

# Der Einfluss kognitiver Basisfertigkeiten auf die Änderung der in PISA gemessenen Lesekompetenz

Carolin Hahnel · Frank Goldhammer · Ulf Kröhne · Anja Schiepe-Tiska ·  
Oliver Lüdtke · Gabriel Nagy

Online publiziert: 9. Mai 2017  
© Springer Fachmedien Wiesbaden 2017

**Zusammenfassung** Die Studie untersucht Zusammenhänge zwischen dem Leseverständnis und basalen Prozessen des Leseverstehens auf Wort- und Satzebene sowie des Arbeitsgedächtnisses 15-jähriger Jugendlicher. Es wurde den Fragen nachgegangen, ob Unterschiede in der Effizienz der betrachteten Teilkomponenten zum

---

C. Hahnel (✉) · Prof. Dr. F. Goldhammer · Dr. U. Kröhne  
Deutsches Institut für Internationale Pädagogische Forschung (DIPF),  
Schloßstraße 29, 60486 Frankfurt am Main, Deutschland  
E-Mail: hahnel@dipf.de

Prof. Dr. F. Goldhammer  
E-Mail: goldhammer@dipf.de

Dr. U. Kröhne  
E-Mail: kroehne@dipf.de

C. Hahnel · Prof. Dr. F. Goldhammer  
Zentrum für internationale Vergleichsstudien (ZIB), Frankfurt am Main, Deutschland

Dr. A. Schiepe-Tiska  
School of Education, Technische Universität München, Arcisstraße 21, 80333 München, Deutschland  
E-Mail: schiepe-tiska@tum.de

Zentrum für internationale Vergleichsstudien (ZIB), München, Deutschland

Prof. Dr. O. Lüdtke · Prof. Dr. G. Nagy  
Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik (IPN),  
Olshausenstraße 62, 24118 Kiel, Deutschland

Prof. Dr. O. Lüdtke  
E-Mail: oluedtke@ipn.uni-kiel.de

Prof. Dr. G. Nagy  
E-Mail: nagy@ipn.uni-kiel.de

Prof. Dr. O. Lüdtke  
Zentrum für internationale Vergleichsstudien (ZIB), Kiel, Deutschland

einen die Lesekompetenz selbst, zum anderen Veränderungen in der Lesekompetenz nach einem Jahr erklären können. Die Studie basiert auf den Leseleistungsdaten des *Programme for International Student Assessment (PISA) 2012* sowie einer längsschnittlichen Nacherhebung in 2013 (PISA Plus). In einer nationalen Begleitstudie 2012 wurden zusätzlich Fähigkeiten zur Worterkennung, semantischen Integration und des Arbeitsgedächtnisses erfasst. Die Ergebnisse zeigen, dass das Leseverständnis in 2012 durch die untersuchten kognitiven Basisfähigkeiten vorhergesagt wurde. Obwohl sich das Leseverständnis nach einem Jahr verbesserte, ließ sich diese Veränderung weder durch basale Lesefähigkeiten der Worterkennung und semantischen Integration noch durch Arbeitsgedächtnisprozesse erklären. Theoretische und praktische Implikationen werden kritisch diskutiert.

**Schlüsselwörter** PISA Lesekompetenz · Veränderung des Leseverständnisses · Worterkennung · Semantische Integration · Arbeitsgedächtnis

## **The impact of cognitive basic skills on the development of reading comprehension measured in PISA**

**Abstract** This study investigates relations between reading comprehension and basic processes of reading on word and sentence level as well as working memory in 15-year-old adolescents. The questions were considered if differences in the efficiency of the investigated components can explain reading competence as well as changes in reading competence after one year. The study is based on reading data of the *Programme for International Student Assessment (PISA) 2012* and a longitudinal follow-up study in 2013 (PISA Plus). Additionally, basic skills in word recognition, semantic integration, and working memory were assessed as part of a national add-on study in 2012. The results show that reading competence in 2012 was predicted by the investigated cognitive basic skills. Although reading comprehension was improved after one year, this change was neither explained by word recognition, semantic integration, nor working memory. Theoretical and practical implications are discussed critically.

**Keywords** Development of reading comprehension · PISA reading competence · Semantic integration · Word recognition · Working memory

