiede zu Gunsten der Jungen, welche in Abbildung 13 zu erkennen sind. Es ist zu betonen, dass die Werte der beiden dargestellten Tests nicht miteinander verglichen werden können, da sich die Anzahl der richtig gelösten Items beim Pseudowörtertest auf eine Dauer von einer Minute und beim sinnentnehmenden Lesen auf acht Minuten bezieht. Darüber hinaus beinhaltet ein Item beim sinnentnehmenden Lesen immer mehr als ein Wort, da es sich um ganze Sätze handelt.

Am MZP13 zeigen die Jungen im Pseudowörtertest mit einem Abstand von etwas mehr als einem mehr richtig gelesenen Pseudowort einen geringen Vorsprung. Dieser bleibt den Jungen über weitere drei MZP mit einem Unterscheid von ca. einem, zwei sowie ca. 1.5 mehr richtig gelesenen Pseudowörtern gegenüber den Mädchen erhalten. Am MZP17 nähern sich die mittleren Werte der beiden Geschlechter jedoch an, wodurch die Differenz auf ca. ein

Wort sinkt. Dennoch ist dieser Vorsprung auf Seiten der Jungen zu verzeichnen, wobei diese im Mittel mit einer leichten Verschlechterung und die Mädchen mit einer Verbesserung in Höhe von 0.45 aus der Förderung herausgehen.

Im sinnentnehmenden Lesen zeigen sich hingegen deutlich größere Geschlechterunterschiede, insbesondere an MZP14 und 15. Am MZP13 führen die Mädchen mit fast drei mehr richtig gelösten Items. Während dieser Wert zu MZP14 um ein Item sinkt, steigt die Anzahl der Jungen verglichen mit der vorherigen Anzahl um knapp sechs Items, wodurch ein vergleichsweise großer Geschlechterunterschied von fast vier Items zustande kommt. Dieser Unterschied steigt bis zu MZP15 auf fast fünf Items an. Es folgt eine Annäherung der beiden Lernverläufe im Mittel, wodurch der Unterschied auf weniger als 0.5 mehr richtig gelöste Items zu Gunsten der Jungen sinkt. Zum MZP17 zeigt sich ein Unterschied von ca. einem mehr richtig gelösten Item für die Jungen. Insgesamt konnten sich die Jungen um fast vier Items verbessern. Bei den Mädchen ist hingegen ein negativer Zuwachs von weniger als -0.2 entstanden. Bei den Jungen und Mädchen konnten in allen vier Tests nur geringe Zuwachsraten von MZP13 bis zu MZP17 im arithmetischen Mittel gemessen werden. Inwiefern die Lernverläufe auf individueller Ebene der SchülerInnen andere Geschlechtertendenzen aufzeigen wird nachfolgend untersucht.

Da bei den wöchentlichen Erhebungen die Lesekompetenz von jeweils 18 SchülerInnen getestet wurde, werden pro Test ebenfalls 18 Lernverläufe betrachtet. Diese werden nach dem Geschlecht getrennt, wodurch entweder zehn oder acht Lernverläufe pro Abbildung dargestellt werden. Trotz der hohen Anzahl an Lernverläufen, insbesondere bei den Mädchen, und einer damit verbundenen verhältnismäßig komplexen Abbildung, wurde sich dazu entschieden, alle Mädchen und alle Jungen jeweils in einer Abbildung darzustellen, anstatt diese aufzuteilen. Auf diese Weise wird ein visuell interindividueller Leistungsvergleich innerhalb einer Abbildung ermöglicht. Zunächst werden die Lernverläufe pro Abbildung hinsichtlich ihrer Stabilität, anschließend bezüglich des erwartungsgemäßen Zuwachses untersucht (vgl. Anhang G-N).

Im Silbentest zeigen die Mädchen auf Grundlage der prozentualen Angaben der Lösungswahrscheinlichkeit über fünf MZP hinweg grundsätzlich stabile Lernverläufe. Die Werte von sieben der zehn Mädchen liegen alle in einem Bereich von 15% um den individuellen Mittelwert. Der Lernverlauf von Schülerin Ta zeigt eine Abweichung, gilt dennoch als stabil (vgl. Anhang K). Lediglich die Lernverläufe von Jo und Ma gelten mit jeweils zwei Abweichungen als variabel. Inwiefern es sich im Silbentest um (Non-) Responder handelt, kann mit der folgenden Abbildung 14 verdeutlicht werden.

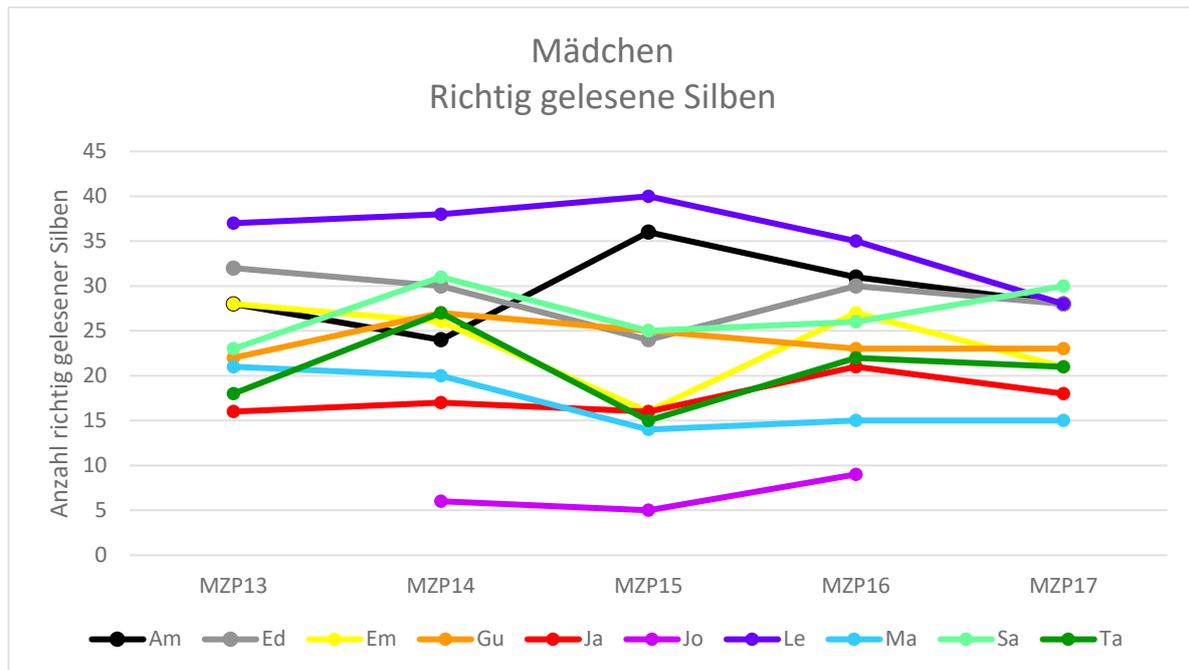


Abbildung 14 Anzahl der richtig gelesenen Silben über fünf MZP der Mädchen

Auf eine Wertebeschriftung der Abbildung 14 wurde aus Gründen der Übersichtlichkeit verzichtet, diese sind jedoch allesamt dem Anhang G zu entnehmen. Es wird deutlich, dass die Mädchen sehr heterogene Leistungen zeigen. Bei der Einordnung, ob es sich bei einer Schülerin oder einem Schüler um Responder oder Non-Responder handelt, wird jedoch primär auf den Zuwachs von MZP13 zu MZP17 geachtet. Für den Silbentest wird ein Zuwachs von sieben mehr richtig gelesenen Items in vier Wochen erwartet. Lediglich Schülerin Sa konnte diesen Wert erreichen und gilt als Responder. Als Non-Responder, jedoch trotzdem mit einem positiven Zuwachs gelten Gu und Ja (+2) sowie Jo und Ta (+3). Schülerin Am las an beiden MZP gleich viele Silben richtig, zeigte jedoch einen Leistungsanstieg an MZP15 und 16. Die Schülerinnen Em, Le und Ma hingegen zeigten sogar negative Zuwächse mit -7, -9 und -6. Im Silbentest gelten über den definierten Zeitraum somit zwei von zehn Lernverläufen als variabel und nur Schülerin Sa gilt als Responder.

Unterbrechungen im Lernverlauf wie in Abbildung 15 kommen durch Abwesenheit der Schüler zum jeweiligen MZP zustande. Im Silbenlesetest der Jungen konnten auf Grundlage der Lösungswahrscheinlichkeit bei zwei Schülern variable Lernverläufe ausgemacht werden. Sechs der acht Lernverläufe gelten mit null Abweichungen vom Intervall von +/- 15% um den Mittelwert als stabil. Lediglich die zwei leistungsschwächsten Jungen Ra und Ti zeigten Abweichungen von fünf und drei Werten. In Anbetracht der Tatsache, dass es sich nur um fünf

MZP handelt, ist die Zahl der Abweichungen als hoch zu bezeichnen. Im Kontext der Förderung handelt es sich jedoch weniger um variable, sondern eher um steigende Lernverläufe, wodurch die variablen Werte auch durch einen hohen Zuwachs zustande kommen können (vgl. Anhang K).

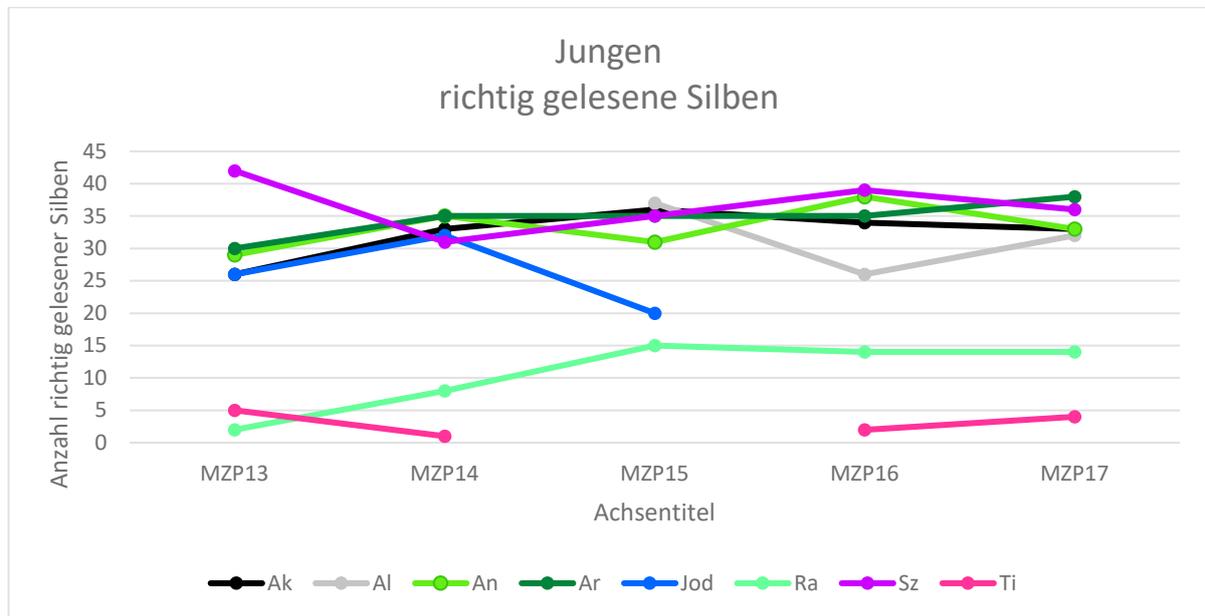


Abbildung 15 Anzahl der richtig gelesenen Silben über fünf MZP der Jungen

In Abbildung 15 wird deutlich, dass ebendiese zwei Jungen mit variablen Lernverläufen, verglichen mit den restlichen sechs Schülern, gleichzeitig auch die schwächsten Leistungen hinsichtlich der Anzahl richtig gelesener Silben erbringen. Betrachtet man den individuellen Zuwachs über vier Wochen, so können die drei Schüler Ak, Ar und Ra mit sieben, acht und 12 mehr richtig gelesenen Silben als Responder betrachtet werden. Die übrigen fünf Schüler gelten als Non-Responder, wobei bei Schüler Al und An dennoch ein positiver Zuwachs von drei und vier Silben zu verzeichnen ist. Die Schüler Jod, Sz und Ti zeigen negative Zuwächse von -6, -6 und -1. Im Silbentest gelten bei den Jungen demzufolge zwei von acht Lernverläufen als variabel und drei Schüler gelten als Responder.

Im Wörterlesetest der Mädchen zeigt sich eine gewisse intra- und interindividuelle Heterogenität der Lösungswahrscheinlichkeit. Lediglich die Lernverläufe von Ed und Em mit null Abweichungen und Le mit einer Abweichung gelten als stabil. Die Lernverläufe von Am, Gu und Sa gelten mit jeweils zwei Abweichungen, die Verläufe von Ja und Ma mit vier sowie die von Jo und Ta mit drei Abweichungen als variabel (vgl. Anhang L).

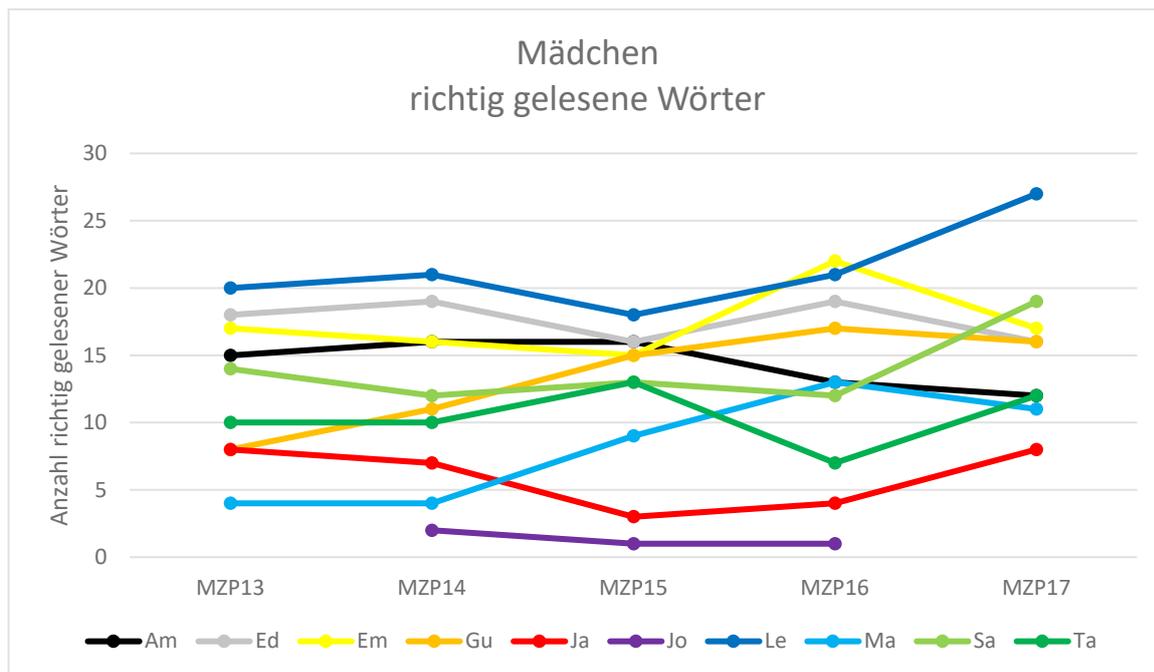


Abbildung 16 Anzahl der richtig gelesenen Wörter über fünf MZP der Mädchen

Abbildung 16 zeigt die Leistungsheterogenität zwischen den Mädchen hinsichtlich der richtig gelesenen Wörter. Um im Wörkertest über vier Wochen als Responder zu gelten, wurde ein Mindestzuwachs von vier mehr richtig gelesenen Wörtern festgelegt. Vier der zehn Schülerinnen konnten diesen Zuwachs erreichen. Dabei handelt es sich um Gu (+8), Le und Ma (+7) sowie um Sa (+5). Die Schülerinnen Em und Ja gelten aufgrund keiner Veränderung als Non-Responder. Schülerin Ta zeigte dennoch einen Zuwachs von +2. Am MZP17 zeigten Am, Ed und Jo mit -3, -2 und -1 schwächere Leistungen verglichen mit denen am MZP13. Insgesamt gelten drei von zehn Lernverläufen der Mädchen beim Wörkertest als stabil, vier von zehn zeigen einen den Erwartungen entsprechenden Zuwachs und gelten als Responder (vgl. Anhang H und L).

Die Jungen zeigen im Wörkertest in ihren Lernverläufen auf Grundlage der prozentualen Lösungswahrscheinlichkeit eine höhere Stabilität als die Mädchen. Fünf Lernverläufe der Schüler gelten mit null Abweichungen vom festgelegten Intervall als stabil. Die Verläufe der Jungen Ak und Ra mit jeweils zwei sowie von Ti mit drei Abweichungen gelten als variabel, wobei die Variabilität im Fall von Ra ggf. einem hohen positiven Zuwachs zugrunde liegt (vgl. Anhang L).

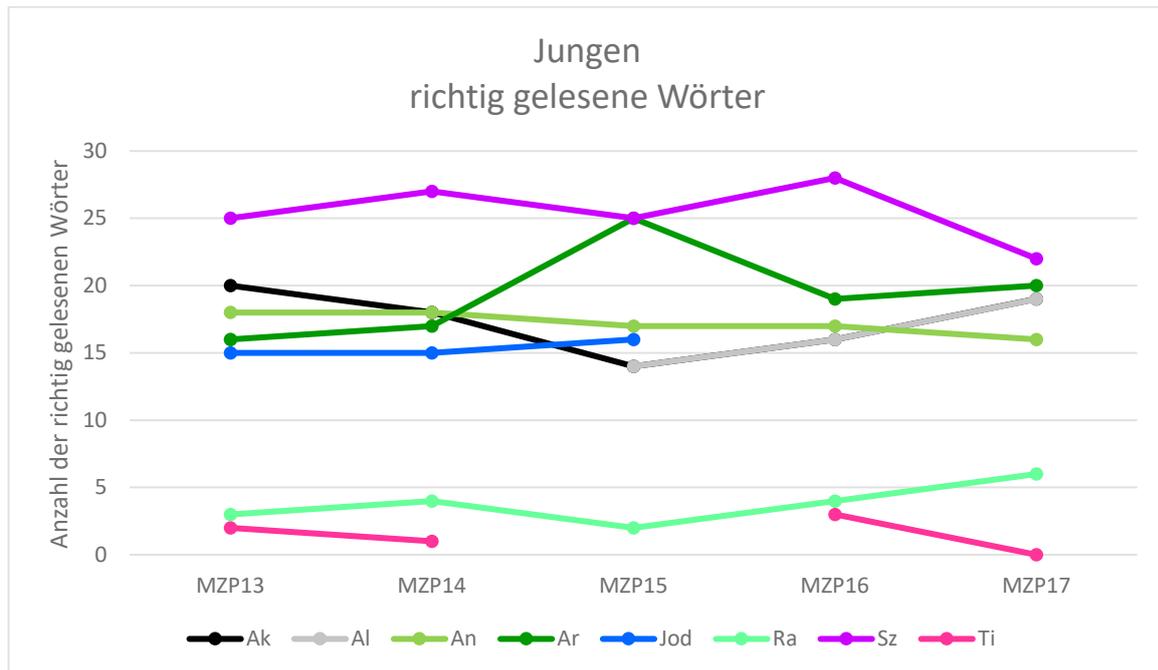


Abbildung 17 Anzahl der richtig gelesenen Wörter über fünf MZP der Jungen

In Abbildung 17 wird erneut deutlich, dass die Schüler Ra und Ti deutlich niedrigere Leistungen verglichen mit den anderen sechs Jungen zeigen. Lediglich Schüler Ar hat einen Zuwachs von vier Wörtern, weswegen er als Responder gilt. Die anderen sieben gelten alle als Non-Responder, da der Leistungszuwachs unter dem festgelegten Wert von vier mehr richtig gelesenen Wörtern liegt. Die Schüler Al und Ra zeigen dennoch einen positiven Zuwachs von drei Wörtern und Schüler Jod von einem Wort. Bei Ak, Sz und Ti hingegen sind negative Zuwächse von -1, -3 und -2 zu verzeichnen. Somit gelten fünf von acht Lernverläufe als stabil und ein Schüler kann als Responder eingestuft werden.

Im Pseudowörtertest zeigen sich die Lernverläufe der Mädchen sehr unterschiedlich. Die eine Hälfte der Mädchen zeigt stabile, die andere Hälfte variable Lernverläufe. Bei den Mädchen mit stabilem Lernverlauf handelt es sich um Jo und Sa mit null Abweichungen vom Bereich um den Mittelwert. Bei Ed, Ja und Le ist jeweils eine Abweichung der Werte vorhanden, bei Ma, Em, Gu, Ta, Am befinden sich mehr Abweichungen vom Mittelwert im Lernverlauf, so dass diese als variabel gelten (vgl. Anhang M).

