

Masterarbeit

Unterschiede in Inklusion und Förderschule

Ein Lernverlaufstest für das sinnentnehmende Lesen der Onlineplattform LEVUMI

vorgelegt von:

Dennis Hoppe

dennis.hoppe@tu-dortmund.de

Lehramt für sonderpädagogische Förderung (LABG 2009)

Matrikelnummer: **160 062**

Betreuende: Prof. Dr. Markus Gebhardt

Prof. Dr. Jörg-Tobias Kuhn

ausgegeben am: 16.01.2018

eingereicht am: 21.02.2018

I Inhaltsverzeichnis

I Inhaltsverzeichnis	I
II Abstract	III
1. Einleitung	1
1.1. Problemstellung	1
1.2. Zielsetzung	2
1.3. Vorgehensweise	3
2. Herausforderung Inklusion	5
2.1. Entstehung eines neuen Leitbildes	5
2.2. Inklusion im deutschen Schulsystem	7
2.3. Gegenwärtiger Stand schulischer Inklusion.....	9
2.4. Inklusion im Förderschwerpunkt Lernen	13
3. Lesen – theoretische Grundlagen	17
3.1. Lesekompetenzen entwickeln.....	18
3.2. Der Prozess des Lesenlernens.....	21
3.3. Sinnentnehmendes Lesen	27
4. Diagnostik im schulischen Kontext	30
4.1. Lernverlaufsdagnostik.....	30
4.2. Curriculum-based measurement (CBM)	32
4.3. Lernverlaufsdagnostik Lesen	33
4.4. Gütekriterien für Mess- und Testverfahren	35
4.5. Bezugsnormen	36
5. Fragestellung und Hypothesen	38
6. Untersuchungsmethode – LeVuMi	40
6.1. Aufbau der Onlineplattform.....	40
6.2. Konstruktion des Tests zum sinnentnehmenden Lesen	42
6.3. Darstellung und Interpretation der Ergebnisse	44
6.4. Analyse der Itemschwierigkeit	45
6.5. Berechnung der statistischen Signifikanz	47
6.6. Stichprobe	49

6.7.	Durchführung.....	50
6.8.	Ergebnisse.....	51
7.	Diskussion	66
8.	Zusammenfassung und Perspektive	78
III	Literaturverzeichnis	IV
IV	Tabellenverzeichnis	XIV
V	Abbildungsverzeichnis	XV
VI	Anhang	XVI

II Abstract

Die Umsetzung des Art. 24 der UN-Behindertenrechtskonvention (2009) bringt für das deutsche Schulsystem erhebliche Veränderungen mit sich. Eine zentrale Herausforderung des neuen Bildungssystems ist dabei die Forderung nach der gemeinsamen Beschulung aller Kinder. Dies geht mit der Erhöhung der Heterogenität der inklusiven Klassen einher und stellt Lehrkräfte im Bereich der Lernverlaufsmessung vor ein neues Aufgabenfeld. Eine Möglichkeit der Abbildung des Lernverlaufs von SchülerInnen mit und ohne sonderpädagogischen Förderbedarf bietet die Online-Lernplattform LEVUMi. Die Plattform bietet für verschiedene Unterrichtsfächer Tests an, die den Lernverlauf von Kindern bzw. ganzen Klassen darstellen können. Die vorliegende Arbeit untersucht anhand des Tests zum *sinnentnehmenden Lesen* die Lern- und Leistungsentwicklung von SchülerInnen der dritten Schulstufe an Grund- und Förderschulen. Dabei werden insbesondere die Entwicklungen der Kinder mit dem Förderschwerpunkt Lernen untersucht. Mögliche Erklärungsansätze für Unterschiede im Lernverlauf werden diskutiert.

The implementation of article 24 of the UN-Convention of the Rights of Persons with Disabilities (2009) results in significant changes for the German school system. A crucial challenge for the new education system is the call for joint teaching of all children. This includes an increase of the heterogeneity of inclusive school classes which presents a new task area for teaching staff dealing with the evaluation of the learning process. One way of portraying the learning process of students with and without special educational needs is given by the online learning platform LEVUMi. It provides tests for various subjects which allow for the depiction of the learning process of individual children or whole classes. This paper will examine the development of learning and reading skills of third grade students at primary schools and special schools by using LEVUMi's test on reading comprehension. The focus will be laid on the analysis of the development of children with learning disabilities. Various explanatory approaches for the differences in the process of learning will be discussed as well.

1 Einleitung

1.1 Problemstellung

Um der Diskriminierung von Menschen mit einer Beeinträchtigung¹ entgegenzuwirken und den Abbau von Barrieren zu beschleunigen, statuierte die *UN-Behindertenrechtskonvention* (2006, UN-BRK) den Art. 24 – das Recht auf *inklusive Bildung* (vgl. Mißling & Ückert 2015). Die Einführung eines inklusiven Bildungssystems stellt die Schule allerdings vor neue Herausforderungen. Das veränderte Bildungssystem ist durch eine hohe Selektivität gekennzeichnet, die zugleich Ausdruck sozialer Ungleichheit ist. Die neue Schulstruktur stellt die methodische und didaktische Gestaltung von Unterricht immer wieder neu auf die Probe. Aber auch das frühzeitige Erkennen schulischer Defizite im Rahmen einer immer differenzierten Unterrichtsgestaltung wurde immens erschwert, denn es ist davon auszugehen, dass die individuellen Lernleistungen der SchülerInnen deutlich unterschiedlicher sein werden als noch im separierten System. Hinzu kommen die ungelösten Strukturprobleme, die die Realisierungschancen einer inklusiven Schule markieren, die nicht nur Kinder mit Beeinträchtigung, sondern alle von Randständigkeit bedrohten SchülerInnen mit einbezieht (vgl. Ellger-Rüttgardt 2016).

Diese veränderten Strukturen der Schule und die Heterogenität der Schülerschaft stellten in der Folge auch an die gängigen diagnostischen Verfahren neue Anforderungen. Die Diagnostik im schulischen Alltag hat das Erfordernis zu erfüllen, SchülerInnen nicht nur auf ihren momentanen Leistungsstand hin zu überprüfen, sondern sie auch in eine individuelle Entwicklung einzuordnen, um folglich Rückschlüsse über gegenwärtige schulische Überwältigungen sowie mögliche benötigte Unterstützungsformen zu ziehen. Es geht also um die Erfassung und Sichtbarmachung von Lernprozessen. Das Erfassen stellt aufgrund der Heterogenität jedoch gewisse Anforderungen, da diese Lernprozesse nicht als statistisches Merkmal zu betrachten sind, sondern vielmehr als Veränderung im Verhaltenspotenzial einer Person. Demzufolge sind, um die Veränderung in einem Verhaltenspotenzial zu messen, nicht nur alter-

In dieser Arbeit wurde darauf geachtet, den Ausdruck „Behinderung“ nicht primär oder nur in direkten Zitaten zu verwenden und durch den Begriff „Beeinträchtigung“ zu ersetzen. Der Begriff Behinderung ist gesellschaftlich zu stark determiniert und gleicht einer Stigmatisierung.

native diagnostische Wege erforderlich (vgl. Börnert 2014), sondern auch eine wiederholte Messung in mindestens zwei Situationen mit identischen Anforderungen. Der Fokus stellt also die Veränderung dar, die Aufschluss darüber geben soll, in welchem Maß eine Person von dem, was zwischen den Messungen war, profitiert hat (vgl. Wilbert 2014). Eine Lösung für diese Problemstellung stellt der Ansatz der *Lernverlaufsdagnostik* dar. Mit ihrer Hilfe können Aussagen über den gegenwärtigen Zustand schulischer Leistungen getroffen werden, um präventiv auf schulische Leistungen zu reagieren. Daher leistet die Lernverlaufsdagnostik einen wichtigen Beitrag zum Gelingen schulischer Inklusion (vgl. Börnert 2014).

Im Rahmen dieser Arbeit wurde daher eine Studie mittels der Onlineplattform LEVUMi (**Lern-Verlaufs-Monitoring**) durchgeführt. Die Online-Plattform www.levumi.de ist ein langfristig angelegtes gemeinsames Forschungsprojekt der Wissenschaftler MARKUS GEBHARDT (Technische Universität Dortmund), KIRSTEN DIEHL (Europa-Universität Flensburg) und ANDREAS MÜHLING (Universität Kiel). Das Projekt stellt auf seiner Online-Plattform mehrere Tests zu verschiedenen Kompetenzbereichen zur Verfügung. Der für dieses Forschungsvorhaben genutzte Test der Plattform ist der Test zum *sinnentnehmenden Lesen*.

1.2 Zielsetzung

Die Zielsetzung dieser Arbeit besteht aus zwei Teilen. Den ersten und kleineren Teil bildet die Berechnung der Itemschwierigkeit des Tests zum sinnentnehmenden Lesen der Onlineplattform LEVUMi. Die einzelnen Items sollen auf ihre Schwierigkeit hin verglichen werden, um so gegebenenfalls ungeeignete Items aus dem Aufgabenpool aussortieren zu können. Zudem soll eine generelle Aussage hinsichtlich des Schwierigkeitsgrades getroffen und dessen Tauglichkeit für Kinder der dritten Klassen bewertet werden. Bei den Stichprobenteilnehmern handelt sich dabei sowohl um Kinder mit als auch ohne sonderpädagogischen Förderbedarf (SPF), die zur Zeit der Untersuchung zum einen inklusiv und zum anderen an Förderschulen unterrichtet wurden.

Der zweite Teil dieser Arbeit befasst sich mit den Ergebnissen des Tests zum sinnentnehmenden Lesen der Kinder mit Förderbedarf, insbesondere denen mit dem Förderschwerpunkt Lernen. Dabei wird zwischen SchülerInnen mit Förderbedarf im

gemeinsamen Unterricht und denen an einer Förderschule differenziert. Ein weiterführendes Ziel soll dabei sein, zu untersuchen, ob die Form der Beschulung Auswirkungen auf das Leistungsniveau hat. Es werden also die Merkmalsausprägungen der Inklusion untersucht. Dabei stellt sich vor allem die Frage, welcher Weg zielführender ist – der inklusive oder der exklusive.

Im Zentrum der Diskussion steht also die Fragestellung, inwieweit sich die schulischen Leistungen von Kindern mit Förderbedarfen durch eine inklusive Beschulung im Vergleich zu jenen der entsprechenden Schüler entwickeln, die eine Förderschule besuchen. Die Ergebnisse der Untersuchung sollen zudem der Weiterentwicklung des Testverfahrens dienen.

1.3 Vorgehensweise

Um eine Grundlage für die Auswertung und Beurteilung der Ergebnisse dieser Untersuchung zu schaffen, geben die Kapitel 2 bis 4 einen theoretischen Überblick über relevante Themen. Dabei sind die drei Themen *Inklusion*, Lernbereich *Lesen* und *Lernverlaufsdagnostik* zentral.

Zum Thema Inklusion wird zum einen die historische Entwicklung dargestellt und zum anderen der aktuelle Stand der schulischen Inklusion. Die Darstellung des gegenwärtigen Standes bezieht sich im letzten Teil des Kapitels vor allem auf den Förderschwerpunkt Lernen, da dieser Bereich bisher am meisten erforscht wurde (vgl. Dederling 2016). Das Kapitel zum Lernbereich Lesen betrachtet den Prozess des Lesenslernens und die Entwicklung von Lesekompetenzen, insbesondere bezogen auf den Entwicklungsbereich des sinnentnehmenden Lesens. Den letzten Teil der theoretischen Grundlagen dieser Arbeit bildet die Lernverlaufsdagnostik. Dabei werden allgemeine Verfahren und Verfahren bezüglich des Lernbereichs Lesen gegeben.

Nach der Darstellung der theoretischen Grundlagen erfolgt die Formulierung der Fragestellungen und Arbeitshypothesen. Die Reihenfolge der Forschungsfragen soll dabei auf das Ziel dieser Arbeit hinführen.

Das nachfolgende Kapitel stellt die Untersuchungsmethode dieser Arbeit vor. Dabei werden sowohl die Onlineplattform LEVUMI allgemein, deren Auswertungsmöglichkeiten und Darstellungsformen, die Konstruktion des Tests zum sinnentnehmen-

den Lesen und die Analyse der Itemschwierigkeit als auch die Berechnung der statistischen Signifikanz beschrieben. Hinzu kommen die Beschreibung der Stichprobe, die Durchführung des Tests und die anschließende Ergebnisdarstellung. Die Ergebnisse werden schließlich im Diskussionsteil gegenübergestellt.

2 Herausforderung Inklusion

Der Begriff *Inklusion* (lat. inclusio = Einschließung, Einschluss) bedeutet vom Wortursprung her „Enthaltensein in etwas“ (vgl. Wember, Stein & Heimlich 2014).

2.1 Entstehung eines neuen Leitbildes

Die Geschichte des Umgangs von Menschen mit Beeinträchtigungen lässt sich in den Kategorien *Inklusion* und *Exklusion* beschreiben. In der Systemtheorie nach LUHMANN (1994) wird der Begriff Inklusion als Gegenbegriff zu Exklusion angesehen. In der Exklusionspolitik bildete der Faschismus, insbesondere in seiner deutschen Ausprägung, den bisherigen Höhepunkt. Der geschichtliche Verlauf der Inklusion stellt sich allerdings anders dar.

Auch wenn dieser zunächst in der Soziologie entstandene systemtheoretische Begriff allgemein zu gesellschaftlichen Bezügen von Beeinträchtigung verwendet wurde, ist er bedeutungsgeschichtlich von einem bildungsbezogenen Inklusionsbegriff zu unterscheiden. Erst durch den erstmalig in den 1980er-Jahren im nordamerikanischen Raum auftretenden Ausdruck „inclusion“ konnte eine Veränderung verzeichnet werden. Die neue Determination zielte auf die uneingeschränkte Gemeinsamkeit von SchülerInnen mit und ohne „special educational needs“ ab und intendierte damit eine Veränderung schulischer Strukturen (vgl. Biewer & Schütz 2016). Vorherrschend sind die zu diesem Zeitpunkt zunehmend divergierenden Begriffsbestimmungen unterschiedlicher bildungspolitischer Akteure, denn fast zeitgleich finden sich die Begriffe „inclusive education“ und „inclusive school“ in Publikationen der britischen Pädagogik wieder (vgl. Biewer & Fasching 2012). Erst ab dem Jahr 2000 fand der Begriff „inclusion“ mit dem Lehnwort Inklusion dann auch Einzug in deutschsprachige Fachdiskussionen (vgl. Biewer & Schütz 2016). Auf nationaler Ebene wurde der Begriff aber unterschiedlich interpretiert. Für LUHMANN (1994) hat Inklusion keine normative Dimension sowie keine unmittelbaren Bezüge zu Bildungszusammenhängen einer der Sonderpädagogik zugeordneten Gruppe von Kindern und Jugendlichen. SPECK (2010) definiert Inklusion als einen Zustand des sozialen Eingeschlossenseins, als Zugehörigkeit zu einer Gemeinschaft und ein Einbezogenensein in lebensrelevante Kommunikationszusammenhänge. Die Autoren HÄCKER & WALM (2015) de-

finieren Inklusion als „[...] menschenrechtsbegründete[s] Programm einer umfassenden, alle gesellschaftlichen Bereiche einschließenden De-Segregation ausgegrenzter, an gesellschaftlicher Teilhabe benachteiligter Individuen und Gruppen“ (S. 11). An der Verbreitung des bildungspolitischen Begriffs auf internationaler Ebene hatte die Erklärung der UNESCO-Konferenz (1994) maßgeblichen Anteil (vgl. Biewer & Schütz 2016). Die darauf im Jahr 2006 beschlossene und 2009 bundesweit in Kraft getretene UN-BRK stellte einen entscheidenden Wendepunkt in einer seit Ende des Zweiten Weltkrieges geführten Menschenrechtsdebatte dar (vgl. Prengel 2016). Inklusion war fortan ein allumfassender und nicht allein auf Menschen mit einer Beeinträchtigung zu beziehender Begriff; er gilt für all jene, die von Bildungsbenachteiligung und -ausschluss bedroht sind (vgl. Ellger-Rüttgardt 2016).

Die UN-BRK (2006) formuliert in 50 Artikeln Ziele für alle Lebensbereiche. Die Schwerpunkte liegen dabei auf Fragen der Nichtdiskriminierung, der Barrierefreiheit im öffentlichen Raum und der rechtlichen Gleichheit sowie auf Fragen zu Frauen mit Behinderungen, zum Schutz vor sexueller Ausbeutung, zum Schutz der Privatsphäre, zum Wohnen, zur Arbeit und zum sozialen Schutz. Für den Bereich Bildung sind die Artikel 4, 7 und insbesondere 24 bedeutsam. Sie spiegeln die Bereiche Partizipation, Gleichberechtigung und inklusive Bildung wider. Artikel 24 (2) führt bezüglich Bildung aus: „An inclusive education system at all levels and lifelong learning“ sei zu gewährleisten, damit „persons with disabilities can access an inclusive, quality and free primary education and secondary education on an equal basis with others in the communities in which they live“ (Committee on the Rights of Persons with Disabilities 2018; OHCHR). Dieses hier für die Inklusion propagierte Ziel der UN-BRK wird aber durch unterschiedliche Deutungen relativiert. Für Deutschland zeigte sich diese Relativierung in der Übersetzung des Begriffs „inclusion“ mit Integration, weshalb die offizielle Übersetzung des englischen Textes ins Deutsche bei Organisationen für die Vertretung von Menschen mit Beeinträchtigung und bei Fachverbänden nach wie vor strittig ist. WOCKEN (2010) hat diesbezüglich verschiedene Interpretationsmöglichkeiten anhand deutschsprachiger Literatur aufgelistet und sie gleichzeitig mit den Interpretationen des Integrationsbegriffs in Verbindung gesetzt. Durch die Anwendung des Integrationsbegriffs leitet er verschiedene Deutungsvarianten des Inklusionsbegriffs ab. So sieht er im Verständnis von Inklusion z. B. Gleichsetzung mit Integration, optimierte oder erweiterte Integration oder auch visionäre Formen von Integration als

Deutungsformen. Ungeachtet der verschiedenen Interpretationen plädiert WOCKEN aber für die Übernahme des Begriffs Inklusion anstelle des bisherigen Integrationsbegriffs. Entscheidend ist für ihn, dass sich die Inklusion auch im internationalen Sprachgebrauch mittlerweile durchgesetzt hat.

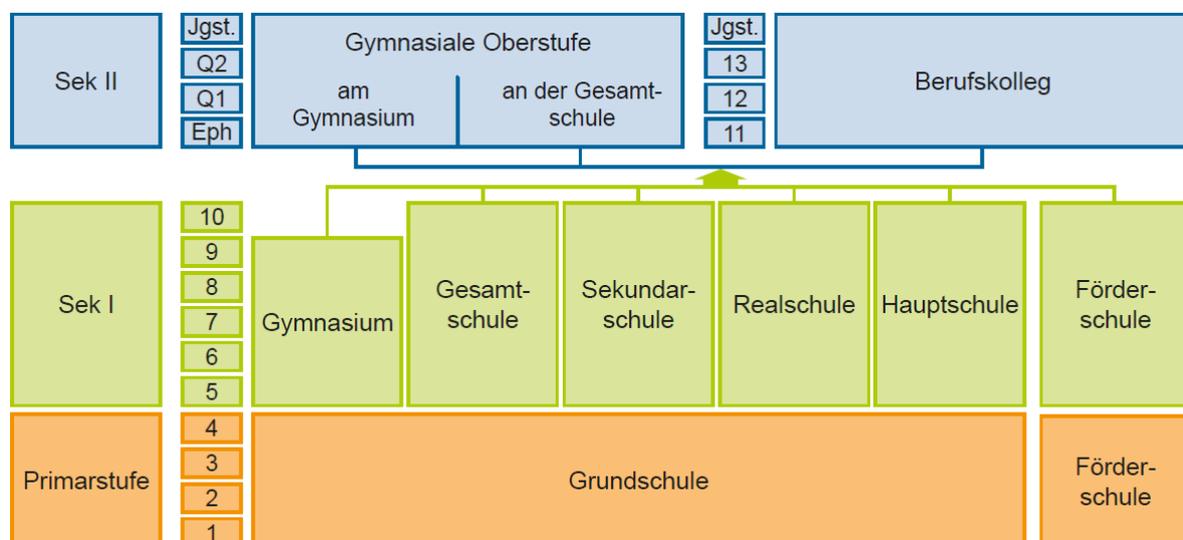
Die unreflektierte Übernahme des Begriffs Inklusion in Kombination mit der fehlenden Angabe der definitorischen Basis führte somit auch zu Missverständnissen in der Bildungspolitik (vgl. Biewer & Fasching 2012). So gibt es Vertreter, die davon überzeugt sind, das gegenwärtige Bildungssystem sei bereits zureichend inklusiv, weil es schon jetzt Bildung für Kinder mit Beeinträchtigung gewährleiste, eben in sogenannten Sonder- oder Förderschulen. Für andere gilt Inklusion erst dann als realisierbar, wenn zuvor das selektive, insbesondere das in der Sekundarstufe gegliederte Bildungssystem vollständig umgestaltet werde (vgl. Preuss-Lausitz 2013). Dementsprechend kann Inklusion nicht als Zustand, sondern vielmehr als Entwicklungsprozess eines Bildungssystems verstanden werden, in dem die Rechte der SchülerInnen vorherrschend sind (vgl. Rioux 2014). Nicht zuletzt wird Inklusion auch auf die gezielte Unterstützung von Begabungen und besonderen Talenten bezogen (vgl. Preuss-Lausitz 2015). Damit schließt letztlich „inclusive education“ alle SchülerInnen ein und damit weit mehr als die von der UN-BRK (2006) genannten „persons with disabilities“. Ungeachtet dessen wäre es aber ein Missverständnis, anzunehmen, spezielle Erziehungs- und Bildungssysteme seien jetzt völlig abzuschaffen. So wird auf die Sicherstellung von speziellen Angeboten (bspw. Gebärdensprache, Brailleschrift, unterstützende Kommunikation etc.) weiterhin Wert gelegt (UN-BRK 2006, Art. 24, 2; Theunissen 2013).

2.2 Inklusion im deutschen Schulsystem

Die Ratifizierung der UN-BRK (2009) hat dazu geführt, dass die einzelnen Bundesländer gravierende Veränderungen im Schulsystem anstreben mussten. Dabei verfolgen die jeweiligen Bundesländer allerdings verschiedene Konzepte. Die gemeinsame Beschulung von Kindern mit und ohne Beeinträchtigung war dabei das einvernehmliche Ziel. Demzufolge gilt eine „spezielle“ Beschulung aktuell als im besonderen Maße begründungspflichtig und wird eher als Ausnahme denn als Regelfall angesehen (vgl. Ahrbeck 2016).

Das Schulsystem im Bundesland Nordrhein-Westfalen ist nach Schulstufen aufgebaut und in Schulformen gegliedert. Die zurzeit gegebenen Schulstufen sind die Primarstufe, die Sekundarstufe I und die Sekundarstufe II. Die Grundschule umfasst die Klassen 1 bis 4 und bildet die gemeinsame Grundstufe des Bildungswesens. Die Sekundarstufe I umfasst an Sekundar-, Gesamt-, Real- und Hauptschulen die Klassen 5 bis 10, an Gymnasien als G8-Bildungsgang die Klassen 5 bis 9. Nach Abschluss der Sekundarstufe I haben die SchülerInnen die Möglichkeit, an das Gymnasium oder die Gesamtschule zu wechseln, um dort das Abitur (Sekundarstufe II) zu machen. Das Abitur kann zudem an berufsbildenden Schulen, dem Berufskolleg oder an einem Weiterbildungskolleg erworben werden (MSW NRW 2018).

Abbildung 1: Das Schulsystem in Nordrhein-Westfalen (IT.NRW 2016, S. 6)



Neben den allgemeinbildenden Schulen steht das System der Förderschulen, das bereits im Primarbereich ansetzt. Förderschulen werden von Kindern besucht, die gemäß § 19 Abs. 1 Schulgesetz NRW besondere Bedarfe haben und gesonderte Unterstützungsleistungen benötigen. Orte der sonderpädagogischen Förderung sind zum einen allgemeine Schulen, an denen SchülerInnen mit sonderpädagogischem Förderbedarf (SPF) gemeinsam mit anderen SchülerInnen unterrichtet werden (Inklusion), und zum anderen Förderschulen im Bereich Grund-, Haupt-, Realschule, Gymnasium, Freie Waldorfschule und Berufskolleg. Förderschulen gliedern sich nach den Förderschwerpunkten Lernen, Sprache, emotionale und soziale Entwicklung, Hören und Kommunikation, Sehen, geistige Entwicklung, körperliche und motorische Entwicklung. Alle Förderschulen (außer Lernen oder geistige Entwicklung) führen zu

denselben Bildungsabschlüssen wie die Regelschulen (vgl. Information und Technik Nordrhein-Westfalen 2016; IT.NRW). Die Schule mit Förderschwerpunkt Lernen führt zum „Abschluss des Bildungsgangs im Förderschwerpunkt Lernen“ (ebd. S. 43), allerdings kann dort auch ein dem Hauptschulabschluss (nach Klasse 9) gleichwertiger Abschluss erreicht werden (vgl. ebd.).

Verschiedene Bundesländer haben zur Realisierung der UN-BRK (2009) im Bereich Bildung ihre Schulgesetze und Verordnungen diesbezüglich geändert (bspw. NRW 9. Schulrechtsänderungsgesetz 2013). So lassen einige Bundesländer bestimmte Förderschulformen (insbesondere Förderschulen Lernen, emotionale und soziale Entwicklung und Sprache) auslaufen bzw. schließen sie oder planen dies. Dadurch bedingt ist die Anzahl der Förderschulen seit dem Schuljahr 2010/11 in Nordrhein-Westfalen um insgesamt 21,5 % zurückgegangen, die der SchülerInnen an dieser Schulform um 20,7 %. Aber nicht nur die Umsetzung der Inklusion hat dazu geführt, dass SchülerInnen mit SPF verstärkt eine Regelschule besuchen, sondern auch die Entwicklung der Gesamtschülerschaft durch den demografischen Wandel, der eine Minderung um 8,5 % seit 2010/11 bewirkt hat. Die Zahl der Schulen insgesamt ist um 6,9 % gesunken (ebd.).

2.3 Gegenwärtiger Stand schulischer Inklusion

Bei der Anwendung der UN-BRK (2009) auf das deutsche Bildungswesen ist das oberste Ziel die Einlösung des Rechts eines jeden beeinträchtigten Menschen auf Bildung. Seit jeher ist der Inklusionsbegriff mit dem Ziel verbunden, allen verschiedenen Angehörigen einer Generation angemessene Grundbildung an einem wohnortnahen, gemeinsamen Ort zukommen zu lassen (vgl. Prengel 2016). Demnach ist inklusive Bildung ein „Transformationsprozess“ (UNESCO 2010, S. 4), der vorrangig zum Ziel hat, dass Schulen und andere Lernzentren alle Kinder – mit oder ohne Beeinträchtigung – aufnehmen und für alle gleichermaßen Lerngelegenheiten schaffen. Zudem ist das Ziel von inklusiver Bildung, Exklusion zu vermeiden. Denn diese entsteht vor allem „[...] durch negative Einstellungen und mangelnde Berücksichtigung von Vielfalt in ökonomischen Voraussetzungen, sozialer Zugehörigkeit, Ethnizität, Sprache, Religion, Geschlecht, sexueller Orientierung und Fähigkeiten. Bildung voll-

zieht sich in formalen und non-formalen Kontexten, in Familien und in den Gemeinden“ (UNESCO 2010, S. 4). Mit Inklusion ist daher ein Vorgang gemeint, der versucht, alle Individuen umfassend in die Gesellschaft einzubinden, sodass eine aktive Teilhabe ermöglicht wird. Inklusion erhält in diesem Verständnis eine prozessuale Bedeutung; das bedeutet, dass sich das Recht auf Bildung an dem Ziel der Teilhabe orientiert (vgl. Biermann & Pfahl 2016). Der Anspruch auf dieses Recht stellt gegenwärtig aus deutscher Sicht zwar eine Selbstverständlichkeit dar, ist aber dennoch auf eine Entwicklung der letzten 50 Jahre zurückzuführen.