### **1 Texte lesen und verstehen: Zur Bedeutung kognitiver Basisfähigkeiten**

Lesen ist eine zentrale Fähigkeit zur Verwendung und selbständigen Aneignung von Wissen und gilt als unverzichtbares kulturelles „Werkzeug“ zur Kommunikation und gesellschaftlichen Teilhabe. Obwohl bereits im frühen Schulalter diese grundlegende Kompetenz erlernt wird, weisen Studien wie das *Programme for International Student Assessment (PISA; Prenzel et al. 2013)* immer wieder auf einen hohen Anteil jugendlicher leseschwacher Schülerinnen und Schüler (SuS) hin. Auch im Erwachsenenalter setzt sich dieser Befund fort, wie das *Programme for the International Assessment of Adult Competencies (PIAAC; OECD 2013b; Rammstedt et al. 2013)*

an einer repräsentativen Stichprobe von 16- bis 65-Jährigen zeigt. Die Analyse von Alterskohorten ergab hier, dass ca. 40 % der jüngeren Erwachsenen (16–34 Jahre) lediglich niedrige Kompetenzstufen im Lesen erreichen, die höchstens einen Basiswortschatz, die Bildung einfacher Inferenzen oder das Verstehen von Sätzen und Paraphrasen erfordern. Viele Studien, die sich mit der Erklärung dieser Leistungsdiskrepanz befassten, haben bereits wichtige Antezedenzen und Faktoren des Lesens identifiziert (z. B. Arrington et al. 2014; Artelt et al. 2001; Bast und Reitsma 1998; Bloom et al. 2008; Boland 1993; Lesaux et al. 2007). Dabei gelten vor allem Längsschnittstudien als ideal, um den Entwicklungsprozess des Lesens über die Zeit hinweg und in Abhängigkeit von potentiellen Einflussfaktoren zu verfolgen. Solche zeit- und kostenaufwendigen Studien finden in der Large-Scale Forschung allerdings selten statt. Im Rahmen von PISA wurden in Deutschland erstmals 2003 Kompetenzen von SuS im Längsschnitt erfasst (PISA-I-Plus; Prenzel et al. 2006). Lesekompetenz wurde jedoch nicht untersucht. Mit der Studie PISA Plus, die 2013 ergänzend zu PISA 2012 durchgeführt wurde, bietet sich nun erneut die Möglichkeit Kompetenzentwicklungen im Allgemeinen und die der Lesekompetenz im Besonderen zu untersuchen.

Während der Beitrag von Nagy et al. 2017b sich vor allem auf den Leseentwicklungsverlauf in Abhängigkeit von Schulform, Geschlecht, Zuwanderungshintergrund und soziokulturellem Hintergrund konzentriert, befasst sich die vorliegende Studie mit kognitiven Teilaspekten, die den Leseprozess konstituieren und so seine Entwicklung beeinflussen können. Bevor ein Text überhaupt verstanden werden kann, müssen einzelne Wörter erlesen und in den inhaltlichen Kontext des Textes eingeordnet werden (Perfetti und Stafura 2014). Diese Leseteilprozesse erfordern jedoch, das Informationen im Gedächtnis behalten und dort flexibel weiterverarbeitet werden können (vgl. Dehn 2008). Es folgt, dass sowohl Defizite in basalen Lesefähigkeiten auf Wort- und Satzebene als auch Begrenzungen in der Fähigkeit zur Gedächtnisaktualisierung die Entwicklung der Lesekompetenz beeinträchtigen können. Die vorliegende Arbeit geht daher den Fragen nach, ob basale Fähigkeiten des Lesens und des Arbeitsgedächtnisses im Jugendalter überhaupt noch prädiktiv für das Leseverständnis und darüber hinaus für seine weitere Entwicklung relevant sind. Hierzu sollen die Daten aus drei Teilstudien – die PISA 2012 Hauptstudie, eine national durchgeführte *Computer-Based Assessment* (CBA) Begleitstudie sowie PISA Plus – Aufschluss geben.

## 2 Kognitive Prozesse in der Konstruktion des Leseverständnisses

### 2.1 Leseverständnis und Prozesse des Leseverstehens

Lesen ist ein komplexer Konstruktionsprozess, der auf der gleichzeitigen Ermittlung von Informationen aus einem Text und der Einordnung ihrer Bedeutung beruht (Groeben 1982; Kintsch und van Dijk 1978). Der Leser ist dabei nicht passiver Rezipient, sondern verarbeitet Texte aktiv durch internale kognitive Prozesse (Groeben 1982), die hierarchisch organisiert und durch Gedächtnisprozesse gestützt sind, um das Textverstehen zu ermöglichen (vgl. Naumann et al. 2010; Richter et al. 2012).

Leseverstehen bezeichnet hierbei den Prozess des verstehenden Lesens, während das Leseverständnis als Produkt des Leseverstehens resultiert (Rost und Schilling 2006). Zentral ist dabei die Aufnahme und Einordnung sprachlicher, semantischer Informationen in einen Kontext. Dabei sind mehrere Ebenen zu differenzieren.