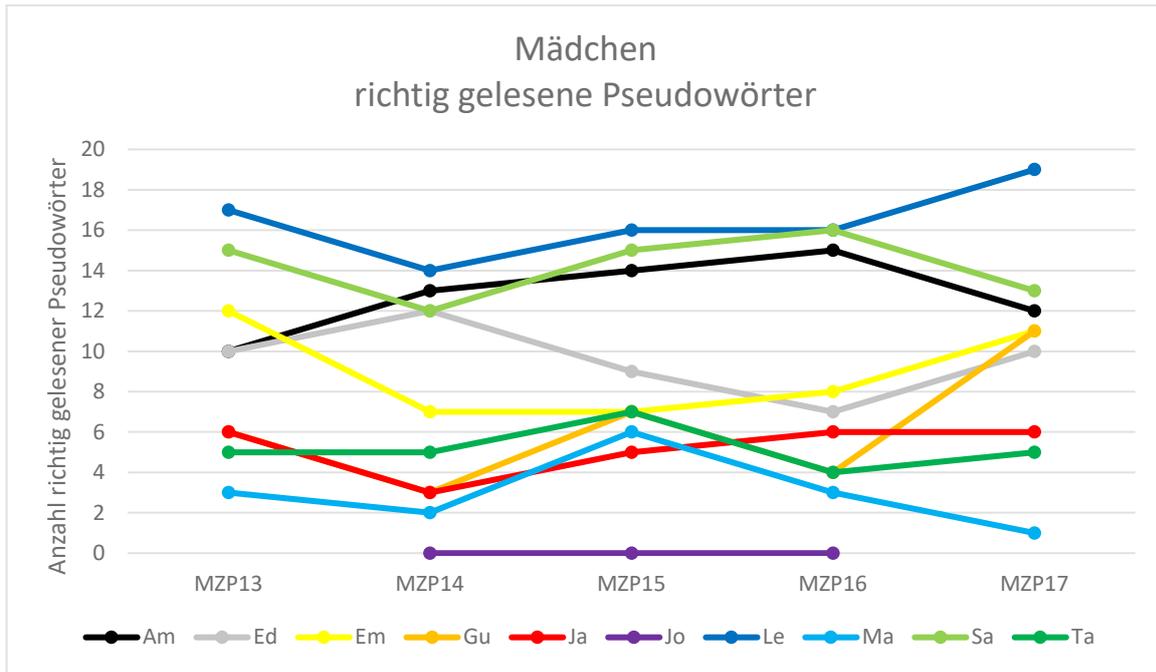


Abbildung 18 Anzahl der richtig gelesenen Pseudowörter über fünf MZP hinweg der Mädchen

In Abbildung 18 zeigen die Lernverläufe mit der Anzahl richtig gelesener Pseudowörter eine große Heterogenität zwischen den Mädchen. Hinsichtlich eines erwarteten Zuwachses von drei Pseudowörtern in vier Wochen kann lediglich Schülerin Gu mit einem Zuwachs von fünf Items als Responder bezeichnet werden. Alle übrigen Mädchen gelten als Non-Responder. Während Am und Le einen positiven Zuwachs von zwei erreichen konnten, blieb die Anzahl der richtig gelesenen Pseudowörter bei Ed, Ja und Ta am MZP17 identisch mit der am MZP13. Em zeigte einen negativen Zuwachs von -1, Ma und Sa in Höhe von -2. Zusammenfassend zeigen fünf Mädchen variable und fünf stabile Lernverläufe. Darüber hinaus gilt lediglich eine Schülerin im Pseudowörtertest als Responder (vgl. Anhang M).

Im Pseudowörtertest zeigt die Mehrzahl der Jungen stabile Lernverläufe. Sowohl die Werte von Al, An, Ra, Sz als auch von Ti befinden sich alle im berechneten Intervall um den Mittelwert. Die Lernverläufe von Ak und Ar sind beide auf Grund von jeweils zwei abweichenden Werten als variabel zu bezeichnen. Der Lernverlauf von Jod ist nicht eindeutig zu bestimmen. Nach Kern (1993) dürfen lediglich 20% der Werte von einem Bereich von  $\pm 15\%$  vom Mittelwert abweichen, was genau einer Abweichung bei fünf MZP, wie im hier analysierten Zeitraum von vier Wochen, entspricht. Für Schüler Jod sind jedoch nur drei Werte vorhanden, da er zu den beiden MZP 16 und 17 nicht anwesend war. Die maximale Abweichung von 20% der Werte wären bei drei Werten weniger als einer. Da sich die Werte in einem Bereich von 33% befinden, wird der Verlauf ebenfalls als variabel bewertet (vgl. Anhang M).

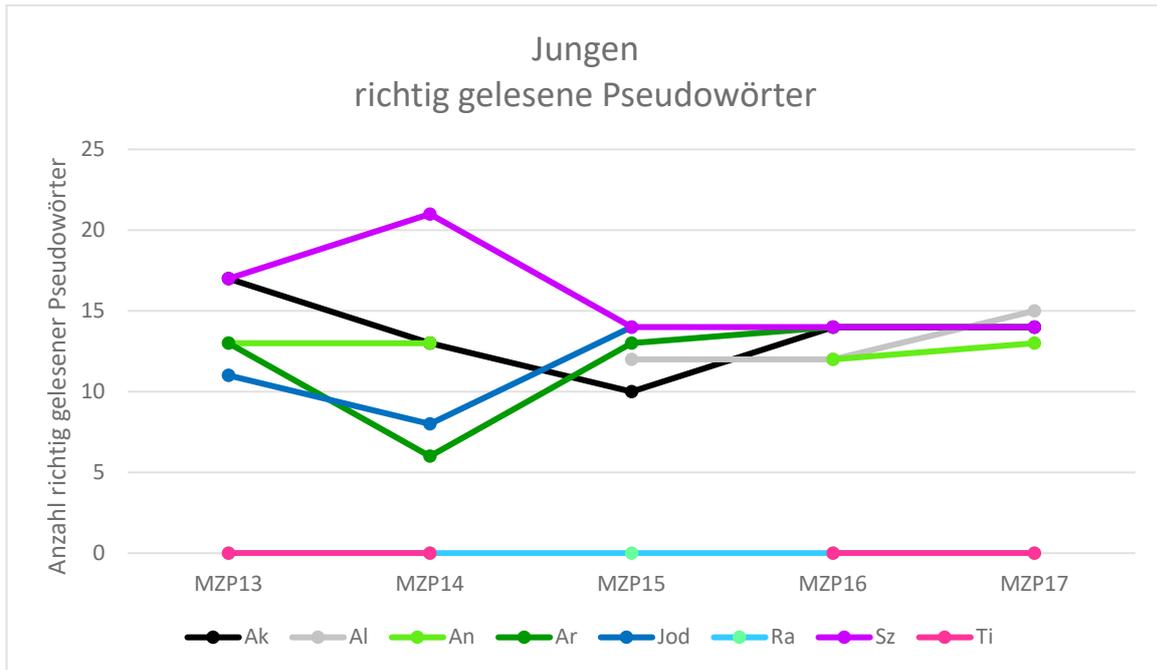


Abbildung 19 Anzahl der richtig gelesenen Pseudowörter über fünf MZP hinweg der Jungen

Hinsichtlich des erwarteten Zuwachses von mindestens drei Pseudowörtern nach vier Wochen, sind zwei der acht Schüler als Responder zu bezeichnen. Dabei handelt es sich um Al mit einem Zuwachs von vier und Jod um drei mehr richtig gelesene Pseudowörter. Es ist jedoch zu erwähnen, dass es sich bei Jod um den Zuwachs innerhalb von zwei Wochen handelt, da er bei MZP16 und 17 nicht anwesend war. Ra und Ti verbesserten sich nicht, wobei zu betonen ist, dass sie sich auf Grund der gleichbleibenden Anzahl von keinem richtig gelesenen Pseudowort auch nicht verschlechtern konnten. Die Leistung von Ar stieg um ein Pseudowort, Ak und Sz zeigten jeweils einen negativen Zuwachs von -3. Insgesamt gelten drei der acht Lernverläufe der Jungen als variabel und zwei Schüler können als Responder bezeichnet werden.

Der Test zum sinnentnehmenden Lesen ist in gewisser Weise gesondert zu betrachten, da dieser mit einer zeitlichen Dauer von acht Minuten nicht als Speedtest gilt und darüber hinaus als schülerzentrierter Test durchgeführt wird, wodurch sich ggf. die motivationalen Rahmenbedingungen von den Tests zur Leseflüssigkeit unterscheiden. Stabilität und Leistungszuwachs werden jedoch nach denselben Prinzipien eingeordnet.

Im Test zum sinnentnehmenden Test der Mädchen sind drei Lernverläufe auf Grundlage der prozentualen Lösungswahrscheinlichkeit als stabil zu bezeichnen. Bei den Schülerinnen Jo und Le weicht keiner der fünf Werte vom errechneten Bereich um den Mittelwert ab, bei Ta kommt es zu einer Abweichung, wodurch immer noch 80% der Werte im Bereich von +/-15% um den Mittelwert liegen. Die übrigen sieben Lernverläufe der Mädchen gelten als variabel.

Bei Schülerin Am, Em, Gu und Sa befinden sich jeweils zwei Werte außerhalb des Bereichs, bei Ja sind es drei und bei Ed liegen alle fünf Werte außerhalb (vgl. Anhang N).

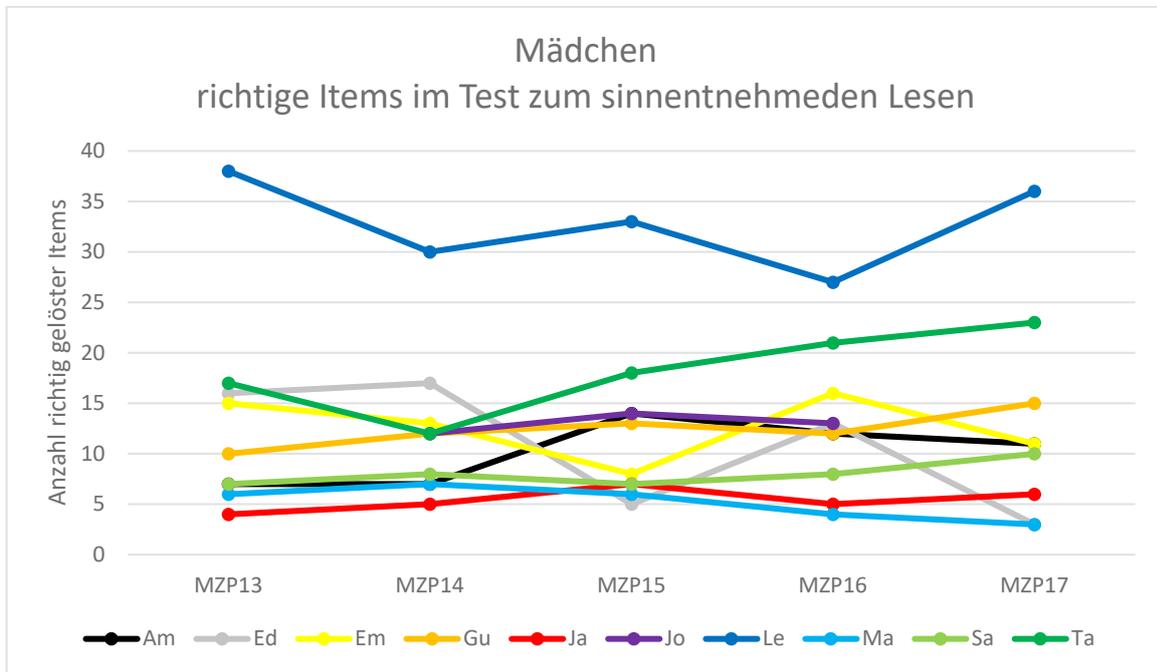


Abbildung 20 Anzahl der richtig gelösten Items im Test zum sinnentnehmenden Lesen über fünf MZP hinweg der Mädchen

In Abbildung 20 zeigt sich eine Häufung der Lernverläufe im unteren Bereich der richtig gelösten Items. Beim sinnentnehmenden Lesen wird ein Zuwachs von vier mehr richtig gelösten Items nach vier Wochen erwartet. Bezüglich der erwartungsgemäßen Zuwächse von MZP13 zu MZP17 können drei der zehn Schülerinnen als Responder bezeichnet werden. Die Schülerinnen Am, Gu und Ta konnten jeweils 4, 5 und 6 Items mehr lösen. Ja, Jo und Sa zeigen mit 2, 1 und 3 mehr gelösten Items einen geringen positiven Zuwachs. Die Schülerinnen Ed (-13), Em (-4), Le (-2) und Ma (-3) zeigten allesamt einen negativen Zuwachs. Somit sind drei von zehn Lernverläufen der Mädchen als stabil zu bewerten und ebenfalls drei von zehn gelten als Responder (vgl. Anhang J).

Im Test zum sinnentnehmenden Lesen der Jungen ist eine besondere intraindividuelle Heterogenität der Lernverläufe auf Grundlage der prozentualen Lösungswahrscheinlichkeit zu erkennen, da alle Lernverläufe als variabel gelten. Bei den Schülern Ak, Al, An, Ar, Jod und Ra befinden sich jeweils zwei der Werte außerhalb des Bereichs um den Mittelwert. Bei Schüler Ti sind es drei, bei Sz vier abweichende Werte (vgl. Anhang N).

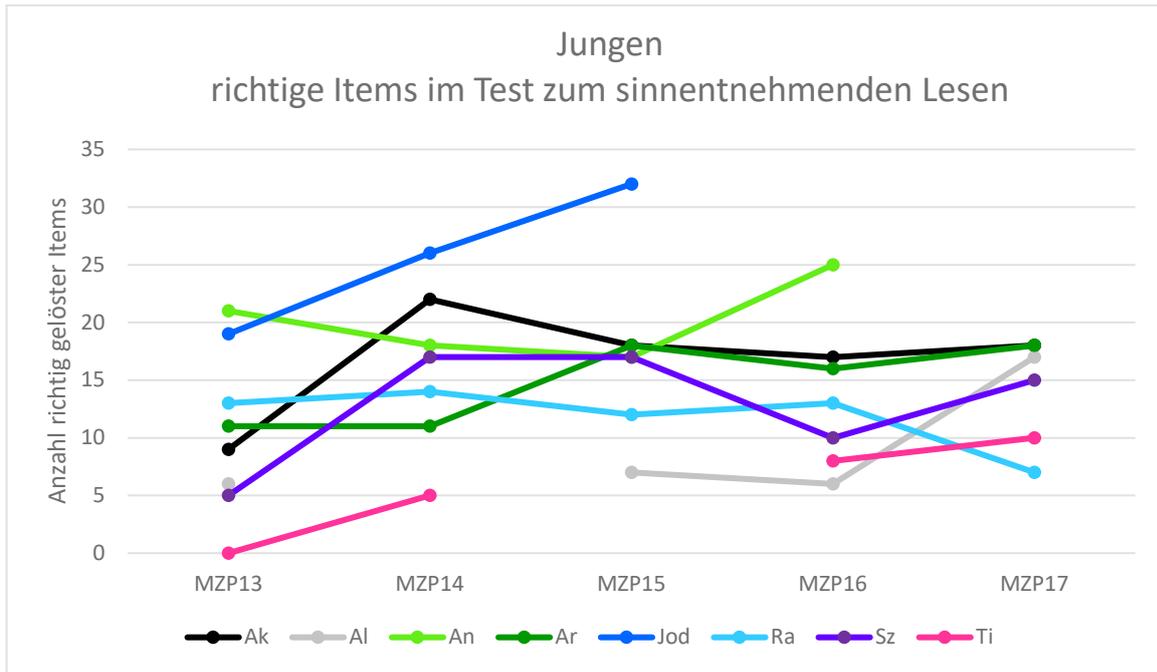


Abbildung 21 Anzahl der richtig gelösten Items im Test zum sinnentnehmenden Lesen über fünf MZP hinweg der Jungen

Bezüglich des erwarteten Zuwachses von vier mehr gelösten Items nach vier Wochen, können alle Schüler bis auf Ra als Responder bezeichnet werden. In Abbildung 21 zeigen sich zum Teil hohe Leistungsanstiege. Bei An, Ar und Ak sind Zuwächse von vier, sieben und neun Items zu verzeichnen. Die Schüler Sz und Ti verbesserten sich um zehn Items, Jod innerhalb von zwei Wochen um 13. Ra als Non-Responder zeigt einen negativen Zuwachs von -6. Zusammengefasst sind alle Lernverläufe der Jungen im sinnentnehmenden Lesen als variabel zu bewerten und sieben von acht Schülern gelten als Responder. Es ist jedoch zu betonen, dass die Lernverläufe bei allen Schülern bis auf Ra sichtbar ansteigen, weswegen die Variabilität möglicherweise durch den teils sehr hohen Zuwachs zu erklären ist.

Anschließend an die geschlechtergetrennte Darstellung der Lernverläufe werden nachfolgend die Unterschiede in der Stabilität, den Lernverläufen selbst und bezüglich des Leistungszuwachses gegenübergestellt. Da sich die Anzahl der Jungen mit acht von der der Mädchen mit zehn unterscheidet, können Unterschiede ggf. nicht 1:1 miteinander verglichen werden. Dennoch werden, falls vorhanden, Geschlechtertendenzen aufgezeigt.

Im Silbentest zeigen sich sowohl bei den Mädchen als auch bei den Jungen zwei variable Lernverläufe. Diese sind demzufolge bei beiden Geschlechtern ähnlich und eher stabil, woraus zu schließen ist, dass Jungen und Mädchen im Silbentest hinsichtlich der individuellen Bezugsnorm konstant ähnliche Leistungen an den fünf MZP zeigen. Die Mittelwerte der Mädchen bewegen sich in einem Bereich von 56-94%. Bei den Jungen ist auffällig, dass sechs

der acht Schüler eine Lösungswahrscheinlichkeit von über 90% zeigen. Der Mittelwert der anderen beiden Jungen liegt bei 28.4% und 26.5%, wodurch sich ein großer Unterschied bezüglich der Leistungen desselben Geschlechts ergibt. Im Wörtest erhöht sich die Anzahl der variablen Lernverläufe der Mädchen auf sieben, bei den Jungen nur geringfügig auf drei. Die Mittelwerte der Mädchen sinken und befinden sich in einem Bereich zwischen 42% und 94%. Als Ausreißer ist Schülerin Jo mit einer Lösungswahrscheinlichkeit von 10% zu betrachten. Beim Wörterlesetest der Jungen befinden sich die Mittelwerte zwischen 72% und 92% wobei sich erneut die mittleren Werte von Ra ( $M = 14.2\%$ ) und Ti ( $M = 29.5$ ) stark von den übrigen absetzen. Im Pseudowörtertest behalten die Jungen das Verhältnis von drei variablen zu fünf stabilen Lernverläufen bei. Fünf der Verläufe der Mädchen gelten als stabil, fünf als variabel. Hinsichtlich der Betrachtung der individuellen Mittelwerte zeigt sich bei den Mädchen ein Wertebereich zwischen 24% und 86% der Lösungswahrscheinlichkeit. Die Auffälligkeiten in den Mittelwerten sind bei den Jungen im Pseudowörtertest dieselben wie im Silben- und Wörtest. Erneut heben sich die Schüler Ra und Ti mit jeweils einer Lösungswahrscheinlichkeit von 0% von den anderen ab. Die Mittelwerte der restlichen sechs Schüler befinden sich in einem Bereich von 62% bis 92%. Die Lernverläufe im sinnentnehmenden Lesen verändern sich in beiden Geschlechtern erneut. Alle acht Lernverläufe der Jungen gelten als variabel, woraus eine intraindividuelle Heterogenität der Werte zu schließen ist. Bei den Mädchen sind drei als stabil und sieben als variabel einzuordnen. Die Mittelwerte der Jungen nähern sich, verglichen mit den Tests zur Leseflüssigkeit an, wodurch sich ein Wertebereich zwischen 19% und 54% ergibt. Hierbei zeigen auch die zuvor deutlich schwächeren Schüler Ra und Ti keine auffällig schwachen Werte. Die Mittelwerte der Mädchen befinden sich innerhalb eines Bereichs zwischen 20% und 52% Lösungswahrscheinlichkeit, wobei sich Schülerin Le als positiver Ausreißer mit einer mittleren Lösungswahrscheinlichkeit von 75.4% von den übrigen Mädchen abhebt. Bei beiden Geschlechtern ist ein Sinken der Lösungswahrscheinlichkeit vom Silben- zum Wörter- und erneut zum Pseudowörtertest zu beobachten.

Auch im Leistungszuwachs über fünf MZP sind geschlechterspezifische Unterschiede zu erkennen. Nach den angepassten zu erwarteten Zuwächsen in Anlehnung an Fuchs (1993) konnten im Silbentest drei von fünf Jungen und eins von zehn Mädchen als Responder identifiziert werden. Bei den Jungen zeigten sich zwei mit einem positiven Zuwachs unterhalb des Grenzwertes von sieben, drei Schüler zeigten jedoch auch negative Zuwächse. Demzufolge konnten sich diese drei Schüler trotz der zusätzlichen Förderung nicht verbessern. Bei den Mädchen zeigen vier der neun Non-Responder trotzdem positive Zuwächse, weitere vier zeigen negative Zuwächse und eine Schülerin veränderte sich weder zum Positiven noch zum Negativen. Im Wörterlesetest sinkt die Anzahl der Responder bei den Jungen auf eins, steigt

bei den Mädchen jedoch auf vier an. Drei der Non-Responder bei den Jungen zeigen einen positiven und vier einen negativen Zuwachs. Daraus ergibt sich, dass vier Jungen keinerlei Fortschritt im Wörterlesen erzielen konnten. Bei den Mädchen zeigte eins der Non-Responder einen positiven Zuwachs, eins veränderte sich nicht und die übrigen drei zeigten am MZP17 schwächere Leistungen als am MZP13.