Bereits 1994 hatte sich die Kultusministerkonferenz (KMK) in den „Empfehlungen zur sonderpädagogischen Förderung in den Schulen in der Bundesrepublik Deutschland“ dafür ausgesprochen, dass ein spezieller Förderbedarf nicht mehr zwangsläufig zu einer Sonderbeschulung führen müsse. Die Integration von Kindern mit besonderen Bedarfen wurde zum vornehmlichen Ziel erklärt (vgl. Ahrbeck 2016). Der Perspektivwechsel, den die KMK zu diesem Zeitpunkt in der bundesdeutschen Sonderpädagogik initiiert hat, hätte nicht grundlegender sein können. Von der Behinderung als individuellem Mangel richtet sich nunmehr der sonderpädagogische Blick auf die Möglichkeiten der Förderung. Nach mehr als 20 Jahren der Veröffentlichung der Empfehlungen stellt sich die Frage, wie sich dieser Minimalkonsens zwischen den Bundesländern bildungspolitisch ausgewirkt hat. Festzuhalten ist, dass die Empfehlungen zu einer Pluralisierung der Orte und Konzepte sonderpädagogischer Förderung beigetragen haben. Es bildete sich eine Vielfalt an innovativen Organisationsformen, die über Förderzentren sowie Förderklassen in der allgemeinen Schule bis hin zu Integrationsklassen reichen und bis heute noch Bestand haben. In der Folge hat sich das sonderpädagogische Kompetenzprofil nachhaltig gewandelt (vgl. Heimlich 2012). Die sich daran anschließende, bereits oben dargestellte Entwicklung wurde allerdings nicht nur von Inklusion und dem demografischen Wandel beeinflusst, sondern auch von einer ganzen Reihe weiterer Parameter, u. a. der soziodemografischen Zusammensetzung der Bevölkerung, die mit unterschiedlichen sozialen Belastungen einhergeht. Für die Genese von Lern-, Sprach- oder Verhaltensstörungen ist dies ein bedeutendes Faktum. Jedoch auch bei körperlichen und Sinnesbeeinträchtigungen haben sich soziale Faktoren als nicht einflusslos erwiesen (vgl. Ahrbeck 2016). Die *European Agency for Development in Special Needs Education* (EADSNE 2009) und

das *Deutsche Institut für Menschenrechte* (Kroworsch 2017) haben in ihren Empfehlungen darauf plädiert, dass eine Inklusion nur gelingen kann, wenn sie von politischen Bildungsträgern unterstützt wird. Die relevanten international geltenden Dokumente plädieren alle auf eine Stärkung der rechtlichen Stellung von Menschen mit Behinderung im Sinne der Selbstbestimmung und gesellschaftlichen Teilhabe. Hierzu zählt das Recht auf Bildung und Teilhabe am Schulsystem. PIJL und FRISSEN (2009) verweisen allerdings auf die Schwierigkeiten der direkten Beeinflussung von Bildungseinrichtungen von außen. Direkte Vorgaben in einer Top-down-Form durch Politik und Bildungsadministration führen häufig nur dazu, dass Schulen dazu gezwungen sind, nur so zu tun, als würden sie diese Vorgaben erfüllen, ohne dass sich in der tatsächlichen Unterrichtspraxis etwas ändert. KATZENBACH und SCHNELL (2012) führen dazu weiter aus, dass dieser Zwang einer Veränderung in den verschiedenen Ländern dazu geführt hat, möglichst viele Entscheidungskompetenzen den Kommunen, Schulen und den vor Ort arbeiteten Akteuren zuzuschreiben. PIJL und FRISSEN (2009) sind der Auffassung, dass die Politik dies zwar von den LehrerInnen verlangen kann, sie aber dennoch wissen lassen sollte, dass sie mit den neuen Anforderungen nicht alleine gelassen werden. Denn eine erfolgreiche Verankerung einer inklusiven Schule ist maßgeblich abhängig von der Zustimmung aller am Prozess beteiligten Personen (vgl. Ellger-Rüttgardt 2016). Der Umsetzungsstatus der Inklusion wird in Deutschland regelmäßig von einem Monitoring erfasst. Das von der Bundesregierung diesbezüglich beauftragte Gremium ist das *Institut für Menschenrechte* in Berlin (vgl. Preuss-Lausitz 2015).

Unterricht in dem Bestreben um Inklusion rechnet vorab mit großer Heterogenität der SchülerInnen innerhalb einer Klasse. Diese Bandbreite kann von hochbegabt bis zu einem sonderpädagogischen Unterstützungsbedarf reichen. Unterricht, der diese Verschiedenheit produktiv zu nutzen versucht, bewegt sich demnach in dem Spannungsfeld von Gemeinsamkeit und Individualität. Daher ist es sehr schwierig, alle Kinder einer inklusiven Klasse gleichermaßen zu erreichen; eine zieldifferente Beschulung ist die Folge (vgl. Rehle 2013). Allerdings stellt inklusiver Unterricht an einer Regelschule für Kinder mit einem SPF eine im Vergleich zum separierten Lernsetting anregende Lernumgebung dar. Das größere Anregungspotenzial ergibt sich nach WOCKEN (2007) zum einen aus den institutionellen Rahmenbedingungen beider

Schularten und zum anderen durch das größere Potenzial aus der sozialen und leistungsbezogenen Schülerzusammensetzung der Klassen und Arbeitsgruppen. Denn gerade die leistungsschwächeren SchülerInnen profitieren von der Leistungsheterogenität, weil in den Lern- und Arbeitsgruppen leistungsstarke Kinder als Vorbilder fungieren und meist auch die Leistungserwartungen der Lehrkräfte an jene SchülerInnen höher sind (vgl. Kocaj et al. 2014). Inwieweit Letzteres als gut zu bewerten ist, ist fraglich. Allerdings profitieren auch Kinder ohne einen sonderpädagogischen Förderbedarf von einer gemeinsamen Beschulung, denn das inklusive Setting stellt eine anregungsreiche Lernumgebung dar. Vor allem leistungsstarke und sehr leistungsschwache SchülerInnen ziehen daraus einen Vorteil. Inklusiver Unterricht stellt somit für alle Lernenden einen sozialen Lern- und Erfahrungsraum dar (vgl. Dederling 2016).

Längst entscheiden nicht mehr schuladministrative Aspekte über den zukünftigen Schulort eines Kindes, sondern Eltern, Kinder und Jugendliche haben ein grundlegendes Recht bei der Auswahl der Schule. Das bedeutet, dass z. B. eine Grundschule allen Kindern mit einer Beeinträchtigung in ihrem Einzugsgebiet zugänglich sein muss und den Auftrag hat, ihnen einen qualitativ hochwertigen Unterricht anzubieten. Die UN entwarf damit das Bild eines Bildungswesens, das prinzipiell allen Menschen mit einer Beeinträchtigung Zugang zu einem System gewähren soll, das ihnen zuvor verwehrt wurde. Dieses Recht wird in der UN-BRK (2009) als ein menschenrechtlich verbürgter Anspruch festgeschrieben. Der Weg zu einem inklusiven Bildungssystem ist also kein auf Anhieb revolutionärer Weg, sondern er ist durch sukzessive Veränderungen gekennzeichnet. Die entscheidende Treibkraft stellt dabei das bereits genannte Wahlrecht der Eltern und Erziehungsberechtigten dar. Deshalb muss eine Schule der Zukunft, die eine „Schule für alle“ sein will, auf die unterschiedlichen Bedürfnisse und Unterstützungsbedarfe der SchülerInnen eingehen, um die Attraktivität zu steigern und eine bessere Alternative zu bieten als das bestehende Schulsystem (vgl. Wagner 2012).

Wenn gegenwärtig in der Debatte um eine „richtige“ oder „angemessene“ Beschulung für Kinder und Jugendliche mit einer Beeinträchtigung der Artikel 24 der UN-BRK zitiert wird, dann dient dies vorrangig als Referenzmodell für die Forderung nach einer Schule für alle, gleich ob mit oder ohne Beeinträchtigung, somit eine *inklusive Schule* (vgl. Ellger-Rüttgardt 2016). An keiner Stelle der Konvention findet sich hingegen die Forderung nach Abschaffung von Förderschulen – eine Feststellung, die

angesichts fälschlicher Interpretationen und Behauptungen in der deutschen Debatte um die inklusive Schule zu unterstreichen ist, denn in Art. 24, Abs. 2b (UN-BRK 2009) ist lediglich ausgeführt, dass die Vertragsstaaten SchülerInnen mit einer Beeinträchtigung den Zugang zu einer inklusiven und wohnortnahen Schule ermöglichen müssen (Ellger-Rüttgardt 2016). HILLENBRAND (2013, S. 366) stellte deshalb heraus: „Der Auftrag zur Etablierung eines inklusiven Bildungssystems wird nach UN-Konvention und der Begrifflichkeit der UNESCO [...] keinesfalls durch die Auflösung der Förderschulen und die Aufnahme aller Schüler mit Behinderung in die Allgemeine Schule erfüllt, sondern durch die Erfüllung der Bedürfnisse aller Lernenden“. Der gemeinsame Unterricht von SchülerInnen mit und ohne sonderpädagogischen Förderbedarf findet inzwischen an den meisten Schulen in Deutschland statt. Dennoch weist Deutschland in Anbetracht der sinkenden Zahlen an Förderschulen im europäischen Vergleich eine besonders hohe Förderschulbesuchsquote (3,72 %) auf (EADSNE 2012), auch wenn diese 2012 noch höher (4,3 %) war. Übertroffen wird die Bundesrepublik dabei nur noch von den beiden Landesteilen Belgiens (6,44 % und 7,06 %) sowie der Slowakei (4,13 %) (ebd.). Einer der Gründe dürfte darin liegen, dass in vielen Ländern Kinder mit einer Beeinträchtigung schulisch deutlich weniger erfasst werden, wie etwa in Schweden, Luxemburg, Spanien, Italien oder England (ebd.). Allerdings darf man sich auch nicht von den Konzepten der traditionell sehr integrationsbereiten und -erfahrenen skandinavischen Länder täuschen lassen, denn auch sie verzichten nicht auf die klassische Förderschule oder weitere spezielle sonderpädagogische Settings. Abgesehen davon sind dort die Grundlagen der sonderpädagogischen Arbeit auch andere. Generell weisen die skandinavischen Länder im Allgemeinen auffallend hohe Quoten eines speziellen Förderbedarfs auf (vgl. Ahrbeck 2016). Finnland bspw. liegt nach der EADSNE (2012) bei 8,3 %.

2.4 Inklusion im Förderschwerpunkt Lernen

SchülerInnen mit dem Förderschwerpunkt Lernen besuchen schon seit einigen Jahren Regelschulen. Spätestens seit den KMK-Empfehlungen (1994) verbreitet sich die Erkenntnis, dass Kinder mit Förderbedarfen in diesem Bereich auch an einer Allgemeinen Schule angemessen unterrichtet werden können. So besuchten im Schuljahr 2015/16 bereits 18 946 Kinder mit dem Förderschwerpunkt Lernen eine Regelschule,

im Schuljahr 2010/11 waren es noch 8 392 (IT.NRW 2016). Mit den Empfehlungen erfolgte aber auch eine Gleichstellung zwischen dem SPF in Förderschulen auf der einen und Allgemeinen Schulen auf der anderen Seite. Es hat sich hier ein erprobtes Praxiskonzept des gemeinsamen Unterrichts herausgestellt, dessen Alltagstauglichkeit sich in den letzten drei Jahrzehnten bewiesen hat. Inzwischen erweist sich dieser Schritt insofern als begrenzt, als die Veränderungswirkungen, bezogen auf Allgemeine Schulen, nicht automatisch eintreten (vgl. Heimlich 2016). Dies lässt sich zum einem mit den knappen Ressourcen der Schulen begründen und zum anderen mit der Einstellung der jeweiligen Lehrkräfte. Eine Analyse der Daten der deutschlandweiten Befragung im *Nationalen Bildungspanel Deutschlands* (NEPS) zeigte, dass LehrerInnen schulischer Inklusion gegenüber „eher skeptisch eingestellt sind“ (Gebhardt et al. 2015, S. 9). Die LehrerInnen der allgemeinbildenden Schulen hatten zur Inklusion eine neutrale bis negative Einstellung (vgl. ebd.). Im Übrigen stellt sich die Frage der Wirksamkeit schulischer Inklusion. Auch wenn es an empirischen Studien bisher noch mangelt, lassen sich Tendenzen erkennen, die etwas über die Gelingensbedingungen aussagen (vgl. Ellger-Rüttgardt 2016).

Erste Zweifel an der Effektivität der „Sonderbeschulung“ von SchülerInnen mit dem Förderschwerpunkt Lernen waren bereits in den Forschungen von ADRIAN KNIEL (1979) zu finden. Integrierte SchülerInnen mit dem Förderschwerpunkt Lernen zeigten positivere Verhaltensweisen und eine positivere Einstellung zur Schule. Die Leistungen der Kinder waren im Vergleich zu separierten Schulen mindestens gleich gut, in einigen Bereichen sogar besser. Aber auch bei SchülerInnen ohne Beeinträchtigung ließen sich zunächst keine leistungsmindernden Effekte nachweisen. Weitere Studien aus den 1990er-Jahren bestätigten diese Ergebnisse (vgl. Heimlich 2016). Jenseits dieser Vergleichsperspektive finden sich für jene Kinder, die an einer Förderschule unterrichtet werden, empirische Hinweise auf ungünstig verlaufende Lernentwicklungen. So wurde ein signifikanter Zusammenhang zwischen der Dauer der Beschulung der SchülerInnen an einer Förderschule und deren Rechtschreib- und Intelligenzleistungen gefunden. Es gilt: Je länger SchülerInnen eine Förderschule besuchen, desto schlechter fallen ihre Leistungen aus (vgl. Dederling 2016; Wocken 2007). Demnach wirken Förderschulen als nicht kompensatorisch. Wenn nun schlussfolgernd die Förderschule für SchülerInnen mit einer Lernbeeinträchtigung

keine Vorteile birgt, so stellt sich gleichwohl die Frage nach der Wirksamkeit des inklusiven Systems in diesem Förderbereich. Im Zentrum der Diskussion steht also die Frage, inwieweit sich die schulischen Leistungen von Kindern mit Förderbedarfen durch eine inklusive Beschulung im Vergleich zu jenen der entsprechenden Schüler entwickeln, die eine Förderschule besuchen. Für die Frage nach deren optimalen Beschulungsform kann auf empirische Befunde zurückgegriffen werden. Die Ergebnisse sind international wie im deutschsprachigen Raum relativ eindeutig. Die Vorteilseffekte der jeweiligen Schulformen stellen sich als eher gering dar, doch im Bereich der kognitiven Kompetenzentwicklung sprechen die Ergebnisse eher für eine inklusive Lernumgebung als für den separierenden Weg (vgl. Spörer et al. 2015). Auch die Ergebnisse der *Bielefelder Längsschnittstudie zum Lernen in inklusiven und exklusiven Förderarrangements* (2017, BiLieF) bestätigen diese Erkenntnis. Die Studie hat die Kompetenzentwicklung von SchülerInnen mit dem Förderschwerpunkt Lernen an Grund- und Förderschulen in den Bereichen Lesen und Rechtschreibung ab der dritten Klasse bis in die ersten Jahre an der weiterführenden Schule verglichen. Bei der Auswertung der Ergebnisse ist ein signifikanter und praktisch bedeutsamer Leistungsvorsprung der inklusiv beschulten SchülerInnen gegenüber den Gleichaltrigen in der Förderschule zu verzeichnen (vgl. Wild et al. 2017). Nach den Befunden des Ländervergleichs weisen Kinder mit einem SPF am Ende der vierten Jahrgangsstufe im inklusiven Setting im Bereich Lesen höhere Kompetenzen auf als SchülerInnen in Förderschulen. Die Unterschiede in diesem Kompetenzbereich liegen zwischen inklusiv und separiert unterrichteten SchülerInnen bei annähernd einem Jahr (vgl. Kocaj et al. 2014); dies gilt auch für den Förderschwerpunkt Lernen (vgl. Dederling 2016). Für diese SchülerInnen liegen zudem repräsentative Daten des IQB Ländervergleichs Primarstufe (2011) vor, der seit einigen Jahren regelmäßig zur Überprüfung der nationalen Bildungsstandards der Länder durchgeführt wird (vgl. Stanat et al. 2012).

Allerdings wird in der Literatur auch auf Grenzen in der Inklusion hingewiesen; so finden sich in verschiedenen Studien Hinweise darauf, dass die gemeinsame Beschulung von SchülerInnen mit und ohne sonderpädagogischen Förderbedarf zwar eine bessere Lern- und Leistungsentwicklung bei den Kindern mit dem Förderschwerpunkt Lernen fördert, die Unterschiede zwischen leistungsstarken und lernschwachen

SchülerInnen jedoch nicht ausgleichen kann (vgl. Dederling 2016) – ein Leistungsrückstand, der folglich nicht von einem inklusiven Lernsetting aufgeholt werden kann. Diese und weiterführende Studien der letzten Jahrzehnte laufen letztendlich auf eine Diskussion zwischen „[...] Schonraum- und Erfahrungsraummodell hin“ (Heimlich 2016, S. 74). Während die positiven Effekte der Förderschulen für das Selbstkonzept sprechen, so kommt dies einer Befürwortung für den Schonraumeffekt der Förderschulen gleich. Kinder mit einer Beeinträchtigung im Bereich Lernen sollten daher so lange wie möglich nicht mit gesellschaftlichen Leistungsanforderungen konfrontiert werden, um gestärkt aus dem Schonraum der gesonderten Beschulung hervorzugehen. Wird die negative Entwicklung des Selbstkonzepts zum Ende des Besuchs der Förderschule in den Vordergrund gestellt, „[...] so lässt sich daraus die Notwendigkeit eines Erfahrungsraumes ableiten, in dem Schüler[Innen] mit Lernbehinderung möglichst frühzeitig und unter pädagogischer Anleitung mit den Problemen der zukünftigen Lebenssituation konfrontiert werden, um eine realistische Selbsteinschätzung zu entwickeln“ (Heimlich 2016, S. 74).

Abschließend gilt es festzuhalten, dass der nationale, aber auch der internationale Forschungsstand zur inklusiven Beschulung eher spärlich ausfällt. Dennoch verweisen die Ergebnisse auf einige Vor- und Nachteile für SchülerInnen mit und ohne sonderpädagogischen Förderbedarf. Vordergründig profitieren beide Schülergruppen hinsichtlich der schulischen und kognitiven Leistungen vom inklusiven Unterricht (vgl. Dederling 2016). Bezogen auf den Förderschwerpunkt Lernen wird deutlich, dass aus bildungswissenschaftlicher Perspektive viele Gründe für den Unterricht an Regelschulen von SchülerInnen mit einer Lernbeeinträchtigung spricht. Um dies zu realisieren, ist aber auch die Einstellung der Lehrkräfte von besonderer Relevanz (vgl. Werning 2016). So muss also eine inklusive Schule, die eine Schule für alle sein will, die unterschiedlichsten Bildungsbedürfnisse und Unterstützungsbedarfe der SchülerInnen diagnostizieren, fördern und erfüllen, um den Kindern eine bessere Alternative zum bestehenden differenzierten Schulsystem zu bieten. Gelingt dies, wird die scheinbare Notwendigkeit einer speziellen Schule sukzessive verringert und die u. a. in der UN-BRK (2006) geforderte Entwicklung eines inklusiven Schulsystems vorangetrieben.

3 Lesen – theoretische Grundlagen

Kinder, die in einem Industrieland aufwachsen, begegnen der Schriftkultur nicht erst mit dem Beginn des schulischen Lesenlernens. In vielen Bereichen ihrer Umwelt begegnen Kleinkinder bereits „literacy events“ (Steinig & Huneke 2007, S. 201), die gleichzeitig Teil ihres Alltages sind. Dieser erste Kontakt, der zudem eine wesentliche Voraussetzung für den Schriftspracherwerb ist, findet sich in Büchern, Memos, Beschriftungen, Einkaufszetteln und Ähnlichem wieder. Der Sinn von Schrift wird hier erkannt und von anderen Zeichen unterschieden. Üblicherweise beginnt dies bereits in der Vorschulzeit und in der Familie zu Hause (vgl. Stöppler & Wachsmuth 2010).

In der kindlichen Entwicklung trägt das Lesen zur sprachlichen, kognitiven und emotionalen Entwicklung bedeutsam bei. In der Schule ist der Umgang mit Texten in jedem Fach eine grundlegende Tätigkeit. Zudem bildet Textverständnis die Grundlage für selbstständiges Lernen auch über das Fach Deutsch hinaus. Keine andere Medientätigkeit fördert die sprachliche Entwicklung eines Kindes so nachhaltig wie das Lesen. Dies begründet sich vor allem durch die Elaboriertheit, die Informationsdichte sowie die Situationsabstraktheit der geschriebenen Sprache gegenüber der gesprochenen. Neben der selbstständigen Beschäftigung mit der Literatur trägt vorrangig die Vorlesesituation, bspw. durch Elternteile, besonders nachhaltig zur Sprachentwicklung bei. Lesen ist jedoch auch eine komplexe kognitive Leistung, denn dem/der LeserIn muss es gelingen, ein auf Buchstaben basierendes Gebilde erst einmal zu entschlüsseln und anschließend immer wieder zu rekonstruieren (vgl. Steck 2008). McELVANY (2008) formuliert diesbezüglich eine Definition, die Lesen als „visuelles Wahrnehmen und Verstehen von Schriftzeichen“ (S. 13) charakterisiert. Für SCHMOTZ & DUTKE (2004) beinhaltet der Begriff des Lesens im Alltagsverständnis die geistige Verarbeitung eines Textes mit dem Ziel, dessen Bedeutung zu erfassen (vgl. König 2009). Es handelt sich also um die Fähigkeit des Dekodierens, bei der die LeserInnen, um den Sinn zu entnehmen, neben der grammatisch-syntaktischen Operation auch ihr gesamtes Vorwissen (Welt- und Sprachwissen) aktivieren müssen (vgl. Steck 2013).

3.1 Lesekompetenzen entwickeln

Dem Kompositum „Lesekompetenzen“ liegt zunächst der Begriff „Lesen“ zugrunde. In seinem prototypischen Kern meint Lesen das Verstehen von kontinuierlichen Texten, im erweiterten das diskontinuierlicher und multimedialer Texte (vgl. Hurrelmann 2011). Nach CHRISTMANN & GROEBEN (1999) handelt es sich beim Lesen in Verbindung mit dem Kompetenzbegriff nicht lediglich um basale, spezielle Fertigkeiten der Kulturtechnik „Lesen“, sondern vielmehr um komplexere Fähigkeiten (vgl. Christmann & Groeben 2009), die auch bestimmte Qualitätsstandards zu erfüllen in der Lage sind. PISA (2000) versteht Lesekompetenzen als ein wichtiges Hilfsmittel für das Erreichen persönlicher Ziele, aber auch als Bedingung für die Weiterentwicklung des eigenen Wissens und der eigenen Fähigkeiten und als Voraussetzung für die Teilhabe am gesellschaftlichen Leben. Dabei orientiert sich das Verständnis von PISA ebenfalls an der angelsächsischen *Literacy*-Konzeption, in der die funktionelle Sicht auf Kompetenzen als basales Kulturwerkzeug zentral ist (vgl. Artelt et al. 2004). Es fragt nach den Basisqualifikationen, die in modernen Gesellschaften für eine in beruflicher und gesellschaftlicher Hinsicht erfolgreiche Lebensführung unerlässlich sind. Das Konzept sieht die mathematischen und naturwissenschaftlichen Kompetenzen und in diesem Sinne eben auch die „Reading Literacy“ (Lesekompetenzen), nicht zuletzt von PISA stark geprägt, als Basisqualifikation an. Lesekompetenzen sind in diesem Verständnis ein basales Kulturwerkzeug, das erforderlich ist für die Bewältigung der charakteristischen Kommunikations- und Handlungsanforderungen, denen ein Individuum bei der Teilhabe an der Gesellschaft und in allen Lebensbereichen begegnet. Die Bezugsnorm stellt in diesem Fall der gesellschaftlich-pragmatische Bedarf dar (vgl. Hurrelmann 2011).

Doch wie lässt sich der komplexe Prozess des Lesens beschreiben und erklären? Worin unterscheiden sich die besseren von den schlechteren LeserInnen? Mit Fragen dieser Art befasst sich die kognitionspsychologische Forschung zu Lesekompetenzen, der es nicht primär um eine psychomotorisch optimierte Erfassung von Lesekompetenzen geht, sondern vielmehr um die Unterschiede zwischen den LeserInnen. Der Ausgangspunkt für die Forschung in diesem Bereich ist immer das Problem, welche der im Rahmen kognitionspsychologischer Modelle ausdifferenzierten Teilprozesse des Lesens von guten LeserInnen effizienter bewältigt werden als von schlech-

ten (vgl. Christmann & Groeben 2009). Die kognitionspsychologische Lesekompetenzforschung nimmt auf der Ebene der Buchstaben- und Worterkennung okulomotorische und perzeptuelle Prozesse an, auf denen Prozesse der Worterkennung, der phonologischen Dekodierung und der lexikalischen Zuordnung aufbauen. Daher finden auf der Satz- und Textebene Prozesse statt, die die semantischen und syntaktischen Bezüge zwischen Teilausdrücken herstellen und durch die Integration von Vorwissen eine Gesamtinterpretation des Textes ermöglichen. Hier schließt sich die Frage der Erklärung an.

Alle modernen Theorien der Lesekompetenz haben eines gemeinsam: Sie sind mehr oder weniger explizit auf hierarchische Modelle des Lesens und Textverstehens zurückzuführen (vgl. Richter & Christmann 2009, nach Dijk & Kintsch 1983, siehe auch Kapitel 3.2). Derartige Modelle gehen davon aus, dass Teilprozesse auf verschiedenen Ebenen, das heißt auf Wort- und Satzebene, in vielfältiger Weise miteinander interagieren können (vgl. Richter & Christmann 2009). Die Einschränkung des Begriffs der Lesekompetenzen auf die Fähigkeiten des Textverstehens ist allerdings zu eng. Lesekompetenz ist vielmehr als Fähigkeit anzusehen, Schriftwerke zu verstehen, „[...] in denen sowohl verbale Informationen in Schriftform von Schriftzeichen (graphemisch) als auch piktorale Informationen in Form von Bildzeichen (graphisch) enthalten sind“ (Schmotz & Dutke 2004, S. 63). Theorien zu Lesekompetenzen lassen sich also danach gruppieren, ob sich die Erklärungsrichtung eher auf eine hergeleitete (bottom up) oder auf eine wissensgeleitete (top down) Verarbeitungsrichtung konzentriert. Die *Theorie der verbalen Effizienz* (Perfetti 1985) nimmt an, dass das allgemeine Lesefähigkeitsniveau von der Effizienz hierarchieniedriger Prozesse abhängt. Das *interaktiv-kompensatorische Modell* (Stanovich 1980) beschäftigt sich mit dem Unterschied im Zusammenwirken hierarchieniedriger und hierarchiehöherer Teilprozesse bei guten und schlechten LeserInnen anhand der Nutzung des Satzkontextes bei der Enkodierung von Wortbedeutungen. Auf der anderen Seite existieren auch schematheoretische Ansätze, die sich die Frage stellen, ob es überhaupt lese-spezifische Fähigkeiten gibt oder ob leseunspezifische Fähigkeitsaspekte wie das inhaltliche Vorwissen die Performanzen beim Leseverstehen dominieren (vgl. Richter & Christmann 2009). In diesem Fall wäre es allerdings fragwürdig, ob an dieser Stelle der Begriff Lesefähigkeit dann überhaupt noch sinnvoll ist.

Das Aufbauen, Fördern und Messen von Lesekompetenzen ist eine komplexe Aufgabe. Dabei handelt es sich nicht um einen rein sprachlichen Prozess, sondern er ist wesentlich durch das Vorwissen und die allgemeine kognitive Disposition der LeserInnen bestimmt. Aus diesem Grund müssen sich Aufgaben, die auf die Förderung ebensolcher Kompetenzen abzielen, nicht nur darauf beschränken Texte lesen zu lassen, sondern sollten vielmehr in komplexe Handlungszusammenhänge eingebettet werden, um das für das Lesen so zentrale Weltwissen in angemessener Weise auf- und auszubauen (vgl. Lindauer & Schneider 2011). Um den Prozess der Leselerntwicklung zu verdeutlichen, haben die Autoren KLICPERA, SCHABMANN und GASTEIGER-KLICPERA (2013) ein Kompetenzmodell entwickelt. Dieses wird hier allerdings nur kurz beschrieben, weil dem Prozess des sinnentnehmenden Lesens für diese Arbeit mehr Aufmerksamkeit zugesprochen wird.