Hinsichtlich der richtig gelesenen Pseudowörter können zwei der acht Jungen als Responder, die anderen sechs als Non-Responder identifiziert werden. Drei der Non-Responder zeigen keine Veränderung beim Vergleich zwischen MZP13 und 17, drei Schüler zeigen jedoch negative Zuwächse. Lediglich ein Mädchen kann, aufgrund ihres Zuwachses, als Responder ausgemacht werden, die übrigen neun gelten als Non-Responder. Von diesen haben jedoch zwei einen positiven Zuwachs, vier zeigen keine Veränderung zum MZP13 und drei zeigen einen negativen Zuwachs. Im sinnentnehmenden Lesen verändern sich die Zuwächse bei beiden Geschlechtern deutlich. Sieben der acht Jungen zeigen deutlich höhere Zuwächse als die erwartete Steigerung um ein mehr richtig gelöstes Item nach vier Wochen. Es zeigen sich mit 4 bis 13 mehr Items wesentlich höhere Zuwächse als bei den Tests zur Leseflüssigkeit. Lediglich ein Schüler zeigt einen negativen Zuwachs. Bei den Mädchen können 6 der 10 Mädchen als Responder ausgemacht werden. Sie zeigen zwar ebenfalls mit einem bis sechs mehr richtig gelösten Items einen höheren Zuwachs als erwartet, liegen dennoch deutlich unter denen der Jungen. Alle vier Non-Responder zeigen einen negativen Zuwachs. Es ist zu betonen, dass sich der Test zum sinnentnehmenden Lesen bezüglich der Lernverläufe und des Zuwachses deutlich von den drei Tests zur Leseflüssigkeit unterscheidet.

4. Gibt es einen Zusammenhang zwischen den Leistungen der SchülerInnen bei den Leseflüssigkeitstests und dem Test zum sinnentnehmenden Lesen unter Berücksichtigung geschlechterspezifischer Unterschiede?

Bei der Auswertung der Daten zur Beantwortung der vierten Forschungsfrage steht nicht die Höhe der Leistung selbst, sondern vielmehr die Entwicklung der Leseleistung im Vergleich der drei Tests zur Leseflüssigkeit und dem Test zum sinnentnehmenden Lesen untereinander im Fokus. Die Erwartungen hinsichtlich eines solchen Zusammenhangs sind auf der steigenden Schwierigkeit der Tests begründet. Beim Silbentest, als leichtester Test, wäre demzufolge die höchste Leistung bezüglich der Lösungswahrscheinlichkeit, der Anzahl bearbeiteter und richtig gelesener Items zu erwarten, welche allesamt im Wörtertest geringer und im Pseudowörtertest nochmals geringer ausfallen würden. Dementsprechend dürften beim Silbentest die niedrigsten Fehlerzahlen vorhanden sein, welche über den Wörter- und Pseudowörtertest

bis zum sinnentnehmenden Lesen als schwierigsten Test ansteigen. Beim sinnentnehmenden Lesen kann aufgrund des anderen Aufgabenformates nur die Lösungswahrscheinlichkeit der vier Tests miteinander verglichen werden, da die Anzahl der bearbeiteten Items in der Leseflüssigkeit jeweils einer Silbe, einem Wort oder einem Pseudowort entspricht und ein Item im sinnentnehmendem Lesen mehrere Wörter in einem Satz und zusätzlich vier Auswahlwörter beinhaltet. Im Zuge dessen werden die jeweiligen Werte der Lösungswahrscheinlichkeit in Prozent, die Anzahl richtig und falsch gelesener Items sowie der Zuwachs zwischen zwei MZP graphisch für jede(n) SchülerIn geschlechtergetrennt dargestellt. Auf diese Weise ist das Verhältnis des Silben- zum Wörter- und zum Pseudowörtertest sowie zum Test zum sinnentnehmenden Lesen direkt aus der Graphik abzulesen. Aufgrund der Anzahl von insgesamt 18 SchülerInnen werden die Leistungen von jeweils zwei Mädchen und Jungen exemplarisch dargestellt. Auf die Werte aller SchülerInnen wird dennoch in gekürzter Form im Fließtext Bezug genommen, welche darüber hinaus dem Anhang entnommen werden können. Die Datengrundlage zur Errechnung des Zuwachses stellen MZP13 und MZP17 dar. Auf diese Weise liegt der Fokus erneut auf dem Zeitraum der Förderung und es liegt ein ausreichend großer Zeitraum zwischen den MZP, um die Leistungen in den vier Tests angemessen miteinander vergleichen zu können. Der Zuwachs ergibt sich aus der Anzahl der mehr richtig gelesenen Wörter nach vier Wochen. Als Datengrundlage für die richtig und falsch gelesenen Items, welche mit Hilfe von gestapelten Balken dargestellt werden, dient jeweils der Mittelwert zweier aufeinanderfolgender MZP. Es werden die individuellen Mittelwerte verwendet, um ggf. Ausreißern, die aufgrund der Tagesform des oder der SchülerIn entstanden sind, entgegenzuwirken. Es werden jedoch nur zwei MZP zur Errechnung verwendet, um den Einfluss von Leistungszuwächsen im Verlauf eines größeren Zeitraums gering zu halten. Bei den gewählten Messzeitpunkten, aus welchen der Mittelwert pro SchülerIn errechnet wird, handelt es sich um MZP14 und 15. Diese MZP wurden gewählt, da an beiden die Mehrzahl der SchülerInnen - bis auf Ti und Al - bei der Erhebung anwesend waren. Für Ti werden MZP13 und 14 verwendet, für Al MZP15 und 16 sowie einmalig für An bei den Pseudowörtern MZP13 und 14. Da die Zuwächse innerhalb von zwei Wochen grundsätzlich eher gering sind, können ohne Weiteres die jeweils umliegenden MZP verwendet werden, sofern diese direkt aufeinander folgen. Die zwei Jungen und Mädchen wurden auf Grundlage des Mittelwerts richtig gelesener Wörter des jeweiligen Geschlechts ausgewählt. Der Mittelwert diente dabei zur Orientierung, um SchülerInnen auszuwählen, die weder die höchsten, noch die niedrigsten Leistungen zeigen. Darüber hinaus sollten sich die zwei SchülerInnen desselben Geschlechts nicht zu sehr ähneln, um dennoch Unterschiede sichtbar zu machen. Auf Grundlage dessen wurden die Mädchen Em und Ta sowie die Jungen An und Al ausgewählt.

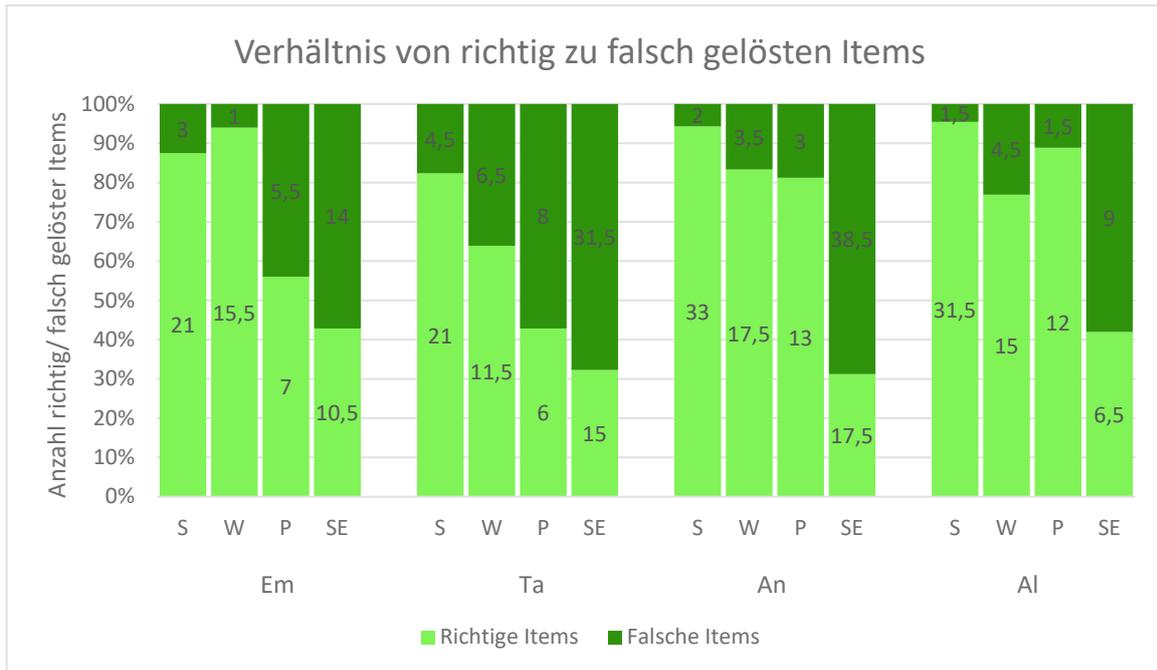


Abbildung 22 Verhältnis von richtig zu falsch gelösten Items aller vier Tests.

Anmerkung: Für eine übersichtliche Darstellung wurden die Tests jeweils mit S für Silben-, W für Wörter- und P für Pseudowörtertest abgekürzt. SE steht für den Test zum sinnentnehmenden Lesen.

In Abbildung 22 werden mehrere Aspekte der Leseleistung von vier SchülerInnen in einem dargestellt. Jede gestapelte Säule setzt sich aus den richtig und den falsch gelösten Items eines oder einer SchülerIn zusammen. Miteinander addiert, ergeben beide Werte die Anzahl der insgesamt gelesenen Items, was wiederum 100% ergibt. Die sich daraus ergebene Lösungswahrscheinlichkeit kann der Y-Achse oder dem Anhang O-R mit den genauen Daten entnommen werden. Für jede(n) der vier SchülerInnen sind vier Säulen zu erkennen, welche für die Werte der vier Levumi-Tests stehen. Nachfolgend werden zunächst die Werte der Mädchen, anschließend die der Jungen mit Verweisen auf die restlichen SchülerInnen ausgewertet.

Bei Schülerin Em sinkt die Anzahl der insgesamt bearbeiteten Items erwartungsgemäß in der Reihenfolge Silben-, Wörter- und Pseudowörterlesen ab. Mit einer Anzahl von 24 liest sie innerhalb einer Minute deutlich mehr Silben als Wörter (16.5). Diese Anzahl verringert sich erneut hinsichtlich der insgesamt gelesenen Pseudowörter (12.5). Diese Entwicklung ist ebenfalls bei Ja mit 26.5 gelesenen Silben, 13 Wörtern und 10.5 Pseudowörtern zu beobachten. Auch die Anzahl der gelesenen Items bei den übrigen SchülerInnen Am, Ed, Gu, Jo, Le, Ma und Sa sinkt. Die Differenz zwischen der Anzahl der gelesenen Silben und Wörter ist jedoch bei allen SchülerInnen deutlich größer als zwischen den gelesenen Wörtern und Pseudowörtern. Bei Schüler An ist diese Entwicklung ebenfalls festzustellen. Er liest insgesamt 35

Silben, 21 Wörter und 16 Pseudowörter. Schüler Al liest 33 Silben, 19.5 Wörter und 13.5 Pseudowörter. Nicht nur bei An und Al, auch bei allen anderen Jungen verringert sich diese Anzahl. Außer bei Schüler Sz ist auch hier der Unterschied zwischen den Silben und Wörtern deutlich größer als zwischen Wörtern und Pseudowörtern. Es ist festzuhalten, dass alle 18 SchülerInnen in der Reihenfolge Silben-, Wörter- und Pseudowörtertest immer weniger Items bearbeiten.

Neben den insgesamt bearbeiteten Items ist das Verhältnis zwischen richtig und falsch gelesenen Items hinsichtlich der Qualität und nicht nur der Quantität des Lesens deutlich aussagekräftiger, was die tatsächliche Leistung der SchülerInnen betrifft. Aufgrund der steigenden Schwierigkeit der aufeinanderfolgenden Tests zur Leseflüssigkeit, kann mit einer steigenden Fehleranzahl und einer damit verbundenen geringeren Lösungswahrscheinlichkeit gerechnet werden. Schülerin Em zeigt eine höhere Lösungswahrscheinlichkeit im Wörtertest als im Silbentest. Bei 21 richtigen Silben liest sie drei falsch, bei 15.5 richtig gelesenen Wörtern liest sie jedoch nur eines falsch. Zum Pseudowörtertest sinkt die Lösungswahrscheinlichkeit deutlich. Bei sieben richtigen Wörtern liest sie 5.5 falsch. Schülerin Ta zeigt hingegen eine erwartungsgemäße Entwicklung. Die Lösungswahrscheinlichkeit zeigt, dass die Anzahl richtig gelesener Items im Verhältnis zu den falsch gelesenen geringer wird. Von 25.5 insgesamt gelesenen Silben las sie 21 richtig und 4.5 falsch, von 18 Wörtern waren 11.5 richtig und 6.5 falsch und beim Pseudowörtertest ändert sich die Tendenz und es werden mit einer Anzahl von acht mehr Items falsch gelesen als richtig (6). Bei den anderen Schülerinnen ist eine ähnliche Entwicklung zu beobachten. Bei allen übrigen Mädchen sinkt die Anzahl der richtig gelesenen Items, wobei auch hier der Abstand zwischen den Silben und Wörtern deutlich größer ist als zwischen den Wörtern und Pseudowörtern. Die Entwicklung der falsch gelesenen Items entwickelt sich jedoch nicht so eindeutig nach diesem Prinzip. Bei Em sind mehr Fehler beim Silben- als beim Pseudowörterlesen vorhanden. Die Fehleranzahl bei den Pseudowörtern ist hingegen höher als bei den Silben und Wörtern, was den Erwartungen entspricht. Ta zeigt erneut eine erwartungsgemäße Entwicklung, indem auch ihre Fehlerzahl vom Silben- zum Wörter- und wiederum zum Pseudowörterlesen ansteigt. Ein ebensolcher Anstieg ist nur noch bei Schülerin Gu zu erkennen. Entgegengesetzt ist die Entwicklung der Fehlerzahl hinsichtlich des Anstiegs vom Silben- bis zum Pseudowörtertest bei Ja und Jo. Bei den übrigen Schülerinnen ist kein Muster bezüglich der Entwicklung der Fehlerzahl zu erkennen.

Bei Schüler An sinkt die Lösungswahrscheinlichkeit ebenfalls. Von 35 gelesenen Silben liest er 33 richtig und zwei falsch, von 21 gelesenen Wörtern liest er 17.5 richtig und 3.5 falsch

und von 16 Pseudowörtern konnte er 13 richtig und 3 falsch lesen. Somit ist bei An ein erwartungsgemäßes Sinken der Lösungswahrscheinlichkeit und der Anzahl richtig gelöster Items zu erkennen. Die Fehlerzahl hingegen steigt nicht mit wachsender Schwierigkeit an. Bei Schüler Al sinkt die Lösungswahrscheinlichkeit in der Reihenfolge Silben, Pseudowörter, Wörter. Von 33 gelesenen Silben liest er 31.5 richtig und 1.5 falsch. Von 19.5 gelesenen Wörtern liest er 15 richtig und 4.5 falsch, bei den Pseudowörtern sind 12 von 13.5 richtig und 1.5 falsch vorgelesen. Damit entspricht Al nur hinsichtlich der Entwicklung der richtig gelesenen Wörter den Erwartungen. Bei den Schülern Ak, Ar, Jod, Ra, Sz und Ti sinkt ebenfalls die Lösungswahrscheinlichkeit. Somit sinkt diese bei allen Jungen außer Al in der Reihenfolge Silben, Wörter und Pseudowörter. Bei den Jungen Ak, Ar, Jod, Ra, Sz und Ti sinken die Werte der richtig gelesenen Items stetig. Sie sinken demzufolge bei allen Jungen in der erwarteten Reihenfolge. Eine ohne Ausnahmen ansteigende Fehlerzahl ist lediglich bei Ar und Jod zu beobachten. Bei den übrigen Jungen ist kein Zusammenhang zwischen Fehlerzahl und Test zu erkennen. Es ist jedoch zu betonen, dass sich die Fehlerzahl eher geringfügig um maximal fünf Items ändert.

Die Werte des Tests zum sinnentnehmenden Lesen werden nachfolgend gesondert betrachtet, da sie sich deutlich vom Aufgabenformat, den benötigten Lesestrategien sowie der Art und Dauer in der Durchführung von den Tests zur Leseflüssigkeit unterscheiden. Grundsätzlich werden verhältnismäßig schwächere Leistungen, verglichen mit den Tests zur Leseflüssigkeit, erwartet, da in diesem Testformat komplexere Leseleistungen gefordert werden. Sowohl bei den in Abbildung 22 dargestellten Mädchen als auch bei den Jungen zeigt sich, dass im sinnentnehmenden Lesen intraindividuell die schwächsten Leistungen in der Lösungswahrscheinlichkeit verglichen mit den anderen drei Tests, zu vernehmen sind. Auch bei den Mädchen Am, Ed, Ja, Sa und Ta ist die Lösungswahrscheinlichkeit jeweils im sinnentnehmenden Lesen am niedrigsten. Bei Gu sind die Werte vom Pseudowörtertest und vom sinnentnehmenden Lesen identisch, bei Jo sind sie sehr ähnlich zum Wörtertest wobei beim Pseudowörtertest von fünf bearbeiteten Items, kein einziges richtig gelesen wurde. Bei Le ist der Wert nahezu identisch zum Wörtertest. Die Anzahl der insgesamt bearbeiteten Items des sinnentnehmenden Tests ist sowohl bei Em und Ta als auch bei Ed, Gu, Jo, Le, und Sa höher als die der anderen Tests (vgl. Abb. 22). Bezüglich der Qualität zeigt sich, dass im intraindividuellen Vergleich der einzelnen Tests beim sinnentnehmenden Lesen gleichzeitig die höchste Anzahl an falsch gelösten Items vorhanden ist. Neben Em und Ta ist auch bei allen anderen Mädchen die Fehlerzahl mit oft großem Abstand höher als in den Tests zur Leseflüssigkeit. Neben diesen teils sehr hohen Werten in der Fehlerzahl, zeigen sich jedoch ebenfalls

unerwartet viele richtig gelöste Items. Die Anzahl der richtig gelösten Items liegt bei Em zwischen dem Wörter- und Pseudowörtertest, bei Ta liegt die Anzahl sogar zwischen denen des Silben- und Wörtertests. Lediglich bei Am handelt es sich um den niedrigsten Wert an richtig gelesenen Wörtern und bei Jo um den höchsten. Bei allen anderen Schülerinnen bewegt sich der Wert der richtig gelösten Items zwischen denen des Silben-, Wörter- und Pseudowörtertests.

Bei Schüler An zeigt sich, verglichen mit den übrigen Tests, eine sehr hohe Anzahl bearbeiteter Items. Bei An, Ak, Jod, Ra und Sz ist dieser Wert sogar individuell jeweils der höchste. Die Schüler Al und Ti bearbeiteten hingegen im sinnentnehmenden Lesen nicht auffällig viele oder wenige Items. Bei den Schülern Ak, Al, An, Ar, Jod, und Sz stellt die Lösungswahrscheinlichkeit im sinnentnehmenden Lesen, verglichen mit den anderen Tests, erwartungsgemäß den niedrigsten Wert dar. Nur bei Ra und Ti ist der Wert, verglichen mit ihren weiteren Werten, relativ hoch und stellt nicht den niedrigsten dar. Bezüglich der Qualität der Bearbeitung ist das bei An und Al in der Abbildung 22 auffällige Verhältnis von richtig und falschen Items, auch im Vergleich mit dem Verhältnis bei den anderen drei Tests, von Bedeutung. An zeigt eine eher unauffällige Anzahl richtig gelesener Items, welche identisch mit dem Wörtertest ist. Bei Al ist die Anzahl der richtig gelösten Items die niedrigste, verglichen mit den anderen Tests. Die Schüler Jod und Ra haben im sinnentnehmenden Lesen auf individueller Ebene die meisten Items richtig gelöst. Die anderen Jungen Ak, Ar, Sz und Ti zeigen im Kontext ihrer eigenen Leistung zwar recht hohe Werte bezüglich der Erwartungen hinsichtlich der Schwierigkeit, sie heben sich jedoch weder im Positiven noch im Negativen besonders von ihnen ab. Die Anzahl der falsch gelesenen Items zeigt bei manchen Schülern jedoch unerwartete Ausprägungen. In Abbildung 22 wird bereits deutlich, dass An nicht nur im Verhältnis zu den richtig gelösten, sondern auch grundsätzlich eine hohe Anzahl an Items falsch gelöst hat. Löste er in den Tests zur Leseflüssigkeit jeweils nur 2, 3.5 und 3 Items falsch, so sind es im sinnentnehmenden Lesen 38.5 Items. Bei Schüler Al ist der Unterschied nicht gleichermaßen stark, dennoch mit 9 Fehlern deutlich höher als 1.5, 4.5 und 1.5 bei den anderen drei Tests. Die Anzahl der falsch gelösten Items ist bei allen Jungen, bis auf Ti, deutlich höher im intraindividuellen Vergleich, bei An, Jod, Ra, Sz ist sie jedoch besonders hoch. Bei den Jungen zeigt sich eine auffällige Quantität bei der Bearbeitung der Tests zum sinnentnehmenden Lesen, jedoch verbunden mit einer hohen Fehlerzahl, was sich in den Lösungswahrscheinlichkeiten widerspiegelt.

Für die graphische Darstellung des Zuwachses wurden erneut je zwei SchülerInnen eines Geschlechts auf der Grundlage des durchschnittlichen Zuwachses im Wörtertest ausgewählt.

Demzufolge werden bei den Mädchen die Zuwächse von Ja und Le und bei den Jungen von Al und Ar in der folgenden Abbildung dargestellt. Darüber hinaus wird jedoch auf die Zuwächse der anderen SchülerInnen zurückgegriffen (vgl. Anhänge O bis R).