Das Modell zur Lesekompetenzentwicklung konstatiert, dass gemäß den theoretischen Konzepten zum Leseprozess bei einem reifen Leser das Worterkennen entweder durch einen direkten Zugriff auf das *mentale Lexikon* erfolgt oder mittels der *phonologischen Rekodierung* (Kapitel 3.1). Die Autoren des Modells betonen, dass der reife Leser beide Zugangswege beherrschen muss, denn neue und bisher unbekannte Wörter können nur über die phonologische Rekodierung erlesen werden. Wörter, deren Schreibweise von der Aussprache abweicht, können hingegen nur über den lexikalischen Zugriff erlesen werden. Die Entwicklung des Lesens läuft daher nach den verschiedenen Phasen ab: der *präalphabetischen Phase*, der *alphabetischen Phase mit geringer Integration* und *voller Integration* und schließlich der *automatisierten und konsolidierten Integration* (vgl. Klicpera, Schabmann & Gasteiger-Klicpera 2013). Die alphabetische Phase mit geringer Integration kann als erste „echte“ (Klicpera, Schabmann & Gasteiger-Klicpera 2013, S. 33) Phase des Lesenslernens bezeichnet werden, denn hier werden die notwendigen Kompetenzen, die zum Lesen benötigt werden, allmählich herausgebildet (vgl. ebd.). Zusammenfassend kann gesagt werden, dass die Autoren den komplexen Prozess des Lesenslernens mit seinen differenten Teilfertigkeiten umfassend beschrieben haben (vgl. Gebhardt et al. 2016a).

Es kann festgehalten werden, dass die Forschung in den letzten vier Jahrzehnten in beiden Bereichen bemerkenswerte Fortschritte erzielt hat. So haben sich einige

Komponenten des Prozesses Lesen als zentrale Faktoren des allgemeinen Leseverständnisses erwiesen. Wiederum andere konnten als Ursachen interindividueller Unterschiede praktisch ausgeschlossen werden (vgl. Richter & Christmann 2009). Dagegen ist es erwiesen, „dass die Leichtigkeit, mit der im individuellen Fall die phonologische Dekodierung bei der Worterkennung vonstattengeht, offenbar auch übergeordnete Verarbeitungsstufen beeinflusst, während die charakteristischen okulomotorischen Muster, die schlechte Leser/innen bei der Worterfassung zeigen, bestenfalls ein Symptom von Leseproblemen darstellen“ (Richter & Christmann 2009, S. 27).

3.2 Der Prozess des Lesenlernens

Lesen lässt sich nach CHRISTMANN und GROEBEN (2006) in einer ersten Annäherung als Fähigkeit auffassen, „[...] visuelle Informationen aus graphischen Gebilden zu entnehmen und deren Bedeutung zu verstehen“ (S. 148). Das zentrale Ziel des Leseerwerbs ist es also, den gelesenen Wörtern ihre Bedeutung zu entnehmen und somit den Inhalt zu konstruieren. Die Entwicklung des Leseverständnisses hängt vor allem von der Entwicklung basaler sprachlicher Fähigkeiten ab, vom Wortschatz sowie von den Ausprägungen des Arbeitsgedächtnisses und von den kognitiven Fähigkeiten (vgl. Lenhard & Artelt 2009). Dieser komplexe Vorgang lässt sich in mehrere Teilprozesse aufgliedern. Zu diesen Prozessen gehört auf der untersten Ebene das Erkennen von Buchstaben und Wörtern sowie die Erfassung von Wortbedeutungen. Auf der mittleren Prozessebene steht die Herstellung semantischer und syntaktischer Relationen zwischen Wortfolgen. Die Textebene beinhaltet eine satzübergreifende Integration von Sätzen zu umfassenden Bedeutungseinheiten sowie den Aufbau einer kohärenten Struktur der globalen Gesamtbedeutung eines Textes (vgl. Christmann & Groeben 2006).

Die SchülerInnen nutzen sowohl beim Lesen als auch beim Schreiben verschiedene Strategien zur Codierung von Gedanken in ein Zeichensystem und zur Re- bzw. Decodierung von Bedeutungen aus niedergelegten Zeichen. Obwohl die sprachwissenschaftlichen Entwicklungsmodelle ihren Fokus vorwiegend auf Schriftzeichen legen, gibt es Ansätze, die darüber hinaus andere Schwerpunkte betrachten (vgl. Terloth & Bauersfeld 2012). HUBLOW (1985) bspw. geht davon aus, dass die zentralen

Aspekte des Leseprozesses auch auf andere, nicht laterale Zeichensysteme angewendet werden können. Lesen ist im Sinne des erweiterten Lesebegriffs das „Wahrnehmen, Deuten, Verstehen von konkreten, bildhaften, symbolhaften oder abstrakten Zeichen und Signalen“ (Terfloth & Bauersfeld 2012).

Bereits im Vorschulalter entwickeln Kinder wichtige Vorläuferfähigkeiten, die eine wesentliche Voraussetzung für den Schriftspracherwerb darstellen. Von zentraler Bedeutung ist dabei die *phonologische Informationsverarbeitung*. Sie umfasst phonologische Operationen im Arbeitsgedächtnis, phonologische Operationen im Zugriff auf das semantische Lexikon und die phonologische Bewusstheit (vgl. Lenhard & Artelt 2009 nach Wagner & Torgesen 1987). Letzteres ist der wichtigste Teilbereich und kennzeichnet die Bewusstheit über die Sprache, also die Befähigung, Wörter in kleinere Einheiten (Silben und Phoneme) zu zerlegen bzw. diese zu Wörtern zusammenzusetzen. Darüber hinaus können Inhaber dieser Fähigkeit, Wortbestandteile austauschen oder Gemeinsamkeiten und Unterschiede in der Lautstruktur und in Wörtern erkennen (vgl. Lenhard & Artelt 2009). Bei der Worterkennung z. B. handelt es sich um Komponenten des Lesens, die beim geübten Leser in der Regel nicht mehr der bewussten Kontrolle unterliegen. Es geht hier also um die grundlegende Lesekompetenz, die der geübte Leser sozusagen als Werkzeug einsetzt, um zum Eigentlichen, dem Inhalt, dem Sinn des Textes und seiner eigenen Stellungnahme zu gelangen (vgl. Hurrelmann 2002). Lesekompetenz stellt eine komplexe Fähigkeit dar, die aus mehreren miteinander interagierenden Teilfähigkeiten besteht. Viele hierbei relevante Komponenten und Faktoren wie Wortschatz, inhaltliches Vorwissen oder die Kapazität des Arbeitsgedächtnisses (Kurzzeitgedächtnis) sind nicht notwendigerweise spezifisch für das Lesen, sondern kommen in ähnlicher Form auch beim Hörverstehen und bei anderen sprachlichen Verstehensleistungen zum Tragen (vgl. Lenhard & Artelt 2009). Das *Programme for International Student Assessment* (2000; PISA) geht von einem sehr spezifischen Begriff der Lesekompetenzen aus, der die *hierarchieniedrigen* und *hierarchiehöheren* kognitiven Prozesse in den Blick nimmt, wie nachfolgend dargestellt (vgl. Hurrelmann 2002).

Hierarchieniedrige Leseprozesse

- Buchstaben- und Worterkennung
- Syntaktische und semantische Analyse von Wortfolgen

Leseverstehen ist untrennbar zusammenhängend mit sprachlichen Fähigkeiten. Diese spielen auf jeder der folgenden Ebenen eine Rolle: Das Dekodieren von Wörtern auf der Ebene der Laut-Buchstaben-Beziehung, der Morphologie und der Wortbedeutung sowie die Decodierung von Sätzen auf der Ebene der Satzstruktur bzw. Syntax und Satzsemantik gehen nicht ohne grundlegende Sprachkenntnisse. Dabei spielen Lesefertigkeiten, wie sie in den ersten Schuljahren erworben werden, eine ebenso bedeutsame Rolle, weil zu der Entwicklung weitere Fähigkeiten und Fertigkeiten führt. Damit ist gemeint, dass bei LeserInnen das Erkennen von Buchstaben, die Zuordnung von Buchstaben zu Lauten, die Synthese der Laute zu Wörtern sowie das Erkennen sogenannter Sichtwörter automatisiert und meist unbewusst abläuft. Diese Prozesse laufen dabei nicht von der Wahrnehmung der Textoberfläche her ab, sondern vom gespeicherten Wissen her (vgl. Lindauer & Schneider 2011).

Da sich die Wortbedeutung aber nicht immer eindeutig bestimmen lässt und häufig der Kontext eine wesentliche Rolle zum Verständnis von Wörtern beiträgt (vgl. Klicpera et al. 2013), stellt der Umfang des Wortschatzes einen wichtigen Prädikator des Leseverständnisses dar. Die Bedeutung des Wortschatzes für das Leseverständnis nimmt im Schulalter stetig zu (vgl. Mayer 2012). Zum Verstehen eines ganzen Satzes reicht die Entschlüsselung von Wörtern allerdings nicht aus. Übereinstimmend gehen kognitionspsychologische Theorien heute davon aus, dass Wortfolgen auf der Grundlage ihrer semantischen Relationen aufeinander bezogen und zu „Prädikat-Argument-Strukturen“ (Richter & Christmann 2009, S. 29). oder Propositionen aufeinander bezogen werden. Dies belegen empirische Studien, die zeigen, dass Sätze mit identischer Oberflächenstruktur, aber differenter propositioneller Tiefenstruktur in Abhängigkeit von Art und Anzahl der Argumente unterschiedlich gut gelesen und verarbeitet werden (vgl. ebd.).

In diesem Zusammenhang sei erwähnt, dass für spracherwerbsverzögerte Kinder mit semantisch-lexikalischen Defiziten bereits auf dieser Ebene Schwierigkeiten anzunehmen sind. Diese Schwierigkeiten müssen sich nicht nur auf Fremdwörter beziehen, sondern können sich schon auf Präpositionswörter, Adverbien, Konjunktionen, Partikel oder Pronomen beziehen (vgl. Mayer 2012). Das prozesshafte Geschehen des hierarchieniedrigen Leseprozesses bildet die theoretische Grundlage für den Test zum sinnentnehmenden Lesen nach LEVUMI.

Wortebene

Der Prozess des Lesens beginnt auf der untersten Stufe mit der Identifikation von Buchstaben und Wörtern und ist primär als visueller Verarbeitungsprozess aufzufassen. Dabei stellt sich zunächst die Frage, ob Buchstaben, Buchstabengruppen oder ganze Wörter die relevanten Einheiten bei der Analyse darstellen (vgl. Christmann & Groeben 1999). Eine eindeutige Antwort auf diese Frage gibt es bis heute nicht. Die diesbezüglichen Erkenntnisse aus der Forschung haben zu sehr differenten und teilweise konkurrierenden Lesemodellen geführt. Die zu diesem Diskurs wahrscheinlich bedeutsamste Theorie zum Wortlesen ist das *Zwei-Wege-Modell* nach COLTHEART (1978). Das Modell differenziert zwei Lesewege: den *direkten* und den *indirekten* (vgl. Gebhardt, Diehl & Mühling 2016a).

Beim direkten Weg kann der Leser einen Kontakt mit dem lexikalischen Eintrag eines Wortes herstellen, und zwar über den orthografischen Kodex des Schriftbildes (vgl. Klicpera et al. 2013). Die Entschlüsselung kann aber nur erfolgen, wenn das entsprechende Wort zuvor schon als Ganzes im Gedächtnis gespeichert ist. Es muss also im mentalen Lexikon eine Repräsentation vorhanden sein (vgl. Gebhardt, Diehl & Mühling 2016a). Eine andere Möglichkeit bietet der Zugang über die Phonemfolge, also über den indirekten Weg. Dies erfolgt über die phonologische Rekodierung, das schrittweise Generieren der Aussprache aus Buchstaben oder Buchstabengruppen (vgl. Klicpera et al. 2013). Synonym zu den Begriffen des direkten und indirekten Weges wird in diesem Kontext auch von lexikalischem und nicht lexikalischem Lesen gesprochen. Die Ausführungen verdeutlichen, dass der/die LeserIn für sich nur zwei Möglichkeiten erschließen kann. Entweder er/sie kann ein Wort erkennen und es als Ganzes aus seinem/i ihrem mentalen Lexikon erlesen, oder das Wort muss mit der phonologischen Rekodierung Buchstabe für Buchstabe ermittelt werden (vgl. Gebhardt, Diehl & Mühling 2016a). Im natürlichen Prozess werden Wörter aber üblicherweise nicht isoliert voneinander gesehen, sondern immer im Satzkontext verarbeitet.

Die systematische Förderung der automatisierten Worterkennung hat im deutschsprachigen Raum keine sehr lange Tradition, sodass derzeit nur wenige empirisch evaluierte Förderprogramme zur Verfügung stehen. Ein Lesetraining mit entsprechender Zielsetzung ist die *Blitzschnelle Worterkennung* (BliWo) nach MAYER (2012). Das Trainingsprogramm versucht, Kinder mit Schwierigkeiten im Lesen beim Erwerb

der automatisierten Worterkennung zu unterstützen (orthografische Lesestrategie) (vgl. Mayer 2012).

Satzebene

Die Identifikation von Wörtern und das Erkennen der Wortbedeutungen sichern allein noch nicht das Verstehen der Bedeutung eines Satzes. In einem weiteren Schritt müssen Wortfolgen aufeinander bezogen und in ein strukturiertes Gesamtgefüge gebracht werden. Dies erfordert eine Analyse nicht nur der semantischen, sondern auch der syntaktischen Relationen der einzelnen Satzelemente. Bei der semantischen Verarbeitung werden Satzelemente im Verarbeitungsprozess auf Grundlage bestimmter Ordnungsprinzipien zu semantischen Bedeutungseinheiten zusammengefasst. Diesbezüglich hat sich seit den 70er-Jahren die Auffassung durchgesetzt, dass der Bedeutungsgehalt von Sätzen in Form von Prädikat-Argument-Strukturen (Propositionen) dargestellt wird. Demnach wird dem Prädikat eine höhere Organisationsfunktion zugesprochen, da es die semantischen Relationen zu den Argumenten festlegt und spezifiziert. Infolgedessen besteht die semantische Satzverarbeitung aus der Extraktion von Prädikat-Argument-Strukturen aus der zugrunde liegenden Satzstruktur. In bestimmten Fällen des Verstehensprozesses reicht die semantische Analyse für eine eindeutige Bedeutungszuordnung nicht aus. In diesem Fall muss sie um eine syntaktische Analyse ergänzt werden (vgl. Christmann & Groeben 2006).

Die Analyse der syntaktischen Struktur eines Satzes erfordert es, dass im Verarbeitungsprozess Wörter und Wortgruppen identifiziert werden, denen eine syntaktische Funktion in Form von Satzgliedern zugewiesen wird. Dabei wäre der einfachste Fall, wenn der/die LeserIn sich auf die Abfolge der Inhaltswörter eines Satzes stützt; Subjekt – Prädikat – Objekt (ebd.).

Uneinigkeit besteht in der Literatur dennoch darüber, ob der Syntax bei der Satzsegmentierung Priorität zukommt oder ob von einem Zusammenspiel von Syntax und Semantik auszugehen ist. Was die Beantwortung dieser Frage angeht, so lassen sich auch hier wieder, wie beim Prozess der Worterkennung, zwei konträre Positionen unterscheiden (ebd.). Die *interaktionistische Syntaxtheorie* geht davon aus, dass die Analyse syntaktischer Strukturen vom semantischen wie pragmatischen Kontext sowie vom Weltwissen beeinflusst wird (vgl. Altmann & Steedman 1988), wohingegen

die syntaktischen und semantischen Teilprozesse weitgehend parallel arbeiten. Hin- gegen geht die autonome Syntaxtheorie davon aus, dass die syntaktische Verarbei- tung zeitlich vor der semantischen Verarbeitung und deshalb unabhängig von dieser erfolgt (vgl. Christmann & Groeben 2006).

Textebene

Die genannten Ausführungen haben verdeutlicht, wie komplex der Prozess der Wor- terkennung sein kann. Allerdings gewährleistet die Fähigkeit der Worterkennung und der semantisch-lexikalischen sowie syntaktisch-morphologischen Verarbeitung maxi- mal ein Verstehen auf der Satzebene (vgl. Mayer 2016) – eine Fähigkeit, die zwar für die Durchführung dieser Studie ausreicht, jedoch nicht als hinreichend betrachtet wer- den kann. Denn um Texte auch sinnentnehmend verarbeiten zu können, müssen Kin- der lernen, „[...] zwischen den Zeilen zu lesen, die nicht explizit genannten Informati- onen zu ergänzen, Schlussfolgerungen zu ziehen, Hintergrundwissen und Weltwis- sen zu integrieren“ (ebd., S. 179). Zudem müssen die Einzelinformationen auf Phra- sen- und Satzebene zusammengefügt und mit dem subjektiven Vorwissen verknüpft werden – eine Fähigkeit, die als durchaus anspruchsvoll betrachtet werden kann, ge- rade für SchülerInnen mit Förderbedarfen. Infolgedessen scheitert dieser Prozess, falls für den/die LeserIn zu wenige kognitive Strategien zur Verfügung stehen, die es ihm/ihr ermöglichen, sich mit dem Gelesenen auseinanderzusetzen (vgl. ebd.).

Hierarchiehöhere Leseprozesse

- satzübergreifende Analyse von Textstrukturen
- Herstellung globaler Kohärenz; Bildung einer Makrostruktur

Texte sind mehr als eine Folge von Sätzen; ihre Struktur ist einerseits an der Ober- fläche (sprachlich und grafisch) explizit markiert, andererseits implizit (z. B. über Welt- wissen) angelegt. Das Verbinden von satzgebundenen Informationen über teilweise weite Textdistanzen hinweg wird als das Herstellen von „globaler Kohärenz“ bezeich- net. Die dafür benötigten kognitiven Prozesse sind nur schwer zu fördern und vor allem zu messen, da sie nämlich in einem hohen Grad mit allgemeinen Dispositionen verknüpft sind. Das bedeutet schlussfolgernd, dass ein mentales Modell über einen

Text nicht nur eine sprachliche Kompetenz ist, sondern dass es sprachliche Fähigkeiten und Fertigkeiten mit anderen Problemlösungsstrategien der menschlichen Kognition verbindet (vgl. Lindauer & Schneider 2011).

Im Stadium des Erwerbs des Lesenlernens stehen diese Prozesse im Mittelpunkt der Aufmerksamkeit der Lehrkraft und des Kindes. Zentral ist dabei das Worterkennen mit den Kompetenzen der Sinnentschlüsselung (Dekodieren) und der Übertragung in die gesprochene Sprache (Rekodieren), das nicht nur in den ersten Schuljahren in hohem Maße auch mit der allgemeinen Lesekompetenz korreliert (vgl. Hurrelmann 2002). „Lesen ist also kein reines Dekodieren auf der Buchstaben-, Wort- und Satzebene, sondern ist ein Prozess, der durch organisiertes Wissen gesteuert wird“, so HURRELMANN (2002, S. 7). Mit der Fähigkeit zur Dekodierung von Schrift entsteht eine Art „innere Übersetzung, innere Verbalisierung“ (ebd.), eine Dynamik zwischen Schrift und Gedanken. Das Kind aktiviert und verändert im Leseprozess eigene Wissensstrukturen, „das Aufgenommene wird zum eigenen Gedanken entwickelt“. Mit dieser Lesefähigkeit entfalten sich sprachliche und begriffliche Kompetenzen; Perspektiven werden erfolgreich differenziert. Lesen ist Konzentration auf das Verstehen (vgl. ebd.).

3.3 Sinnentnehmendes Lesen

Die genannten Ausführungen haben bereits verdeutlicht, dass Lesen mehr beinhaltet als eine einfache Zeichen-Laut-Zuordnung. Der Leser muss vor allem die Fähigkeit der Rekodierung (Erlesen von Worten) und der Dekodierung (Erkennen ganzer Wörter) entwickelt haben, um einem Wort eine Bedeutung zuordnen zu können. Die beiden Fähigkeiten bilden wiederum die Voraussetzung für das *sinnentnehmende Lesen*. Der Ausdruck sinnentnehmendes Lesen wird in der gängigen Literatur auch synonym zu dem Begriff *sinnverstehendes Lesen* verwendet. Die weiteren Ausführungen fokussieren Ersteres.

Das *sinnentnehmende Lesen* von Sätzen und Texten basiert auf dem *Simple View of Reading*-Ansatz. Es setzt sich aus „[...] einem intakten Hörsprachverständnis und [...] durchschnittlichen Fähigkeiten im Bereich der Worterkennung“ (Mayer 2016, S. 179) zusammen. Der Simple View of Reading, der sprachlichen Fähigkeiten und Fertigkeiten nicht nur auf semantisch-lexikalischer und auf grammatischer Ebene

fokussiert, sondern auch die Worterkennung als zentralen Faktor sieht, muss deshalb durch eine kognitive Komponente ergänzt werden. Diese Komponente soll die aktive Auseinandersetzung mit dem Gelesenen ermöglichen und die Konstruktion des Gesamtbildes eines Textes erlauben. Kindern mit einer unauffälligen Leseentwicklung gelingt es meist, auf der Grundlage der phonologischen Rekodierung Wörter durch ausreichende Leseerfahrung zu verstehen. Leseschwachen SchülerInnen macht eben dieser Schritt Probleme. Sie kennzeichnen sich häufig dadurch, dass sie das Lesen als Worterkennungsaufgabe verstehen und nicht auf semantischer Ebene versuchen, den Satz als Ganzes zu verstehen. Dies endet dann in einem „Wort-für-Wort-Lesen“ (ebd., S. 180). Die automatisierte Worttrennung spielt aber auf dem Weg zum kompetenten Leser eine entscheidende Rolle. Deshalb benötigen diese SchülerInnen explizite Hilfe, um das Wortlesen zugunsten eines phrasen- und satzübergreifenden Lesens aufzugeben, das es ihnen ermöglicht, „neue Informationen in das Gelesene zu integrieren und so gegebenenfalls die bis dahin aufgebaute Bedeutungskonstruktion zu verändern oder zu ergänzen“ (ebd., S. 180). Automatisierte Leseprozesse stellen eine zentrale Voraussetzung für die hierarchiehöheren Prozesse (Kapitel 3.2) des sinnentnehmenden Lesens auf Satz- und Textebene dar (vgl. Mayer 2012). Deshalb sollte gerade zu Beginn des Leseerwerbs und parallel zum phonologischen Rekodieren geübt werden (vgl. Jungjohann et al. 2017). Denn erst, wenn sich basale Leseprozesse auf Wortebene weitgehend automatisiert haben, stehen dem/der LeserIn ausreichend kognitive Ressourcen für die komplexen Prozesse des Leseverständnisses zur Verfügung (vgl. Mayer 2012).

Gelingt dem/der LeserIn die Interpretation der Wortbedeutungen auf der Basis des Kontextes, so befindet sich das Individuum bereits auf der Ebene von Satzbestandteilen oder ganzen Sätzen. Von dem Prozess des sinnentnehmenden Lesens auf Satzebene, das heißt mit der Extraktion von Propositionen, führt der Weg auf die nächsthöhere Ebene, die des Textes bzw. dessen Kodierung. Der/die LeserIn muss nun satzübergreifend Bedeutungsstrukturen herstellen, indem er/sie die Inhalte einzelner Sätze aufeinander bezieht, diese integriert und einen sinnvollen Zusammenhang herstellt. Dieser Prozess wird allerdings nicht weiter beschrieben, weil dies nicht nur den Rahmen dieser Arbeit übersteigen würde, sondern auch, weil es thematisch

über die Grundlage des verwendeten Textes führen würde. Sinnentnehmendes Lesen entwickelt sich demnach von der Wort- über die Satz- bis hin zur Textebene (vgl. Hesse & Latzko 2011).

4 Diagnostik im schulischen Kontext

Diagnostik dient der Vorhersage zukünftiger Leistungen. GROHNFELDT (2009, S. 17) gibt eine grundlegende Definition des umfassenden Begriffs vor. Er umschreibt sie als einen „[...] Vorgang der Unterscheidung und Beurteilung“, einen „Prozess des Durchschauens (,dia‘)“ sowie das „[...] Erkennen (,gnosis‘) des dahinter Liegenden (sic!) [...]“. Die Diagnostik stellt sich darum die Frage, inwieweit ein Individuum in der Lage ist, zukünftige Anforderungen zu meistern, welche Potenziale vorliegen und ob eine Förderung notwendig ist (vgl. Lenhard 2013). GROHNFELDT (2009) führt diesbezüglich weiter aus, dass die Methoden des Diagnostizierens u. a. Beobachtungen, Testungen und Interviews umfassen. Im schulischen Kontext können mithilfe wissenschaftlicher Theorien und Modelle der kindlichen Entwicklung somit Ergebnisse gedeutet werden.

4.1 Lernverlaufsdagnostik

Die *Lernverlaufsdagnostik* stellt eine Möglichkeit der langfristigen Überprüfung von SchülerInnen sowie Klassen über mehrere Wochen hinweg dar (vgl. Strathmann & Klauer 2010). Dadurch lassen sich Rückschlüsse ziehen, inwiefern sich entsprechende Fähigkeiten entwickelt haben. Die sich entwickelnden Fähigkeiten lassen sich auch als Kompetenzen bzw. Expertisen bezeichnen und sind somit deutlich weniger statisch als die angenommenen Konstrukte, die den meisten Intelligenztests und Persönlichkeits- sowie Begabungsdagnostiken zugrunde liegen. Demnach ergibt sich die Leistung eines Individuums durch die aus dem Lernprozess gewonnenen Kompetenzen und die Bearbeitung eines bestimmten Aufgabenformates. Kompetenzen bauen sich somit aus der Summe unzähliger Lerninstanzen, Aufgaben und Anforderungen an die jeweilige Person aus, die gleichermaßen das Ausmaß an Kompetenzen der Person darstellen. Die Annahme von Kompetenzen als sich entwickelnde Expertisen ist grundlegend für theoretische Modelle der Lernverlaufsdagnostik, vor deren Hintergrund eine Überprüfung der Leistung als Folge einer pädagogischen Intervention erst sinnvoll wird (vgl. Wilbert 2014). Die Erhebung solcher Kompetenzen erfolgt in der Regel über Tests.

Großflächige und vergleichende Tests gehören immer mehr zum Monitoring von Bildungssystemen. Solche Tests können z. B. für die Diagnose von Leistungsunterschieden zwischen schwächeren und stärkeren SchülerInnen (bspw. zwischen Förder- und Regelschule) verwendet werden. Diese Tests werden typischerweise großflächig angelegt, um nicht einzig für einen bestimmten Moment zu dienen (vgl. Schneider & Lindauer 2011). Ein zentrales Unterscheidungsmerkmal einer Klassifizierung besteht darin, ob ein Schulleistungstest (= *formell*) von fachpsychologischen Testkonstruktionen ausgearbeitet und die Gütekriterien empirisch nachgewiesen werden können und ob die anderen Tests (= *informelle*) von Lehrkräften selbst konstruiert werden. Die formellen und wissenschaftlich ausgearbeiteten Verfahren können zunächst das Ziel haben, die Leistungen von SchülerInnen zu prüfen und anschließend mit den Leistungen einer großen Stichprobe in Bezug zu setzen. Ziel dessen ist es, eine Aussage über die Leistung eines/r einzelnen Schülers/In im Vergleich zur Leistung aller SchülerInnen dieser Alterskohorte treffen zu können. Mit einer anderen Form könnte überprüft werden, wie gut SchülerInnen in Bezug auf ein Kriterium (= *sachliche Bezugsnorm*) gelernt haben. Dabei kann ein Kriterium sich auch auf schulische Kompetenzen, bspw. Lesekompetenzen, beziehen, woraufhin mit dem Test festgestellt wird, auf welcher Niveaustufe der/die Schüler/In diese Kompetenz besitzt (vgl. Hesse & Latzko 2011). Schulleistungstests stellen eine diagnostische Alternative zur traditionellen Leistungsmessung in der Schulpraxis dar. Die einschlägige Literatur verwendet teilweise verschiedene Umschreibungen dieser Testform wie „standardisierte“, „normorientierte“ oder „kriteriumsorientierte Schulleistungstests“ (ebd.).

Eine alternative Bezeichnung für Lernverlaufsdiagnostik ist der Begriff der *formativen Evaluation*, die KLAUER (2014) als eine Evaluation des laufenden Lernprozesses bezeichnet. Das Ziel ist es, nicht nur den Prozess des Lernfortschritts zu analysieren und zu dokumentieren, sondern den erlangten Fortschritt der SchülerInnen durch Rückmeldung an Beteiligte und sie selbst zurückzugeben (vgl. Börnert 2014). Lehrkräfte erkennen so frühzeitig Lernlücken bzw. erwartungswidrige Leistungen, die den weiteren Lernerfolg erschweren können (vgl. Mühling, Gebhardt & Diehl 2017). Denn „[...] nur durch eine bewusste Wahrnehmung von sich entwickelnden Schwierigkeiten, (sic!) können bewusste Veränderungen vorgenommen werden“ (Börnert 2014, S. 3). Zusätzlich erhält die Lehrkraft Rückmeldung über den Erfolg und die Wirksamkeit ihres Unterrichts und ihrer Förderung (vgl. Mühling, Gebhardt & Diehl 2017).

Der *formativen Evaluation* steht das Konzept der *summativen Evaluation* gegenüber. Summative Evaluation ist eine am Ende des Lernprozesses stehende Form der Evaluation. Einen ersten Ansatz formativer schulischer Diagnostik stellt das sogenannte „Curriculum-based Measurement“ (CBM) dar, das im nachfolgenden Kapitel genauer beschrieben wird (vgl. Klauer 2014).

4.2 Curriculum-based measurement (CBM)

Der Ansatz *Curriculum-based measurement* (CBM) ist in den 1970er-Jahren in den USA entstanden und auf das Team von STANLEY DENO (2003) zurückzuführen (vgl. Börnert 2014). Der Ansatz wurde für die Sonderpädagogik entwickelt und betitelt alle Testungen und Untersuchungen, welche die Beherrschung der im aktuellen Unterricht tatsächlich vermittelten Inhalte erfassen sollte (vgl. Klauer 2011). Mithilfe von *curriculum-based Tests* (vgl. Deno 2003) können die Lernfortschritte der SchülerInnen in differenten Lernbereichen zeitnah beurteilt werden. Das heißt, die Lernfortschrittmessung dient nicht primär der Klassifikation oder dem Einteilen in Leistungsgruppen oder gar der Selektion, sondern der Feststellung des Lernfortschrittes in einem Zeitraum (vgl. Breitenbach 2014). So lösen SchülerInnen innerhalb einer zeitlichen Vorgabe, meist nur wenigen Minuten, möglichst viele Aufgaben eines Tests zu einem Kompetenzbereich wie bspw. dem Lesen. Das heißt, es werden Paralleltestformen, meist nach mehreren Wochen, durchgeführt, welche so konstruiert sind, „dass sie entweder aus einem gemeinsamen Itempool (Aufgabenpool) gezogen werden, oder aus ähnlich schweren Items aufgebaut sind“ (Diehl, Gebhardt & Mühling 2016a, S. 3). Nur wenn die Voraussetzung erfüllt ist, dass alle Paralleltests gleich schwierig sind, dürfen die Ergebnisse der Messzeitpunkte (MZP) in einer Lernverlaufsmessung zusammengefasst, interpretiert und dargestellt werden (vgl. Voß 2013).

Konstruiert werden CBM überwiegend nach der *Klassischen Testtheorie* (KTT) (vgl. Deno 2003). WILBERT und LINNEMANN (2011) betonen aber, dass es aufgrund der Überprüfung von Dimensionalität, Messinvarianz und Änderungssensibilität nicht mehr der aktuelle Stand der Wissenschaft ist. Die Autorengruppe um LEVUMI betont daher, dass sie aus diesem Grund zur Skalierung der Tests die Konzepte der *Item Response-Theory* (IRT) verwendet (vgl. Diehl, Gebhardt & Mühling 2016b). Das IRT

ist entgegen mancher Auffassung nicht als Alternative zur KTT zu betrachten. Das IRT modelliert das Zustandekommen einer Antwort auf eine Aufgabe, Frage oder Feststellung (*Item*). Anders als die KTT fokussiert sie sich auf die einzelnen Items, nicht auf den Test als solchen (vgl. Moosbrugger 2012).

Tatsächlich sind deutsche Schulleistungstests so konzipiert, dass sie sich typischerweise an den Lehrplänen soweit orientieren, als sie mithilfe eines Tests den Lernerfolg am Ende eines Lernprozesses summativ feststellen (vgl. Klauer 2014), „wohingegen das curriculumbasierte Messen den Lernfortschritt laufend dokumentiert, also formativ evaluiert und damit den Lernverlauf abbildet“ (Breitenbach 2014, S. 102). Insofern hebt sich das Verfahren von den klassischen Schulleistungstests ab. Ziel sollte stattdessen sein, was hier und jetzt gelehrt und gelernt worden ist (vgl. Klauer 2011). Demnach erfasst curriculumbasiertes Messen laut KLAUER (2006) auf zügige Weise den Lehr-Lern-Erfolg, und zwar im förderdiagnostischen Sinne. CBM kann somit als Vorläufer der heutigen *Lernverlaufsdagnostik* angesehen werden. „Bei der Lernverlaufsdagnostik geht es darum, statt einer punktuellen Testerhebung regelmäßig in relativ kurzfristigen Abständen die Entwicklung einer Kompetenz über längere Zeit hinweg durch geeignete Tests zu verfolgen“ (Klauer 2011, S. 207). Dabei soll es sich um immer neue Tests handeln, die aber stets die Ausprägung ein und derselben Kompetenz messen (vgl. ebd.). Der gemeinsame Kern soll im Folgenden durch den Begriff Lernverlaufsdagnostik konstatiert werden.

4.3 Lernverlaufsdagnostik Lesen

Lesen zu können, zählt unumstritten zu den wichtigsten Kulturtechniken in unserer Gesellschaft. Lesen ist eine komplexe Fähigkeit, die sich aus mehreren Teilfertigkeiten und Prozessen zusammensetzt, deren Entwicklung weit über den Primarbereich hinausgeht. Der Leselernprozess ist vor allem kein abgeschlossener oder zeitlich begrenzter Vorgang, wie die Ergebnisse der PISA-Studie zwischen den Jahren 2000 und 2006 verdeutlichen. Komplexe Lesefertigkeiten und -fähigkeiten bilden sich teilweise erst nach der Grundschule aus. Deshalb ist es von Bedeutung, den Leselernprozess über den Primarbereich hinaus diagnostisch zu begleiten (vgl. Hesse & Latzko 2011).

Die Diagnostik im Zusammenhang mit dem Aufbau von Lesekompetenzen wurde im deutschsprachigen Raum lange unter dem Begriff der „Legasthenie“ thematisiert. Ziel war es in erster Linie, SchülerInnen mit einer Leser-Rechtschreibschwäche zu diagnostizieren. Allerdings wurde Legasthenie nur dann identifiziert, wenn die Leistung des jeweiligen Individuums beim Lesen und Schreiben unter einen Prozentrang von 15 fiel und der Intelligenzquotient gleichzeitig größer als 85 war. Daraus ergab sich eine Diskrepanz, denn viele SchülerInnen mit einer ungünstigen Lese-Rechtschreib-Entwicklung, die nicht gleichzeitig die geforderte Intelligenzabweichung aufwiesen, wurden von jeweiligen Fördermaßnahmen ausgeschlossen. Unter dieser kritischen Perspektive kam es zu einer Intensivierung der Forschungsbemühungen zur Leseentwicklung. Dabei wurde vor allem darauf geachtet, Leselernprozesse von allen Kindern mit in den Fokus zu nehmen, unabhängig davon, ob sie der Legastheniekategorie zugeordnet wurden oder nicht. Ergebnisse der darauf folgenden Studien haben in erster Linie dargestellt, dass die traditionelle Diagnostik Defizite in der Leselernentwicklung von Kindern zeitlich zu spät identifiziert hat. Als zentrales Ergebnis hat sich schließlich gezeigt, dass Merkmale der phonologischen Informationsverarbeitung die ausschlaggebende Variable für einen erfolgreichen Schriftspracherwerb aller Heranwachsenden sind (ebd.)

Für den deutschsprachigen Raum stehen ab Ende der ersten Klasse zahlreiche Verfahren zur Verfügung, die die Lesekompetenzen, orthografische Fähigkeiten sowie das Leseverständnis objektiv, valide und in den meisten Fällen auch ökonomisch überprüfen (vgl. Mayer 2016). Zur Erfassung der Lesekompetenzen wird in der Grundschule häufig der KNUSPELS-L (Marx 1998) genutzt. Dabei handelt es sich um einen Lesetest für GrundschülerInnen vom Ende der ersten bis Ende der vierten Schulstufe. Die Bezeichnung KNUSPELS geht dabei auf die Knuspel-Fabelwesen zurück, die die Kinder im Test als Leitfigur begleiten, um eine kindgerechte Testatmosphäre zu schaffen (vgl. Hesse & Latzko 2011). Dieses Vorgehen verfolgen auch JUNGJOHANN, GEBHARDT, DIEHL und MÜHLING (2017) in ihrem Programm mit dem Drachen LEVUMI. Der SLRT II, eine Weiterentwicklung des *Salzburger Lese-Rechtschreibtests* (SLRT), erfasst aufseiten des Lesens die Genauigkeit sowie Geschwindigkeit des phonologischen Rekodierens und der direkten Worterkennung. Für die Ermittlung der Lesegeschwindigkeit bietet der Test normierte Vergleichswerte für das erste und zweite Schulhalbjahr zwischen der ersten und sechsten Klasse (vgl. Mayer

2016). Weitere Verfahren zur Überprüfung des Lernverlaufs der Lesekompetenz ist sowohl die *Lernfortschrittsdiagnostik Lesen* (LDL) als auch die *Verlaufsdagnostik sinnerfassenden Lesens* (VSL) nach WALTER (vgl. 2014). WALTER hat mit dem VSL ein Instrument zur längsschnittlichen Erfassung von Lesekompetenz entwickelt. Eingesetzt werden kann der Test dabei in den Klassen zwei bis sechs. Der LDL wurde ebenfalls von WALTER entwickelt und dient insbesondere der Verlaufsmessung allgemeiner Lesefähigkeit. Eingesetzt werden kann der Test sowohl an Grundschulen als auch an Haupt- und Förderschulen. Vorteilhaft ist, dass beide Tests sich durch eine sehr kurze Durchführungszeit (vier bzw. zwei Minuten) auszeichnen und zufriedenstellende psychometrische Eigenschaften hinsichtlich Validität und Reliabilität aufweisen (vgl. Börnert 2014).

4.4 Gütekriterien für Mess- und Testverfahren

Um eine möglichst hohe Kontrolle über einen Test zu erhalten, wird in der Testmethodologie mit verschiedenen Verfahren die *Testgüte* ermittelt. Die zur Bestimmung der Güte eines Tests benötigten Kriterien werden *Testgütekriterien* genannt. Diese Kriterien stellen Anforderungen an den Forschenden, die verwendeten Messinstrumente und somit an die gesamte Untersuchung (vgl. Sikora 2015).

Die Gütekriterien zu Testverfahren wurden im Rahmen der *Klassischen Testtheorie* (KTT) entwickelt und können ebenso auf empirische Forschungen im Allgemeinen übertragen werden. Es werden drei Hauptgütekriterien unterschieden, die im Nachfolgenden genauer beschrieben werden (ebd.).

Objektivität

Als objektiv gilt ein Test, wenn identische Erhebungsbedingungen und Auswertungsbedingungen bestehen. Die Berechnung dieses Maßes hat das Ziel, die Auswertung des Tests zu objektivieren (vgl. Schneider & Lindauer 2011). Die Objektivität bezeichnet infolgedessen den Grad einer Messung, in welchem ihre Ergebnisse unabhängig von der Person des Messenden sind. Im Einzelnen werden drei Formen der Objektivität unterschieden: die Durchführungsobjektivität, die Auswertungsobjektivität und die Interpretationsobjektivität (vgl. Sacher 2014).

Reliabilität

Die Reliabilität eines Tests sagt etwas darüber aus, wie genau und sicher er ein bestimmtes Merkmal misst. Dabei werden verschiedene Methoden verwendet, um den Test zu quantifizieren. In der Testpsychologie und auch bei der Konstruktion von schulischen Leistungstests wird die Reliabilität mit der Wiederholungs-, Paralleltest- und Halbierungsmethode überprüft (vgl. Sacher 2014). Bei der gängigsten Form, der Halbierungsmethode, werden die Testaufgaben in zwei zufällig zusammengestellte gleich große Teile halbiert, die dann miteinander korreliert werden. Ist das Ergebnis nur eine geringe Korrelation, so muss davon ausgegangen werden, dass die Items ein bestimmtes Merkmal nicht präzise genug erfassen (vgl. Schneider & Lindauer 2011).

Validität

Validität ist das Kernkriterium der Testgüte. Die Validität bzw. Gültigkeit stellt fest, ob ein Test das misst, was er messen soll. Gewöhnlich werden fünf Aspekte der Validität unterschieden: die Inhaltsvalidität, die Prognosevalidität, die Übereinstimmungsgültigkeit, die Konstruktvalidität und die Testfairness (vgl. Sacher 2014).

Neben den Hauptgütekriterien werden häufig noch weitere Nebengütekriterien aufgeführt: die der *Ökonomie*, der *Nützlichkeit* und *Zumutbarkeit* sowie der *Akzeptanz* (vgl. Sacher 2014; Sikora 2015).

4.5 Bezugsnormen

Um etwas zu bewerten, muss das Werturteil gewissermaßen in einer Norm justiert werden, von der her sich bestimmt, was als gut oder schlecht anzusehen ist. Grundsätzlich werden drei Bezugsnormen unterschieden:

Die **soziale Norm**: Einzelleistungen werden dabei nach ihrem Verhältnis zur Leistung einer Gruppe beurteilt.

Die **kriteriale Norm**: Der Beurteilung liegen fachliche bzw. sachliche Anforderungen zugrunde, die unabhängig von der Gruppenleistung formuliert werden.

Die **individuelle Norm**: Ausschlaggebend für die Bewertung ist der Lernfortschritt, der sich in der vorliegenden Leistung zeigt (vgl. Sacher 2014).

5 Fragestellung und Hypothesen

Basierend auf den theoretischen Grundlagen wurde im Rahmen dieser Arbeit eine Normierungsstudie für ein neu entwickeltes Lernverlaufsdiagnostikverfahren zum sinnentnehmenden Lesen, der Onlineplattform LEVUMI, durchgeführt. Dabei stehen folgende Fragestellungen im Zentrum des Interesses:

- 1) Wie ist die Itemschwierigkeit des Tests zum sinnentnehmenden Lesen für SchülerInnen der dritten Schulstufe zu bewerten?
- 2) Wie verändern sich die Ergebnisse aller SchülerInnen vom ersten zum zweiten Messzeitpunkt?

H₀ Es gibt beim Lernverlaufstest zum sinnentnehmenden Lesen (LeVuMi) keinen Unterschied zwischen den Leistungen der SchülerInnen vom ersten zum zweiten Messzeitpunkt.

H_A Es gibt beim Lernverlaufstest zum sinnentnehmenden Lesen (LeVuMi) einen Unterschied zwischen den Leistungen der SchülerInnen vom ersten zum zweiten Messzeitpunkt.

- 3) Wie verändern sich die Ergebnisse der an dieser Studie teilnehmenden SchülerInnen mit dem Förderschwerpunkt Lernen vom ersten zum zweiten Messzeitpunkt?

H₀ Es gibt beim Lernverlaufstest zum sinnentnehmenden Lesen (LeVuMi) keinen Unterschied zwischen den Leistungen der SchülerInnen mit dem Förderschwerpunkt Lernen vom ersten zum zweiten Messzeitpunkt.

H_A Es gibt beim Lernverlaufstest zum sinnentnehmenden Lesen (LeVuMi) einen Unterschied zwischen den Leistungen der SchülerInnen mit dem Förderschwerpunkt Lernen vom ersten zum zweiten Messzeitpunkt.

- 4) Wie verändern sich die Ergebnisse der SchülerInnen mit dem Förderschwerpunkt Lernen, die inklusiv beschult werden, gegenüber den SchülerInnen mit dem Förderschwerpunkt Lernen an einer Förderschule?

H_0 Beim Test zum sinnentnehmenden Lesen (LeVuMi) erzielen die inklusiv beschulten SchülerInnen mit dem Förderschwerpunkt Lernen vom ersten zum zweiten Messzeitpunkt keine besseren Ergebnisse.

H_A Beim Test zum sinnentnehmenden Lesen (LeVuMi) erzielen die inklusiv beschulten SchülerInnen mit dem Förderschwerpunkt Lernen vom ersten zum zweiten Messzeitpunkt bessere Ergebnisse.

H_0 Beim Test zum sinnentnehmenden Lesen (LeVuMi) erzielen die SchülerInnen mit dem Förderschwerpunkt Lernen (die eine Förderschule besuchen) vom ersten zum zweiten Messzeitpunkt keine besseren Ergebnisse.

H_A Beim Test zum sinnentnehmenden Lesen (LeVuMi) erzielen die SchülerInnen mit dem Förderschwerpunkt Lernen (die eine Förderschule besuchen) vom ersten zum zweiten Messzeitpunkt bessere Ergebnisse.

5) Erzielen die SchülerInnen mit dem Förderschwerpunkt Lernen, die inklusiv beschult werden, bessere Ergebnisse als SchülerInnen mit dem Förderschwerpunkt Lernen an einer Förderschule?

H_0 Es gibt beim Lernverlaufstest zum sinnentnehmenden Lesen (LeVuMi) keinen Unterschied zwischen den Leistungen der SchülerInnen mit dem Förderschwerpunkt Lernen, die inklusiv beschult werden, und denen mit dem Förderschwerpunkt Lernen, die eine Förderschule besuchen.

H_A Es gibt beim Lernverlaufstest zum sinnentnehmenden Lesen (LeVuMi) einen Unterschied zwischen den Leistungen der SchülerInnen mit dem Förderschwerpunkt Lernen, die inklusiv beschult werden, und denen mit dem Förderschwerpunkt Lernen, die eine Förderschule besuchen.

6 Untersuchungsmethode – LeVuMi

Den bisherigen Ausführungen ist zu entnehmen, welchen Stellenwert die Lernverlaufsdagnostik für ein inklusives Setting hat. Allerdings stellt die Verlaufsdagnostik durch häufige Überprüfungen, den hohen Zeitaufwand und ständige Auswertungen Lehrkräfte vor große Herausforderungen. Zudem ergibt sich durch Relevanz auch ein enormer methodischer Anspruch. Das Erhebungs-konstrukt muss daher gewissen methodischen Standards entsprechen wie bspw. den o. g. klassischen Gütekriterien.

Zur Beantwortung der im vorherigen Kapitel genannten Fragestellungen und zur Abhilfe der genannten Problematik wird die Lernverlaufsdagnostik der Online-Lernplattform LeVuMi (**Lern-Verlaufs-Monitoring**) hinzugezogen. Das methodische Vorgehen der Lernplattform wird im Nachfolgenden beschrieben.

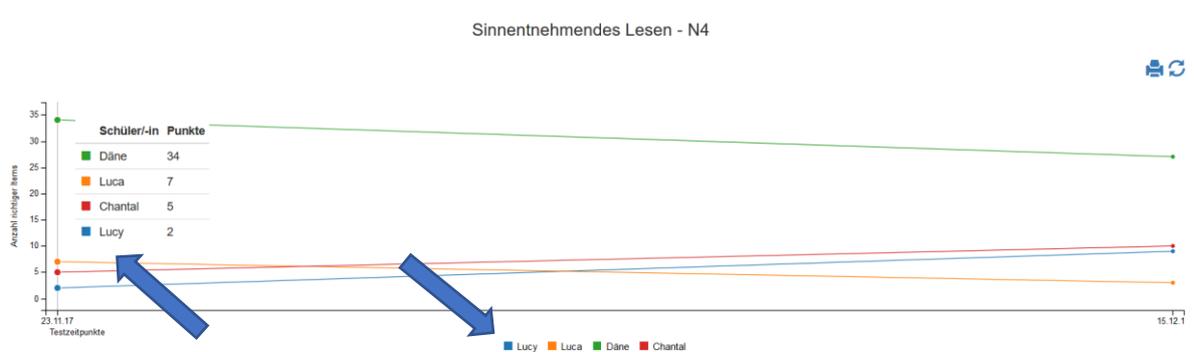
6.1 Aufbau der Onlineplattform

Die Online-Plattform www.Levumi.de ist ein langfristig angelegtes gemeinsames Forschungsprojekt der Wissenschaftler MARKUS GEBHARDT (Technische Universität Dortmund), KIRSTEN DIEHL (Europa-Universität Flensburg) und ANDREAS MÜHLING (Universität Kiel). Die Lernplattform ist komplett kostenfrei und setzt lediglich den Browser *Mozilla Firefox* voraus. Eine ständige Onlineverbindung ist selbstverständlich Voraussetzung (vgl. Gebhardt, Diehl & Mühling 2016b). Das Ziel der Plattform LeVuMi ist es einerseits, „[...] die Forschung zur Lernverlaufsdagnostik zu vertiefen und andererseits, der Schulpraxis ein praktikables Online-Instrument zur Lernverlaufsmessung anzubieten“ (Gebhardt, Diehl & Mühling 2016b, S. 447). Ziel ist außerdem eine Optimierung sowie Erweiterung der Testauswahl über lange Sicht.

Die Anwendung eines computergestützten Onlinetests für die Lernverlaufsdagnostik bietet gegenüber einer papierbasierten Version erhebliche Vorteile. So übernimmt der Computer für die Lehrkraft die sehr zeitaufwändige Phase der Auswahl der Aufgaben, der Auswertung und Darstellung. Ebenso verlangt die Onlinebasierung keine bestimmten Anforderungen an das lokale Gerät, mit Ausnahme des Browsers. Zudem sind die Daten jederzeit von jedem Gerät aus abrufbar. Darüber hinaus hat die Lehrkraft aber auch die Möglichkeit, den Test als Paper-Pencil-Test durchzuführen, allerdings müssen die Ergebnisse dann von Hand in das System eingegeben werden (vgl. Gebhardt, Diehl & Mühling 2016b). Die äußere Gestaltung der Plattform

LEVUMi verzichtet weitestgehend auf dekorative Elemente, die bei der Testdurchführung ablenken könnten. Das System ist lediglich auf die notwendigen Funktionen – sowohl für Testleiter als auch SchülerInnen – begrenzt. Die Einarbeitung in das Programm ist abhängig von den individuellen Fähigkeiten der Testleitung. Die Plattform bietet aber neben den Lernverlaufstest auch deren Auswertung mit qualitativen und quantitativen Daten an, was die Attraktivität für Lehrkräfte deutlich steigert. Die Ergebnisse des jeweiligen Tests werden in einem Lernverlaufsgraphen dargestellt, der neben den unterschiedlichen MZP auch eine Klassen- sowie Schüleransicht bietet (vgl. Gebhardt, Diehl & Mühling 2016a & b).

Abbildung 2: Klassenübersicht mit zwei Messzeitpunkten



In der Schülerübersicht kann die Lernentwicklung jedes/r einzelnen/r Schülers/in (durch eine Farbe von den anderen abgehoben) betrachtet und mit den Ergebnissen der MitschülerInnen verglichen werden. Die Ergebnisse werden als Prozentränge (ordinale Datendarstellung) dargestellt (vgl. Gebhardt, Diehl & Mühling 2016a).

Zum aktuellen Zeitpunkt umfasst LEVUMi im Kompetenzbereich Mathematik Tests zu den Lernbereichen *Zahlen lesen* und *Zahlenstrahl* sowie für den Kompetenzbereich Deutsch die Bereiche *Leseflüssigkeit*, *Rechtschreibung*, *Wortschatz* und *sinnentnehmendes Lesen*. Die Tests zum Kompetenzbereich Lesen in LEVUMi orientieren sich am Lehrplan der Grundschule für das Fach Deutsch und am Kieler Leseaufbau (KLA). Der KLA kommt ursprünglich aus der Lesetherapie und wird insbesondere bei SchülerInnen mit Lernschwierigkeiten genutzt. Der KLA gliedert sich nach linguistischen Kriterien in 14 Schwierigkeitsstufen auf, die von LEVUMi anhand von Niveaustufen (N0 bis N4) abgebildet werden (vgl. Jungjohann et al. 2017). Die verschiedenen Niveaustufen unterscheiden sich hinsichtlich des Anspruchs bzw. Schweregrades und der erwarteten Kompetenzentwicklung, wobei die Niveaustufe N0 nur für SchülerInnen mit einem sonderpädagogischen Förderbedarf vorgesehen ist (ebd.).

2016a). Für den Test zum sinnentnehmenden Lesen stehen aktuell allerdings nur die Stufen N2 und N4 zur Verfügung. Die für die Durchführung dieses Projekts ausgewählte Stufe war N4.

Bei den bisher abgebildeten Tests auf der Plattform LEVUMi handelt es sich um Geschwindigkeitstest (Speedtests). Das bedeutet, dass der jeweilige Test eine Zeitbeschränkung hat und dass die Lösungsgeschwindigkeit der bearbeiteten Items vom/ von dem/ der SchülerIn ermittelt und abgebildet wird. Der Ablauf der Zeit wird während des Test allerdings nicht angezeigt, um die SchülerInnen nicht einem überflüssigen Druck auszusetzen.

Da das Projekt LEVUMi noch am Anfang seiner Entwicklung steht, liegen noch keine ausführlichen Daten zu den Gütekriterien vor. Zudem sind viele Fragen bezüglich der Umsetzung und Anwendung der Lernverlaufsdagnostik noch offen. Zwar wurden die verschiedenen Tests von LeVuMi alle aus reliablen Tests entwickelt, jedoch können aufgrund der geringen Stichprobengröße noch keine fundierten Aussagen bezüglich Objektivität, Reliabilität oder Validität getätigt werden. Zudem ist der Bereich der Online-Lernverlaufsdagnostiken im deutschsprachigen Raum bisher wenig untersucht, was konkrete Aussagen zusätzlich erschwert (vgl. Jungjohann et al. 2017).

6.2 Konstruktion des Tests zum sinnentnehmenden Lesen

LEVUMi hat sein Konzept im Bereich der Leseförderung nach Bausteinen aufgebaut. Der Aufbau nach Förderbausteinen macht die Arbeit mit dem System für Lehrkräfte transparenter und verdeutlicht, welche Ansprüche hinter den einzelnen Übungsaufgaben stehen bzw. zu welchen Förderzielen sie passen. Zudem erläutern sie einerseits die unterschiedlichen Entwicklungsbereiche, die im Leseanfangsunterricht wichtig sind, und spezifizieren andererseits die Förderung (vgl. Jungjohann et al. 2017).