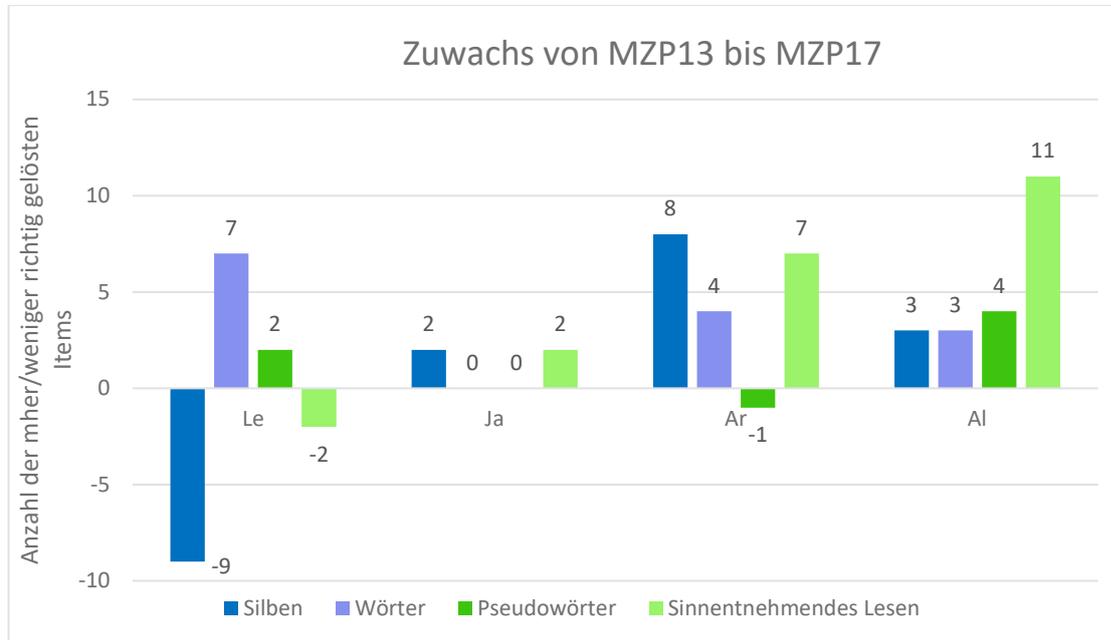


Abbildung 23 Leistungszuwachs von MZP13 bis MZP17 in allen vier Tests

In Abbildung 23 ist der Leistungszuwachs von MZP13 zu MZP17 für je zwei SchülerInnen eines Geschlechts exemplarisch dargestellt. Der Zuwachs ergibt sich aus den mehr richtig gelösten Items nach vier Wochen. Die Zuwächse können sowohl positiv, negativ als auch unverändert sein. Daraus ergeben sich die Säulen im negativen Bereich und die optisch fehlenden Säulen bei einer unveränderten Leistung. Die angenommene ansteigende Schwierigkeit der Tests in der Reihenfolge Silben-, Wörter-, Pseudowörtertest und sinnentnehmendes Lesen legt die Erwartung nahe, dass dementsprechend sinkende Zuwächse zu verzeichnen sind. Bei Schülerin Le zeigt sich ein fast erwartungsgemäßes Bild. Während beim Wörtertest ein Zuwachs von sieben zu verzeichnen ist, liegt dieser beim Pseudowörtertest nur noch bei zwei und beim sinnentnehmenden Lesen im negativen Bereich bei -2. Lediglich der negative Zuwachs von -9 Items beim Silbenlesen hebt sich von den anderen Werten ab. Während Ja am MZP17 2 Silben mehr gelesen hat als am MZP13, verändert sich die Anzahl richtig gelester Wörter und Pseudowörter nicht. Im sinnentnehmenden Test zeigt sie jedoch einen Zuwachs von 2 Items. Bei Betrachtung der übrigen Schülerinnen zeigt sich keinerlei Tendenz hinsichtlich eines höheren Zuwachses im Silbenlesen und eines niedrigeren im Pseudowörterlesen und dem sinnentnehmenden Lesen. Es ist jedoch anzumerken, dass nur fünf Mädchen einen positiven Zuwachs im Silbentest erreichen. Vier zeigen negative Zuwächse und

eine Schülerin zeigt keine veränderte Leistung. Im Wörkertest sind ebenfalls fünf Zuwächse positiv, drei negativ und zwei Schülerinnen veränderten sich in den vier Wochen nicht. Beim Pseudowörkertest ist der Zuwachs von drei Mädchen positiv, von dreien negativ und vier zeigten keine Veränderungen in der Anzahl richtig gelöster Items. Im sinnentnehmenden Test zeigen sechs Schülerinnen positive und vier negative, wobei jedoch die intraindividuelle Heterogenität im sinnentnehmenden Lesen zu beachten ist.

Schüler Ar zeigt größtenteils den Erwartungen entsprechende Zuwächse. Während bei ihm ein Zuwachs von acht Silben, vier Wörtern sowie einem Abfall von einem Pseudowort vorhanden ist, zeigt sich der Zuwachs im sinnentnehmenden Lesen in Höhe von sieben überraschend hoch. Schüler Al zeigt einen Zuwachs von je drei Silben und Wörtern sowie vier Pseudowörtern. Auch bei ihm ist der Zuwachs im sinnentnehmenden Lesen mehr als dreimal so hoch wie im Silben-, und Wörterlesen. Der verhältnismäßig hohe Zuwachs im sinnentnehmenden Lesen zeigt sich ebenso bei den Jungen Ak, Jod, Sz und Ti. Bei den Schülern An und Ar hebt sich dieser Zuwachs nicht auffällig von den anderen Tests ab. Auch bei den Jungen sind einige negative Zuwächse zu verzeichnen. Im Silbentest zeigen fünf Schüler positive und drei negative Zuwächse. Im Wörkertest sind vier Zuwächse positiv und vier negativ. Beim Pseudowörkertest können nur zwei Schüler einen positiven Zuwachs erreichen, drei verändern sich in ihrer Leistung nicht und zwei zeigen einen negativen Zuwachs. Im sinnentnehmenden Lesen zeigen sich nicht nur besonders hohe, sondern auch sieben positive Zuwächse. Lediglich ein Schüler zeigt einen negativen Zuwachs. Unterschiede zwischen den Geschlechtern zeigen sich hierbei nicht.

## 7 Diskussion

In der nachfolgenden Diskussion wird sich interpretativ mit den zuvor ausgewerteten und primär deskriptiv dargestellten Daten auseinandergesetzt. Die Forschungsfragen dienen als zentrale Strukturierung dieses Kapitels, im Rahmen dessen sie beantwortet werden. Dabei werden die Ergebnisse kritisch hinsichtlich der Forschungslage auf Grundlage des Theorie- teils diskutiert. Ebenso werden die Grenzen der vorliegenden Forschung aufgezeigt und das sich daraus ergebende Forschungsinteresse wird dargelegt. Grundsätzlich ist zu betonen, dass es sich bei der vorliegenden Grundgesamtheit um eine kleine Stichprobe handelt, wes- wegen keine Rückschlüsse auf SchülerInnen der zweiten Klasse im Allgemeinen gezogen werden können.

1. Inwiefern gibt es Unterschiede zwischen der ersten und letzten Messung hinsichtlich des Leistungszuwachses der SchülerInnen unter Berücksichtigung des Geschlechts?

Der Zuwachs zwischen den MZP 1 und 18 ist, genauso wie die Leseleistung an sich, nicht nur aus einer Perspektive zu betrachten. Es gibt unterschiedliche Bereiche, die wiederum zu verschiedenen Deutungen der Schülerschaft führen. Grundsätzlich ist zu betonen, dass im Mittel in allen Tests und Bereichen bei beiden Geschlechtern über den Zeitraum von 21 Wo- chen ein positiver Leistungszuwachs vorhanden ist. Lediglich im sinnentnehmenden Lesen kommt es bei den Jungen in der Lösungswahrscheinlichkeit zu einem negativen Zuwachs und im Bereich falsch gelöster Items schwanken die Mittelwerte sowohl im positiven als auch im negativen Bereich.

Nach 21 Wochen zeigt sich mit wenigen Ausnahmen grundsätzlich ein höherer Leistungs- wuchs auf Seiten der Mädchen. Je nach dem welcher Bereich im Fokus des Interesses steht, zeigen sich leicht variierende Ergebnisse. Betrachtet man den Zuwachs an richtig gelesenen Items, so zeigen die Mädchen in dieser Hinsicht in allen vier Tests höhere Zuwächse. Außer- dem zeigt sich die Tendenz, dass die Jungen ihre Fehler im Wörtest stärker reduzieren konnten als die Mädchen, wobei jedoch eine grundsätzlich höhere Fehlerzahl der Jungen zu Beginn zu beobachten ist. Im Silbentest ist diesbezüglich kein Unterscheid zwischen den Geschlechtern zu erkennen, im sinnentnehmenden Lesen nimmt die Fehlerzahl bei beiden Geschlechtern zu. Nimmt man Bezug zu den Ergebnissen der wöchentlich getesteten Schü- lerInnen von Wenner (2018) hinsichtlich der Zuwächse über acht MZP hinweg, zeigt sich dort ebenfalls, dass bei den Mädchen bezüglich der Anzahl richtig gelesener Items in allen drei Tests zur Leseflüssigkeit höhere Zuwächse zu verzeichnen sind.

Neben den Leistungszuwächsen zeigen sich auch Geschlechterunterschiede in der Leseleistung selbst, welche sich jedoch nach 21 Wochen verändern. Die Leseleistung wurde an dieser Stelle ausschließlich anhand der Lösungswahrscheinlichkeit bestimmt. Am MZP1 zeigen sich die Jungen in 3/4 Tests den Mädchen überlegen, am MZP18 zeigen die Mädchen in 3/4 Tests stärkere Leistungen. Setzt man nun die Zuwächse und die Leistung selbst zueinander in Beziehung, kann der Wandel von den Jungen hin zu den Mädchen als leistungsstärkeres Geschlecht durch die höheren Zuwächse von Seiten der Mädchen erklärt werden.

Verglichen mit den zuvor dargestellten Studien, wie KIM oder IGLU, bestätigen die im Rahmen dieser Forschung erhobenen Daten die Ergebnisse dieser Studien teilweise. Auf Grund der geringen Forschungslage bezüglich der hierarchieniedrigen Leseleistungen, zeigen die vorliegenden Daten jedoch auch eine weitere Perspektive. Im Rahmen von IGLU und PISA wurde deutlich, dass die geschlechterspezifischen Unterschiede zu Gunsten der Mädchen mit zunehmender Klassenstufe größer werden. Grundsätzlich handelt es sich jedoch um eher geringe Unterschiede, wie es auch die vorliegenden Daten zeigen. Da sich beispielsweise die IGLU auf GrundschülerInnen der vierten Jahrgangsstufe und auf hierarchiehöhere Prozesse bezieht, kann dies bedeuten, dass die Geschlechtsunterschiede in der zweiten Klasse ggf. noch nicht signifikant sind und darüber hinaus möglicherweise eine andere Tendenz zu Gunsten der Jungen zeigt. Im gesamten Messzeitraum zeigt sich, dass die Mädchen im Mittel größere Lernzuwächse erreichen, woraufhin diese 21 Wochen später bessere Leistungen als die Jungen erzielen. Unter Vorbehalt könnte dies beinhalten, dass in den Klassenstufen eins und zwei nur geringe Unterschiede in der Lesekompetenz zwischen den Geschlechtern zu verzeichnen sind, welche sich jedoch aufgrund der höheren Zuwächse der Mädchen vergrößern. Falls die Mädchen auch zukünftig größere Zuwächse als die Jungen erreichen, ergibt sich unweigerlich eine immer größere Leistungsdifferenz zwischen den Geschlechtern.

Aus den vorliegenden Ergebnissen kann unter Vorbehalt ein Anstieg des Geschlechtsunterschiedes vermutet werden. Die vorliegenden Studien werden insofern bestätigt, als dass die Mädchen, wie erwartet, zum letzten MZP bessere Leistungen als die Jungen zeigen, jedoch mit dem Zusatz, dass sich diese Geschlechtertendenz erst im Verlauf der Grundschulzeit manifestiert und nicht ab Schuleintritt vorhanden ist. Gleichzeitig zeigt sich ein Unterschied zu den Ergebnissen von Niklas und Schneider (2012), welche bereits im ersten Schuljahr höhere Leistungen in der Lesekompetenz auf Seiten der Mädchen ergaben. Insofern zeigen sich in den hier vorliegenden Daten unerwartete Ergebnisse. Eine formative Erhebung der Lesekompetenz einer größeren Stichprobe über die gesamten vier Jahre der Grundschulzeit mitsamt den hierarchieniedrigen und höheren Prozessen wäre bezüglich der Frage wann Geschlechterunterschiede und wie konstant diese bestehen von Interesse.

Bezüglich der durchschnittlichen Höhe des Zuwachses an mehr gelesenen Wörtern ist anzumerken, dass bei getesteter Grundgesamtheit ein sehr geringer Leistungszuwachs zu vermerken ist. Vergleicht man den Zuwachs der Jungen von 2.38 und den der Mädchen von 4.67 mehr richtig gelesenen Wörtern nach 21 Wochen mit dem erwarteten Zuwachs von 1.46 Wörtern pro Woche nach Fuchs et al. (1993), also insgesamt 30.66 mehr Wörter nach 21 Wochen, so wird deutlich, wie gering der Leistungszuwachs der getesteten SchülerInnen ist. Nach Walter (2011) können 25% der GrundschülerInnen als Non-Responder ausgemacht werden, die keinen Fortschritt innerhalb eines Schuljahres erzielen. Wie viel Prozent der vorliegenden Stichprobe tatsächlich keinen Zuwachs zeigen, wurde im Rahmen der Auswertung der Mittelwerte nicht betrachtet. Da der mittlere Zuwachs jedoch deutlich unter den durchschnittlichen Erwartungswerten nach Fuchs et al. (1993) liegt, kann ein großer Anteil an Non-Respondern, bzw. SchülerInnen mit einem sehr geringen Zuwachs, vermutet werden. Eine mögliche Erklärung für die geringen Zuwächse beider Geschlechter kann beispielsweise der Migrationshintergrund oder der sozioökonomische Status der SchülerInnen sein. Dieser kann insofern einen Einflussfaktor für die Lesekompetenz darstellen, als dass mit einem familiären Migrationshintergrund oftmals das Erlernen der deutschen Sprache als Zweitsprache verbunden ist. Darüber hinaus stellt sich die Frage, welche Sprache im außerschulischen Bereich gesprochen wird und ob umfangreiche und abwechslungsreiche Sprachangebote zur Verfügung stehen. Ebenso das außerschulische Lesen kann ggf. Verzögerungen im Schriftspracherwerb kompensieren. Inwiefern diese Faktoren jedoch auf die Schülerschaft zutreffen, wurde im Rahmen dieser Forschungsarbeit nicht umfassend erhoben. Eine Erhebung von außerschulischen Rahmenbedingungen, die auf die Qualität und Quantität der Sprach- und Leseangebote abzielen wäre zukünftig ggf. von Interesse.

Im Rahmen der t-Tests mit einem Signifikanzniveau von 5% zeigen sich keine signifikanten Geschlechterunterschiede, weder zu MZP1 noch zu MZP18 in keinem der vier Tests. Im Kontext dessen sind die sehr hohen Standardabweichungen zu erwähnen. Die damit verbundene Heterogenität zeigt sich auch am Ende der Primarstufe, beispielsweise in den Kompetenzstufen der IGLU. Im Kontext der Unterrichtsentwicklung resultieren aus den bereits in der zweiten Klasse vorhandenen Leistungsunterschiede im Lesen die Notwendigkeit von differenzierten Lernangeboten und zusätzlicher Förderung, um einer Manifestierung von Schwierigkeiten beim Schriftspracherwerb entgegenzuwirken.

2. Inwieweit sind zu zwei Messzeitpunkten geschlechterspezifische Unterschiede zu erkennen?

- 2a) Geschlechterspezifische Unterschiede zum Messzeitpunkt 11 vor Beginn der Förderung mit dem Unterricht als primären Einflussfaktor
- 2b) Geschlechterspezifische Unterschiede zum Messzeitpunkt 17 nach der Förderung mit dieser als zusätzlichen Einflussfaktor

Am MZP11 zeigen die Jungen im Silben-, Wörter- und im Pseudowörtertest in drei von vier Bereichen der Lösungswahrscheinlichkeit, der Anzahl bearbeiteter und der Anzahl richtig gelesener Items, höhere Leistungen als die Mädchen, wobei jedoch ihre Fehlerzahl ebenfalls höher ist. Somit zeigt sich eine konstant höhere Leistung der Jungen in den Leseflüssigkeitstest, jedoch nicht im sinnentnehmenden Lesen. Am MZP17 zeigen jedoch die Mädchen in allen Leseflüssigkeitstests grundsätzlich bessere Leistungen als die Jungen. Die Fehlerzahl der Mädchen liegt in allen vier Tests unter der der Jungen. Im Bereich der bearbeiteten Items zeigen die Jungen in allen Tests bessere Leistungen, bezüglich der richtig gelösten Items zeigen sich geringe Geschlechtsunterschiede zu Gunsten der Mädchen. Im sinnentnehmenden Lesen erreichen die Jungen in allen Bereichen bis auf die Fehlerzahl höhere Werte. Insgesamt zeigt sich eine deutliche Verbesserung der Mädchen nach bzw. am Ende der Interventionsphase. Die in der Literatur vorzufindende Geschlechtertendenz zu Gunsten der Mädchen scheint sich auch in der Gruppe der leseschwächeren SchülerInnen im Verlauf des Messzeitraums zu entwickeln. Demzufolge entspricht die Tendenz, dass die Mädchen bessere Leseleistungen als die Jungen zeigen, zum Ende des Messzeitraumes dem Forschungsstand und passt ebenfalls zu den Ergebnissen der vorherigen Forschungsfrage mit den SchülerInnen der ganzen Klassen als Datengrundlage. Es stellt sich jedoch die Frage, inwiefern die Förderung den Zuwachs möglicherweise geschlechterbedingt beeinflusst und ob die Mädchen ggf. mehr von der Förderung profitieren als die Jungen.

Auffällig bei Betrachtung der Leistungen in den verschiedenen Bereichen der Tests nach dem Geschlecht getrennt ist, dass die Jungen durchweg mehr Items bearbeiten und auch tendenziell mehr Items falsch beantworteten. Diese Ergebnisse stimmen mit der mittleren Fehleranzahl nach Wenner (2018) über acht MZP überein. Dies kann Aufschluss über die Art des Lesens geben. Daraus ergibt sich, dass die Jungen ihren Fokus möglicherweise eher auf die Quantität statt die Qualität beim Lesen legen. Sie lösen zwar auch mehr Items richtig, weswegen die Lösungswahrscheinlichkeit nicht zwingend niedriger ausfällt, dennoch zeigt sich eine schnellere Bearbeitung der Items von den Jungen als von den Mädchen. Diese Tendenz in der Art der Bearbeitung kann durch die Beobachtungen der Testleitung bestätigt werden. Oftmals zeigte sich eine Art wettbewerbsorientiertes Testverhalten der Jungen, welche mög-

lichst viele Items bearbeiten wollten. Diese Beobachtungen wurden nicht auf Basis wissenschaftlicher Rahmenbedingungen durchgeführt, wodurch es sich mehr um eine beiläufige Auffälligkeit während der Testsituation handelt. Dennoch stützen diese die Ergebnisse. Dass die Jungen zu MZP11 grundsätzlich höhere Leistungen zeigen sowie die durchweg höheren Werte in der Fehleranzahl deckt sich ebenfalls mit den zeitlich früher erhobenen Daten von Wenner (2018).

Betrachtet man die gesamten 18 SchülerInnen und deren geschlechterspezifischen Ausprägungen vor und nach der Förderung, stellt sich jedoch die Frage, inwiefern die übrigen SchülerInnen der Klassen, wären sie im selben Zeitraum ebenfalls wöchentlich jedoch ohne zusätzliche Förderung getestet worden, ähnliche Ausprägungen zeigen würden. Die wöchentliche Testung einer Kontrollgruppe hätte ggf. den Einfluss der Förderung auf das Geschlecht aufzeigen können. Dennoch zeigt sich dieselbe Tendenz wie in der ersten Forschungsfrage. Auf Grundlage dessen lässt sich die Forschungsfrage dahingehend beantworten, als dass in den Tests pro MZP eindeutig Geschlechterunterschiede festgemacht werden können, die sich ähnlich in der Literatur zeigen. Beispielsweise die IGLU kann mit dem Zusatz bestätigt werden, dass sich der Geschlechterunterschied zugunsten der Mädchen in vorliegender Stichprobe erst im Verlauf des zweiten Schuljahres ausprägt.

Obwohl es sich bei den 18 SchülerInnen bereits um die Leseschwächsten aus den drei Klassen handelt, zeigen sich dennoch hohe Standardabweichungen, besonders beim sinnentnehmenden Lesen am MZP11. Aus der großen Heterogenität der Leseleistungen ergibt sich ein Bedarf an differenziertem Unterricht und bedürfnisorientierter Förderung, um den unterschiedlichen Förderbedarfen der Schülerschaft gerecht zu werden.

### 3. Inwieweit zeigen sich geschlechterspezifische Unterschiede im Lernverlauf während der Förderung über fünf Messzeitpunkte hinweg?