Abbildung 3: Förderbausteine und Entwicklungsbereiche von LEVUMi (vgl. Jungjohann et al. 2017)

6. Miteinander Lesen

- lautes Lesen für die Leseflüssigkeit

5. Inhalte verstehen

- sinnentnehmendes Lesen von Anfang an

4. Wörter kennen und erkennen

- automatisierte Worterkennung

3. Richtiges Lesen

- phonologisches Rekodieren

2. Buchstaben-Laut-Beziehung

- Graphem-Phonem-Korrespondenz

1. Vorläuferfähigkeiten

- Phonologisches Bewusstsein

Der 5. Förderbaustein bei LEVUMi stellt den Bereich „Inhalte verstehen“ dar. Hier steht das semantisch-lexikalische Lesen im Vordergrund. Jede/r SchülerIn erhält für sich einen einzelnen Test, der anhand eines Aufgabenpools (Itempools) ermittelt wird. Bei der ersten Durchführung ist die Reihenfolge der Items noch identisch. Das System wählt dann ab dem zweiten MZP zufällig aus, sodass mehrere Paralleltests gebildet werden. Das heißt, dass jede/r SchülerIn einen eigenen Test erhält, allerdings mit einer unterschiedlichen Reihenfolge. Aus diesem Grund ist es entscheidend, dass sich die Items in ihrer inhärenten Schwierigkeit gleichen. Wäre das nicht der Fall, könnte es passieren, dass ein/e SchülerIn zu einem MZP weniger Aufgaben löst als in einem vorherigen oder späteren Test (ebd.). Des Weiteren wurde bei der Konstruktion der Tests und somit bei der Erstellung der Items von der Autorengruppe um LEVUMi darauf geachtet, lehrplanferne Aufgabestellungen zu konzipieren, um einen bundesland- und schulartunabhängigen Einsatz zu ermöglichen. Zudem wurde darauf geachtet, die Tests für SchülerInnen mit und ohne sonderpädagogischen Förderbedarf oder Migrationshintergrund zu konstruieren (vgl. Gebhardt, Diehl & Mühling 2016a).

Wie bereits erwähnt, bietet LEVUMi aktuell zwei Tests für den mathematischen Bereich und vier für den deutschen Bereich an. Der für diese Studie genutzte Test ist

der des *sinnentnehmenden Lesens* mit der Niveaustufe N4. Der Test ist schülerzentriert und wird im Schülerzugang eigenständig von den Kindern durchgeführt. Er besteht aus 63 Items, die innerhalb von 480 sec. bearbeitet werden müssen. Die SchülerInnen müssen während des Tests Sätze vervollständigen, indem sie sinnentnehmend lesen und aus vier möglichen Items das Richtige einsetzen.

Beispiel: Ein Lama hat vier **Beine**/Daumen/Kamele/Bücher.

Der Test kann jedoch auch länger als 480 sec. dauern, insofern die Bearbeitung des letzten Items in die letzten Sekunden fällt. Infolgedessen hört der Test erst nach der Bearbeitung des letzten angezeigten Items auf. Die zu bearbeitenden Items sind:

aber, Anstatt, Auf, Augen, außer, Baby, backt, Beine, Bett, bevor, Biene, Bilder, Blumen, böse, Büro, dicke, Durch, Enten, fleißig, Freunde, Frösche, für, gut, Hase, Haus, hungrig, In, isst, kauft, lustigen, mit, Nachdem, neues, runde, sammeln, scheint, Schere, schläft, schmeckt, schnelle, Schuhe, Schwester, Sobald, Sonne, spitz, spricht, süß, Tür, über, unter, Vogel, von, Während, wartest, Wasser, weder, wegen, weil, Wenn, wohne, Zwischen

Die dargestellten Worte sind maximal zweisilbig, verwenden alle Buchstaben des KLA und haben Konsonantencluster inne. Die Items sind nochmals nach Kategorien mit unterschiedlichen Schwierigkeitsgraden eingeteilt, aus denen das System für jede/n SchülerIn einen Test gleicher Schwere konstruiert.

6.3 Darstellung und Interpretation der Ergebnisse

Die Ergebnisse werden bei LEVUMI, wie bereits oben beschrieben, in Prozenträngen in einem Lernverlaufsgraphen (Abbildung 2) dargestellt. Für die Erstellung der Prozentränge (25, 50, 75) wird die Verteilung der Klasse in vier gleich große Gruppen (Quartile) vorgenommen. So bildet der Prozentrang 50 (auch Median) den Wert, den der/die mittlere SchülerIn zu dem angegebenen Testzeitpunkt hatte. Vorteilhaft ist bei der Verwendung des Prozentrangs 50, dass er weniger sensibel auf schiefe Verteilungen reagiert als der Mittelwert der jeweiligen Klasse. Bei einer Normalverteilung

der Ergebnisse innerhalb der Testdurchführung einer Klasse entspricht der Mittelwert dem Prozentrang 50 (vgl. Gebhardt, Diehl & Mühling 2016a).

Neben den Graphen werden aber auch für jede Messung qualitative und quantitative Ergebnisse erfasst. So werden pro Testzeitpunkt die richtig gelösten Items, die falschen Items, die Anzahl der gelösten Items, die Anzahl der falsch gelösten Items sowie die Lösungswahrscheinlichkeiten festgehalten. Dabei gibt die Lösungswahrscheinlichkeit (in Prozent) an, wie hoch der Anteil der Anzahl der gelösten Items im Vergleich zur Anzahl der falsch gelösten Items war. Würde ein/e SchülerIn bspw. eine Lösungswahrscheinlichkeit von 100 % besitzen, bedeutete dies, dass er/sie alle Items, die im Test innerhalb des Zeitfensters tatsächlich präsentiert wurden, gelöst habe, wohingegen bei einer Lösungswahrscheinlichkeit von 50 % theoretisch nur die Hälfte dieser Items gelöst wurde. Durch diesen Hinweis kann bestimmt werden, ob der/die jeweilige SchülerIn die Items zwar langsam liest, aber dafür genau (ebd.). Darüber hinaus werden die einzelnen gelösten sowie nicht gelösten Items gelistet. So ist bezüglich des Tests zum sinnentnehmenden Lesen zu erkennen, ob ein/e Schüler/In das jeweilige Item *richtig* oder *falsch* gelöst bzw. *nicht bearbeitet* hat.

6.4 Analyse der Itemschwierigkeit

Um den Lernverlauf der SchülerInnen möglichst exakt zu erheben, müssen an das diagnostische Verfahren, in diesem Fall den Test zum sinnentnehmenden Lesen, möglichst hohe Ansprüche gestellt werden. Dazu gehört neben den traditionellen Gütekriterien auch die Berechnung der Schwierigkeit der Items (vgl. Moosbrugger & Kelava 2012). Im Idealfall besteht ein Test, der einen Lernverlauf diagnostiziert, „[...] aus einer Reihe exakt gleichschwerer (sic!) Test mit exakt gleicher Trennschärfe“ (Wilbert & Linnemann 2011, S. 228). Präziser ausgedrückt bedeutet das, dass es sich um *Paralleltestungen* handeln soll – als Tests, die truescore- und fehlervarianzhomogen sind (vgl. ebd.). Daher müssen zur Erhebung des Lernverlaufs die Items eines Tests ein identisches Anforderungsniveau besitzen, das heißt, sie müssen homogen sein (vgl. Klauer 2014). Ansonsten können keine Aussagen über den tatsächlichen Lernverlauf der SchülerInnen gemacht werden.

Um ein identisches Anforderungsniveau zu gewährleisten, wäre die logische Konsequenz, ein und denselben Test mehrfach zu verwenden. Da der Test aber schon

alleine durch die eine regelmäßige Wiederholung für die SchülerInnen leichter werden würde, ist es fraglich, ob er bis zum Ende des Testzeitraumes noch dasselbe messen würde. Dementsprechend muss der neue Test so konstruiert sein, dass er stets dasselbe Leistungsspektrum abdeckt und dabei immer gleich schwer bleibt (vgl. Klauer 2014). Um diesem Problem Abhilfe zu schaffen, greift die Onlineplattform LEVUMi auf die bereits beschriebene Paralleltestmethode zurück, bei der aus einem Pool von Aufgaben (Items) ein zufälliger Test konstruiert wird (vgl. Gebhardt, Diehl & Mühling 2016b). Nur so ist es zulässig, die Ergebnisse miteinander zu vergleichen, um Rückschlüsse auf Leistung und Lernverlauf der SchülerInnen zu ziehen (vgl. Voß 2013).

Ermittelt wird die Schwierigkeit der Items zum sinnentnehmenden Lesen in dieser Studie nach den Autoren MOOSBRUGGER und KELAVA (2012). Sie benennen als Maß der Schwierigkeit den *Schwierigkeitsindex* P_i . Der Index „[...] eines Items i ist der Quotient der bei diesem Item tatsächlich erreichten Punktsomme aller n Probanden und der maximal erreichbaren Punktsomme“ (Moosbrugger & Kelava 2012, S. 76). Im Fall dieses Tests wird die erreichte Punktsomme aller n Probanden durch die Anzahl aller richtig eingesetzten Items dargestellt. Die durch das System LEVUMi maximal errechnete Punktsomme stellt die Anzahl der richtig gegebenen Antworten für das abgebildete Item dar. Bei einer einfachen Kodierung der Items (richtig gelöst = 1, falsch gelöst = 0) ergibt sich infolgedessen (ebd.):

$$P_i = \hat{x}_i * 100 = \frac{\text{Anzahl richtiger Antworten}}{\text{Anzahl gesamter Antworten}} * 100$$

Durch das Multiplizieren mit 100 kann der Schwierigkeitsindex P_i Werte zwischen 0 und 100 erreichen. Dabei gilt: Je höher der erreichte Wert pro Item, desto leichter ist das Item zu lösen. Damit spiegelt „[...] die numerische Höhe des Schwierigkeitsindex P_i [...] die Leichtigkeit des Items i [wider] und nicht [dessen] Schwierigkeit“ (Moosbrugger & Kelava 2012, S. 76), was ihn vom Schwierigkeitsindex der *Item Response-Theory* (IRT) unterscheidet. Bei der IRT ist der Schwierigkeitsparameter so definiert, „[...] dass seine numerische Höhe tatsächlich die Schwierigkeit kennzeichnet“ (Moosbrugger & Kelava 2012, S. 77). Eine Auswertung nach der IRT würde allerdings den Rahmen dieser Arbeit übersteigen. Eine Auswertung des Schwierigkeitsindex nach

der *Speedtest*-Variante kann leider auch nicht erfolgen, da hierfür nicht nur die Anzahl der richtigen (R-Antworten) und falschen (F-Antworten) gelösten Items benötigt wird, sondern auch die Anzahl derer, die entweder ausgelassen (A-Antworten) oder aufgrund von Zeitmangel (U-Antworten) nicht bearbeitet wurden. Letztere bildet das System LeVuMi ungünstigerweise identisch ab.

Items von mittlerer Schwierigkeit ($P_i = 50$) sind grundsätzlich zu bevorzugen, weil sie am ehesten die Grundlage für differenzierte Aussagen über die Merkmalsausprägungen bieten. Ist es gewünscht, dass ein Test sowohl sehr hohe als auch sehr niedrige Merkmalsausprägungen besitzt, so sollten auch Items mit Schwierigkeiten von $5 \leq P_i \leq 20$ (schwer) bzw. $80 \leq P_i \leq 95$ (leicht) enthalten sein. Je nach Zielsetzung des Tests sollten daher sehr schwere oder sehr niedrige Items aussortiert bzw. gleichmäßig verteilt werden. Es ergibt sich dem Dreischnitt folgend demnach für die numerische Abgrenzung nachstehende Einteilung der Kennwerte (ebd.):

Leichte Items: $100 \geq P_i \geq 67$

Mittelschwere Items: $66 \geq P_i \geq 34$

Schwere Items: $33 \geq P_i \geq 1$

6.5 Berechnung der statistischen Signifikanz

Die wahrscheinlich am häufigsten verwendeten Verfahren in der sonderpädagogischen Forschung zur Berechnung der statistischen Signifikanz ist die *Varianzanalyse* und der *t*-Test. Diese Verfahren werden benötigt, wenn Unterschiede zwischen den Kennwerten (Mittelwerten) einer Gruppe und gegebenen Werten oder Werten anderer Gruppen nur zufällig oder systematisch, das heißt *signifikant* sind. Der *t*-Test wird für den Vergleich von max. zwei Gruppen herangezogen, die Varianzanalyse hingegen findet bei mehr als drei Gruppen Verwendung (vgl. Kuhl & Sinner 2015) und wird daher aufgrund der Zielsetzung dieser Arbeit nicht benötigt.

Der *t*-Test prüft aufgrund eines empirischen Ereignisses, welche der im Vorfeld aufgestellten Hypothesen (*Null-* und *Alternativhypothese*) angenommen werden kann. Eine Nullhypothese nimmt an, dass es keinen Unterschied zwischen zwei Gruppen oder Messwerten gibt. Die Alternativhypothese hingegen nimmt an, dass es einen

Unterschied gibt (vgl. Rasch et al. 2014). Das Ergebnis eines t -Tests wird wie folgt dargestellt:

$$t(df) = t\text{-Wert}, p = \text{Signifikanz}$$

Dabei steht t für die Prüfgröße, die aus den Mittelwerten, den Standardfehlern und der Stichprobengröße berechnet wurde. Df steht für die *Freiheitsgrade*. Sie bestimmen das Aussehen der t -Verteilung und damit auch, ab wann ein Wert signifikant wird. Die Angabe *Sig. (2-seitig)* stellt den wichtigsten Wert dar, die *statistische Signifikanz* (ebd.).

Die Art der Abhängigkeit der betrachteten Stichproben entscheidet über die Wahl des richtigen t -Tests. Die für diese Arbeit verwendeten Verfahren des t -Tests werden im Nachfolgenden beschrieben.

t -Test für zwei abhängige Stichproben

Sollen zwei Werte miteinander verglichen werden, die in einer Stichprobe (zu unterschiedlichen Zeitpunkten) erhoben wurden, wird der t -Test für *abhängige* oder *verbundene* Stichproben durchgeführt. Dabei wird jedem Wert in einer Stichprobe der Wert in einer anderen Stichprobe zugewiesen. Die Literatur spricht hier von Wertpaaren (vgl. Sedlmeier & Renkewitz 2013). Der t -Test für abhängige Stichproben prüft also, ob die gefundene Mittelwertdifferenz zwischen den Stichproben „bedeutsam von der für die entsprechenden Populationen vermuteten Mittelwertdifferenz abweicht“ (ebd. S. 402). Daher ist der Unterschied zwischen zwei abhängigen Stichproben der Mittelwert der Differenz aller n Wertpaare. Somit wird sich die Stichprobengröße n , die für diesen Test verwendet wird, auf die Anzahl der Wertpaare und nicht der Messwerte beziehen. Der t -Wert wird daher wie folgt berechnet (ebd.):

$$t = \frac{\bar{x}_{diff} - \mu_{diff}}{\hat{\sigma}_{\bar{x}_{diff}}}$$

t-Test für zwei unabhängige Stichproben

Soll überprüft werden, ob sich die Mittelwerte zweier unabhängiger Gruppen (A und B) signifikant voneinander unterscheiden, wird der *t*-Test für *unabhängige* Stichproben verwendet. Diese Unabhängigkeit bedeutet, dass die Werte einer der Gruppen keinerlei Vorhersage über die der anderen erlauben. Für die Berechnung dieses Mittelwertunterschieds wird nachfolgende Formel genutzt (ebd.):

$$t = \frac{(\hat{x}_A - \hat{x}_B) - (\mu_A - \mu_B)}{\hat{\sigma}_{\hat{x}_A - \hat{x}_B}}$$

Da in dieser Arbeit u. a. die Leistungen der SchülerInnen, die inklusiv beschult werden, mit denen, die eine Förderschule besuchen, verglichen werden, bietet sich der *t*-Test für unabhängige Stichproben als ideal an.

6.6 Stichprobe

Die Stichprobe dieser Arbeit setzt sich aus vier Grundschulen und zwei Förderschulen mit dem Förderschwerpunkt Lernen zusammen. Durch die leider geringe Rückmeldung seitens der Schulen musste ein großer Bezirk ausgewählt werden, womit sich der Kreis der teilnehmenden Schulen auf die Städte Dortmund, Bochum und Herne erstreckte. Somit beläuft sich die Gesamtzahl dieser Studie auf $n=203$ SchülerInnen, wobei aufgrund von Abwesenheit insgesamt an beiden MZP $n=198$ anwesend waren. Von diesen $n=198$ SchülerInnen waren $n=99$ (50 %) männlich und $n=99$ (50 %) weiblich. Die SchülerInnen waren zu den MZP zwischen 8 und 11 Jahren alt. Sie besuchten alle die dritte Schulstufe.

Von der Gesamtstichprobe hatten $n=25$ (12,6 %) einen diagnostizierten Förderschwerpunkt und $n=117$ (59,1 %) eine Migrationsgeschichte. Bei den SchülerInnen mit einem zugeschriebenen Förderschwerpunkt hatten anteilig an der Gesamtstichprobe $n=1$ (0,5 %) den Förderschwerpunkt *geistige Entwicklung*, $n=19$ (9,6 %) den Förderschwerpunkt *Lernen* und $n=5$ (2,5 %) einen *anderen* Förderbedarf außer den genannten. Darüber hinaus haben $n=47$ (23,7 %) SchülerInnen eine Förderung im Fach Deutsch erhalten.

Leider gestaltete sich die Suche nach weiteren Förderschulen mit dem Förderschwerpunkt Lernen als sehr schwierig. Dies ist zum einen durch die mittlerweile sehr

geringe Dichte an Schulen dieser Form begründet und zum anderen durch die Sorgen der Lehrkräfte, dass die Anforderungen des Tests zu hoch sind. Inwieweit diese Sorge begründet ist, wird in den nachfolgenden Kapiteln erläutert. Aufgrund dieser Tatsache beläuft sich die Anzahl der SchülerInnen an der Gesamtstichprobe mit dem Förderschwerpunkt Lernen, die eine Förderschule besuchen, auf $n=8$ (4,1 %) und der SchülerInnen, die inklusiv an einer Grundschule beschult werden, auf $n=11$ (5,5 %). Die Ergebnisse zu dieser Stichprobe werden in Kapitel 6.8. genauer dargestellt.

6.7 Durchführung

Die Durchführung der Tests gestaltete sich an den einzelnen Schulen sehr unterschiedlich. Dies lässt sich zum einem damit begründen, dass die Schulen multimedial teilweise sehr unterschiedlich ausgestattet sind, und zum anderen damit, dass die Absprache innerhalb der Klassenteams unvollständig war. So waren teilweise nicht alle Lehrkräfte der zu testenden Klassen innerhalb einer Schule über den Testzeitpunkt informiert. Zudem erfüllten nicht alle Schulen die Standardvoraussetzungen des im Vorgesprächs mit den Lehrkräften betonten und benötigten Browsers *Mozilla Firefox*, der folglich zu Beginn der Testungen zunächst installiert werden musste – ein erheblicher, nicht eingeplanter Zeitaufwand.

Vorbereitung:

Die Kontaktaufnahme gestaltet sich sehr unterschiedlich – teilweise telefonisch bzw. über den Postweg, über die Schulleitung oder über das Klassenteam selbst. Da die Testungen an zwei MZP im Abstand von drei Wochen erfolgen mussten, war es notwendig, zwei mögliche Zeitpunkte zu finden, die in den Stunden- und Zeitplan der jeweiligen Klasse passten – ein Vorhaben, das sich aufgrund der verschiedenen Aktivitäten der SchülerInnen und LehrerInnen nicht immer als leicht erwiesen hat.

Da jede/r SchülerIn einen eigenen Zugang zum System erhält, mussten vor der ersten Testung alle teilnehmenden Kinder auf www.Levumi.de angemeldet werden. Diesbezüglich mussten von den Lehrkräften im Vorfeld bestimmte Daten der Kinder eingeholt werden, die für die anschließende Ergebnisbeurteilung von Bedeutung sind. Die einzuholenden Informationen sind Alter (Monat und Jahr), Migrationshintergrund,

Geschlecht und Förderschwerpunkt (*Lernen, geistige Entwicklung, Deutsch* oder *Andere*)

Durchführung

Die Testdurchführung war an den einzelnen Schulen ebenfalls sehr verschieden. Bedingt durch die teilweise sehr unterschiedlich ausgestatteten Computerräume der Schulen, war es nicht möglich, immer eine einheitliche Gruppengröße zu testen. Dementsprechend orientierte sich die Anzahl der gleichzeitig getesteten SchülerInnen an der Anzahl der verfügbaren Computerplätze. Die jeweiligen Kinder wurden also zum vereinbarten Zeitpunkt vom Testleiter in einer nicht beliebigen Reihenfolge aus dem Klassenverband geholt und in den Computerraum begleitet. Da der Test ein Einzeltest ist, wurde jedem Kind ein Computerplatz zugewiesen. Aus Gründen der Zeitersparnis wurden die Kinder vorher vom Testleiter bereits im System angemeldet. Um einen Überblick über die bisherigen Erfahrungen der Kinder mit dem Medium Computer zu bekommen, wurde eine kurze Vorstellungsrunde in der jeweiligen Kleingruppe durchgeführt. Dies diente einerseits dazu, den Kindern den Ablauf des Tests zu erläutern, und andererseits der Reduzierung ihrer Nervosität. Die Erläuterung des Vorhabens erfolgte mittels der *Wörtlichen Instruktionen zum sinnentnehmenden Lesen* (2017; Anhang 1.) und des Verweises auf die Möglichkeit, während der Durchführung bei Unklarheiten jederzeit Fragen zu stellen.

Der Test selbst beginnt mit einem einführenden Beispielsatz, bei dem das Kind aus vier vorgegeben Wörtern das passende auswählen muss, um den Satz zu vervollständigen. Damit ist gewährleistet, dass der Ablauf des Tests verstanden wurde.

Nachbereitung:

Die Nachbereitung erfolgte durch die Kontrolle der Übernahme der Testdaten zum jeweiligen Zeitpunkt sowie die Übermittlung der Ergebnisse der einzelnen Klassen an die jeweiligen Lehrkräfte nach dem zweiten MZP.

6.8 Ergebnisse

Im Rahmen dieser Arbeit wurde der Test zum sinnentnehmenden Lesen der Online-Plattform LEVUMi an verschiedenen Schulen durchgeführt. Die Ergebnisse werden

nachfolgend beschrieben. Die Auswertung dieser Studie wurde mit den Programmen Microsoft EXCEL und IBM SPSS 25 durchgeführt. Die Auswertung erfolgt zunächst quantitativ durch die Analyse der Ergebnisse des Tests. Neben den Werten der Gesamtstichprobe werden zusätzlich die Werte von zwei Subgruppen betrachtet (Förderschwerpunkt Lernen inklusiv beschult – Förderschwerpunkt Lernen an einer Förderschule). Eine qualitative Ergebnisdarstellung erfolgt auf Grundlage von Beobachtungen während der Testdurchführung.

Forschungsfrage 1:

Wie ist die Itemschwierigkeit des Tests zum sinnentnehmenden Lesen für SchülerInnen der dritten Schulstufe zu bewerten?

Die Berechnung der Schwierigkeit der einzelnen Items wurde bereits in Kapitel 6.1.3. beschrieben. Die Darstellung der Ergebnisse erfolgt zunächst in Tabellen, die die Anzahl der *richtigen Lösungen* (N), die Anzahl der *bearbeiteten Items* (N), den *Mittelwert* (M) und den errechneten *Schwierigkeitsindex* (P_i) abbilden.

Da Tests zur Lernverlaufsdagnostik theoretisch durch den Lernzuwachs der SchülerInnen immer leichter werden (vgl. Klauer 2014), wäre die Berechnung der Itemschwierigkeit nach mehreren Testdurchläufen ein Fehler. Daher machen KLAUER und STRATHMANN (2010) den Vorschlag, die Itemschwierigkeit an zwei aufeinanderfolgenden Testzeitpunkten zu berechnen – jedoch immer unter dem Vorbehalt des wahrscheinlich geringen Lernzuwachses der Kinder; ein Vorschlag, der für diese Arbeit umgesetzt wird. Die Ergebnisse der gesamten Stichprobe des ersten und zweiten MZP werden daher im Nachfolgenden dargestellt.

Tabelle 1: Itemschwierigkeit des Tests zum sinnentnehmenden Lesen (N4)

Messzeitpunkt 1				
Item	N (Lösungen)	N (gesamt)	M (Mittelwert)	P_i
Augen	152	187	0,81	81
fleißig	142	180	0,79	79
Sobald	121	184	0,66	66
Blumen	156	185	0,84	84

schmeckt	135	183	0,24	24
In	45	185	0,83	83
Schere	150	181	0,79	79
Haus	146	185	0,62	62
wegen	112	180	0,65	65
bevor	116	178	0,80	80
wohne	144	178	0,85	85
schnelle	151	177	0,71	71
aber	125	176	0,85	85
böse	146	171	0,89	89
Baby	156	176	0,65	65
Zwischen	113	175	0,60	60
spricht	101	167	0,90	90
Bett	155	172	0,74	74
von	127	172	0,80	80
scheint	136	171	0,56	56
Nachdem	94	168	0,86	86
runde	142	166	0,80	80
Freunde	137	172	0,88	88
gut	137	158	0,88	88
Vogel	143	163	0,54	54
Anstatt	86	160	0,86	86
lustigen	136	158	0,61	61
Durch	94	154	0,80	80
Beine	119	149	0,83	83
sammeln	119	143	0,89	89
hungrig	124	140	0,52	52
weder	72	137	0,52	52
Sonne	114	129	0,88	88

unter	96	127	0,76	76
schläft	90	120	0,75	75
Schuhe	92	124	0,74	74
backt	88	112	0,79	79
über	41	112	0,37	37
Bilder	85	107	0,79	79
Wenn	81	107	0,76	76
süß	80	99	0,81	81
Frösche	83	98	0,85	85
Schwester	72	87	0,83	83
dicke	63	85	0,74	74
Biene	53	84	0,63	63
Während	49	81	0,60	60
für	47	74	0,64	64
isst	55	66	0,83	83
Enten	48	61	0,79	79
Auf	38	58	0,66	66
wartest	23	59	0,39	39
Wasser	53	59	0,90	90
neues	53	57	0,99	99
außer	30	56	0,53	53
spitz	37	47	0,79	79
weil	26	47	0,55	55
kauft	36	49	0,73	73
Hase	42	52	0,81	81
Tür	43	50	0,86	86
mit	20	48	0,82	82
Büro	38	44	0,86	86
M	94	127	0,74	74

Insgesamt zeigten die Ergebnisse zum sinnentnehmenden Lesen nach LEVUMI, dass der Test zum ersten Messzeitpunkt einen durchschnittlichen Schwierigkeitsindex von $P_i=74$ ($M=.74$) hatte, was im oberen Bereich der *mittleren* Schwierigkeit für Items liegt. Für die Gesamtstichprobe ergaben sich Werte von $P_i=24$ ($M=.24$) bis $P_i=99$ ($M=.99$). Bei der Betrachtung der Schwierigkeit der einzelnen Items haben sich die Items „spricht“ ($P_i=90$), „Wasser“ ($P_i=90$) und „neues“ ($P_i=99$) als besonders *leicht* erwiesen. Für die Items „Vogel“ ($P_i=54$), „weil“ ($P_i=55$) und „scheint“ ($P_i=56$) konnte eine *mittlere* Schwierigkeit nachgewiesen werden. Lediglich für ein Item („schmeckt“ $P_i=24$) konnte ein Schwierigkeitsindex im Bereich *schwer* ermittelt werden. Das nächstschwerere Item („über“ $P_i=37$) liegt schon wieder im Bereich der *mittleren* Schwierigkeit.

Bei der Betrachtung der Anzahl der richtig gelösten und der insgesamt bearbeiteten Items konnte festgestellt werden, dass durchschnittlich $N=127$ Items während des ersten MZP bearbeitet wurden, wovon $N=94$ richtig gelöst wurden. Die Items „Blumen“ ($N=156$) und „Baby“ ($N=156$) wurden dabei am häufigsten richtig gelöst, das Item „mit“ ($N=20$) am wenigsten. Aufseiten der insgesamt bearbeiteten Items kann konstatiert werden, dass die Bearbeitungszahlen sukzessiv abnehmen. Das ist damit zu erklären, dass die SchülerInnen in LEVUMI zwar bei der ersten Testung alle die gleiche Reihenfolge der Aufgaben (Items) bekommen, aber aufgrund der Speedtest-Variation nicht alle Items bearbeiten konnten.