Im Rahmen der dritten Forschungsfrage wurden die Lernverläufe sowohl im Mittel als auch individuell betrachtet sowie hinsichtlich der Stabilität und eines erwartungsgemäßen Zuwachses eingeordnet. In allen vier Tests zeigen die Jungen bezüglich der Anzahl richtig gelöster Items bei der Mehrzahl der MZP höhere Werte als die Mädchen.

Auf individueller Ebene zeigt sich je nach Test eine höhere intra- und interindividuelle Heterogenität. Die Mädchen zeigen in den Leseflüssigkeitstests über fünf MZP während der durchgeführten Förderung grundsätzlich mehr variable Lernverläufe als die Jungen. Daraus ergibt sich eine hohe Heterogenität im intraindividuellen Vergleich. Trotz einer gewissen Variabilität der Werte einzelner Jungen, zeigen diese konstantere Werte als die Mädchen. Diese

Ergebnisse zeigen eine ähnliche Tendenz wie auch die Forschung von Wenner (2018). Auch in dieser zeigen die Mädchen tendenziell eine höhere Variabilität in den Lernverläufen. Im sinnentnehmenden Lesen gelten bei den Jungen alle Lernverläufe als variabel, bei den Mädchen sind es 7/10. Aus den insbesondere bei den Jungen stark schwankenden Werten um mehr als +/- 15% um den Mittelwert resultiert ein Absinken der Zuverlässigkeit eines Wertes zu einem MZP. Auf Grundlage dessen sollten einzelne Messungen im Kontext der angrenzenden MZP betrachtet werden.

Die auf Grundlage von Literatur dargestellten Einflussfaktoren auf die Lesekompetenz können an dieser Stelle abgewandelt auf den Test zum sinnentnehmenden Lesen bezogen werden. Motivation stellt einen bedeutsamen Einflussfaktor auf den Schriftspracherwerb dar. Jedoch auch im Kontext von punktueller Leistungsmessung ist Motivation ein Kernelement für hohe Leistung. Die Tests zur Leseflüssigkeit finden in einer 1:1-Situation statt, wobei der Testleiter permanent beim Schüler ist. Es kann vermutet werden, dass allein die Anwesenheit im Nahbereich der SchülerInnen für eine höhere Anstrengungsbereitschaft förderlich ist. Es handelt sich dabei um eine Erhöhung der extrinsischen Motivation. Darüber hinaus dauern die Tests zur Leseflüssigkeit nur je eine Minute, wohingegen der Test zum sinnentnehmenden Lesen ohne TestleiterIn durchgeführt wird und acht Minuten dauert. Auf diese Weise wäre eine mögliche Erklärung für die stark variierenden Verlaufskurven, dass das Maß der extrinsischen Motivation durch den oder die TestleiterIn deutlich geringer ist. Neben der Motivation kann auch die Konzentration der SchülerInnen, welche jedoch nicht erhoben wurde, als einflussreicher Faktor vermutet werden. Es wäre möglich, was durch die Beobachtungen der Testleitung bestätigt wird, dass die Durchführung der Tests zum sinnentnehmenden Lesen, welche immer mit allen sechs SchülerInnen in einem Raum durchgeführt wurden, die Konzentration beeinträchtigen. Blicke zum und Gespräche mit dem oder der TischnachbarIn wurden zwar zeitnah durch die Testleitung unterbunden, können jedoch trotzdem einen Einfluss auf die Konzentration der SchülerInnen gehabt haben. Neben den äußerlichen Einflussfaktoren und der Testdauer, ist auch das Testformat selbst teils kritisch zu betrachten. Es beinhaltet die Optionen, dass ein Item sowohl nicht beantwortet werden kann, wodurch es zu mehr falschen Items kommt, als auch, dass das richtige Wort erraten werden kann, ohne den Satz tatsächlich zu lesen. Neben der grundsätzlich die Leistung prägenden Tagesverfassung stellen die genannten Aspekte stark variierende Einflussfaktoren dar, wodurch möglicherweise die intraindividuell heterogenen Leistungen zu erklären sind. Auch die Beobachtungen der Testleiterin bekräftigen die Erklärungsansätze, da es oftmals nötig war, die SchülerInnen daran zu erinnern, weiterzuarbeiten und die Items nicht zu überspringen. Ebenso stiegen die

Unterhaltungen oder das Versunkensein in den eigenen Gedanken mit der Bearbeitungsdauer weiter an.

Hinsichtlich des erwartungsgemäßen Zuwachses auf Grundlage der Richtwerte von Fuchs et al. (1993) wurde geprüft, welche SchülerInnen als (Non-)Responder zu werten sind. Grundsätzlich zeigt sich, dass ein Großteil der SchülerInnen den erwarteten Zuwachs nicht erreicht hat. Darüber hinaus handelt es sich bereits um an leistungsschwächere SchülerInnen angepasste Zuwächse und es fand in dem hier betrachteten Zeitraum zusätzlich zweimal pro Woche eine Förderung statt. Betrachtet man die Tests zur Leseflüssigkeit, so kann eine gewisse Tendenz zu Gunsten der Jungen ausgemacht werden. Im Silbentest gelten im vorliegenden Zeitraum der Förderung 3/8 Jungen und nur 1/10 Mädchen als Responder. Im Wörtertest sind es 1/8 Jungen und 4/10 Mädchen. Im Pseudowörtertest gelten 2/8 Jungen als Responder und erneut nur 1/10 Mädchen. Es zeigt sich, dass der Großteil der SchülerInnen keinen nennenswerten Zuwachs durch Unterricht und Förderung erzielen konnte. Diese grundsätzlich eher geringen Zuwächse können zum einen dadurch zustande kommen, dass es sich um die Gruppe leseschwacher SchülerInnen handelt, auf der anderen Seite besteht die Möglichkeit, dass der Zeitraum der Förderung nicht lang genug war. Torgesen (2005) spricht von der Notwendigkeit mehrmonatiger Förderung bei schwerwiegenden Leseproblemen, was möglicherweise den Zuwachs der SchülerInnen erhöht hätte.

Umso unerwarteter zeigt sich der hohe Anteil an Respondern bei den Jungen im Sinnentnehmenden Lesen. Von ihnen können 7/8 und von den Mädchen 3/10 als Responder ausgemacht werden. An dieser Stelle ist erneut auf die sehr variablen Lernverläufe im sinnentnehmenden Lesen zu verweisen. Da die Responder auf Grundlage eines Vergleichs von zwei MZP identifiziert wurden, kann gefolgert werden, dass die Zuwächse ggf. an anderen MZP deutlich abweichende Werte beinhalten. Es stellt sich demzufolge die Frage, inwiefern die Ergebnisse hinsichtlich des Zuwachses im sinnentnehmenden Lesen als zuverlässig gelten. Ähnlich wie bei Wenner (2018) zeigt sich, dass nicht nur die leistungsschwächsten SchülerInnen als Non-Responder gelten. Dies birgt eine besondere Gefahr einer längerfristigen Stagnation der Lesekompetenz, da mit Verweis auf die Wait-to-fail Problematik, SchülerInnen oftmals erst in ihrem Bedarf an zusätzlichen Fördermaßnahmen identifiziert werden, sobald diese eine erhebliche Leistungsdifferenz zu ihren MitschülerInnen zeigen. Sofern Defizite in der Lesekompetenz bezüglich der sozialen Bezugsnorm nicht gravierend sind, wird womöglich kein Handlungsbedarf von Seiten der Lehrkräfte wahrgenommen. Lediglich Schülerin Gu und Schüler Ar gelten in jeweils 3/4 Tests als Responder. Nur die SchülerInnen Sa, Al, Jod und Ak gelten in der Hälfte der Tests als Responder. An dieser Stelle ist jedoch zu betonen,

dass die SchülerInnen Le und Sz mit oftmals den höchsten Werten bei den richtig gelesenen Items auffallen. Beispielsweise bei 40 richtig gelesenen Silben von Le zu MZP 15 oder Sz mit 42 richtig gelesenen Silben am MZP 13 stellt sich die Frage, in welchem Ausmaß eine mögliche Leistungssteigerung pro Minute noch im realistischen Bereich ist oder ob der erwartungsgemäße Zuwachs von sieben richtigen Silben mehr nach fünf Wochen überhaupt möglich ist. Damit verbunden sollte auch die mit dem Silbentest N4 messbare Leistungsgrenze beachtet werden. Es zeigt sich die Relevanz einer umfangreichen Auswertungskompetenz. Die SchülerInnen sollten im Kontext ihrer Persönlichkeit und ihres Lernverlaufs betrachtet werden. Darüber hinaus besteht die Notwendigkeit, dass sich Lehrkräfte über die Grenzen der Tests wie beispielsweise dem Deckeneffekt bewusst sind.

Grundsätzlich zeigen sich in den Tests zur Leseflüssigkeit eher niedrige Zuwächse trotz der zusätzlichen Förderung. Darüber hinaus zeigt sich eine geringfügige Tendenz zu Gunsten der Jungen hinsichtlich einer höheren Stabilität im Lernverlauf sowie bezüglich höherer Zuwächse bei den richtig gelösten Items und einer damit verbundenen stärkeren Reaktion auf Unterricht und Förderung. An dieser Stelle ist das Ergebnis von denen der ersten und zweiten Forschungsfrage abzugrenzen. Trotz der hier geringen Tendenz hinsichtlich eines höheren Zuwachses der Jungen widersprechen diese Ergebnisse nicht den vorherigen. Die Ergebnisse der ersten und zweiten Forschungsfrage beziehen sich einmal auf die Mittelwerte der gesamten Stichprobe der drei Klassen und einmal auf die Mittelwerte der 18 wöchentlich getesteten SchülerInnen. Bei der Auswertung dieser Forschungsfrage wurden die SchülerInnen jedoch auf individueller Ebene betrachtet und weiterführend durch einen Vergleich mit dem erwartungsgemäßen Zuwachs als Responder oder Non-Responder benannt. Die genaue Größe der Zahlenwerte, welche den Mittelwert beeinflussen, wurden hier nicht näher betrachtet. Ebenfalls sind die Grenzwerte bezüglich des nötigen Zuwachses, um als Responder zu gelten, teils kritisch zu betrachten. Da es keine vergleichbare Erhebung mit Richtwerten für den Zuwachs leseschwacher SchülerInnen der zweiten Klasse hinsichtlich hierarchieniedriger Lesekompetenz gibt, wurden die Werte auf Grundlage der Daten von Fuchs et al. (1993) angepasst. Bei einer Orientierung an aktuelleren und passgenauen Daten bestünde die Möglichkeit eines abweichenden Verhältnisses von Respondern zu Non-Respondern.

Bei den SchülerInnen mit auffallend schwachen Leseleistungen zeigt sich die Heterogenität in der Art der Schwierigkeiten deutlich. Die SchülerInnen Jo, Ti und Ra zeigen einen besonderen Abstand zu den anderen 15 SchülerInnen. Trotz der drei ähnlich schwachen Leistun-

gen zeigen sich bei genauer Analyse der Werte sowie des Leseverhaltens bei der Testdurchführung ganz unterschiedliche Lesestrategien sowie Ausprägungen der Leistung. Schülerin Jo beispielsweise liest bei den Leseflüssigkeitstests eher weniger Items und davon viele falsch. Bei der Durchführung fiel auf, dass sie lange bei einem Item verweilt und mehrmals erneut am Wortanfang beginnt. Auf Grundlage dessen kann vermutet werden, dass Jo primär die nicht-lexikalische Strategie des Lesens nutzt und hauptsächlich phonologisch rekodiert. Sie scheint die Graphem-Phonem-Korrespondenz noch nicht automatisiert zu haben. Die lange Bearbeitung eines Items und das wiederholte Beginnen am Wortanfang kann ggf. mit einer geringen Kapazität des Arbeitsgedächtnisses und Problemen im artikulatorischen Zusammenführen von gelesenen Einheiten begründet werden. Dies erklärt zum Teil die Lösungswahrscheinlichkeit von 0% im Pseudowörtertest. Aus je mehr Silben ein Wort besteht, desto größere Kapazitäten des Arbeitsgedächtnisses sind nötig um die gelesenen Segmente zusammenzufügen.

Bei Schüler Ti zeigen sich zwar ebenfalls niedrige Leseleistungen, jedoch in anderer Form. Bei den Tests zur Leseflüssigkeit zeigt sich trotz der lehrerzentrierten Durchführung ein hohes Maß an Ablenkung durch grundsätzlich als eher gering zu bezeichnende Störfaktoren. Es scheint als würden Silben und Wörter teilweise ganzheitlich im Sinne der lexikalischen Strategie erfasst werden, wobei es dennoch zum Vorlesen falscher Items kam. An dieser Stelle kann vermutet werden, dass Ti insbesondere Wörter auf Grundlage der ersten Silbe identifiziert und durch den Kontext benennt. Auf diese Weise kamen gehäuft Fehler wie <Flasche> statt <falsche> oder <Freunde> statt <Freude> zustande. Darüber hinaus können geringe Leistungen hinsichtlich der Aufmerksamkeit und Motivation, auf Grundlage von Beobachtungen der Testleiterin, als stark beeinflussend für die erhobene Lesekompetenz vermutet werden.

Schüler Ra hingegen, als ebenfalls besonders leistungsschwacher Schüler, zeigt teilweise eine mit den leistungsstarken SchülerInnen vergleichbare Anzahl bearbeiteter Items, welche jedoch mit sehr hohen Fehlerzahlen verbunden sind. Aus den Beobachtungen während der Durchführung ist zu schließen, dass er oftmals nicht phonologisch rekodiert und Items versucht ganzheitlich zu Lesen, wobei er sich bei den Wörtern und insbesondere den Pseudowörtern an markanten Graphemen orientiert und darüber hinaus die restlichen errät. Die Orientierung an markanten graphischen Merkmalen eines Wortes kann jedoch ebenfalls ein Zeichen für ein Verweilen im Bereich des logographischen Lesens sein. Schüler Ra sollte demzufolge ggf. Förderung in der Stärkung der Phonem-Graphem-Korrespondenz erhalten, um auf diese Weise nach und nach mehr Segmente oder ganze Wörter ins mentale Lexikon

zu übertragen. Dadurch kann das direkt-lexikalische Lesen, im Sinne von parallel genutzten Strategien, ausgeweitet werden.

Um SchülerInnen durch lernverlaufsdagnostische Verfahren wie Levumi betreffend ihrer erwartungsgemäßen Zuwächse und ggf. im Kontext von sozialen oder kriterialen Bezugsnormen einzuordnen, ist eine umfangreiche Erhebung solcher Daten mit großen Stichproben von Nöten. Auf diese Weise kann Lehrkräften eine grobe Orientierung für die Einordnung von Ist-Ständen sowie für individuelle Zielsetzungen ermöglicht werden. Darüber hinaus ist jedoch zu betonen, dass die reine Identifikation von Non-Respondern keine Auswirkungen auf den Zuwachs der SchülerInnen beinhaltet. Erst mit der Analyse der formativen Daten, verknüpft mit Beobachtungen des Leseverhaltens und den verwendeten Lesestrategien im Unterricht und ggf. auch während den Tests zur Leseflüssigkeit, kann der individuelle Stand der Lesekompetenz umfangreich erfasst werden. Von ebensolcher Relevanz ist die anschließende Förderung, die die individuellen Förderbedarfe in den Fokus stellt.

4. Gibt es einen Zusammenhang zwischen den Leistungen der SchülerInnen bei den Leseflüßigkeitstests und dem Test zum sinnentnehmenden Lesen in Abhängigkeit zur Testschwierigkeit unter Berücksichtigung geschlechterspezifischer Unterschiede?

Die Ergebnisse entsprechen den Erwartungen hinsichtlich des Zusammenhangs der genannten Variablen teilweise. Bei allen SchülerInnen konnte ein erwartungsgemäßer Abfall der Anzahl bearbeiteter Items in der Reihenfolge der zunehmenden Schwierigkeit bei dem Silben-, Wörter- und Pseudowörtertest ausgemacht werden. Sowohl bei den Jungen als auch bei den Mädchen ist zu erkennen, dass die Differenz zwischen der Anzahl der bearbeiteten Silben und Wörter deutlich größer ist als zwischen den bearbeiteten Wörtern und Pseudowörtern. Dieser Unterschied in der Differenz kann damit begründet werden, dass die einzelnen Items bei den Silben deutlich kürzer sind als die mehrsilbigen Items der Wörter und Pseudowörter. Bei allen Jungen und fast allen Mädchen setzt sich diese Entwicklung auch hinsichtlich der richtig gelesenen Items fort. Der Abstand zwischen dem Silben- und Wörtertest sowie zwischen dem Wörter- und Pseudowörtertest bei den richtig gelesenen Items verhält sich sehr ähnlich wie bei den insgesamt bearbeiteten Items. Es bestätigt sich, dass der Silbentest am leichtesten ist. Darüber hinaus erscheint der Unterschied in der Schwierigkeit des Wörter- und Pseudowörtertests nicht in dem Maße ausgeprägt. Auch in der Lösungswahrscheinlichkeit sinken die Werte bei 7/8 Jungen und bei 9/10 Mädchen erwartungsgemäß. Lediglich bei Al und Sa zeigen sich unerwartete Werte. In Bezug auf die Fehleranzahl zeigen bei beiden Geschlechtern jeweils nur zwei SchülerInnen einen den Erwartungen entsprechenden Anstieg der Fehleranzahl bei ansteigender Schwierigkeit der drei Tests. Lediglich die Schüler

Ar und Jod sowie die Schülerinnen Ta und Gu erhöhen ihre Fehlerzahl kontinuierlich. Bei allen anderen SchülerInnen zeigen sich bei der Fehleranzahl keine eindeutigen Muster. Da es sich grundsätzlich um zahlenmäßig niedrige Fehlerwerte handelt zeigt sich, dass die Fehleranzahl nicht so sehr von der Schwierigkeit der Tests geprägt ist wie die Werte der anderen Bereiche. Hinsichtlich der Zuwächse wurden nach Testschwierigkeit sinkende Werte erwartet. Jedoch zeigen sich sowohl bei den Jungen als auch bei den Mädchen in den drei Tests zur Leseflüssigkeit keine erkennbaren Strukturen. Lediglich die Anzahl der SchülerInnen mit positivem Zuwachs bezüglich der richtig gelesenen Items zeigt sich bei beiden Geschlechtern im Pseudowörtertest geringfügig niedriger als im Silben- und Wörertest. Somit wird trotz einiger weniger Ausreißer eine erwartungsgemäße Entwicklung der Leseflüssigkeit in den Bereichen insgesamt bearbeitete, richtig gelöste Items und der Lösungswahrscheinlichkeit angemessen an die Schwierigkeit der drei Tests zur Leseflüssigkeit sichtbar. Geschlechterspezifische Unterschiede zeigen sich an dieser Stelle nicht.

Es wurde erwartet, dass sich die Leistungen des sinnentnehmenden Lesens verglichen mit den Tests zur Leseflüssigkeit in deutlich schwächeren Leistungen ausprägen, da es sich um hierarchiehöhere Prozesse beim Lesen handelt. Es wurden zwar alle Bereiche des sinnentnehmenden Lesens ausgewertet, wirklich vergleichbar mit den Werten der anderen Tests ist jedoch lediglich die Lösungswahrscheinlichkeit. Dies hängt damit zusammen, dass einem Item im sinnentnehmenden Lesen nicht eine Silbe oder ein (Pseudo-)Wort zukommt, sondern mehrere Wörter, da es sich um ganze Sätze handelt. Darüber hinaus beträgt die Dauer des Tests statt einer, acht Minuten. Trotz der längeren Dauer ist nicht zu erwarten, dass im sinnentnehmenden Lesen mehr Items als im Wörertest ernsthaft bearbeitet werden können. Dies ist der Fall, da bei korrekter Bearbeitung eines Items im sinnentnehmenden Lesen bis zu 14 Wörter mindestens einmal gelesen werden müssen, um das richtige Wort für die Lücke zu identifizieren.

Betrachtet man die Lösungswahrscheinlichkeit des sinnentnehmenden Lesens im Kontext der Tests zur Leseflüssigkeit, so handelt es sich bei fast allen SchülerInnen um den niedrigsten Prozentwert. Dieser Wert scheint zum größten Teil realistisch, jedoch bei Ra und Ti ist dieser Wert im Kontext der anderen Tests auffällig hoch, insbesondere, da beide bei den Pseudowörtern eine Lösungswahrscheinlichkeit von 0% zeigen. Diese Auffälligkeit könnte als Zeichen für das Erraten des richtigen Wortes gesehen werden, da besonders Ti, Ra und Jo als die in allen Tests am leistungsschwächsten ausgemacht werden können. An dieser Stelle wird erneut deutlich, dass der Test zum sinnentnehmenden Lesen unterschiedliche Merkmale

oder Funktionen beinhaltet, welche die Reliabilität und Validität des Tests beeinflussen können. Noch auffälliger heben sich die Werte im sinnentnehmenden Lesen hervor, wenn die Anzahl der bearbeiteten und richtig gelösten Items betrachtet werden. Es zeigt sich eine nicht erwartungsgemäße und unrealistisch hohe Anzahl bearbeiteter Items, da diese bei vielen SchülerInnen so hoch wie beim Silbentest oder noch erheblich höher sind. Beispielsweise Jo, Ra und Sz bearbeiten im Durchschnitt der beiden gewählten MZP 61 Items in den 8 Minuten des Tests. Diese Anzahl kann rein rechnerisch bezüglich der aus mehreren Wörtern bestehenden Sätze nicht korrekt bearbeitet worden sein.

Auch die Anzahl an SchülerInnen mit hohem positivem Zuwachs zeigt sich bei den Mädchen mit 6/10 und bei den Jungen mit 7/8 verglichen mit den Tests zur Leseflüssigkeit am höchsten. Somit kann die Forschungsfrage insofern beantwortet werden, als dass die Werte der Leseflüssigkeitstests sich erwartungsgemäß hinsichtlich der Testschwierigkeit verhalten, jedoch ohne auffällige Geschlechterunterschiede. Der Test zum sinnentnehmenden Lesen zeigt auffällig hohe Werte in den bearbeiteten und richtig gelösten Items entgegen den Erwartungen. Bezüglich eines geschlechterspezifischen Unterschieds, kann lediglich ein Unterschied in der durchschnittlichen Anzahl bearbeiteter Items festgemacht werden. Die Mädchen bearbeiten im Mittel 35 und die Jungen 40 Items, was erneut die eher quantitativ orientierte Testbearbeitung von Seiten der Jungen unterstützt. Interessant wäre eine nähere Analyse der Reaktionszeiten des sinnentnehmenden Lesens. Dafür wäre jedoch eine Mindestdauer pro Item bei gewissenhafter Bearbeitung nötig, um die tatsächlichen Werte ermitteln zu können. Dies wurde aus ökonomischen Gründen und einem anderen Schwerpunkt der Forschungsfragen im Rahmen dieser Forschungsarbeit nicht näher betrachtet.

Bei Betrachtung aller Ergebnisse wurde deutlich, dass die Lesekompetenz ein umfassendes Konstrukt bestehend aus verschiedenen Teilbereichen ist. Die an dieser Stelle erhobenen bzw. errechneten Daten zu den Bereichen Lösungswahrscheinlichkeit, insgesamt bearbeitete, richtig und falsch gelöste Items ergeben oftmals unterschiedliche Tendenzen zu Gunsten eines Geschlechts. Die Bereiche sind jedoch insofern voneinander zu differenzieren, als dass sie mehr oder weniger aussagekräftig für die tatsächliche Leseleistung sind. Die Fehlerzahl, sofern dieser Wert allein betrachtet wird, gibt keinen konkreten Aufschluss über die Leseleistung. Erst im Kontext der richtig gelesenen Wörter, woraus sich die Lösungswahrscheinlichkeit ergibt, kann die Qualität des Lesens abgeleitet werden. Ebenso verhält es sich mit den insgesamt bearbeiteten Items. Diese zeigen zwar die Lesegeschwindigkeit an, jedoch nicht wie viele Items richtig gelöst wurden. Die Lösungswahrscheinlichkeit, die durchaus meh-

rere Aspekte berücksichtigt, ist jedoch aufgrund der prozentualen Darstellung nicht aussagekräftig für die Lesegeschwindigkeit, da auch fünf richtig gelesene Items pro Minute eine 100%ige Lösungswahrscheinlichkeit ergeben. Daraus ergibt sich, dass auch die geschlechterspezifischen Unterschiede je nach analysiertem Bereich unterschiedliche Tendenzen ergeben. Somit ist es nötig die Lesekompetenz auf verschiedenen Ebenen zu betrachten, um umfassende Aussagen über SchülerInnen treffen zu können.

## 8 Zusammenfassung und Ausblick

In vorliegender Forschungsarbeit zeigten sich sowohl erwartungsgemäße als auch überraschende Ergebnisse in Bezug auf geschlechtsspezifische Unterschiede in der Lesekompetenz. Vorab ist jedoch zu betonen, dass sich die Ergebnisse nur auf die getesteten SchülerInnen beziehen sowie, dass sie wegen der eher geringen Stichprobengröße nicht auf die Leseleistung von Zweitklässlern im Allgemeinen übertragen werden können.

Die Fragestellung der Arbeit ist insofern zu beantworten, als dass in verschiedenen Bereichen der Leseleistung Geschlechterunterschiede vorhanden sind. Diese Unterschiede zeigen sich am letzten MZP sowie am Ende der Förderung der 18 SchülerInnen erwartungsgemäß zu Gunsten der Mädchen. Entgegen der Erwartungen zeigen die Jungen jedoch am ersten MZP sowie kurz vor der Förderung stärkere Leseleistungen. Damit geht ein höherer Leistungszuwachs der Mädchen einher, welcher ausreichend war, um zum Ende stärkere Leistungen zu erbringen. Darüber hinaus zeigen sich wie erwartet sinkende Leistungen der SchülerInnen beim Anstieg der Schwierigkeit eines Tests zur Leseflüssigkeit. Die Ergebnisse des Tests zum sinnentnehmenden Lesen sind je nach Bereich von unerwartet hohen Leistungen geprägt, die primär nicht zu den Leistungen der Leseflüssigkeitstests eines Kindes passen. Ebenfalls zeigte sich bei allen SchülerInnen -sowohl bei den leistungsstärkeren als auch bei den leistungsschwächsten- ein grundsätzlich geringer Leistungszuwachs durch Unterricht und Förderung. Es wurde eine hohe Leistungsheterogenität auf inter- und intraindividuelle Ebene sichtbar. Insgesamt zeigen die Jungen in ihrer Testbearbeitung ein eher quantitativ orientiertes Verhalten, was eine höhere Anzahl insgesamt bearbeiteter Items jedoch auch eine höhere Fehlerzahl beinhaltet. Die Mädchen hingegen legen ihren Fokus weniger auf die Schnelligkeit, sondern mehr auf die Qualität des Lesens.

Aus der Erhebung mit der Online-Plattform Levumi und der anschließenden Auswertung der Daten resultieren verschiedene Konsequenzen für LehrerInnen im Rahmen der Unterrichts- und Förderplanung sowie der Diagnose selbst. Die 18 wöchentlich getesteten SchülerInnen zeigen grundsätzlich niedrige Leseleistungen, die auch von den Lehrkräften identifiziert wurden. Demzufolge ist die Diagnosekompetenz der LehrerInnen der drei Klassen als gut zu bewerten. Bei der Auswertung der quantitativen Daten zeigte sich jedoch die Notwendigkeit einer umfassenden Schülerbetrachtung bezüglich des Schülerverhaltens und des vorangegangenen Lernverlaufs sowie der Analyse unterschiedlicher Teilbereiche. Es wurde insbesondere deutlich, dass die Leseleistung im Kontext verschiedener Werte zu betrachten ist. Bezüglich der Unterrichtsplanung ergibt sich aus den Ergebnissen, dass die gewählten Rahmenbedingungen für viele SchülerInnen zu einem nicht ausreichenden Leistungszuwachs

führen. Um Bildungsbenachteiligung auf Grund der Geschlechtszugehörigkeit entgegenzuwirken, ist es nötig, dass sich Lehrkräfte dieser Tendenz bewusst sind und ggf. im Rahmen der Unterrichtsgestaltung darauf reagieren. Beispielsweise könnte im Bereich der Lesemotivation angesetzt werden. Dabei wären sowohl geschlechtsneutrale Themen im Unterricht selbst von Bedeutung als auch vielfältige Leseangebote, die von Jungen und Mädchen im Kontext der intrinsischen Motivation selbst ausgewählt werden.

Individuelle Förderung im Kontext heterogener Lerngruppen, insbesondere im gemeinsamen Lernen von Kindern mit und ohne sonderpädagogischem Förderbedarf, stellt eine große Herausforderung dar, ist jedoch das Recht jeden Kindes, welches auch im Schulgesetz NRW (2005) verankert ist. Auf Grundlage verlaufdiagnostischer Verfahren kann der individuelle Leistungsstand erfasst werden, der anschließend in der Unterrichtsplanung berücksichtigt werden kann. Dabei bietet die Online-Plattform Levumi eine für LehrerInnen ökonomische Möglichkeit, die Leseleistungen der SchülerInnen dauerhaft in verschiedenen Bereichen zu erheben. Die Nutzung von Levumi zeigte sich auch in der vorliegenden Forschung als unkompliziert, wobei die Auswertung der Daten im Schulalltag größere Herausforderungen mit sich bringt. Durch die Plattform wird den Lehrkräften jedoch ein umfassendes Angebot visueller Darstellungen des Lernverlaufs angeboten, außerdem wird die Auswertung quantitativer und qualitativer Daten ermöglicht. Inwiefern auf Basis dieser Daten eine individuellere Förderung durchgeführt wird, hängt jedoch vom Interesse und Engagement der jeweiligen Lehrkraft ab. Bei der Nutzung von Levumi, mit dem Ziel den eigenen Unterricht zu evaluieren und diesen mehr an die Bedürfnisse der SchülerInnen anzupassen, kann der *Wait-to-fail-Problematik* entgegengewirkt werden, indem Schwierigkeiten im Lesen frühzeitig erkannt werden.

Um lernverlaufdiagnostische Maßnahmen in Kombination mit passgenauen Fördereinheiten mehr an Grundschulen zu etablieren, ist es nötig, den Mehrwert der formativen Diagnostik für LehrerInnen und SchülerInnen an Schulen zu kommunizieren. Weiterführende Studien, welche die Effektivität von passgenauer Förderung auf Grundlage von LVD aufzeigen, können die Relevanz der Auseinandersetzung mit ebensolchen Instrumenten voranbringen. Um Vergleichswerte im Rahmen von Alters- oder Jahrgangsnormen bezüglich der Leistung selbst und des wöchentlichen Zuwachses zu erhalten, wären weiterführende Studien bezüglich der Lesekompetenz über die gesamte Grundschulzeit relevant. Auf diese Weise können Schülerleistungen nicht nur im Kontext der sozialen Bezugsnorm der Klasse, sondern auch auf Grundlage repräsentativer Referenzwerte eingeordnet werden. Davon ausgehend können auch individuell passende Ziele für SchülerInnen formuliert werden.

**III Literaturverzeichnis**

- Bos, W.; Hornberg, S.; Arnold, K.-H.; Faust, G.; Fried, L.; Lankes, E.-M. et al. (2008): IGLU-E 2006. Die Länder der Bundesrepublik Deutschland im nationalen und internationalen Vergleich. Münster: Waxmann.
- Bos, W.; Valtin, R.; Hußmann, A.; Wendt, H.; Goy, M. (2017): IGLU 2016: Wichtige Ergebnisse im Überblick. In: A. Hußmann, H. Wendt, W. Bos, A. Bremerich-Vos, D. Kasper, E.-M. Lankes et al. (Hg.): IGLU 2016. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich. Münster, New York: Waxmann, S. 13–28.
- Bremerich-Vos, A.; Stahns, R.; Hußmann, A. & Schurig, M. (2017): Förderung von Leseflüssigkeit und Leseverstehen im Leseunterricht. In: A. Hußmann, H. Wendt, W. Bos, A. Bremerich-Vos, D. Kasper, E.-M. Lankes et al. (Hg.): IGLU 2016. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich. Münster, New York: Waxmann, 279–296.
- Bremerich-Vos, A.; Tarelli, I.; Valtin, R. (2012): Das Konzept von Lesekompetenz von IGLU 2011. In: W. Bos, I. Tarelli, A. Bremerich-Vos und K. Schwippert (Hg.): IGLU 2011. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich. Münster: Waxmann, S. 69–90.
- Casale, G.; Grosche, M.; Hennemann, T.; Huber, C. (2015): Testgütekriterien der Verlaufsdagnostik von Schülerverhalten im Förderschwerpunkt Emotionale und soziale Entwicklung. In: *Heilpädagogische Forschung* (1), S. 37–54.
- Coltheart, M. (2005): Modeling Reading: The dual-Route Approach. In: M. Snowling und C. Hulme (Hg.): *The science of reading. A handbook*. Malden, Mass.: Blackwell Publ (Blackwell handbooks of developmental psychology), S. 6–23.
- Deno, S. L. (2003a): Curriculum-Based Measures: Development and Perspectives. In: *Assessment for effective intervention* (28, 3&4), S. 3–12.
- Deno, S. L. (2003b): Developments in Curriculum-Based Measurement. In: *The journal of special education* (37, 3), S. 184–192.
- Diehl, K.; Hartke, B. (2007): Curriculumnahe Lernfortschrittsmessungen. In: *Sonderpädagogik* (4), S. 195–211.
- Dummer-Smoch, L.; Hackethal, R. (2016): *Kieler Leseaufbau*. 9. Auflage. Kiel: Veris-Verl.
- Dürscheid, C. (2012): *Einführung in die Schriftlinguistik*. 4., überarb. und aktualisierte Aufl. Göttingen: Vandenhoeck & Ruprecht.
- Feierabend, S.; Plankenhorn, T.; Rathgelb, T. (2017): *KIM-Studie 2016. Kindheit, Internet, Medien. Basisstudie zum Medienumgang 6- bis 13-Jähriger in Deutschland*. Stuttgart: Medienpädagogischer Forschungsverbund Südwest.
- Fischer, B. (2015): *Wahrnehmungs- und Blickfunktionen bei Lernproblemen. Besser werden im Schreiben; Lesen; Rechnen*. 1. Aufl. s.l.: Centaurus Verlag Media (Reihe Psychologie, v.41).

- Förster, N.; Souvignier, E. (2014): Learning progress assessment and goal setting: Effects on reading achievement, reading motivation and reading self-concept. In: *Learning and instruction* (32), S. 91–100.
- Frith, U. (1985): Beneath the surface of developmental dyslexia. In: K. E. Patterson, J. C. Marshall und M. Coltheart (Hg.): *Surface dyslexia. Neuropsychological and cognitive studies of phonological reading*. London: Lawrence Erlbaum Associates, S. 301–330.
- Fuchs, D.; Fuchs, L. S.; Compton, D. L. (2004): Identifying Reading disabilities by Responsiveness-to-Instruction. Specifying Measures and Criteria. In: *Learning Disability Quarterly* (27), S. 216–227.
- Fuchs, D.; Fuchs, L. S.; Hamlett, C. L.; Walz, L.; Germann, G. (1993): Formative evaluation of academic progress: How much growth can we expect? In: *School Psychology Review* 1993 (22, 1), S. 27–48.
- Gebhardt, M.; Diehl, K.; Mühling, A. (2016a): Lern-Verlaufs-Monitoring. LEVUMI Lehrerhandbuch. Online verfügbar unter <https://bit.ly/2OKJN05>, zuletzt aktualisiert am 05.08.2018.
- Gebhardt, M.; Diehl, K.; Mühling, A. (2016b): Online-Lernverlaufsmessung für alle Schülerinnen und Schüler in inklusiven Klassen. In: *Zeitschrift für Heilpädagogik* (10), S. 444–453.
- Gebhardt, M., Heine, J.-H., Zeuch, N. & Förster, N. (2015). Lernverlaufsdagnostik im Mathematikunterricht der zweiten Klasse. Raschanalysen zur Adaptation eines Testverfahrens für den Einsatz in inklusiven Klassen. *Empirische Sonderpädagogik*, (3), 206-222.
- Gebhardt, M., Sälzer, C., Mang, J., Müller, K. & Prenzel, M. (2015). Performance of Students with Special Educational Needs in Germany. Findings from PISA 2012. *Journal of Cognitive Education and Psychology*, 14(3). 343-356 <https://doi.org/10.1891/1945-8959.14.3.343>.
- Groth, K.; Hasko, S.; Bruder, J.; Kunze, S.; Schulte-Körne, G. (2013): Interventionseffekte bei Lese-Rechtschreibstörung: Evaluation von zwei Förderkonzepten unter besonderer Betrachtung methodischer Aspekte. In: *Lernen und Lernstörungen* (2, 3), S. 161–175.
- Grünke, M. (2012): Auswertung von Daten aus kontrollierten Einzelfallstudien mit Hilfe von Randomisierungstests. In: *Empirische Sonderpädagogik* (3/4), S. 247–264.
- Günther, K. B. (1986): Ein Stufenmodell der Entwicklung kindlicher Lese- und Rechtschreibstrategien. In: H. Brügelmann (Hg.): *ABC und Schriftsprache. Rätsel für Kinder, Lehrer und Forscher*. Konstanz: Faude, S. 32–35.
- Heinzel, C.; Bartsch, V.; Eckert, I.; Weinfurter, L. (2015): Programm zum sinnentnehmenden Lesen. ProSL. Auf der Grundlage der blitzschnellen Worterkennung (BliWo). Dortmund: Verl. Modernes Lernen.
- Huber, C.; Grosche, M. (2012): Das response-to-intervention Modell als Grundlage für einen inklusiven Paradigmenwechsel in der Sonderpädagogik. In: *Zeitschrift für Heilpädagogik* (8).
- Hurrelmann, B. (2002): Leseleistung-Lesekompetenz. In: *Praxis Deutsch* (176), S. 6–18.
- Jain, A.; Spieß, R. (2012): Versuchspläne der experimentellen Einzelfallforschung. In: *Empirische Sonderpädagogik* (2/3), S. 211–245.

- Jungjohann, J.; Gebhardt, M.; Diehl, K.; Mühling, A. (2017): Förderansätze im LESEN mit LEVUMI. Online verfügbar unter <https://bit.ly/2Ogc1Q5>.
- Jungjohann, J. & Gebhardt, M. (2018). Lernverlaufsdiagnostik im inklusiven Anfangsunterricht Lesen. Verschränkung von Lernverlaufsdiagnostik, Förderplanung und Wochenplanarbeit. In F. Hellmich, Görel & M.F. Löper (Hrsg.), *Inklusive Schul- und Unterrichtsentwicklung*. Stuttgart: Kohlhammer, S.160-173.
- Jungjohann, J., DeVries, J. M., Gebhardt, M. & Mühling, A. (2018). Levumi: A Web-Based Curriculum-Based Measurement to Monitor Learning Progress in Inclusive Classrooms. In K. Miesenberger & G. Kouroupetroglou (Hrsg.), *Computers Helping People with Special Needs* (Lecture Notes in Computer Science, Bd. 10896). Cham: Springer International Publishing, S.369-378.
- Kazdin, A. E. (1982): Single-case research designs. New York: Oxford University Press.
- Kern, H. J. (1997): Einzelfallforschung. Eine Einführung für Studierende und Praktiker. Weinheim: Beltz.
- Klauer, K. Josef (2006): Erfassung des Lernfortschritts durch curriculumbasierte Messung. In: *Heilpädagogische Forschung* (1), S. 16–26.
- Klauer, K. Josef (2011): Lernverlaufsdiagnostik. Konzept, Schwierigkeiten und Möglichkeiten. In: *Empirische Sonderpädagogik* (3), S. 207–224.
- Klauer, K. Josef (2014): Formative Leistungsdiagnostik: Historischer Hintergrund und Weiterentwicklung zur Lernverlaufsdiagnostik. In: M. Hasselhorn, W. Schneider und U. Trautwein (Hg.): *Lernverlaufsdiagnostik*. Göttingen, Bern, Wien, Paris: Hogrefe (Tests und Trends, Neue Folge Band12), S. 1–17.
- Klicpera, C.; Schabmann, A.; Gasteiger-Klicpera, B. (2010): Legasthenie - LRS. Modelle, Diagnose, Therapie und Förderung; mit 100 Übungsfragen. 3., aktualisierte Aufl. München: Reinhardt.
- Klicpera, C.; Schabmann, A.; Gasteiger-Klicpera, B.; Schmidt, B. (2017): Legasthenie - LRS. Modelle, Diagnose, Therapie und Förderung: mit 100 Übungsfragen. 5., überarbeitete und erweiterte Auflage. München, Basel, Stuttgart: Reinhardt.
- Kostorz, P. (2016): Bewertungsmaßstäbe und Bezugsnormen bei der Notenvergabe unter der Lupe des Schulrechts – Was ist pädagogisch sinnvoll, was juristisch möglich? In: *RdJB - Recht der Jugend und des Bildungswesens* (2), S. 270–289.
- Kruse, G. (2011): Das Lesen trainieren: Zu Kompetenzen von Leseunterricht und Leseübung. In: A. Bertschi-Kaufmann und T. Graber (Hg.): *Lesekompetenz - Leseleistung - Leseförderung. Grundlagen, Modelle und Materialien*. 4. Auflage. Zug, Seelze: Klett und Balmers Verlag; Klett-Kallmeyer, S. 176–188.
- Kuhl, J., Gebhardt, M., Bienstein, P., Käßler, C., Quinten, S., Ritterfeld, U., Tröster, H. & Wember, F. (2017). Implementationsforschung als Voraussetzung für eine evidenzbasierte sonderpädagogische Praxis. *Sonderpädagogische Förderung*, 62(4), 383-393.
- Lenhard, W. (2013): Leseverständnis und Lesekompetenz. Grundlagen - Diagnostik - Förderung. Stuttgart: W. Kohlhammer.