Tabelle 2: Itemschwierigkeit des Tests zum sinnentnehmenden Lesen (N4)

Messzeitpunkt 2				
Item	N (Lösungen)	N (gesamt)	Mittelwert	P_i
Augen	111	146	0,76	76
fleißig	114	136	0,84	84
Sobald	90	146	0,66	66
Blumen	109	145	0,75	75
schmeckt	95	138	0,69	69
In	33	142	0,23	23
Schere	113	143	0,79	79
Haus	114	144	0,79	79

wegen	112	144	0,78	78
bevor	96	146	0,66	66
wohne	105	136	0,77	77
schnelle	112	138	0,81	81
aber	111	148	0,75	75
böse	117	134	0,87	87
Baby	127	151	0,84	84
Zwischen	103	150	0,69	69
spricht	72	129	0,56	56
Bett	128	144	0,89	89
von	112	148	0,76	76
scheint	111	138	0,80	80
Nachdem	80	141	0,57	57
runde	126	151	0,83	83
Freunde	113	150	0,75	75
gut	117	137	0,85	85
Vogel	123	147	0,84	84
Anstatt	74	147	0,50	50
lustigen	121	143	0,85	85
Durch	81	145	0,56	56
Beine	125	146	0,86	86
sammeln	116	141	0,82	82
hungrig	108	131	0,82	82
weder	55	139	0,40	40
Sonne	129	145	0,90	90
unter	111	146	0,76	76
schläft	112	147	0,76	76
Schuhe	116	149	0,78	78
backt	110	140	0,79	79

über	53	151	0,35	35
Bilder	119	153	0,78	78
Wenn	99	138	0,72	72
süß	105	134	0,78	78
Frösche	122	145	0,84	84
Schwester	114	150	0,76	76
dicke	86	132	0,65	65
Biene	100	149	0,67	67
Während	96	151	0,64	64
für	104	143	0,73	73
isst	108	137	0,79	79
Enten	118	145	0,81	81
Auf	106	149	0,71	71
wartest	49	140	0,35	35
Wasser	121	148	0,82	82
neues	114	134	0,85	85
außer	86	144	0,59	59
spitz	115	140	0,82	82
weil	83	147	0,56	56
kauft	98	135	0,72	72
Hase	122	148	0,82	82
Tür	120	147	0,81	81
mit	78	147	0,53	53
Büro	121	144	0,84	84
M	104	143	0,73	73

Die Ergebnisse zum zweiten Messzeitpunkt zeigen, dass der Test einen durchschnittlichen Schwierigkeitsindex von $P_i=73$ ($M=.73$) erreicht, was ebenfalls im oberen Bereich der mittleren Schwierigkeit für Items liegt. Für die Gesamtstichprobe ergaben

sich bei der zweiten Messung Werte von $P_i=23$ ($M=.23$) bis $P_i=90$ ($M=.90$). Bei der Betrachtung der Schwierigkeit der einzelnen Items haben sich diesmal die Items „Sonne“ ($P_i=90$), „Bett“ ($P_i=89$) und „böse“ ($P_i=87$) als besonders leicht erwiesen. Für die Items „spricht“ ($P_i=56$), „Durch“ ($P_i=56$) und „weil“ ($P_i=56$) konnte eine mittlere Schwierigkeit nachgewiesen werden. Das Item „weil“ wurde beim ersten Messzeitpunkt ebenfalls im mittleren Bereich abgebildet. Items, die einen Schwierigkeitsindex von *schwer* erreichen, konnten auch beim zweiten Messzeitpunkt bis auf ein einziges Item („In“ $P_i=23$) nicht festgestellt werden.

Bei der Betrachtung der Anzahl der richtig gelösten und der insgesamt bearbeiteten Items zum zweiten Messzeitpunkt können im Vergleich zum ersten Messzeitpunkt Unterschiede verzeichnet werden. So konnte festgestellt werden, dass durchschnittlich $N=143$ Items während des zweiten Messzeitpunktes bearbeitet wurden, wovon $N=104$ richtig gelöst wurden. Die Items „Sonne“ ($N=129$), „Bett“ ($N=128$) und „Baby“ ($N=127$) wurden diesmal am häufigsten richtig gelöst, das Item „In“ ($N=33$) am wenigsten. Zu der Anzahl der insgesamt bearbeiteten Items muss festgehalten werden, dass beim zweiten und bei jedem weiteren Messzeitpunkt die Reihenfolge der Aufgaben (Items) aufgrund der Paralleltestform variiert und damit nicht sukzessiv abnimmt.

Forschungsfrage 2:

Wie verändern sich die Ergebnisse aller SchülerInnen vom ersten zum zweiten Messzeitpunkt?

H_0 Es gibt beim Lernverlaufstest zum sinnentnehmenden Lesen (LEVUMI) keinen Unterschied zwischen den Leistungen der SchülerInnen vom ersten zum zweiten Messzeitpunkt.

H_A Es gibt beim Lernverlaufstest zum sinnentnehmenden Lesen (LEVUMI) einen Unterschied zwischen den Leistungen der SchülerInnen vom ersten zum zweiten Messzeitpunkt.

Beim Vergleich der Ergebnisse beider Messzeitpunkte kann festgehalten werden, dass die SchülerInnen bei der ersten Testung einen Mittelwert von $.73$ ($SD=.20$) und

bei der zweiten ebenfalls einen Wert von .73 (SD=.23) erreichten haben. Die Auswertung ergab nur eine geringe Mittelwertdifferenz von .06. Die Kinder haben zudem im Durchschnitt beim ersten Mal N=28 Items richtig gelöst (bearbeitet N=39) und beim zweiten Mal N=32 (bearbeitet N=44). Beim Vergleich der beiden Messzeitpunkte der Gruppen ergab der *t*-Test für abhängige Stichproben für das Ergebnis $t(197) = .30$, $p = 0.77$. Dieser Unterschied konnte nicht als signifikant nachgewiesen werden, demzufolge muss die Nullhypothese (H_0) angenommen und die Alternativhypothese (H_A) verworfen werden.

Die nachfolgend dargestellten Tabellen werden sowohl für die Beantwortung der dritten als auch der vierten Forschungsfrage herangezogen. Die Ergebnisse beziehen sich auf die Gruppe der an dieser Studie teilnehmenden SchülerInnen mit dem Förderschwerpunkt Lernen und inklusiv beschult sowie auf die, die den Förderschwerpunkt Lernen haben und in einer Förderschule unterrichtet werden.

Tabelle 3: Ergebnisse der Gruppe der inklusiv beschulten SchülerInnen mit dem Förderschwerpunkt Lernen

Schüler	Messzeitpunkt 1		Messzeitpunkt 2	
	N (Lösungen)	Mittelwert	N (Lösungen)	Mittelwert
SchülerIn 1	5	0,13	18	0,30
SchülerIn 2	31	0,74	18	0,38
SchülerIn 3	17	0,46	10	0,61
SchülerIn 4	12	0,20	13	0,21
SchülerIn 5	15	0,58	16	0,43
SchülerIn 6	24	0,75	26	0,61
SchülerIn 7	13	0,68	18	0,78
SchülerIn 8	6	0,5	14	0,58
SchülerIn 9	17	0,28	13	0,21
SchülerIn 10	10	0,31	18	0,69
SchülerIn 11	15	0,44	19	0,33

M	15	0,46	17	0,43
----------	-----------	-------------	-----------	-------------

Tabelle 4: Ergebnisse der Gruppe der SchülerInnen mit dem Förderschwerpunkt Lernen, die eine Förderschule besuchen

Schüler	Messzeitpunkt 1		Messzeitpunkt 2	
	N (Lösungen)	Mittelwert	N (Lösungen)	Mittelwert
SchülerIn 12	2	1,00	1	0,41
SchülerIn 13	7	0,78	3	0,60
SchülerIn 14	34	0,76	27	0,56
SchülerIn 15	5	1,00	10	0,52
SchülerIn 16	4	0,44	3	0,54
SchülerIn 17	2	0,50	4	0,56
SchülerIn 18	8	0,22	9	0,57
SchülerIn 19	3	0,27	4	0,66
M	8	0,62	8	0,55

Forschungsfrage 3:

Wie verändern sich die Ergebnisse der an dieser Studie teilnehmenden SchülerInnen mit dem Förderschwerpunkt Lernen vom ersten zum zweiten Messzeitpunkt?

H_0 Es gibt beim Lernverlaufstest zum sinnentnehmenden Lesen (LeVuMi) keinen Unterschied zwischen den Leistungen der SchülerInnen mit dem Förderschwerpunkt Lernen vom ersten zum zweiten Messzeitpunkt.

H_A Es gibt beim Lernverlaufstest zum sinnentnehmenden Lesen (LeVuMi) einen Unterschied zwischen den Leistungen der SchülerInnen mit dem Förderschwerpunkt Lernen vom ersten zum zweiten Messzeitpunkt.

Beim Vergleich der Messzeitpunkte der insgesamt an der Studie teilgenommen habenden SchülerInnen mit dem Förderschwerpunkt Lernen ($n=19$) konnte für den ersten Messzeitpunkt ein Mittelwert von .53 ($SD=.26$) und für den zweiten Messzeitpunkt ein Wert von .48 ($SD=.18$) festgestellt werden. Die SchülerInnen haben zudem bei

der ersten Testung durchschnittlich $N=12$ Items und bei der zweiten $N=13$ Items richtig beantwortet. Die Kinder haben durchschnittlich beim ersten Mal $M=.26$ Aufgaben bearbeitet und bei der zweiten Testung $M=.33$. Der t -Test für abhängige Stichproben ergab das Ergebnis $t(18) = .72$, $p = 0,48$. Dieser Unterschied konnte nicht als signifikant nachgewiesen werden. Deshalb muss die Nullhypothese (H_0) angenommen und die Alternativhypothese (H_A) verworfen werden.

Forschungsfrage 4:

Wie verändern sich die Ergebnisse der SchülerInnen mit dem Förderschwerpunkt Lernen, die inklusiv beschult werden, gegenüber den SchülerInnen mit dem Förderschwerpunkt Lernen an einer Förderschule?

H_0 Beim Test zum sinnentnehmenden Lesen (LEVUMI) erzielen die inklusiv beschulten SchülerInnen mit dem Förderschwerpunkt Lernen vom ersten zum zweiten Messzeitpunkt keine besseren Ergebnisse.

H_A Beim Test zum sinnentnehmenden Lesen (LEVUMI) erzielen die inklusiv beschulten SchülerInnen mit dem Förderschwerpunkt Lernen vom ersten zum zweiten Messzeitpunkt bessere Ergebnisse.

H_0 Beim Test zum sinnentnehmenden Lesen (LEVUMI) erzielen die SchülerInnen mit dem Förderschwerpunkt Lernen (die eine Förderschule besuchen) vom ersten zum zweiten Messzeitpunkt keine besseren Ergebnisse.

H_A Beim Test zum sinnentnehmenden Lesen (LEVUMI) erzielen die SchülerInnen mit dem Förderschwerpunkt Lernen (die eine Förderschule besuchen) vom ersten zum zweiten Messzeitpunkt bessere Ergebnisse.

Zunächst werden die Ergebnisse vom ersten zum zweiten Messzeitpunkt der SchülerInnen mit dem Förderschwerpunkt Lernen, die inklusiv beschult werden, verglichen, danach die der SchülerInnen mit dem Förderschwerpunkt Lernen, die eine Förderschule besuchen.

Beim Vergleich des ersten und zweiten Messzeitpunktes der inklusiv beschulten SchülerInnen mit dem Förderschwerpunkt Lernen ($n=11$) konnte für den ersten Messzeitpunkt ein Mittelwert von $.46$ ($SD=.21$) und für den zweiten Messzeitpunkt ein Wert

von .43 (SD=.21) festgestellt werden. Die SchülerInnen haben zudem bei der ersten Testung durchschnittlich N=15 Items und bei der zweiten N=16 Items richtig beantwortet. Der t -Test für abhängige Stichproben ergab das Ergebnis $t(10) = .45$, $p = 0,67$. Dieser Unterschied konnte nicht als signifikant nachgewiesen werden. Infolgedessen muss die Nullhypothese (H_0) angenommen und die Alternativhypothese (H_A) verworfen werden.

Bei der Gruppe der SchülerInnen mit dem Förderschwerpunkt Lernen, die eine Förderschule (N=8) besuchen, konnte für den ersten Messzeitpunkt ein Mittelwert von .62 (SD=.31) und für den zweiten Testzeitraum ein Wert von .55 (SD=.07) errechnet werden. Die Teilnehmer dieser Subgruppe haben sowohl bei der ersten als auch bei der zweiten Testung jeweils N=8 richtige Antworten gegeben. Beim Vergleich der beiden Testzeitpunkte ergab der t -Test für abhängige Stichproben das Ergebnis $t(7) = .54$, $p = 0,61$. Dieser Unterschied hat sich ebenfalls als nicht signifikant erwiesen. Demgemäß muss die Nullhypothese (H_0) angenommen und die Alternativhypothese (H_A) verworfen werden. Der Vergleich der beiden Subgruppen untereinander erfolgt im nächsten Abschnitt.

Forschungsfrage 5:

Erzielen die SchülerInnen mit dem Förderschwerpunkt Lernen, die inklusive beschult werden, bessere Ergebnisse als SchülerInnen mit dem Förderschwerpunkt Lernen an einer Förderschule?

H_0 Es gibt beim Lernverlaufstest zum sinnentnehmenden Lesen (LEVUMI) keinen Unterschied zwischen den Leistungen der SchülerInnen mit dem Förderschwerpunkt Lernen, die inklusiv beschult werden, und denen mit dem Förderschwerpunkt Lernen, die eine Förderschule besuchen.

H_A Es gibt beim Lernverlaufstest zum sinnentnehmenden Lesen (LEVUMI) einen Unterschied zwischen den Leistungen der SchülerInnen mit dem Förderschwerpunkt Lernen, die inklusiv beschult werden, und denen mit dem Förderschwerpunkt Lernen, die eine Förderschule besuchen.

Beim Vergleich des ersten Messzeitpunktes der Gruppe der inklusiv beschulten Kinder mit dem Förderschwerpunkt Lernen ($n=11$; $M=.46$, $SD=.21$) und den Kindern mit

dem Förderschwerpunkt Lernen, die eine Förderschule ($n=8$; $M=.62$, $SD=.31$) besuchen, ergab der t-Test für unabhängige Stichproben das Ergebnis $t(17) = -1,37$, $p = 0,21$. Dieser Unterschied konnte nicht als signifikant nachgewiesen werden. Beim Vergleich des zweiten Messzeitpunktes der beiden Subgruppen ($n=11$; $M=.42$, $SD=.21$ und $n=8$; $M=.55$, $SD=.07$) ergab der Test das Ergebnis $t(17) = -1,50$, $p = 0,13$. Es konnte ebenfalls keine Signifikanz nachgewiesen werden. Daraus resultiert, dass in beiden Fällen die Nullhypothese (H_0) angenommen und die Alternativhypothese (H_A) verworfen werden.

Anschließend erfolgt eine Darstellung der beobachteten Ergebnisse. Dabei werden sowohl die Erfahrungen der Testleistung mit LEVUMI als auch die Rückmeldungen der SchülerInnen benannt.

Bei der Kontaktaufnahme mit den Schulen wurden verschiedene Wege (E-Mail, Post, Telefon) gewählt, die sich nicht immer als günstig erwiesen. Die meisten Rückmeldungen wurden auf dem telefonischen Weg erzielt. Die Gründe für eine Nichtteilnahme am Programm lagen nicht per se an mangelndem Interesse, sondern an einem bereits vollen Programm der Schulen und der Arbeitsbelastung der Lehrkräfte. Tatsächliches Interesse, das Programm auch nach der Testreihe selbst weiterführen zu wollen, konnte lediglich bei zwei Lehrkräften festgestellt werden. Gleichwohl wurden auch viele Lehrkräfte davon abgeschreckt, dass es sich um ein online- bzw. computerbasiertes Programm handelt, dessen Durchführung sie sich nicht zutrauen. Daher bestand bei den meisten Lehrkräften auch kein Interesse an einer Vor- oder Einführung in das Programm.

Die SchülerInnen zeigten sich alle begeistert, wenn es darum ging, dass sie etwas am Computer durchführen durften. Auch wenn einige Kinder noch keinen Kontakt mit dem Medium Computer hatten, so fanden sie sich doch schnell in das Bedienungsfeld ein. Benannte Kinder erfuhren von der Testleitung daher eine gesonderte Einführung in LEVUMI. Es konnte beobachtet werden, dass auch die Kinder mit noch nicht sehr ausgeprägten digitalen Kompetenzen sehr schnell mit der Bedienung der Maus zurechtkamen.

Das Einführungsbeispiel zu Beginn des Tests zum sinnentnehmenden Lesen führte häufig zu Problemen. Durch das Einsetzen eines Items hat sich der Bildausschnitt verschoben und der Bedienbutton zum Starten wurde aus dem Sichtfenster geschoben. Den wenigsten SchülerInnen war die Scroll-Funktion der Maus bekannt, womit sich der Beginn der Tests verzögert hat.

Während der Testungen ließen sich bei den SchülerInnen teilweise verschiedene Strategien zum Erlesen der Items erkennen. In den meisten Fällen haben sie sich die Sätze mit den einzusetzenden Items laut vorgelesen. Dies störte die Gruppe aber insgesamt nicht. Dass SchülerInnen gegenseitig voneinander abgesehen haben, konnte kaum festgestellt werden. Wenn der Blick auf den Nachbarbildschirm erfolgte, dann nur zu der Feststellung, bei welchem Item der/die MitschülerIn denn sei. Die Rückmeldung der SchülerInnen war durchweg gleich. Alle wollten gerne die Ergebnisse sehen, sprich wie gut oder wie schlecht sie abgeschnitten haben. Zudem wurde die Testleistung immer wieder nach der noch zur Verfügung stehenden Zeit gefragt. Vereinzelt wurde die Testleistung bei verschiedenen Items nach Unterstützung gefragt; dies wurde jedes Mal verneint. Die SchülerInnen der Grundschule betonten auf Rückfragen der Testleitung immer wieder, als wie leicht sie den Test empfanden. Die teilnehmenden FörderschülerInnen betonten, wie schwer sie den Test fanden. Des Weiteren ist vielen SchülerInnen aufgefallen, dass der Test bei der zweiten Durchführung dieselben Items enthält wie bei der ersten Testung.

Leider musste aber auch die Erfahrung gemacht werden, dass einige SchülerInnen Spaß daran hatten, schnellstmöglich weiterzuklicken, um das Programm frühzeitig zu beenden – ein Vorgehen, das sich letztendlich auf die Auswertung der Ergebnisse auswirkt. Ebenso konnte beobachtet werden, dass viele SchülerInnen ein Item zur Vervollständigung des vorgegebenen Satzes angeklickt haben, aber das Item nicht richtig getroffen haben und dann vorzeitig auf „weiter“ geklickt haben. So konnte in vielen Fällen zwar festgestellt werden, dass das richtige Item angeklickt wurde, aber das System es als „nicht bearbeitet“ erfasst hat. Dies stellt ebenfalls einen entscheidenden Punkt bei der Interpretation der Ergebnisse dar.

Die Betrachtung und Auswertung der Ergebnisse der SchülerInnen stellt sich über den Lerngraphen auf www.Levumi.de als unproblematisch dar. Auch wenn der Graph durch eine große Schüleranzahl sehr unübersichtlich wirkt, so besteht die Möglichkeit

der Auswahl eines/r einzelnen Schülers/in mit dessen/deren Lernverlauf. Diese Option besteht jedoch nur beim Onlineverfahren. Den Lehrkräften wurden die Ergebnisse immer per E-Mail in dem Dateiformat .pdf übermittelt, worauf die Rückmeldung der Undeutbarkeit ab einer bestimmten Klassengröße erfolgte.

7 Diskussion

Folgender Abschnitt nimmt Bezug auf die im vorherigen Kapitel dargestellten Ergebnisse. Zudem werden die in Kapitel 5 dargestellten Fragestellungen und Hypothesen aufgegriffen und sukzessive diskutiert.

Forschungsfrage 1:

Wie ist die Itemschwierigkeit des Tests zum sinnentnehmenden Lesen für SchülerInnen der dritten Schulstufe zu bewerten?

Die Schwierigkeit eines Tests bildet für die Lernverlaufsdagnostik ein entscheidendes Element. Es muss gewährleistet sein, dass ein Test nicht nur dasselbe erfasst, sondern auch stets gleich schwer ist. Bei der Betrachtung der quantitativen Ergebnisse bezüglich der Itemschwierigkeit fällt zunächst der sehr hohe Schwierigkeitsindex (Tabelle 1) auf. Der Test hat zum ersten Messzeitpunkt einen Index von $P_i=74$ ($M=.74$) erreicht und zum zweiten einen Index von $P_i=73$ ($M=.73$). Bezugnehmend auf die Ausführungen von MOOSBRUGGER und KELAVA (2012, Kapitel 6.1.3) kann der Test zum sinnentnehmenden Lesen durchaus als leicht betrachtet werden. Durch die minimale Abweichung der Indexe zwischen den beiden MZP bleibt der Test scheinbar auch gleich schwer und kann nicht nur als leicht, sondern auch als fair bezeichnet werden. Die Fairness des Tests wurde aber im Rahmen dieser Arbeit nicht weiter untersucht. Bei der Beurteilung der Itemschwierigkeit sollte der erste Messzeitpunkt als der aussagekräftigere verstanden werden, denn für den ersten Zeitpunkt hält LEVUMI bei der Auswahl der Items eine feste Reihenfolge ein. Daraus resultiert, dass bestimmte Aufgaben (Items) alle SchülerInnen bearbeitet haben.

Da dies ein Speedtest ist, ist bei den Ergebnissen des ersten Messzeitpunktes zu erkennen, dass die letzten Items immer weniger bearbeitet wurden. In der Konsequenz haben es nicht alle SchülerInnen geschafft, alle Items zu bearbeiten bzw. den Test bis zum Ende durchzuführen. Aufgrund der Paralleltestform mischt das System die Items bei der zweiten und jeder weiteren Testung neu durch, bleibt dabei allerdings immer gleich schwer. Das bedeutet aber auch, dass SchülerInnen, die bei der ersten Testdurchführung nicht bis zum Ende gekommen sind, also nicht alle Items bearbeitet haben, bei der zweiten Testung eventuell Items bekommen, die sie beim ersten Mal nicht hatten. Das System muss den Ergebnissen zufolge also auch wieder

Items des gleichen Schwierigkeitsgrades ausgewählt haben. Dennoch stellt sich die Frage nach der Homogenität des Tests über einen längeren Zeitraum hinweg, bspw. ein Jahr. Da davon auszugehen ist, dass die SchülerInnen i. d. Regel einen Leistungszuwachs erfahren, wird auch der Test von MZP zu MZP immer leichter werden.

Bei der Betrachtung der Schwierigkeitsausprägungen aller Items (Tabelle 1 und 2) ist auffällig, dass insgesamt sehr hohe Werte erreicht wurden. Das heißt, der Test ist – wie bereits erwähnt – als leicht zu bezeichnen. Lediglich für zwei Items wurden Werte im Bereich schwer ermittelt. Der Argumentation von MOOSBRUGGER und KELAVA (2012) folgend, sollte ein Test je nach Zielsetzung auch Items im Bereich $5 \leq P_i \leq 20$ (schwer) bzw. $80 \leq P_i \leq 95$ (leicht) (Kapitel 6.1.3) gleichermaßen abbilden. Dies ist aber nicht der Fall, denn der Test verfügt überwiegend über Items der *leichten* Schwierigkeitsstufe. Daher könnte der Test für die dritte Schulstufe als zu leicht bewertet werden. Dennoch sollte eine Bewertung nicht voreilig geschehen. Denn um eine konkrete Aussage der Tauglichkeit des Tests für die dritte Schulstufe machen zu können, sollten auch die Subgruppen innerhalb dieser Stichprobe beachtet werden. Wie stellt sich die Schwierigkeit für SchülerInnen mit einem Migrationshintergrund dar? Wie für Kinder mit einem Förderschwerpunkt? Gibt es altersbezogene Leistungsunterschiede?

Für die meisten Tests werden in der Regel Items mit einer mittleren Schwierigkeit bevorzugt, um in einem mittleren Leistungsbereich gut differenzieren zu können. Das Ziel des Programms LEVUMi ist es allerdings, Leistungsschwächen aufzudecken – gerade deswegen, weil das Programm für den Einsatz im inklusiven Unterricht und an Förderschulen gedacht ist –, und diese Leistungsschwächen liegen bekanntlich im unteren Leistungsbereich. Das heißt, dass das Programm ein sehr breites Leistungsspektrum bedienen muss, um noch fair für SchülerInnen mit einem Förderbedarf zu sein. So ergibt es durchaus Sinn, dass LEVUMi vom Anspruch des Schwierigkeitsgrades her sehr *leicht* gehalten ist, um für jene SchülerInnen noch durchführbar und vor allem lösbar zu sein und keine demotivierende Wirkung zu erzeugen. Schließlich ist der Kern des Programms LEVUMi die Abbildung des Lernverlaufs jedes/r einzelnen Schülers/In, um einerseits eine Verbesserung bzw. Verschlechterung zurückzumelden und Fördermaßnahmen zu optimieren. Andererseits muss die Itemschwierigkeit noch für weitere Gruppen berechnet werden, denn es ist nicht auszuschließen, dass der Test für andere Gruppen deutlich schwieriger oder leichter ist. Zudem muss bei

der Interpretation der Ergebnisse auch beachtet werden, dass falsche Antworten bei einer kleineren Stichprobe sehr viel mehr Auswirkungen auf das Gesamtergebnis haben als bei einer größeren. So muss festgehalten werden, dass aufgrund der geringen Stichprobengröße und des Umstandes, dass zum jetzigen Zeitpunkt für den Test des sinnentnehmenden Lesens noch keine weiteren Berechnungen vorliegen, weitere Untersuchungen mit größeren Stichproben durchgeführt werden müssen, um eine endgültige Aussage treffen zu können. Eine pauschalisierte Aussage auf Basis dieser Studie wäre nicht ratsam.

Forschungsfrage 2:

Wie verändern sich die Ergebnisse aller SchülerInnen vom ersten zum zweiten Messzeitpunkt?

Die Ergebnisse der an dieser Studie teilgenommen habenden SchülerInnen haben sich auf den ersten Blick grundlegend nicht geändert. Sowohl bei der ersten als auch bei der zweiten Testung wurde ein Mittelwert von .73 errechnet. Gerade das Ergebnis der zweiten Testung kann als gut bewertet werden, denn den Ausführungen der vorherigen Forschungsfrage folgend hat ein Teil der Kinder aufgrund des randomisierten Testverfahrens bei der zweiten Testung Items bearbeitet, die zum ersten Messzeitpunkt aus Zeitgründen und infolgedessen auch aus Leistungsgründen nicht mehr bearbeitet werden konnten. Aufseiten der gelösten Items konnte ein Anstieg von vier zusätzlichen Items (1. N=28, 2. N=32) verzeichnet werden. Auch dies ist prinzipiell als gut zu bewerten. Dennoch kann die Forschungsfrage, inwieweit sich die Ergebnisse der an dieser Studie teilgenommen habenden Kinder verändert haben, nur mit *bedingt* beantwortet werden. Dies wirkt in Anbetracht der höheren Anzahl an gelösten Items als befremdlich. Es ist also fraglich, warum die Mittelwerte der beiden Messzeitpunkte nur eine so geringe Differenz aufweisen. Diese Frage ist damit zu beantworten, dass die SchülerInnen bei der zweiten Testung zwar mehr Items richtig gelöst, aber auch mehr Items insgesamt (1. N=39, 2. N=44) bearbeitet haben. Dies führt zu keinem oder nur einem sehr geringen Unterschied von .06 im Mittelwert. Dennoch reichen diese Differenz und der geringe Unterschied in den Standardabweichungen (1. SD=20, 2. SD=23) nicht aus, um mit dem *t*-Test für abhängige Stichproben einen

signifikanten Unterschied zu finden. Eine weitere Begründung dafür könnte die geringe Stichprobengröße von $n=189$ SchülerInnen darstellen. Demzufolge muss, wie in der Ergebnisdarstellung bereits ausgeführt, die Nullhypothese – *Es gibt beim Lernverlaufstest zum sinnentnehmenden Lesen (LeVuMi) keinen Unterschied zwischen den Leistungen der SchülerInnen vom ersten zum zweiten Messzeitpunkt* – angenommen werden. Dies wirft zwei weitere Fragen auf: 1) Warum ist kein Anstieg der Leistung zu verzeichnen? 2) Worauf ist die höhere Anzahl gelöster Items zurückzuführen?