- Lindauer, T.; Schneider, H. (2011): Lesekompetenz ermitteln: Aufgaben im Unterricht. In: A. Bertschi-Kaufmann und T. Graber (Hg.): Lesekompetenz - Leseleistung - Leseförderung. Grundlagen, Modelle und Materialien. 4. Auflage. Zug, Seelze: Klett und Balmer Verlag; Klett-Kallmeyer.
- Marx, A. (2017): Kinder und Jugendliche mit Zuwanderungshintergrund. In: M. Philipp (Hg.): Handbuch. Schriftspracherwerb und weiterführendes Lesen und Schreiben. 1. Auflage. Weinheim, Basel: Beltz Juventa, S. 332–346.
- McElvany, N.; Kessels, Ursula, Schwabe, Franziska; Kasper, D. (2017): Geschlecht und Lesekompetenz. In: A. Hussmann, H. Wendt, W. Bos, A. Bremerich-Vos, D. Kasper, E.-M. Lankes et al. (Hg.): IGLU 2016. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich. Münster, New York: Waxmann, S. 177–194.
- McElvany, N.; Schneider, C. (2009): Förderung von Lesekompetenz. In: W. Lenhard und W. Schneider (Hg.): Diagnostik und Förderung des Leseverständnisses. Göttingen: Hogrefe, S. 151–183.
- Moosbrugger, H. (2012): Item-Response-Theory (ITR). In: H. Moosbrugger und A. Kelava (Hg.): Testtheorie und Fragebogenkonstruktion. 2., aktualisierte und überarbeitete Auflage. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg, S. 228–271.
- Moosbrugger, H.; Kelava, A. (2012): Qualitätsanforderungen an einen psychologischen Test. In: H. Moosbrugger und A. Kelava (Hg.): Testtheorie und Fragebogenkonstruktion. 2., aktualisierte und überarbeitete Auflage. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg, S. 8–25.
- Mühling, A.; Gebhardt, M.; Diehl, K. (2017): Formative Diagnostik durch die Onlineplattform Levumi. In: *Informatik Spektrum* (40, 6), S. 556–561.
- Müller, B.; Križan, A.; Hecht, T.; Richter, T.; Ennemoser, M.: Leseflüssigkeit im Grundschulalter: Entwicklungsverlauf und Effekte systematischer Leseförderung. In: *Lernen und Lernstörungen* (2), S. 131–146.
- Nation, K. (2005): Children's Reading Comprehension Difficulties. In: M. Snowling und C. Hulme (Hg.): The science of reading. A handbook. Malden, Mass.: Blackwell Publ (Blackwell handbooks of developmental psychology), S. 248–266.
- Niklas, F.; Schneider, W. (2012): Die Anfänge geschlechtsspezifischer Leistungsunterschiede in mathematischen und schriftsprachlichen Kompetenzen. In: *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie* (44, 3), S. 123–138.
- Philipp, M. (2013): Wie motiviert lese und schreibe ich? Und wenn ja: auf wie viele Arten und Weisen? Zur Klärung und Klärungsbedürftigkeit der vielschichtigen Lese- und Schreibmotivation. Hg. v. leseforum.ch. Online verfügbar unter <https://bit.ly/2LMyyIV>.
- Philipp, M. (2017a): Lese- und Schreibkompetenz. In: M. Philipp (Hg.): Handbuch. Schriftspracherwerb und weiterführendes Lesen und Schreiben. 1. Auflage. Weinheim, Basel: Beltz Juventa, S. 36–50.
- Philipp, M. (2017b): Lese- und Schreibmotivation: Begriffsklärungen, Domänenspezifik und Einflussfaktoren. In: M. Philipp (Hg.): Handbuch. Schriftspracherwerb und weiterführendes Lesen und Schreiben. 1. Auflage. Weinheim, Basel: Beltz Juventa, S. 138–155.

- Richter, T.; Müller, B. (2017): Entwicklung hierarchieniedriger Leseprozesse. In: M. Philipp (Hg.): Handbuch. Schriftspracherwerb und weiterführendes Lesen und Schreiben. 1. Auflage. Weinheim, Basel: Beltz Juventa, S. 51–66.
- Rißling, J.-K.; Mahlau, K.; Hartke, B.; Petermann, F. (2014): Effektivität schulischer Sprachförderung bei Erstklässlern mit Sprachentwicklungsstörungen. In: *Zeitschrift für Erziehungswissenschaft* (17, 3), S. 543–562.
- Rosebrock, C.; Nix, D.; Rieckmann, C.; Gold, A. (2011): Leseflüssigkeit fördern: Lautleseverfahren für die Primar- und Sekundarstufe. 1. Auflage. Seelze: Klett Kallmeyer.
- Sassenroth, M. (2000): Schriftspracherwerb. Entwicklungsverlauf, Diagnostik und Förderung. 4., unveränd. Aufl. Bern: Haupt (Beiträge zur Heil- und Sonderpädagogik, 12).
- Schneider, W. (2017): Lesen und Schreiben lernen. Wie erobern Kinder die Schriftsprache? Berlin, Heidelberg: Springer.
- Schulgesetz des Landes Nordrhein-Westfalen (SchulG NRW) vom 15.2.2005 (GV. NRW. S.102) zuletzt geändert durch das Gesetz am 13.11.2012 (GV. NRW. S.512). Online verfügbar unter: <https://bit.ly/2NDGuX3> [23.09.2018].
- Schründer-Lenzen, A. (2013): Schriftspracherwerb. 4., völlig überarb. Aufl. Wiesbaden: Springer VS.
- Stecker, P. M.; Fuchs, L. S.; Fuchs, D. (2005): Using curriculum-based measurement to improve student achievement: Review of research. In: *Psychology in the schools* (42, 5), S. 795–819.
- Strathmann, A. M.; Klauer, K. Josef (2010): Lernverlaufsdagnostik: Ein Ansatz zur längerfristigen Lernfortschrittsmessung. In: *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie* (42, 2), S. 111–122.
- Tarelli, I.; Valtin, R.; Bos, W.; Bremerich-Vos, A.; Schwippert, K. (2012): Wichtige Ergebnisse im Überblick. In: W. Bos, I. Tarelli, A. Bremerich-Vos und K. Schwippert (Hg.): IGLU 2011. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich. Münster: Waxmann, S. 11–25.
- Torgesen, J. K. (2005): Recent Discoveries on Remedial Interventions for Children with Dyslexia. In: M. Snowling und C. Hulme (Hg.): *The science of reading. A handbook*. Malden, Mass.: Blackwell Publ (Blackwell handbooks of developmental psychology), S. 521–549.
- Voss, A. (2006): Leseverständnis. Theorie und Empirie. In: *Schulmanagement* (5), S. 8–10.
- Voß, S. (2014): Curriculumbasierte Messverfahren im mathematischen Erstunterricht. Zur Güte und Anwendbarkeit einer Adaption US-amerikanischer Verfahren im deutschen Schulsystem. 1. Aufl. Saarbrücken: Südwestdeutscher Verlag für Hochschulschriften.
- Walter, J. (2008): Curriculumbasiertes Messen (CBM) als lernprozessbegleitende Diagnostik: Reliabilität und Veränderungssensibilität eines robusten Indikators zur Lernfortschrittsmessung beim Lesen. In: *Heilpädagogische Forschung* (2), S. 62–79.
- Walter, J. (2009): Theorie und Praxis Curriculumbasierten Messens (CBM) in Unterricht und Förderung. In: *Zeitschrift für Heilpädagogik* (5), S. 162–170.

- 
- Walter, J. (2010): Lernfortschrittsdiagnostik am Beispiel der Lesekompetenz (LDL): Messtechnische Grundlagen sowie Befunde über zu erwartende Zuwachsraten während der Grundschulzeit. In: *Heilpädagogische Forschung* (4), S. 162–176.
- Walter, J. (2011): Die Messung der Entwicklung der Lesekompetenz im Dienste der systematischen formativen Evaluation von Lehr- und Lernprozessen. In: *Zeitschrift für Heilpädagogik* (6), S. 204–217.
- Walter, J. (2014): Lernfortschrittsdiagnostik Lesen (LDL) und Verlaufsdagnostik sinnerfassendes Lesen (VSL): Zwei Verfahren als Instrumente einer formativ orientierten Lese-diagnostik. In: M. Hasselhorn, W. Schneider und U. Trautwein (Hg.): *Lernverlaufsdagnostik*. Göttingen, Bern, Wien, Paris: Hogrefe (Tests und Trends, Neue Folge Band12), S. 165–201.
- Wenner, I. (2018): Entwicklung von Mädchen und Jungen über acht Messzeitpunkte in der Leseflüssigkeit hinweg. Eine Studie mit der Online-Plattform Levumi. Masterarbeit (unveröffentlicht). Technische Universität Dortmund.
- Wilbert, J. (2014): Instrumente zur Lernverlaufsmessung: Gütekriterien und Auswertungsherausforderungen. In: M. Hasselhorn, W. Schneider und U. Trautwein (Hg.): *Lernverlaufsdagnostik*. Göttingen, Bern, Wien, Paris: Hogrefe (Tests und Trends, Neue Folge Band12), 281-308.
- Wilbert, J.; Linnemann, M. (2011): Kriterien zur Analyse eines Tests zur Lernverlaufsdagnostik. In: *Empirische Sonderpädagogik* (3), S. 225–242.
- Zöller, I.; Roos, J. (2009): Einfluss individueller Merkmale und familiärer Faktoren auf den Schriftspracherwerb. In: J. Roos und H. Schöler (Hg.): *Entwicklung des Schriftspracherwerbs in der Grundschule. Längsschnittanalyse zweier Kohorten über die Grundschulzeit*. 1. Aufl. Wiesbaden: VS Verlag für Sozialwissenschaften / GWV Fachverlage GmbH Wiesbaden.

## IV Abbildungsverzeichnis

Abbildung 1 Förderbausteine und Entwicklungsbereiche LEVUMI (Jungjohann et al., 2017, S.9).....	37
Abbildung 2 Ablauf der übergeordneten Studie (eigene Darstellung).....	44
Abbildung 3 Leistungszuwachs der Jungen und Mädchen im Silbentest nach 21 Wochen.....	50
Abbildung 4 Zuwachs der Lösungswahrscheinlichkeit nach 21 Wochen.....	51
Abbildung 5 Leistungszuwachs der Jungen und Mädchen im Wörkertest nach 21 Wochen.....	52
Abbildung 6 Leistungszuwachs der Jungen und Mädchen im Pseudowörkertest nach 21 Wochen.....	53
Abbildung 7 Leistungszuwachs der Jungen und Mädchen im Test zum sinnentnehmenden Lesen nach 21 Wochen .....	54
Abbildung 8 Mittelwerte der Jungen und Mädchen im Silben- und Wörkertest zum MZP11.....	56
Abbildung 9 Mittelwerte der Jungen und Mädchen des Pseudowörkertests und des Tests zum sinnentnehmenden Lesen am MZP11 .....	58
Abbildung 10 Mittelwerte der Jungen und Mädchen im Silben- und Wörkertest am MZP17.....	60
Abbildung 11 Mittelwerte der Jungen und Mädchen des Pseudowörkertests und des Tests zum sinnentnehmenden Lesen am MZP17 .....	62
Abbildung 12 Mittelwerte der richtig gelesenen Silben und Wörter über fünf MZP hinweg .....	65
Abbildung 13 Mittelwerte der Jungen und Mädchen der richtig gelösten Items im Pseudowörkertest und im Test zum sinnentnehmenden Lesen über fünf MZP .....	66
Abbildung 14 Anzahl der richtig gelesenen Silben über fünf MZP der Mädchen .....	68
Abbildung 15 Anzahl der richtig gelesenen Silben über fünf MZP der Jungen .....	69
Abbildung 16 Anzahl der richtig gelesenen Wörter über fünf MZP der Mädchen .....	70
Abbildung 17 Anzahl der richtig gelesenen Wörter über fünf MZP der Jungen .....	71
Abbildung 18 Anzahl der richtig gelesenen Pseudowörter über fünf MZP hinweg der Mädchen.....	72
Abbildung 19 Anzahl der richtig gelesenen Pseudowörter über fünf MZP hinweg der Jungen .....	73
Abbildung 20 Anzahl der richtig gelösten Items im Test zum sinnentnehmenden Lesen über fünf MZP hinweg der Mädchen .....	74
Abbildung 21 Anzahl der richtig gelösten Items im Test zum sinnentnehmenden Lesen über fünf MZP hinweg der Jungen .....	75
Abbildung 22 Verhältnis von richtig zu falsch gelösten Items aller vier Tests.....	79
Abbildung 23 Leistungszuwachs von MZP13 bis MZP17 in allen vier Tests .....	83

**V Tabellenverzeichnis**

Tabelle 1 Übersicht der Förderzeiträume nach Klassen .....	45
Tabelle 2 Übersicht der Stichprobengröße von MZP1 und MZP18 .....	49
Tabelle 3 Übersicht zum Mittelwert, der Standardabweichung und des Medians des Silbentests am MZP11.....	57
Tabelle 4 Übersicht zum Mittelwert, der Standardabweichung und des Medians des Wörtertests am MZP11 .....	57
Tabelle 5 Übersicht zum Mittelwert, der Standardabweichung und des Medians des Pseudowörtertests am MZP11.....	59
Tabelle 6 Übersicht zum Mittelwert, der Standardabweichung und des Medians des Tests zum sinnentnehmenden Lesen am MZP11.....	59
Tabelle 7 Übersicht zum Mittelwert, der Standardabweichung und des Medians des Silbentests am MZP17.....	61
Tabelle 8 Übersicht zum Mittelwert, der Standardabweichung und des Medians des Wörtertests am MZP17 .....	61
Tabelle 9 Übersicht zum Mittelwert, der Standardabweichung und des Medians des Pseudowörtertests am MZP17.....	63
Tabelle 10 Übersicht zum Mittelwert, der Standardabweichung und des Medians des Tests zum sinnentnehmenden Lesen am MZP17.....	63

## VI Anhang

## Anhang A

	Ergebnisse der vier Tests am MZP1							
	Silben		Wörter		Pseudowörter		Sinnentnehmendes Lesen	
	J	M	J	M	J	M	J	M
<b>Lösungs-wahr-scheinlichkeit</b>								
N	40	26	40	25	40	26	40	26
Min.	0	42%	0	38%	0	7%	0	15%
Max.	100%	100%	100%	100%	100%	96%	97%	96%
M	80.57%	84.43%	76.79%	74.09%	65.5%	63.45%	54.65%	51.98%
SD	28.88%	16.70%	24.88%	19.50%	28.48%	26.22%	21.72%	21.86%
Md	92.85%	92.15%	83.67%	77.78%	73.21%	59.17%	54.79%	50.15%
<b>Insgesamt gele-sene Items</b>	J	M	J	M	J	M	J	M
Min.	7	9	5	5	6	6	7	4
Max.	51	46	40	37	30	26	54	60
M	29.45	26.15	24	19.52	17.3	15.62	23.43	20.88
SD	11.505	9.711	9.795	8.496	6.040	5.224	11.075	12.323
Md	31	24	22.5	17	16.5	14.5	21	17
<b>Richtig gelesene Items</b>	J	M	J	M	J	M	J	M
Min.	0	5	0	2	0	1	0	3
Max.	50	45	40	37	30	24	46	29
M	25.73	23.08	19.75	15.64	11.98	10.81	13.7	10.38
SD	13.721	11.328	10.667	9.937	7.141	6.905	9.757	7.020
Md	27.5	21.5	20	13	12.5	8.5	11	9
<b>Falsch gelesene Items</b>	J	M	J	M	J	M	J	M
Min.	0	0	0	0	0	1	1	1
Max.	21	11	14	11	17	13	25	51
M	3.73	3.08	4.25	3.88	5.05	4.81	9.73	10.5
SD	4.461	2.938	3.514	2.571	3.823	3.099	5.866	10.542
Md	2	2.5	3.5	4	4.5	5	8	7

## Anhang B

## Ergebnisse der vier Tests am MZP18

	Silben		Wörter		Pseudowörter		Sinnentnehmen- des Lesen	
	J	M	J	M	J	M	J	M
<b>Lösungs-wahr- scheinlichkeit</b>								
N	39	29	39	29	39	29	29	27
Min.	31%	42%	17%	0%	0%	0%	5%	17%
Zuwachs↑	+31%	.	+17%	-38%	.	-7%	+5%	+2%
Max.	100%	100%	100%	100%	100%	100%	92%	100%
Zuwachs↑	.	.	.	.	.	+4%	-5%	+4%
M	86.46%	88.79%	80.80%	81.45%	71.29%	67.1%	51.13	61.98%
Zuwachs ↑	+5.89%	+4.36%	+4.01%	+7.36%	+5.76%	+3.65%	-3.52%	+10%
SD	17.02%	13.40%	19.08%	22.69%	24.98%	24.44%	23.12%	26.29%
Md	92.86%	93.48%	84.38%	90.48%	76.47%	71.43%	47.06%	67.39%
<b>Insgesamt gele- sene Items</b>	J	M	J	M	J	M	J	M
Min.	11	14	11	7	5	7	8	9
Zuwachs↑	+4	+5	+6	+2	-1	+1	+1	+5
Max.	49	51	41	39	27	29	61	61
Zuwachs↑	-2	+5	+1	+2	-3	+3	+7	+1
M	34.05	32.52	26.64	23.93	18.26	17.79	32.38	31.48
Zuwachs↑	+7.6	+6.37	+2.64	+4.41	+0.96	+2,17	+8.95	+10,6
SD	9.15	9.52	8.37	7.63	5.41	5.56	14.8	12.40
Md	36	29	26	23	20	17	31	32
<b>Richtig gelesene Items</b>	J	M	J	M	J	M	J	M
Min.	5	9	2	0	0	0	1	3
Zuwachs↑	+5	+4	+2	-2	/	-1	+1	/
Max.	49	50	40	39	27	28	42	41
Zuwachs↑	-1	+5	/	+2	-3	+4	-4	+12
M	30.38	29.48	22.13	20.31	13.51	12.66	16.21	19
Zuwachs↑	+4.65	+6.4	+2.38	+4.67	+1.53	+1.85	+2.51	+8.62
SD	11.37	10.95	9.67	9.12	6.43	7.19	9.48	10.19
Md	33	28	22	19	14	12	16	18
<b>Falsch gelesene Items</b>	J	M	J	M	J	M	J	M
Min.	0	0	0	0	0	0	2	0
Zuwachs↑	.	.	.	.	.	-1	+1	-1
Max.	15	15	21	14	22	15	58	46
Zuwachs↑	-6	+4	+7	+3	+5	+2	+33	-5
M	3.67	3.03	4.51	3.62	4.74	5.14	16.17	12,48
Zuwachs↑	-0.06	-0.05	+0.26	-0.26	-0.31	+0.33	+6.44	+1.98
SD	3.9	3.33	4.18	3.72	4.04	3.47	13.13	11.80
Md	2	2	3	2	4	4	11	8

## Anhang C

---

 Ergebnisse des t-Tests der Levumi-Tests am MZP1
 

---

	df	T	Signifikanz p
Silben	64	.618	.539
Wörter	63	-.461	.646
Pseudowörter	64	-.294	.769
Sinnentnehmendes Lesen	64	-.486	.629

*Anmerkung:* Ein positiver T-Wert steht für eine bessere Leistung zu Gunsten der Mädchen, ein negativer Wert für eine bessere Leistung der Jungen.

---

 Ergebnisse des t-Tests der Levumi-Tests am MZP18
 

---

	df	T	Signifikanz p
Silben	66	.611	.543
Wörter	66	.127	.900
Pseudowörter	66	-.689	.493
Sinnentnehmendes Lesen	54	1.642	.106

*Anmerkung:* Ein positiver T-Wert steht für eine bessere Leistung zu Gunsten der Mädchen, ein negativer Wert für eine bessere Leistung der Jungen.

## Anhang D

## Ergebnisse der vier Tests am MZP11

	Silben		Wörter		Pseudowörter		Sinnentnehmen- des Lesen	
	J	M	J	M	J	M	J	M
<b>Lösungs-wahr- scheinlichkeit</b>								
N	7	8	7	8	7	8	7	8
Min.	11%	33%	14%	33%	0%	0%	15%	19%
Max.	100%	93%	96%	94%	93%	91%	57%	87%
M	82.27%	76.04%	73.84%	66.78%	66.04%	53.29%	38.53%	44.89%
SD	31.57%	20.07%	28.62%	22.29%	30.92%	32.41%	16.20%	29.98
Md	92.86%	84.24%	82.61%	70.12%	71.43%	61.65%	42.11%	28.86%
<b>Insgesamt gele- sene Items</b>	J	M	J	M	J	M	J	M
Min.	28	9	17	6	12	4	14	9
Max.	41	46	28	29	26	22	61	54
M	34.14	26.36	22.29	17	17.71	13	44.29	30
SD	5.699	10.676	3.817	6.547	4.751	5.237	20.197	14.716
Md	36	28,5	23	17,5	17	12	57	29
<b>Richtig gelesene Items</b>	J	M	J	M	J	M	J	M
Min.	4	3	4	2	0	0	8	3
Max.	41	42	24	25	16	17	24	31
M	28	21.75	16	12.13	10.86	7.63	14.71	12.25
SD	12.342	11.461	6.583	6.937	5.669	5.476	6.157	9.161
Md	26	23.5	17	12.5	14	8	13	10
<b>Falsch gelesene Items</b>	J	M	J	M	J	M	J	M
Min.	0	2	1	1	1	1	6	3
Max.	32	9	24	10	26	14	52	44
M	6.14	4.88	6.29	4.88	6.86	5.38	29.57	17.75
SD	11.452	2.167	8.098	2.949	8.591	3.815	18.365	15.078
Md	2	4.5	4	4	4	5	33	13.5

## Anhang E

## Werte der vier Tests am MZP17

	Silben		Wörter		Pseudowörter		Sinnentnehmen- des Lesen	
	J	M	J	M	J	M	J	M
<b>Lösungs-wahr- scheinlichkeit</b>								
N	7	9	7	9	7	9	6	9
Min.	31%	60%	0%	55%	0%	8%	11%	12%
Max.	97%	97%	88%	96%	100%	86%	85%	71%
M	74.94%	80.97%	59.90%	79.14%	57.19%	59.21%	39.60%	35.66%
SD	28.22%	11.79%	34.38%	15.30%	40.50%	25.35%	26.57%	18.54%
Md	89.19%	84.85%	76.19%	84.62%	70%	71.43%	33.72%	34.88%
<b>Insgesamt gele- sene Items</b>	J	M	J	M	J	M	J	M
Min.	13	25	8	13	6	13	20	15
Max.	40	34	27	28	27	22	61	59
M	34.14	28.89	22.29	19.22	17.43	15.89	44	34
SD	9.63	3.55	6.63	4.92	6.45	3.41	15.75	14.93
Md	37	27	24	21	19	14	43.5	27
<b>Richtig gelesene Items</b>	J	M	J	M	J	M	J	M
Min.	4	15	0	8	0	1	7	3
Max.	38	30	22	27	15	19	18	36
M	27.14	23.56	14.57	15.33	10	9.78	14.7	13.11
SD	12.9	5.22	8.28	5.57	6.86	5.22	4.62	10.6
Md	33	23	19	16	14	11	16	11
<b>Falsch gelesene Items</b>	J	M	J	M	J	M	J	M
Min.	1	1	3	1	0	3	3	9
Max.	23	10	21	10	27	12	54	36
M	7	5.33	7.71	3.98	7.43	6.11	29.83	20.89
SD	7.59	2.92	6.16	3.06	8.94	3.26	18.93	9.03
Md	4	5	5	2	6	4	29.5	21

## Anhang F

## Anpassung der Erwartungswerte nach Fuchs et al. (1993)

	Silben	Wörter	Pseudowörter	Sinnentnehmendes Lesen
CBM Zuwachs pro Woche	.	+1.46	.	+0.39
SD-Zuwachs	.	+0.69	.	+0.24
Zuwachs mit Förde- rung nach vier Wo- chen	.	+8.6	.	+2.52
Angepasster Zu- wachs mit Förde- rung in vier Wochen (Zuwachs:2)	+6.66	+4.3	+3.12	+4.032*

*Anmerkung:* \*Da sich die Angaben nach Fuchs (1993) beim sinnentnehmenden Lesen auf eine Dauer von 2.5 Minuten beziehen, wurde der Wert im letzten Schritt an den achtminütigen Levumi-Test angepasst.