Letzteres ist sicherlich darauf zurückzuführen, dass den Kindern das Programm bei der zweiten Testung bereits bekannt war. Den SchülerInnen war die Testleitung bereits bekannt und sie hatten keine Hemmung vor dem zum ersten Zeitpunkt noch unbekanntem Test. Es war deutlich zu erkennen, dass die SchülerInnen mit dem Programm besser zurechtkamen. Bei den SchülerInnen, die noch keine oder nur wenig ausgeprägte digitale Kompetenzen hatten, konnte ebenfalls ein sicherer Umgang mit dem Computer festgestellt werden.

Zur Beantwortung der ersten Frage kann gesagt werden, dass der geringe Zeitraum von nur drei Wochen zwischen der ersten und zweiten Testung nicht ausreicht, um einen herausragend Leistungsanstieg zu verzeichnen. Dafür müsste bei diesem Test ein größerer Zeitraum gewählt werden, um tatsächlich einen Anstieg zu verzeichnen. Dennoch ist eine Stagnation der Leistungen auch sicherlich als gut zu bewerten.

Es kann aber auch geschlussfolgert werden, dass LEVUMI nicht zu einem Anstieg der Leseleistungen bei SchülerInnen der dritten Schulstufe im genannten Zeitraum führt. Oder aber der Schwierigkeitsindex von $P_i=74$ ($M=.74$) ist zu gering für die SchülerInnen der dritten Stufe. Nichtsdestoweniger muss auch beachtet werden, dass LEVUMI zwar auch auf die Förderung von Lesekompetenz ausgelegt ist, in erster Linie aber eine diagnostische Methode ist, um den Lernverlauf von SchülerInnen zu einem Fachbereich abzubilden. Daher sollte dem Anstieg von Lernleistungen durch das Programm selbst nicht zu viel Beachtung geschenkt werden. Zudem sollte das Programm auch mit Tagesschwankungen von Kindern sowie weiteren äußeren Einflüssen rechnen, denen schließlich alle Beteiligten unterliegen.

Trotzdem sollten Lehrkräfte aus den Ergebnissen didaktische Konsequenzen ziehen. Sie müssen sich infolgedessen die Frage stellen, warum ihre SchülerInnen in

den drei Wochen nicht besser geworden sind. Diese Frage ist sicherlich nicht einfach zu beantworten und würde den Rahmen dieser Arbeit übersteigen, muss aber bei der Analyse der Ergebnisse der einzelnen Klassen von den Lehrkräften herangezogen werden.

Wie bereits ausgeführt, stellt die Homogenität der Testschwierigkeit eine unabdingbare Voraussetzung für eine Lernverlaufsdiagnostik dar. Die Tests müssen in der Lage sein, Veränderungen zu erfassen; das bedeutet, sie müssen Leistungs- oder -abnahmen abbilden können. Um dies zu gewährleisten, reicht es jedoch nicht aus, nur die Ergebnisse der Gesamtstichprobe zu betrachten; es muss vor allem auf den/die einzelne/n SchülerIn geschaut werden, um Rückschlüsse ziehen zu können – eine Ansicht, die auch WILBERT (2014) teilt.

Letztlich kann festgehalten werden, dass die Kinder nur insofern besser geworden sind, als sie zwischen dem ersten und zweiten Messzeitpunkt mehr Items bearbeiten haben; dieser Anstieg verläuft aber gleichmäßig zur Anzahl der gelösten Items. Das heißt, die festgestellte Veränderung bezieht sich in erster Linie auf das ausgeprägtere Lern- und Arbeitstempo.

Forschungsfrage 3:

Wie verändern sich die Ergebnisse der an dieser Studie teilnehmenden SchülerInnen mit dem Förderschwerpunkt Lernen vom ersten zum zweiten Messzeitpunkt?

Die Gruppe der Schülerinnen mit dem Förderschwerpunkt Lernen beläuft sich bei dieser Stichprobe auf $n=19$. Ein Teil der SchülerInnen wurde zu diesem Zeitpunkt inklusiv beschult ($N=11$), und die anderen haben eine Förderschule mit dem Förderschwerpunkt Lernen ($n=8$) besucht. Hier konnte zum ersten Messzeitpunkt ein Mittelwert von $.53$ errechnet werden und zum zweiten Messzeitpunkt von $.48$. Es muss also zunächst festgehalten werden, dass die SchülerInnen schlechter geworden sind. Daraus lässt sich schlussfolgern, dass die Kinder in der Konsequenz auch bei der Anzahl der richtig gelösten Items schlechter geworden sein müssten – das ist aber nicht der Fall. Bei der ersten Testung haben sie durchschnittlich $N=12$ Aufgaben gelöst und beim zweiten Mal $N=13$. Woher kommt dann der schlechtere Mittelwert? Dies lässt sich ähnlich begründen wie beim Vergleich des ersten und zweiten Messzeitpunktes bzgl. der gesamten Stichprobe. Kinder haben durchschnittlich beim ersten Mal $M=.26$

Aufgaben bearbeitet und bei der zweiten Testung $M=.33$. Sie haben also beim zweiten Mal mehr Aufgaben geschafft. Auch dies kann sicherlich damit begründet werden, dass die Kinder beim zweiten MZP deutlich sicherer waren als noch beim ersten MZP. Dennoch sind die Ergebnisse der richtig gelösten Items nicht deutlich nach oben gegangen. Der Grund liegt in der höheren Fehlerquote. Sie haben zwar mehr bearbeitet, aber dafür auch mehr Fehler gemacht. Durch die randomisierte Ziehung der Items haben die SchülerInnen teils neue, unbekannte Aufgaben bekommen. Dennoch hat der t -Test für abhängige Stichproben ergeben, dass es keine signifikanten Unterschiede gibt. Die Nullhypothese – *Es gibt beim Lernverlaufstest zum sinnentnehmenden Lesen (LeVuMi) keinen Unterschied zwischen den Leistungen der SchülerInnen mit dem Förderschwerpunkt Lernen vom ersten zum zweiten Messzeitpunkt* – muss demnach angenommen werden. Ebenso wie bei der vorherigen Forschungsfrage stellt sich hier die Frage der Begründung. Diesbezüglich auffällig sind bei der Ergebnisbetrachtung die Werte der Standardabweichung (1. $SD=.26$, 2. $SD=.18$). Auf Basis dieser Werte hätte SPSS eigentlich einen Unterschied finden müssen, dennoch muss das Ergebnis ($p=0,48$) als nicht signifikant gewertet werden.

Dass die Ergebnisse der SchülerInnen mit dem Förderschwerpunkt Lernen im Verhältnis zur Gesamtstichprobe offenkundig schlechter sind, war zu erwarten. Aber das sollte weder das Ziel dieser Arbeit noch der vorliegenden Fragestellung sein. Es sollte vielmehr aufgezeigt werden, dass auch SchülerInnen mit einem Förderbedarf ihre Leistungen in einem kurzen Zeitraum ausbauen und erweitern können. Selbstverständlich liegen Kinder mit besonderen Bedarfen auf einer anderen Leistungsebene als Kinder ohne Förderbedarfe, aber das muss nicht in jedem Kompetenzbereich sein. So kann ein/e SchülerIn in einem Bereich (bspw. Mathematik) gute Leistungen erbringen, während er/ sie in einem Anderen (bspw. Deutsch) Förderung benötigt. Schließlich finden sich bei der Betrachtung der Ergebnisse der SchülerInnen 1-19 mit dem Förderschwerpunkt Lernen (Tabelle 4 und 5) teilweise sehr hohe Ergebnisse – ein Diskussionspunkt, der in den beiden nachfolgenden Fragestellungen genauer dargestellt wird.

Forschungsfrage 4:

Wie verändern sich die Ergebnisse der SchülerInnen mit dem Förderschwerpunkt Lernen, die inklusiv beschult werden, gegenüber den SchülerInnen mit dem Förderschwerpunkt Lernen an einer Förderschule?

Zunächst werden die inklusiv beschulten SchülerInnen betrachtet, dann die Kinder an der Förderschule. Anschließend wird die jeweilige Entwicklung der beiden Subgruppen diskutiert.

Die Subgruppe der inklusiv beschulten Kinder beläuft sich in dieser Untersuchung auf $n=11$. Diese Kinder haben zum ersten MZP einen Mittelwert von .46 erreicht und beim zweiten von .43. Richtig gelöst haben die Kinder bei der ersten Testung $N=15$ Items und bei der zweiten $N=16$ Items. Die Begründung des gesunkenen Mittelwertes bei einer gleichzeitig angestiegenen Anzahl der richtig gelösten Antworten lässt sich wie bei der Auswertung der Gesamtstichprobe damit erklären, dass die SchülerInnen zwar mehr richtige Antworten gegeben, aber gleichzeitig auch mehr Items nicht beantwortet oder bearbeitet haben. Dies begründet den gesunkenen Mittelwert. Der t -Test hat auch hier keinen Unterschied zwischen beiden MZP ermitteln können. Somit muss die Nullhypothese – *Beim Test zum sinnentnehmenden Lesen (LeVuMi) erzielen die inklusiv beschulten SchülerInnen mit dem Förderschwerpunkt Lernen vom ersten zum zweiten Messzeitpunkt keine besseren Ergebnisse* – ebenfalls angenommen werden.

Bei der zweiten Subgruppe, den SchülerInnen mit dem Förderschwerpunkt Lernen, die eine Förderschule ($n=8$) besuchen, sehen die Ergebnisse etwas anders aus. Die Kinder haben beim ersten MZP einen Mittelwert von .55 und beim zweiten einen Wert von .62. Sie sind also insgesamt besser geworden. Richtig gelöst haben sie an beiden Messzeitpunkten jeweils $N=8$ Items. Dennoch hat die Paaranalyse keinen signifikanten Unterschied feststellen können. Die Nullhypothese – *Beim Test zum sinnentnehmenden Lesen (LeVuMi) erzielen die SchülerInnen mit dem Förderschwerpunkt Lernen (die eine Förderschule besuchen) vom ersten zum zweiten Messzeitpunkt keine besseren Ergebnisse* – wird demnach angenommen.

Es stellt sich die Frage, warum der t -Test für verbundene Stichproben keinen Unterschied finden konnte. Zum einen liegt dies wahrscheinlich an der sehr geringen Gruppengröße und zum anderen an dem tatsächlich sehr geringen Unterschied in

den Leistungen. Wie bereits zuvor diskutiert, stellt sich die Frage, ob ein immenser Leistungsunterschied überhaupt verzeichnet werden muss, oder ob es nicht schon als zielführend und positiv bewertet werden kann, dass die Kinder keinen Leistungsabfall hatten, zumal der Test schließlich auch Leistungsdefizite diagnostizieren soll. Dies allerdings soll nicht Gegenstand weiterer Ausführungen werden, denn es soll vielmehr der Blick auf einzelne SchülerInnen dieser Gruppen gerichtet werden.

Bei den inklusiv beschulten SchülerInnen können enorme Leistungsunterschiede festgestellt werden (Tabelle 3). S6 erreichte beim ersten MZP einen Mittelwert von .75 und damit den höchsten Wert zum Untersuchungszeitpunkt, wohingegen S1 den niedrigsten ($M=.13$) hatte. Beim Vergleich beider Messzeitpunkte ist eines sehr auffällig, nämlich dass teilweise Kinder, die beim ersten MZP gut waren, beim zweiten MZP schlechter geworden sind (bspw. S2 $M=.74 \rightarrow M=.38$; S5 $M=.58 \rightarrow M=.43$) – umgekehrt aber genauso (bspw. S1 $M=.08 \rightarrow M=.13$; S10 $M=.31 \rightarrow M=.69$). Das spricht zum einen wieder für die randomisierte Ziehung von Items, aber zum anderen auch für das Leistungsspektrum der SchülerInnen mit dem Förderschwerpunkt Lernen. SchülerInnen mit Lernschwierigkeiten haben häufig Probleme damit, ihr Leistungsniveau zu halten bzw. in einer entsprechenden Situation abzurufen. Daher sind extreme Leistungsanstiege oder -abfälle, auch in kurzen Zeitabständen, nicht ungewöhnlich. Insgesamt aber haben einige SchülerInnen (S2, S5, S6, S7) auch Ergebnisse auf dem Niveau ihrer MitschülerInnen erreichen können. Ob und in welchen Bereichen sich die jeweilige Lernschwierigkeit zeigt, wurde jedoch nicht mit erhoben.

Ein vergleichbares Bild ist auch bei den SchülerInnen an der Förderschule ($n=8$) zu erkennen, allerdings mit deutlich weniger bearbeiteten Aufgaben. Die Kinder haben teilweise nur zwei oder drei Items richtig gelöst (MZP1 S12, S17, S3, MZP2 S16) und auch insgesamt weniger bearbeitet. Zudem hat ihre Leistung vom ersten MZP ($M=.62$) zum zweiten MZP ($M=.55$) abgenommen. Ein heterogenes Bild von Leistungen bietet sich auch bei dieser Gruppe.

Es kann also festgehalten werden, dass in beiden Gruppen sowohl Leistungsanstiege als auch -abfälle verzeichnet werden können. Bei einer Gegenüberstellung der MZP der beiden Gruppen muss dennoch gesagt werden, dass beide Gruppen zwar insgesamt schlechter geworden sind, dass aber der Leistungsabfall der SchülerInnen, die an einer Förderschule unterrichtet werden, noch größer ist. Auch an dieser Stelle sind die Gründe wieder vielfältig. Einerseits werden die SchülerInnen bei der

zweiten Testung wieder unbekannte Items erhalten haben, andererseits muss an dieser Stelle aber auch über ein grundlegend differentes Leistungsniveau von Kindern, die eine Förderschule besuchen, und denen, die inklusiv geschult werden, diskutiert werden. Eine generelle Aussage kann und sollte sicherlich nicht auf einer Gruppe von nur 19 SchülerInnen fußen, soll aber Bestandteil der nachfolgenden und für diese Arbeit wichtigsten Forschungsfrage sein.

Forschungsfrage 5:

Erzielen die SchülerInnen mit dem Förderschwerpunkt Lernen, die inklusiv beschult werden, bessere Ergebnisse als SchülerInnen mit dem Förderschwerpunkt Lernen an einer Förderschule?

Zentral ist für die Fragestellung, inwieweit sich Leistungsunterschiede zwischen der Gruppe der SchülerInnen mit einem Förderschwerpunkt, die inklusiv beschult werden, und jenen, die eine Förderschule besuchen, zeigen. Darüber hinaus sollen generelle Merkmalsausprägungen jener Beschulungsformen diskutiert werden.

In der Ergebnisdarstellung (Kapitel 6.4) wurde bereits der Vergleich der Subgruppen zum ersten und zweiten MZP dargestellt. Zu beiden MZP konnte die Berechnung mit SPSS keinen signifikanten Unterschied berechnen. Das heißt, dass auch im letzteren Fall die Nullhypothese – *Es gibt beim Lernverlaufstest zum sinnentnehmenden Lesen (LEVUMI) keinen Unterschied zwischen den Leistungen der SchülerInnen mit dem Förderschwerpunkt Lernen, die inklusiv beschult werden, und denen mit dem Förderschwerpunkt Lernen, die eine Förderschule besuchen* – angenommen werden muss. Die Gründe dafür sind sicherlich identisch mit den vorherigen Ausführungen. Zum einen ist die Gruppengröße sehr unterschiedlich und zum anderen sind die Ergebnisausprägungen nicht groß genug. Dennoch sind bei der Ergebnisauswertung der SchülerInnen beider Gruppen die differenten Anzahlen an gelösten Items prägnant. So haben die inklusiv beschulten SchülerInnen wesentlich mehr Items richtig gelöst als die Gruppe an der Förderschule – ein Ergebnis, das in Anbetracht der beschriebenen Beobachtungen während der Testungen nicht außergewöhnlich ist. Die inklusiven Kinder sind mit dem Programm LEVUMI augenscheinlich besser zurechtgekommen. In den bisherigen Ausführungen der vorherigen Forschungsfragen stan-

den vor allem die tatsächlich numerisch erzielten Werte im Vordergrund. Diese Ergebnisse sollen Anlass dazu sein, darüber zu diskutieren, warum es zu solchen Unterschieden zwischen den SchülerInnen kommen kann und welche Beschulungsform eventuell die zielführendere für den/die SchülerIn sein könnte.

Inklusion ist in unserer Gesellschaft ein allgegenwärtiger Begriff, der nicht nur prägend für das heutige Schulsystem sein soll, sondern zur Teilhabe in allen Bereichen des Lebens führen soll. Die UN-BRK (2006) hat dazu geführt, das schulische System in Deutschland neu zu konstruieren und didaktische Maßnahmen zu ergreifen, die letztlich gewinnbringend für alle sein sollen. Die mit der Inklusion einhergehende Heterogenität führte zu einer grundlegenden Veränderung der Struktur in Schulen. Diese veränderte Struktur führt nicht nur zu einer höheren Arbeitsbelastung der Lehrkräfte durch ein immer weiter ansteigendes Aufkommen von Differenzierungsanforderungen, sondern verlangt auch nach neuen diagnostischen Verfahren. Die Plattform LEVUMi bietet eine Möglichkeit, diesen neuen Anforderungen gerecht zu werden.

Auch wenn die Zahlen der inklusiv beschulten Kinder jährlich steigen (Kapitel 2.2) und der Abbau der Förderschulen immer mehr zunimmt, stellt sich die Frage, welche Auswirkungen das neue System auf die einzelnen Schüler hat. Ist der inklusive Weg wirklich der richtige? Oder sind einige SchülerInnen an der Förderschule weiterhin gut aufgehoben? Fragen, die diese Arbeit nicht beantworten kann, über die man aber zumindest diskutieren kann.

In Kapitel 2.4. wurde bereits beschrieben, dass es bisher nur sehr wenige empirische Nachweise zur Umsetzung der Inklusion gibt. Bezüglich des Förderschwerpunktes Lernen konnten positive Effekte im Rahmen einer inklusiven Beschulung festgestellt werden (ebenfalls Kapitel 2.4). Insbesondere konnte festgestellt werden, dass SchülerInnen, die eine Förderschule besuchen, mit der Zeit immer schlechter wurden. Dieser Argumentation folgend stellt sich die Frage, warum dann noch Förderschulen für den Förderschwerpunkt Lernen existieren.

Die beschriebene Leistungsabnahme von FörderschülerInnen im Verlauf der Zeit kann auf die Ergebnisse der Subgruppe der Förderschüler aus dieser Studie übertragen werden, jedoch auch auf die Gruppe der Grundschüler mit einem Förderbedarf, was die Argumentation wieder relativiert. Hinzu kommt, dass der Zeitraum von drei

Woche zwischen den MZP nicht kongruent zum zeitlichen Aspekt der genannten Argumentation genutzt werden kann. Dennoch: Fortschritte haben letztlich beide Gruppen nicht gemacht. Trotzdem bleibt unumstritten, dass sich von allen Förderungsbedarfen der Bereich der Lernschwierigkeiten wahrscheinlich am leichtesten in das bestehende Schulsystem inkludieren lässt. Denn für lernbeeinträchtigte SchülerInnen ist es zweifellos ein Gewinn, im gemeinsamen Unterricht beschult zu werden. Gleichwohl muss auch gesagt werden, dass eine inklusive Beschulung auch nicht für jede/n SchülerIn automatisch Gelingensbedingungen darstellt. So muss anhand der festgestellten Ergebnisse der Gruppe der Förderschulkinder gesagt werden, dass sie – zumindest was den Kompetenzbereich Lesen angeht – an der jeweiligen Förderschule besser aufgehoben sind. Ihre Leistungen weichen zu stark ab, und das nicht nur von der Subgruppe der inklusiv beschulten SchülerInnen. Zudem werden einige SchülerInnen, nicht zuletzt aufgrund unabdingbarer Fördernotwendigkeiten, weiterhin auf spezielle pädagogische Settings angewiesen sein. Das bedeutet: Es ist immer notwendig, den Einzelfall in den Blick zu nehmen, um dann zu entscheiden, welche Beschulungsform die beste ist. Ausgehend vom Einzelfall muss bei der Betrachtung der Ergebnisse der SchülerInnen aber noch ein weiterer Punkt beachtet werden: der des Zeitpunktes der Zuschreibung eines Förderbedarfs. In der Regel ist davon auszugehen, dass die SchülerInnen einer Klasse, sofern keine Differenzierung vonnöten ist, nach den gleichen Methoden und mit denselben Medien unterrichtet werden. Sobald aber ein/e SchülerIn Defizite oder deutlich bessere Leistungen als die MitschülerInnen aufweist, werden Differenzierungsmaßnahmen ergriffen. Wenn bei einer Leistungsabnahme daraufhin ein Förderschwerpunkt zugeschrieben wird, könnte vorsichtig von einer „langsameren“ Beschulung gesprochen werden. SchülerInnen, die bereits zu Beginn ihrer schulischen Karriere eine Förderschule besucht haben, könnten sofort eine ihnen angemessenere Beschulung erfahren und müssten somit auch keine demotivierenden leistungsbezogenen Erfahrungen machen. Dies könnte ebenfalls eine Begründung für den Leistungsunterschied zwischen beiden Gruppen darstellen. Die schulischen Leistungen eines/r SchülerIn sind aber nicht nur abhängig vom Zeitpunkt der Zuschreibung eines Förderbedarfs und von der Beschulungsform, sondern auch zu einem großen Teil von der Einstellung der Lehrkraft (Kapitel 2.4).

Denn die Entwicklung einer inklusiven Schule kann nur als Prozess verstanden werden, der nur durch in ihm aktiv und kooperierende Akteure, insbesondere der Lehrkräfte, realisierbar ist.

Bezüglich des sinnentnehmenden Lesens muss gesagt werden, dass der Test, zumindest für die Gruppe der Förderschulkinder, zu schwer ist – unabhängig der Argumentation von MOOSBRUGGER und KELAVA (2012). Der Test hatte auf die teilnehmenden Kinder eher eine demotivierende Wirkung als dass er Defizite aufdecken konnte, um etwaige Fördermaßnahmen zu initiieren oder zu optimieren. Der vorhandene Test mit der Niveaustufe N2 wäre wahrscheinlich zielführender.

Ungeachtet der Ergebnisse dieser Studie sollte dennoch das langfristige Ziel die Implementierung eines inklusiven Schulsystems sein, das guten inklusiven Unterricht produziert. Ebenso sollte die Aufrechterhaltung der Förderschulen weiterhin als notwendig erachtet werden, denn es wird auch weiterhin einen gewissen Prozentsatz der Schülerschaft geben, für die der inklusive Weg nicht der richtige ist.

8 Zusammenfassung und Perspektive

Ziel der vorliegenden Arbeit war zum einen die Berechnung der Itemschwierigkeit des Tests zum sinnentnehmenden Lesen aus LEVUMI und zum anderen der diesbezügliche Vergleich der Ergebnisse der SchülerInnen. Dabei wurde vor allem die Entwicklung des Lernverlaufs der Kinder mit dem Förderschwerpunkt Lernen, die eine Förderschule oder eine Grundschule besuchen, fokussiert. Die Untersuchung wurde zu zwei MZP an verschiedenen Schulen in Herne, Bochum und Dortmund durchgeführt. Bedingt durch die sehr heterogene Klassenstruktur fielen die Ergebnisse der einzelnen SchülerInnen sehr unterschiedlich aus, insbesondere bei den Kindern mit einem sonderpädagogischen Förderbedarf (SPF).

Neben der Berechnung der Itemschwierigkeit war das Thema inklusive Bildung zentral. Die aus der Ratifizierung der UN-BRK resultierende inklusive Bildung fordert die gemeinsame Beschulung aller Kinder. Dazu zählt die Beschulung von Kindern sowohl mit als auch ohne SPF. Die damit einhergehende Erhöhung der Heterogenität der einzelnen Klassen erforderte auch neue Wege bei der Diagnose von Lernverläufen. Da sich das Abbilden von Lernverläufen über einen längeren Zeitraum schon immer als sehr zeitaufwendig erwiesen hat, wurde in der Vergangenheit meist darauf verzichtet. Die Entwicklung von computerbasierten Schulleistungstests hat sich hier als sehr hilfreich erwiesen. Die in dieser Arbeit vorgestellte Online-Plattform LEVUMI ist ein Beispiel für ein computergestütztes Verfahren, das in der Lage ist, den Lernverlauf einzelner SchülerInnen oder ganzer Klassen darzustellen. Es ist unbestreitbar, dass den Schulleistungstests ein wichtiger Stellenwert bei Fragen der Objektivierung von Leistungen der SchülerInnen zukommt, allerdings ist es unbefriedigend, dass diese Systeme in der schulischen Praxis kaum eingesetzt werden – und dies liegt nicht immer an Unkenntnis über die Existenz solcher Programme.

Der aus dem Repertoire von LeVuMi verwendete Test für diese Untersuchung war der des sinnentnehmenden Lesens. Der Test hat eine zeitliche Begrenzung und verwendet Items aus verschiedenen Kategorien. Die Itemschwierigkeit kann für die Teilnehmer dieser Stichprobe insgesamt als leicht bezeichnet werden. Zu jedem MZP wurde jeweils nur ein Item mit dem Schwierigkeitsindex von $33 \geq P_i \geq 1$ festgestellt. Eine Darstellung des Schwierigkeitsgrades nach den Kategorien (Kapitel 6.1.1) des Tests wurde allerdings nicht übernommen. Zum einen würde es den Rahmen dieser Arbeit übersteigen und zum anderen wird aufgrund der Ergebnisse der einzelnen

Items keine Notwendigkeit gesehen, zumindest nicht, was die Aussortierung schwerer Items angeht. Dennoch sind die Ergebnisse bezüglich der teilnehmenden SchülerInnen sehr unterschiedlich ausgefallen. Dies lässt sich vor allem durch die Heterogenität der Stichprobe erklären. Die Ergebnisse aller an der Untersuchung teilnehmenden SchülerInnen waren zwischen dem ersten und zweiten MZP fast gleich. Da sich diese Arbeit insbesondere auf die Merkmalsausprägungen der Kinder mit dem Förderschwerpunkt Lernen konzentriert hat, wurden die Ergebnisse schülerweise betrachtet. Auch wenn die Auswertung mit SPSS keine Signifikanzen festgestellt hat, konnten in den Einzelnachweisen der SchülerInnen Unterschiede festgestellt werden. Diese Unterschiede betreffen vor allem die Differenz der Werte zwischen den inklusiv beschulten Kindern mit dem Förderschwerpunkt Lernen und denen an der Förderschule. Eine mögliche Begründung – die der inklusiven Beschulungsform – wurde dabei diskutiert. Die Arbeit ist zu dem Schluss gekommen, dass das inklusive Schulsystem maßgeblichen Anteil an den schulischen Leistungen von Kindern hat, kann dies aber nur mit den Ergebnissen dieser Untersuchung und dem Test zum sinnentnehmenden Lesen nach LeVuMi nicht belegen.

Fakt ist: Inklusive Bildung stellt eine besondere Herausforderung für die Institution Schule dar. Sie ist kein Thema mehr, das am Rande der Gesellschaft steht, sondern es ist zentral, um qualitativ hochwertige Bildung für alle Lernenden zu erreichen und um eine inklusivere Gesellschaft zu entwickeln. Zudem ist sie wesentlich, um soziale Gerechtigkeit zu erreichen (vgl. UNESCO 2010). Die Gewährleistung dieser Anforderung ist dabei maßgeblich von der Zustimmung aller am Prozess beteiligten Personen abhängig (vgl. Ellger-Rüttgardt 2016). Inklusive Bildung ist eine Vision, die möglicherweise niemals ganz realisierbar ist, aber vielleicht sollte das auch nicht das Ziel sein. Inklusion sollte vielmehr als Chance gesehen werden, für mehr Gemeinsamkeit von Kindern mit und ohne Beeinträchtigung zu sorgen und gleichzeitig dazu beizutragen, Diskriminierung zu vermeiden. Und gerade deswegen sollte bei der Wahl der richtigen Schulform für ein Kind – ob mit oder ohne SPF – immer auf den Einzelfall geschaut und dann entschieden werden, welche Beschulungsform am sinnvollsten ist.

III Literaturverzeichnis

Ahrbeck, Bernd (2016): *Inklusion. Eine Kritik*. 3. Aufl. Stuttgart: Kohlhammer (Brennpunkt Schule).

Altmann, Gerry; Steedman, Mark (1988): *Interaction with context during human sentence processing*. In: *Cognition* 30 (3), S. 191–238.

Artelt, Cordula; Stanat, Petra; Schneider, Wolfgang; Schiefele, Ulrich; Lehmann, Rainer (2004): *Die PISA-Studie zur Lesekompetenz: Überblick und weiterführende Analysen*. In: Ulrich Schiefele, Cordula Artelt, Wolfgang Schneider und Petra Stanat (Hg.): *Struktur, Entwicklung und Förderung von Lesekompetenz. Vertiefende Analysen im Rahmen von PISA 2000*. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 139–168.

Beauftragte der Bundesregierung für die Belange von Menschen mit Behinderungen (Hg.) (2017): *Die UN-Behindertenrechtskonvention. Übereinkommen über die Rechte von Menschen mit Behinderungen*. Online verfügbar unter https://www.behindertenbeauftragte.de/SharedDocs/Publikationen/UN_Konvention_deutsch.pdf?__blob=publicationFile&v=2, zuletzt geprüft am 15.02.2018.

Biermann, Julia; Pfahl, Lisa (2016): *Menschenrechtliche Zugänge und inklusive Bildung*. In: Gottfried Biewer, Ingeborg Hedderich, Judith Hollenweger und Reinhard Markowetz (Hg.): *Handbuch Inklusion und Sonderpädagogik*. Bad Heilbrunn: Klinkhardt, S. 199–207.

Biewer, Gottfried; Fasching, Helga (2012): *Von der Förderung zum inklusiven Bildungssystem - Perspektive der Schulentwicklung*. In: Ulrich Heimlich und Joachim Kahlert (Hg.): *Inklusion in Schule und Unterricht. Wege zur Bildung für alle*. Stuttgart: Kohlhammer (Praxis Heilpädagogik: Handlungsfelder), S. 117–152.

Biewer, Gottfried; Schütz, Sandra (2016): *Inklusion*. In: Gottfried Biewer, Ingeborg Hedderich, Judith Hollenweger und Reinhard Markowetz (Hg.): *Handbuch Inklusion und Sonderpädagogik*. Bad Heilbrunn: Klinkhardt, S. 123–131.

Börnert, Moritz (2014): *Lernverlaufsdiagnostik. Definition, Einsatzbereiche und Perspektiven für die pädagogische Praxis*. Universität Potsdam, Potsdam. Zentrum

für empirische Inklusionsforschung. Online verfügbar unter https://www.uni-potsdam.de/fileadmin01/projects/inklusion/PDFs/ZEIF-Blog/B%C3%B6rnert_2014_Lernverlauf.pdf, zuletzt geprüft am 12.01.2018.

Breitenbach, Erwin (2014): *Psychologie in der Heil- und Sonderpädagogik*. Unter Mitarbeit von Annett Kuschel. Stuttgart: Kohlhammer (Nachbarwissenschaften der Heil- und Sonderpädagogik, 1).

Bundesministerium für Arbeit und Soziales, Referat Information, Monitoring, Bürgerservice, Bibliothek (Hg.) (2011): *Übereinkommen der Vereinten Nationen über die Rechte von Menschen mit Behinderungen*. Online verfügbar unter http://www.bmas.de/SharedDocs/Downloads/DE/PDF-Publikationen/a729-un-konvention.pdf?__blob=publicationFile, zuletzt geprüft am 18.02.2018.

Christmann Ursula; Groeben, Norbert (2006): *Psychologie des Lesens*. In: Bodo Franzmann, Klaus Hasemann, Dietrich Löffler und Erich Schön (Hg.): *Handbuch Lesen*. Im Auftrag der Stiftung Lesen und der Deutschen Literaturkonferenz. Unter Mitarbeit von Georg Jäger, Wolfgang R. Langenbucher und Ferdinand Melichar. 2. Aufl. Baltmannsweiler: Schneider Verl. Hohengehren, S. 145–223.

Christmann Ursula; Groeben, Norbert (2009): *Anforderungen und Einflussfaktoren bei Sach- und Informationstexten*. In: Norbert Groeben und Bettina Hurrelmann (Hg.): *Lesekompetenz. Bedingungen, Dimensionen, Funktionen*. 3. Aufl. Weinheim, München: Juventa (Lesesozialisation und Medien / Herausgegeben von Bettina Hurrelmann und Gisela Wilkending), S. 150–173.

Dedering, Kathrin (2016): *Zusammen oder getrennt? Inklusive und separierende Beschulung im Licht der empirischen Forschung*. In: Bertelsmann Stiftung (Hg.): *Inklusion kann gelingen! Forschungsergebnisse und Beispiele guter schulischer Praxis*. Gütersloh: Verlag Bertelsmann Stiftung, S. 34–49.

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

Deutsche UNESCO-Kommission e.V. (Hg.) (2010): *Inklusion: Leitlinien für die Bildungspolitik*. Online verfügbar unter <https://www.unesco.de/fileadmin/medien/Dokumente/Bildung/InklusionLeitlinienBildungspolitik.pdf>, zuletzt geprüft am 20.02.2018.

- Ellger-Rüttgardt, Sieglind** (2016): *Inklusion. Vision und Wirklichkeit*. Stuttgart: Kohlhammer.
- European Agency for Development in Special Needs Education (EADSNE)** (Hg.) (2012): *Special Needs Education. Country Data 2012*. Online verfügbar unter <http://www.europe-an-agency.org/publications/ereports/sne-country-data-2012/sne-country-data-2012.pdf>, zuletzt geprüft am 17.01.2018.
- Gebhardt, Markus; Diehl, Kirsten; Mühling, Andreas** (2016a): *Lern-Verlaufs-Monitoring. LEVUMI Lehrerhandbuch*. Version 1.1. Unter Mitarbeit von Christine Engert-Seitz, Ute Haid, Roman Heinz, Katrin Scheler und Sabine Thoma. Online verfügbar unter https://eldorado.tu-dortmund.de/bitstream/2003/35765/2/CBM_Lehrerhandbuch%20LEVUMI_final_1.1%20September.pdf, <http://dx.doi.org/10.17877/DE290R-17792>.
- Gebhardt, Markus; Diehl, Kirsten; Mühling, Andreas** (2016b): *Online-Lernverlaufsmessung für alle Schülerinnen und Schüler in inklusiven Klassen*. In: *Zeitschrift für Heilpädagogik* (10), S. 444–453.
- Gebhardt, Markus; Marco G. P. Hessels; Nusser, Lena; Schwab, Susanne** (2015): *Einstellungen und Selbstwirksamkeit von Lehrerinnen und Lehrern zur schulischen Inklusion in Deutschland – eine Analyse mit Daten des Nationalen Bildungspanels Deutschlands (NEPS)*. In: *Empirische Pädagogik* 29 (2), S. 211–229.
- Grohnfeldt, Manfred** (2009): *Diagnostik, Prävention und Evaluation in der Sprachheilpädagogik*. In: Manfred Grohnfeldt (Hg.): *Lehrbuch der Sprachheilpädagogik und Logopädie*. 2. Aufl. Stuttgart: Kohlhammer (Diagnostik, Prävention und Evaluation, 3), S. 17–29.
- Häcker, Thomas; Walm, Maik** (Hg.) (2015): *Inklusion als Entwicklung. Konsequenzen für Schule und Lehrerbildung*. Bad Heilbrunn: Klinkhardt.
- Heimlich, Ulrich** (2012): *Einleitung: Inklusion und Sonderpädagogik*. In: Ulrich Heimlich und Joachim Kahlert (Hg.): *Inklusion in Schule und Unterricht. Wege zur Bildung für alle*. Stuttgart: Kohlhammer (Praxis Heilpädagogik : Handlungsfelder), S. 9–26.

- Heimlich, Ulrich** (2016): *Gemeinsamer Unterricht im Rahmen inklusiver Didaktik*. In: Ulrich Heimlich und Franz B. Wember (Hg.): *Didaktik des Unterrichts im Förderschwerpunkt Lernen*. Ein Handbuch für Studium und Praxis. 3. Aufl. Stuttgart: Kohlhammer, S. 69–80.
- Hesse, Ingrid; Latzko, Brigitte** (2011): *Diagnostik für Lehrkräfte*. 2. Aufl. Opladen: Verlag Barbara Budrich.
- Hillenbrand, Clemens** (2013): *Inklusive Bildung in der Schule. Probleme und Perspektiven für die Bildungsberichterstattung*. In: *Zeitschrift für Heilpädagogik* (9), S. 359–369.
- Hurrelmann, Bettina** (2002): *Leseleistung - Lesekompetenzen. Folgerungen aus PISA, mit einem Plädoyer für ein didaktisches Konzept des Lesens als kultureller Praxis*. In: *Praxis Deutsch* 29 (176), 6-18.
- Hurrelmann, Bettina** (2011): *Modelle und Merkmale der Lesekompetenzen*. In: Andrea Bertschi-Kaufmann und Tanja Graber (Hg.): *Lesekompetenz - Leseleistung - Leseförderung*. Grundlagen, Modelle und Materialien. 4. Aufl. Seelze: Klett und Balmer Verlag (Lehren lernen), S. 18–28.
- Information und Technik Nordrhein-Westfalen** (Hg.) (2016): *NRW (ge)zählt: Schulen in Nordrhein-Westfalen. Ergebnisse der Schulstatistik – Ausgabe 2016*. Geschäftsbereich Statistik, zuletzt geprüft am 14.01.2018.
- Jungjohann, Jana; Gebhardt, Markus; Diehl, Kirsten; Mühling, Andreas** (2017): *Förderansätze im LESEN mit LEVUMI*. Unter Mitarbeit von Christine Engert-Seitz, Ute Haid, Roman Heinz, Katrin Scheler und Sabine Thoma. Online verfügbar unter https://eldorado.tu-dortmund.de/bitstream/2003/36024/1/F%c3%b6rderans%c3%a4tze_Lehrerhandbuch%20LEVUMI.PDF, <http://dx.doi.org/10.17877/DE290R-18042>.
- Katzenbach, Dieter; Schnell, Irmtraud** (2012): *Strukturelle Voraussetzungen inklusiver Bildung*. In: Vera Moser (Hg.): *Die inklusive Schule. Standards für die Umsetzung*. Stuttgart: Kohlhammer Verlag (Schulpädagogik), S. 21–39.
- Klauer, Karl Josef** (2006): *Erfassung des Lernfortschritts durch curriculumbasierte Messung*. In: *Heilpädagogische Forschung* 32 (1), S. 16–26.

- Klauer, Karl Josef** (2011): *Lernverlaufsdagnostik – Konzept, Schwierigkeiten und Möglichkeiten*. In: *Empirische Sonderpädagogik* 3 (3), S. 207–224.
- Klauer, Karl Josef** (2014): *Formative Leistungsdiagnostik: Historischer Hintergrund und Weiterentwicklung zur Lernverlaufsdagnostik*. In: Marcus Hasselhorn, Wolfgang Schneider und Ulrich Trautwein (Hg.): *Lernverlaufsdagnostik*. Göttingen: Hogrefe (Tests und Trends, 12), S. 2–17.
- Klicpera, Christian; Gasteiger-Klicpera, Barbara; Schabmann, Alfred; Schmidt, Barbara** (2013): *Legasthenie - LRS. Modelle, Diagnose, Therapie und Förderung*. 4. Aufl. München: Ernst Reinhardt Verlag (UTB).
- Kocaj, Aleksander; Kuhl, Poldi; Kroth, Anna J.; Pant, Hans Anand; Stanat, Petra** (2014): *Wo lernen Kinder mit sonderpädagogischem Förderbedarf besser? Ein Vergleich schulischer Kompetenzen zwischen Regel- und Förderschulen in der Primarstufe*. In: *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 66 (2), S. 165–191.
- König, Philip** (2009): *Förderung der Lesekompetenz durch kooperative und selbstgesteuerte Lernformen*. Hamburg: Igel Verlag (Gesellschaftswissenschaften).
- Kroworsch, Susann** (2017): *Das Recht auf inklusive Bildung. Allgemeine Bemerkung Nr. 4 des UN-Ausschusses für die Rechte von Menschen mit Behinderungen*. Hg. v. Deutsches Institut für Menschenrechte. Berlin. Online verfügbar unter http://www.institut-fuer-menschenrechte.de/fileadmin/user_upload/Publikationen/Information/Information_Nr_12_Das_Recht_auf_inklusive_Bildung_bf.pdf, zuletzt geprüft am 18.02.2018.
- Kuhl, Jan; Sinner, Daniel** (2015): *t-Test*. In: Katja Koch und Stephan Ellinger (Hg.): *Empirische Forschungsmethoden in der Heil- und Sonderpädagogik. Eine Einführung*. Göttingen: Hogrefe, S. 153–158.
- Lenhard, Wolfgang** (2013): *Leseverständnis und Lesekompetenz. Grundlagen - Diagnostik - Förderung*. Stuttgart: Kohlhammer (Lehren und Lernen).
- Lenhard, Wolfgang; Artelt, Cordula** (2009): *Komponenten des Leseverständnisses*. In: Wolfgang Lenhard und Wolfgang Schneider (Hg.): *Diagnostik und Förderung des Leseverständnisses*. Göttingen: Hogrefe (Tests und Trends, 7), S. 1–17.

- Lindauer, Thomas; Schneider, Hansjakob** (2011): *Lesekompetenzen ermitteln: Aufgaben im Unterricht*. In: Andrea Bertschi-Kaufmann und Tanja Graber (Hg.): Lesekompetenz - Leseleistung - Leseförderung. Grundlagen, Modelle und Materialien. 4. Aufl. Seelze: Klett und Balmer Verlag (Lehren lernen), S. 109–125.
- Luhmann, Niklas** (1994): *Inklusion und Exklusion*. In: Helmut Berding (Hg.): Nationales Bewusstsein und Identität. Frankfurt am Main: Suhrkamp, S. 15–46.
- Mayer, Andreas** (2012): *Blitzschnelle Worterkennung (BliWo). Grundlagen und Praxis*. 2. Aufl. Dortmund: Borgmann Media.
- Mayer, Andreas** (2016): *Lese-Rechtschreibstörungen (LRS). Ursachen, Diagnose und Förderung*. Stuttgart: UTB; Ernst Reinhardt (UTB, UTB-Band-Nr. 8662).
- McElvany, Nele** (2008): *Förderung von Lesekompetenz im Kontext der Familie*. Münster: Waxmann.
- Mißling, Sven; Ückert, Oliver** (2015): *Das Recht auf inklusive Bildung in der Schule. Stand der Umsetzung von Art. 24 UN-BRK in der deutschen Schulgesetzgebung nach fünf Jahren*. In: *Recht der Jugend und des Bildungswesens* 62 (1), S. 63–78.
- Moosbrugger, Helfried** (2012): *Item-Response-Theorie (IRT)*. In: Augustin Kelava und Helfried Moosbrugger (Hg.): Testtheorie und Fragebogenkonstruktion. 2. Aufl. Berlin: Springer (Springer-Lehrbuch), S. 227–274.
- Mühling, Andreas; Gebhardt, Markus; Diehl, Kirsten** (2017): *Formative Diagnostik durch die Onlineplattform LEVUMI*. In: *Informatik Spektrum* 40 (6), S. 556–561. DOI: 10.1007/s00287-017-1069-7.
- Pijl, Sip Jan; Frissen, Paul H.A.** (2009): *What Policymakers Can Do to Make Education Inclusive*. In: *Educational Management Administration & Leadership* 37 (3), S. 366–377.
- Prengel, Annedore** (2016): *Inklusive Pädagogik und ihre Bedeutung für relationale Prozesse*. In: Rolf Göppel und Bernhard Rauh (Hg.): Inklusion. Idealistische forderung, individuelle forderung, institutionelle: Kohlhammer Verlag, S. 30–45.

- Preuss-Lausitz, Ulf** (2013): *Muss eine inklusive ‚Schule für alle‘ die Auflösung des gegliederten Schulsystems zur Folge haben? Zu Anforderungen an die Zukunftsfähigkeit unserer Schulen.* In: Mathias Brodkorb und Katja Koch (Hg.): *Inklusion – Ende des gegliederten Schulsystems? Zweiter Inklusionskongress M-V. Dokumentation.* Schwerin, S. 19–50.
- Preuss-Lausitz, Ulf** (2015): *Implementation schulischer Inklusion in Deutschland. Herausforderungen und Perspektiven.* In: Thomas Häcker und Maik Walm (Hg.): *Inklusion als Entwicklung. Konsequenzen für Schule und Lehrerbildung.* Bad Heilbrunn: Klinkhardt, S. 85–100.
- Rasch, Björn; Friese, Malte; Hofmann, Wilhelm Johann; Neumann, Ewald** (2014): *Quantitative Methoden 1. Einführung in die Statistik für Psychologen und Sozialwissenschaftler.* 4. Aufl. Berlin, Heidelberg: Springer Berlin Heidelberg (Springer-Lehrbuch).
- Rehle, Cornelia** (2013): *Inklusiver Unterricht - (wie) geht das?* In: Klaus Metzger und Erich Weigl (Hg.): *Inklusion - eine Schule für alle.* 3. Aufl. Berlin: Cornelsen Scriptor (Lehrerbücherei Grundschule), S. 42–53.
- Richter, Tobias; Christmann Ursula** (2009): *Lesekompetenzen: Prozessebenen und interindividuelle Unterschiede.* In: Norbert Groeben und Bettina Hurrelmann (Hg.): *Lesekompetenz. Bedingungen, Dimensionen, Funktionen.* 3. Aufl. Weinheim, München: Juventa (Lesesozialisation und Medien / Herausgegeben von Bettina Hurrelmann und Gisela Wilkending), S. 25–58.
- Rioux, Marcia** (2014): *Disability Rights in Education.* In: Lani Florian (Hg.): *The SAGE handbook of special education.* 2. Aufl. London: Sage, S. 131–147.
- Sacher, Werner** (2014): *Leistungen entwickeln, überprüfen und beurteilen. Bewährte und neue Wege für die Primar- und Sekundarstufe.* 6. Aufl. Bad Heilbrunn: Klinkhardt, Julius.
- Schmotz, Wolfgang; Dutke, Stephan** (2004): *Kognitionspsychologische Grundlagen der Lesekompetenzen: Mehrebenenverarbeitung anhand multipler Informationsquellen.* In: Ulrich Schiefele, Cordula Artelt, Wolfgang Schneider und Petra

Stanat (Hg.): Struktur, Entwicklung und Förderung von Lesekompetenz. Vertiefende Analysen im Rahmen von PISA 2000. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 61–99.

Sedlmeier, Peter; Renkewitz, Frank (2013): *Forschungsmethoden und Statistik für Psychologen und Sozialwissenschaftler*. 2. Aufl. München: Pearson.

Sikora, Simon (2015): *Messinstrumente*. In: Katja Koch und Stephan Ellinger (Hg.): Empirische Forschungsmethoden in der Heil- und Sonderpädagogik. Eine Einführung. Göttingen: Hogrefe, S. 76–80.

Speck, Otto (2010): *Schulische Inklusion aus heilpädagogischer Sicht*. München: Reinhardt Verlag.

Spörer, Nadine; Maaz, Kai; Vock, Miriam; Schröder-Lenzen; Agi et al. (2015): *Lernen in der inklusiven Grundschule: Zusammenhänge zwischen fachlichen Kompetenzen. Sozialklima und Facette des Selbstkonzepts*. In: *Unterrichtswissenschaft* 43 (1), S. 22–35.

Stanat, Petra; Pant, Hans Anand; Böhme, Katrin; Richter, Dirk (2012): *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik. Ergebnisse des IQB-Ländervergleichs 2011*. Münster: Waxmann.

Steck, Andrea (2013): *Förderung des Leseverstehens in der Grundschule. Fortbildungsbausteine für Lehrkräfte*. 2. Aufl. Baltmannsweiler: Schneider Verl. Hohengehren.

Steinig, Wolfgang; Huneke, Hans-Werner (2007): *Sprachdidaktik Deutsch. Eine Einführung*. 3. Aufl. Berlin: Erich Schmidt Verlag (Grundlagen der Germanistik, 38).

Stöppler, Reinhilde; Wachsmuth, Susanne (2010): *Förderschwerpunkt geistige Entwicklung. Eine Einführung in didaktische Handlungsfelder*. Paderborn: Schöningh (Aspekte moderner Sonderpädagogik).

Strathmann, Alfons M.; Klauer, Karl Josef (2010): *Lernverlaufsdagnostik. Ein Ansatz zur längerfristigen Lernfortschrittsmessung*. In: *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie* 42 (2), S. 111–122.

- Terfloth, Karin; Bauersfeld, Sören** (2012): *Schüler mit geistiger Behinderung unterrichten. Didaktik für Förder- und Regelschule; mit ... 33 Tabellen*. München, Basel: Reinhardt (UTB).
- Theunissen, Georg** (2013): *Die UN-Konvention, Artikel 24: ein Kommentar*. In: Klaus Metzger und Erich Weigl (Hg.): *Inklusion - eine Schule für alle*. 3. Aufl. Berlin: Cornelsen Scriptor (Lehrerbücherei Grundschule), S. 21–28.
- Voß, Stefan** (2013): *Curriculumbasierte Messverfahren im mathematischen Erstunterricht. Zur Güte und Anwendbarkeit einer Adaption US-amerikanischer Verfahren im deutschen Schulsystem*. Universität Rostock, Rostock. Philosophische Fakultät.
- Wagner, Michael** (2012): *Die inklusive Schule für alle - wirklich eine Schule für alle?* In: Cornelius Breyer, Günther Fohrer, Walter Goschler, Maunela Heger, Christina Kießling und Christoph Ratz (Hg.): *Sonderpädagogik und Inklusion*. Oberhausen: Athena, S. 117–126.
- Walter, Jürgen** (2014): *Lernfortschrittsdiagnostik Lesen (LDL) und Verlaufsdagnostik sinnerfassenden Lesens (VSL). Zwei Verfahren als Instrumente einer formativ orientierten Lesediagnostik*. In: Marcus Hasselhorn, Wolfgang Schneider und Ulrich Trautwein (Hg.): *Lernverlaufsdagnostik*. Göttingen: Hogrefe (Tests und Trends, 12), S. 165–201.
- Wember, Franz B.; Stein, Roland; Heimlich, Ulrich** (Hg.) (2014): *Handlexikon Lernschwierigkeiten und Verhaltensstörungen*. Stuttgart: Kohlhammer.
- Werning, Rolf** (2016): *Lernen*. In: Gottfried Biewer, Ingeborg Hedderich, Judith Hollenweger und Reinhard Markowetz (Hg.): *Handbuch Inklusion und Sonderpädagogik*. Bad Heilbrunn: Klinkhardt, S. 229–234.
- Wilbert, Jürgen** (2014): *Instrumente zur Lernverlaufsmessung: Gütekriterien und Auswertungsherausforderungen*. In: Marcus Hasselhorn, Wolfgang Schneider und Ulrich Trautwein (Hg.): *Lernverlaufsdagnostik*. Göttingen: Hogrefe (Tests und Trends, 12), S. 281–308.
- Wilbert, Jürgen; Linnemann, Markus** (2011): *Kriterien zur Analyse eines Tests zur Lernverlaufsdagnostik*. In: *Empirische Sonderpädagogik* (3), S. 225–242.

Wild, Elke; Lütje-Klose, Birgit; Schwinger, Malte; Gorges, Julia; Neumann, Philip (2017): *Bielefelder Längsschnittstudie zum Lernen in inklusiven und exklusiven Förderarrangements (BiLieF)*. Technical Report, zuletzt aktualisiert am 2017.

Wocken, Hans (2007): *Fördert Sonderschule? Eine empirische Rundreise durch Schulen für "optimale" Förderung*. In: Irene Demmer-Dieckmann und Annette Textor (Hg.): *Integrationsforschung und Bildungspolitik im Dialog*. Bad Heilbrunn: Klinkhardt, S. 35–60.

Wocken, Hans (2010): *Integration und Inklusion. Ein Versuch die Integration vor der Abwertung und die Inklusion vor Träumerei zu retten*. In: Anne-Dore Stein, Imke Niediek und Stefanie Krach (Hg.): *Integration und Inklusion auf dem Weg ins Gemeinwesen*. Bad Heilbrunn: Verlag Julius Klinkhardt, S. 204–234.

Internetquellen:

MSW NRW (2018): Online verfügbar unter <https://www.schulministerium.nrw.de/docs/Schulsystem/Schulformen/index.html>, zuletzt geprüft am 12.02.2018.

Committee on the Rights of Persons with Disabilities (OHCHR) (2018): Online verfügbar unter <https://www.un.org/development/desa/disabilities/convention-on-the-rights-of-persons-with-disabilities/article-24-education.html>, zuletzt geprüft am 10.02.2018.

IV Tabellenverzeichnis

Tabelle 1: Itemschwierigkeit des Tests zum sinnentnehmenden Lesen (N4)	52
Tabelle 2: Itemschwierigkeit des Tests zum sinnentnehmenden Lesen (N4)	55
Tabelle 3: Ergebnisse der Gruppe der inklusiv beschulten SchülerInnen mit dem Förderschwerpunkt Lernen	59
Tabelle 4: Ergebnisse der Gruppe der SchülerInnen mit dem Förderschwerpunkt Lernen, die eine Förderschule besuchen	60

V Abbildungsverzeichnis

Abbildung 1: Das Schulsystem in Nordrhein-Westfalen (IT.NRW 2016, S. 6)	8
Abbildung 2: Klassenübersicht mit zwei Messzeitpunkten	41
Abbildung 3: Förderbausteine und Entwicklungsbereiche von LeVuMi (vgl. Jungjohann et al. 2017)	43

VI Anhang

1. Wörtliche Instruktionen des zum sinnentnehmenden Lesen

„Der kleine Drache LeVuMi möchte heute gerne wissen, wie gut du schon lesen kannst. LeVuMi hat viele Sätze mitgebracht, in denen immer ein Wort fehlt. LeVuMi fragt sich, ob du herausfinden kannst, welches Wort in den Satz passt?“

Dafür arbeiten wir gleich zusammen am Computer. Am Computer siehst du nacheinander verschiedene Sätze. In jedem Satz fehlt immer genau ein Wort. Deine Aufgabe ist es, das richtige Wort herauszufinden. Dir werden dazu immer vier Wörter vorgeschlagen. Ein Wort passt und drei Wörter sind falsch. Mit der Maus kannst du die Wörter auswählen und weiterklicken. Du hast 7 Minuten lang Zeit, so viele Sätze zu lesen, wie du schaffst. Konzentriere dich gut, damit du keine Fehler machst. In drei Wochen komme ich noch einmal wieder und du darfst noch einmal mit LeVuMi lesen. Dann zeigt LeVuMi dir, ob du etwas besser lesen kannst als heute.“

Ortswechsel zum Computer. Entweder den Log-In gemeinsam mit den Kindern durchführen. Alternativ können die Schüleraccounts auch schon vorher geöffnet werden. Bitte überprüfen Sie, dass jedes Kind am richtigen Account sitzt und arbeitet. Wenn alle Kinder die Übersichtsseite sehen, kann mit der Erklärung fortgefahren werden.

„Hier siehst du ein Beispiel. Der Satz lautet: Ein ... kann fahren. Dazu siehst du vier Wörter zur Auswahl. Hier sind es: Auge, Essen, Sonntag, Auto. Welches Wort passt in den Satz?“

Eine Schülerin oder einen Schüler antworten lassen.

„Genau. Ein Auto kann fahren. Wenn du weißt, welches Wort richtig ist, kannst du es mit der Maus auswählen. Dazu musst du es nur anklicken. Danach siehst du das Wort in der Lücke. Wenn du dich umentscheidest, kannst du danach auch ein anderes Wort auswählen. Dafür klickst du das andere Wort wieder mit der Maus an. Dir werden immer vier Wörter vorgeschlagen. Ein Wort passt in den Satz. Die anderen drei sind falsch und ergeben keinen Sinn. Wenn du dir sicher bist, dass du das richtige Wort ausgewählt hast und es in der Lücke steht, klickst du auf ‚weiter‘. Danach zeigt dir LeVuMi einen neuen Satz an, in dem wieder ein Wort fehlt. Denke daran, dass du

7 Minuten Zeit hast. Der Computer achtet für dich auf die Zeit. Die Zeit ist um, wenn du LeVuMi in groß siehst. Hast du noch eine Frage?“

Fragen der Schülerinnen und Schüler klären. „Dann kannst du jetzt auf Start klicken. Danach zeigt LeVuMi dir direkt den ersten Satz.“