## Anhang G

## Individuelle Anzahl richtig gelesener Silben von MZP13 bis MZP17

	MZP13	MZP14	MZP15	MZP16	MZP17	M	Zuwachs	(Non-) Responder
Richtige Silben der Mädchen								
Am	28	24	36	31	28	294	0	NR
Ed	32	30	24	30	28	28.8	-2	NR
Em	28	26	16	27	21	23.6	-7	NR
Gu	22	27	25	23	23	24	2	NR
Ja	16	17	16	21	18	17.6	2	NR
Jo	.	6	5	9	.	6.7	3	NR
Le	37	38	40	35	28	35.6	-9	NR
Ma	21	20	14	15	15	17	-6	NR
Sa	23	31	25	26	30	27	7	R
Ta	18	27	15	22	21	20.6	3	NR
Richtige Silben der Jungen								
Ak	26	33	36	34	33	32.4	7	R
Al	29	.	37	26	32	31	3	NR
An	29	35	31	38	33	33.2	4	NR
Ar	30	35	35	35	38	34.6	8	R
Jod	26	32	20	.	.	26	-6	NR
Ra	2	8	15	14	14	10.6	12	R
Sz	42	31	35	39	36	36.6	-6	NR
Ti	5	1	.	2	4	3	-1	NR

Anmerkung: Fehlende Werte sind durch die Abwesenheit von SchülerInnen zum jeweiligen MZP entstanden.

## Anhang H

## Individuelle Anzahl richtig gelesener Wörter von MZP13 bis MZP17

	MZP13	MZP14	MZP15	MZP16	MZP17	M	Zuwachs	(Non-) Responder
Richtige Wörter der Mädchen								
Am	15	16	16	13	12	14.4	-3	NR
Ed	18	19	16	19	16	17.6	-2	NR
Em	17	16	15	22	17	17.4	0	NR
Gu	8	11	15	17	16	13.4	8	R
Ja	8	7	3	4	8	6	0	NR
Jo	.	2	1	1	.	1.3	-1	NR
Le	20	21	18	21	27	21.4	7	R
Ma	4	4	9	13	11	8.2	7	R
Sa	14	12	13	12	19	14	5	R
Ta	10	10	13	7	12	10.4	2	NR
Richtige Wörter der Jungen								
Ak	20	18	14	16	19	17.4	-1	NR
Al	16	.	14	16	19	16.25	3	NR
An	18	18	17	17	16	17.2	-2	NR
Ar	16	17	25	19	20	19.4	4	R
Jod	15	15	16	.	.	15.3	1	NR
Ra	3	4	2	4	6	3.8	3	NR
Sz	25	27	25	28	22	25.4	-3	NR
Ti	2	1	.	3	0	1.5	-2	NR

Anmerkung: Fehlende Werte sind durch die Abwesenheit von SchülerInnen zum jeweiligen MZP entstanden.

## Anhang I

## Individuelle Anzahl richtig gelesener Pseudowörter von MZP13 bis MZP17

	MZP13	MZP14	MZP15	MZP16	MZP17	M	Zuwachs	(Non-) Res-ponder
<b>Richtige Pseudowörter Mädchen</b>								
Am	10	13	14	15	12	12.8	2	NR
Ed	10	12	9	7	10	9.6	0	NR
Em	12	7	7	8	11	9	-1	NR
Gu	6	3	7	4	11	6.2	5	R
Ja	6	3	5	6	6	5.2	0	NR
Jo		0	0	0	.	0	0	NR
Le	17	14	16	16	19	16.4	2	NR
Ma	3	2	6	3	1	3	-2	NR
Sa	15	12	15	16	13	14.2	-2	NR
Ta	5	5	7	4	5	5.2	0	NR
<b>Richtige Pseudowörter Jungen</b>								
Ak	17	13	10	14	14	13.6	-3	NR
Al	11	.	12	12	15	12.5	4	R
An	13	13	.	12	13	12.75	0	NR
Ar	13	6	13	14	14	12	1	NR
Jod	11	8	14	.	.	11	3	R
Ra	0	0	0	0	0	0	0	NR
Sz	17	21	14	14	14	16	-3	NR
Ti	0	0	.	0	0	.0	0	NR

Anmerkung: Fehlende Werte sind durch die Abwesenheit von SchülerInnen zum jeweiligen MZP entstanden.

## Anhang J

Individuelle Anzahl richtig gelöster Items im sinnentnehmenden Lesen von MZP13 bis MZP17

	MZP13	MZP14	MZP15	MZP16	MZP17	M	Zuwachs	(Non-) Responder
Richtige Items der Mädchen								
Ed	16	17	5	13	3	10.8	-13	NR
Em	15	13	8	16	11	12.6	-4	NR
Gu	10	12	13	12	15	12.4	5	R
Ja	4	5	7	5	6	5.4	2	NR
Jo	.	12	14	13	.	13	1	NR
Le	38	30	33	27	36	32.8	-2	NR
Ma	6	7	6	4	3	5.2	-3	NR
Sa	7	8	7	8	10	8	3	NR
Ta	17	12	18	21	23	18.2	6	R
Richtige Items der Jungen								
Ak	9	22	18	17	18	16.8	9	R
Al	6	.	7	6	17	9	11	R
An	21	18	17	25	.	20.25	4	R
Ar	11	11	18	16	18	14.8	7	R
Jod	19	26	32	.	.	25.67	13	R
Ra	13	14	12	13	7	11.8	-6	NR
Sz	5	17	17	10	15	12.8	10	R
Ti	0	5	.	8	10	5.57	10	R

Anmerkung: Fehlende Werte sind durch die Abwesenheit von SchülerInnen zum jeweiligen MZP entstanden.

## Anhang K

## Lösungswahrscheinlichkeit im Silbentest über fünf MZP

	MZP13	MZP14	MZP15	MZP16	MZP17	M	Intervall	Abweichungen	Stabilität
<b>Lw der Mädchen</b>									
Am	90	75	95	84	85	85.80	73-99	0	stabil
Ed	100	94	86	97	85	92.40	79-100	0	stabil
Em	97	87	89	90	78	88.20	75-100	0	stabil
Gu	85	87	86	82	92	86.40	73-99	0	stabil
Ja	73	65	59	68	67	66.40	56-76.36	0	stabil
Jo	.	55	38	75	.	56	48-64	2	variabel
Le	93	97	95	90	97	94.40	80-100	0	stabil
Ma	75	80	52	60	60	65.40	56-75	2	variabel
Sa	92	91	83	87	88	88.20	75-100	0	stabil
Ta	67	93	68	92	78	79.60	68-92	1	stabil
<b>Lw der Jungen</b>									
Ak	90	94	97	97	89	93.40	79-100	0	stabil
Al	97	.	95	96	97	96.30	82-100	0	stabil
An	91	95	94	90	83	90.60	77-100	0	stabil
Ar	94	95	95	95	95	94.80	81-100	0	stabil
Jod	90	100	100	.	.	96.70	82-100	0	stabil
Ra	8	22	38	36	38	28.40	24-33	5	variabel
Sz	98	97	97	100	92	96.80	82-100	0	stabil
Ti	42	11	.	22	31	26.50	23-31%	3	variabel

*Anmerkung:* Fehlende Werte sind durch die Abwesenheit von SchülerInnen zum jeweiligen MZP entstanden. Alle Werte bis auf die Anzahl der Abweichungen sind Angaben in Prozent.

## Anhang L

	Lösungswahrscheinlichkeit im Wörtest über fünf MZP					M	Intervall	Abweichungen	Stabilität
	MZP13	MZP14	MZP15	MZP16	MZP17				
Lw der Mädchen									
Am	75	80	80	52	55	68.40	58-79	2	variabel
Ed	100	86	80	86	73	85	72-98	0	stabil
Em	94	100	88	96	94	94.40	80-100	0	stabil
Gu	53	65	75	85	76	70.80	60-81	2	variabel
Ja	47	50	25	33	57	42.40	36-49	4	variabel
Jo	.	33	13	14	.	10	8.5-12	3	variabel
Le	80	75	75	78	96	80.80	69-93	1	stabil
Ma	36	25	60	72	85	55.60	47-64	4	variabel
Sa	74	57	65	67	90	70.60	60-81	2	variabel
Ta	71	56	72	50	86	67	57-77	3	variabel
Lw der Jungen									
Ak	74	86	58	73	70	72.20	61-83	2	variabel
Al	84	.	74	80	79	79.25	67-91	0	stabil
An	90	86	81	85	76	83.60	71,1-96	0	stabil
Ar	67	74	86	76	83	77.20	66-89	0	stabil
Jod	88	88	89	.	.	88.30	75-100	0	stabil
Ra	12	14	8	15	22	14.20	12-16	2	variabel
Sz	96	90	96	93	88	92.60	79-100	0	stabil
Ti	25	17	.	75	0	29.25	25-34	3	variabel

## Anhang M

Lösungswahrscheinlichkeit im Pseudowörtertest über fünf MZP									
	MZP13	MZP14	MZP15	MZP16	MZP17	M	Intervall	Abweichungen	Stabilität
<b>Lw der Mädchen</b>									
Am	67	87	70	65	57	69.20	59-80	2	variabel
Ed	83	80	69	58	71	72.20	61-83	1	stabil
Em	75	54	58	57	79	64.60	55-74	3	variabel
Gu	40	21	39	29	73	40.40	34-47	3	variabel
Ja	46	30	45	46	46	42.60	36-49	1	stabil
Jo	.	0	0	0	.	0	0	0	stabil
Le	85	64	80	76	86	78.20	67-90	1	stabil
Ma	27	18	46	23	8	24.40	21-28	3	variabel
Sa	94	80	88	94	76	86.40	73-99	0	stabil
Ta	36	36	50	29	36	37.40	32-43	2	variabel
<b>Lw der Jungen</b>									
Ak	94	76	63	74	70	75.40	64-87	2	variabel
Al	92	.	86	92	100	92.50	79-100	0	stabil
An	81	81	.	71	87	80	68-92	0	stabil
Ar	65	38	72	64	74	62.60	53-72	2	variabel
Jod	100	67	93	.	.	86.70	73-99	1	variabel
Ra	0	0	0	0	0	0	0	0	stabil
Sz	85	95	82	82	70	82.80	70-95	0	stabil
Ti	0	0	.	0	0	0	0	0	stabil

*Anmerkung:* Fehlende Werte sind durch die Abwesenheit von SchülerInnen zum jeweiligen MZP entstanden. Alle Werte bis auf die Anzahl der Abweichungen sind Angaben in Prozent.

## Anhang N

	Lösungswahrscheinlichkeit im Test zum sinnentnehmenden Lesen über fünf MZP								
	MZP13	MZP14	MZP15	MZP16	MZP17	M	Intervall	Stabil/	variabel
<b>Lw der Mädchen</b>									
Am	54	24	74	55	55	52.40	45-60	2	variabel
Ed	55	49	17	45	12	35.60	30-41	5	variabel
Em	65	46	38	50	42	48.20	41-55	2	variabel
Gu	24	35	25	36	35	31	26-36	2	variabel
Ja	17	17	30	21	22	21.40	18-25	3	variabel
Jo	.	20	23	21	.	21.30	18-25	0	stabil
Le	79	73	79	75	71	75.40	64-87	0	stabil
Ma	35	44	40	27	20	33.20	28-38	3	variabel
Sa	13	20	22	24	25	20.80	18-24	2	variabel
Ta	30	31	33	34	39	33.40	28-38	1	stabil
<b>Lw der Jungen</b>									
Ak	31	58	51	57	51	49.60	42-57	2	variabel
Al	46	50	.	35	85	54	46-62	2	variabel
An	41	35	28	41	.	36.25	31-42	2	variabel
Ar	31	37	53	40	43	40.80	35-47	2	variabel
Jod	38	46	68	.	.	50.70	43-58	2	variabel
Ra	21	23	20	21	11	19.20	16-22	2	variabel
Sz	8	28	28	16	25	21	18-24	4	variabel
Ti	0	36	.	38	22	24	20-28	3	variabel

*Anmerkung:* Fehlende Werte sind durch die Abwesenheit von SchülerInnen zum jeweiligen MZP entstanden. Alle Werte bis auf die Anzahl der Abweichungen sind Angaben in Prozent.

## Anhang O

## Individuelle Mittelwerte von MZP14 und MZP15 der Mädchen

Schülerin	Am				Ed				Em				Gu				Ja			
	S	W	P	SE	S	W	P	SE	S	W	P	SE	S	W	P	SE	S	W	P	SE
Zuwachs MZP13-17	0	-3	2	4	-2	-2	0	-13	-7	0	-1	-4	2	8	5	5	2	0	0	2
Bearbeitete Items	35	20	17.5	24	30	21	14	32	24	16.5	12.5	24.5	30	18.5	16	43	26.5	13	10.5	26
M Richtige Items	30	16	13.5	10.5	27	17.5	10.5	11	21	15,5	7	10.5	26	13	5	12,5	16.5	5	4	6
M Falsche Items	5	4	4	13.5	3	3.5	3.5	21	3	1	5.5	14	4	5.5	11	30.5	10	8	6.5	20
M LW in %	85	80	78.5	49	90	83	74.5	33	88	94	56	42	86.5	70	30	30	62	37.5	37.5	23.5

*Anmerkung:* Für jede(n) SchülerIn werden die Daten für alle vier Tests im Mittel der MZP14 und 15 dargestellt. Dabei bedeutet S=Silben-, W=Wörter-, P=Pseudowörtertest und SE= Sinnentnehmendes Lesen. Die Zuwächse ergeben sich aus den mehr richtig gelösten Items von MZP13 bis MZP18, bei Abwesenheit errechnet sich der Zuwachs aus anderen MZP, welche im Fließtext vermerkt sind.

## Anhang P

## Individuelle Mittelwerte von MZP14 und MZP15 der Mädchen

Schülerin	Jo				Le				Ma				Sa				Ta			
	S	W	P	SE	S	W	P	SE	S	W	P	SE	S	W	P	SE	S	W	P	SE
Zuwachs MZP13-17	3	-1	0	1	-9	7	2	-2	-6	7	-2	-3	7	5	-2	3	3	2	0	6
M Bearbeitete Items	12	7	5	61	40.5	26	21	41.5	26	15.5	12	15.5	32	20.5	16	36	25.5	18	14	46.5
M Richtige Items	5.5	1.5	0	13	39	19.5	15	31.5	17	6.5	4	6.5	28	12.5	13.5	7.5	21	11.5	6	15
M Falsche Items	6.5	5.5	5	48	1.5	6.5	6	10	9	9	8	9	4	8	2.5	28.5	4.5	6.5	8	31.5
M LW in %	46.5	23	0	21.5	96	75	72	76	66	42.5	32	42	87	61	84	21	80.5	64	43	32

*Anmerkung:* Für jeden SchülerIn werden die Daten für alle vier Tests im Mittel der MZP14 und 15 dargestellt. Dabei bedeutet s=Silben-, W=Wörter-, P=Pseudowörtertest und SE= Sinnentnehmendes Lesen. Die Zuwächse ergeben sich aus den mehr richtig gelösten Items von MZP13 bis MZP18, bei Abwesenheit errechnet sich der Zuwachs aus anderen MZP, welche im Fließtext vermerkt sind.

## Anhang Q

## Individuelle Mittelwerte von MZP14 und MZP15 der Jungen

Schüler	Ak				Al				An				Ar			
	S	W	P	SE	S	W	P	SE	S	W	P	SE	S	W	P	SE
Zuwachs MZP13-17	7	-1	-3	9	3	3	4	11	4	-2	0	4	8	4	-1	7
M Bearbeitete Items	36	22.5	16.5	36.5	33	19.5	13.5	15.5	35	21	16	56	37	26	17	32
M Richtige Items	34.5	16	11.5	20	31.5	15	12	6.5	33	17.5	13	17.5	35	21	9.5	14.5
M Falsche Items	1.5	6,5	5	16.5	1.5	4.5	1.5	9	2	3.5	3	38.5	2	5	7.5	17.5
M LW in %	95.5	72	69.5	54.5	95.5	77	89	48	94.5	83.5	81	31.5	95	80	55	45

*Anmerkung:* Für jeden SchülerIn werden die Daten für alle vier Tests im Mittel der MZP14 und 15 dargestellt. Dabei bedeutet s=Silben-, W=Wörter-, P=Pseudowörtertest und SE= Sinnentnehmendes Lesen. Die Zuwächse ergeben sich aus den mehr richtig gelösten Items von MZP13 bis MZP18, bei Abwesenheit errechnet sich der Zuwachs aus anderen MZP, welche im Fließtext vermerkt sind.

## Anhang R

## Individuelle Mittelwerte von MZP14 und MZP15 der Jungen

Schüler	Jod				Ra				Sz				Ti			
	S	W	P	SE	S	W	P	SE	S	W	P	SE	S	W	P	SE
Zuwachs MZP13-17	-6	1	3	13	12	3	0	-6	-6	-3	-3	10	-1	-2	0	10
M Bearbeitete Items	26	17.5	13.5	52	38	26.5	25	61	34	28	19.5	61	10.5	7	6	8.5
M Richtige Items	26	15,5	11	29	11.5	3	0	13	33	26	17.5	17	3	1.5	0	2.5
M Falsche Items	0	2	2.5	23	26.5	23.5	25	48	1	2	2	44	7.5	5.5	6	6
M LW in %	100	88.5	80	57	30	11	0	21.5	97	93	88.5	28	26.5	21	0	18

*Anmerkung:* Für jeden SchülerIn werden die Daten für alle vier Tests im Mittel der MZP14 und 15 dargestellt. Dabei bedeutet s=Silben-, W=Wörter-, P=Pseudowörtertest und SE= Sinnentnehmendes Lesen. Die Zuwächse ergeben sich aus den mehr richtig gelösten Items von MZP13 bis MZP18, bei Abwesenheit errechnet sich der Zuwachs aus anderen MZP, welche im Fließtext vermerkt sind.

**Eidesstattliche Versicherung**

Bacha, Melina  
Name, Vorname

168056  
Matr.-Nr.

Ich versichere hiermit an Eides statt, dass ich die vorliegende Masterarbeit mit dem Titel *Geschlechterspezifische Unterschiede in der Entwicklung von Leseflüssigkeit und sinnentnehmendem Lesen hinsichtlich zusätzlicher Fördermaßnahmen im Rahmen einer Längsschnittuntersuchung in drei Grundschulklassen der Jahrgangsstufe zwei in Levumi* selbstständig und ohne unzulässige fremde Hilfe erbracht habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie wörtliche und sinngemäße Zitate kenntlich gemacht. Die Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

\_\_\_\_\_  
Ort, Datum

\_\_\_\_\_  
Unterschrift

**Belehrung:**

Wer vorsätzlich gegen eine die Täuschung über Prüfungsleistungen betreffende Regelung einer Hochschulprüfungsordnung verstößt und/oder eine falsche eidesstattliche Versicherung abgibt, handelt ordnungswidrig. Die Ordnungswidrigkeit kann mit einer Geldbuße von bis zu 50.000,00 € geahndet werden. Zuständige Verwaltungsbehörde für die Verfolgung und Ahndung von Ordnungswidrigkeiten ist der Kanzler/die Kanzlerin der Technischen Universität Dortmund. Im Falle eines mehrfachen oder sonstigen schwerwiegenden Täuschungsversuches kann der Prüfling zudem exmatrikuliert werden. (§ 63 Abs. 5 Hochschulgesetz - HG -) Die Technische Universität Dortmund wird gfls. elektronische Vergleichswerkzeuge (wie z.B. die Software „turnitin“) zur Überprüfung von Ordnungswidrigkeiten in Prüfungsverfahren nutzen. Die oben stehende Belehrung habe ich zur Kenntnis genommen:

\_\_\_\_\_  
Ort, Datum

\_\_\_\_\_  
Unterschrift