Auf Wortebene müssen Informationen aus einem Text (wieder-)erkannt und in ihrer Bedeutung identifiziert werden. Hochwertige mentale Wortrepräsentationen ermöglichen dem Leser dabei das schnelle und erfolgreiche Erlesen von Wörtern (*lexical quality hypothesis*, LQH; Perfetti 2007). Die Stabilität und Präzision, mit der solche Wortrepräsentationen im mentalen Lexikon definiert sind, bestimmt die Geschwindigkeit, Genauigkeit und Leichtigkeit, mit der ein Wort abgerufen und identifiziert werden kann. Dadurch werden sie zum limitierenden Faktor im Verstehensprozess (Perfetti und Stafura 2014). Durch Übung kann der Abruf von Wortrepräsentation jedoch zunehmend automatisiert werden. Automatisierung betrifft hierbei nicht nur die Geschwindigkeit, mit der ein Wort erkannt wird, sondern auch dessen Exaktheit (Samuels und Flor 1997). Erfolgreiche Prozesse der Worterkennung präzise und schnell, stellen sie geringere Anforderungen an kognitive Ressourcen des Lesers, die in der Folge für hierarchiehöhere Verstehensprozesse beim Lesen zur Verfügung stehen (Walczyk 2000).

Auf Satzebene müssen erlesene Wörter korrekt kombiniert bzw. semantisch abgeglichen werden, wobei parallele Integrationsvorgänge bei der Verarbeitung von Sätzen eine wichtige Rolle spielen (Groeben 1982; Perfetti und Stafura 2014). Semantische Integration betrifft folglich nicht nur die Verbindung von Textinformationen in einem Satz, sondern auch deren Eingliederung in das individuelle Wissenssystem eines Lesers. Indem der Lesende konstruktiv Inferenzen zu gegebenen sprachlichen Informationen hinzufügt, erhalten diese erst ihren Sinngehalt. Dabei können entgegengesetzte Prozessrichtungen ablaufen: Beispielsweise müssen Worte zur Identifikation von Satzpropositionen erkannt werden (*bottom-up*), können aber wiederum durch die Elaboration des semantischen Kontextes leichter identifiziert werden (*top-down*).

Auf Textebene setzen diese Strukturierungs- und Konstruktionsprozesse ebenso an, wobei die Beziehung zwischen sprachlicher Information, Textsituation und Hintergrundwissen des Lesers noch stärker im Vordergrund tritt. Das Erkennen und Rekodieren von Sätzen kann als Zwischenstadium des Textverstehens angesehen werden, in dem Satzstrukturen und deren propositionale Bedeutungen in eine integrierte Textrepräsentation gefügt werden (Perfetti und Stafura 2014). Durch das Anreichern dieser Textrepräsentation mit Wissens-elementen, Assoziationen und Erfahrungen konstruiert der Leser ein kohärentes Modell seines Verständnisses (Kintsch 1998). Dabei sind Fähigkeiten zum kontinuierlichen Überwachen der Textstruktur und zur kritischen Reflexion und Bewertung eines Textes essentiell (Groeben 1982; Kendeou et al. 2014). Hierzu müssen kontextrelevante Text- und Wissensinformationen nicht nur im Arbeitsgedächtnis behalten, sondern auch parallel und flexibel weiterverarbeitet werden (Daneman und Carpenter 1980; Oberauer und Kliegl 2010). Das Arbeitsgedächtnis stellt insofern eine fundamentale kognitive Ressource im Leseprozess dar (Arrington et al. 2014; Feldman Barrett et al. 2004; Engle und Conway 1998).

## 2.2 Die Entwicklung des Leseverständnisses

Entwicklungspsychologisch zeigen Längsschnittstudien, dass sich das Leseverständnis vor allem zu Beginn des Lesenlernens rasant entwickelt (Farnia und Geva 2013; Lesaux et al. 2007; Parrila et al. 2005). Während sich in der Grundschule starke Leistungszuwächse zeigen, fallen diese zwischen höheren Klassenübergängen wesentlich geringer aus, wie auch die Längsschnittuntersuchungen von van Gelderen et al. (2007) und die Kohortenvergleiche von Tilstra et al. (2009) für niederländische und englisch-sprachige SuS der Sekundarstufe zeigen. Für den Übergang von Klasse 9 auf Klasse 10 quantifizierten Bloom et al. (2008) den Leistungszuwachs im Lesen für US-amerikanische SuS auf durchschnittlich 0,19 Standardabweichungen für verschiedene Maße des Leseverständnisses (Range: 0,04–0,32).

Die individuellen Verläufe der Leseentwicklung sind zwischen SuS im Allgemeinen über die Zeit sehr stabil (Roeschl-Heils et al. 2003). Boland (1993) fand beispielsweise, dass sich niederländische SuS bereits zu Beginn der Grundschule in ihrer Leseleistung unterschieden und diese Unterschiede auch für ihre spätere Leseleistung prädiktiv waren (mit latenten, standardisierten Regressionskoeffizienten von 0,85 bis 0,88 zwischen den Klassenstufen 2 bis 6). Ähnliche Befunde berichten Klicpera und Schabmann (1993) für deutschsprachige SuS von der 2. bis zur 8. Klasse. Dabei deuteten hohe Interkorrelationen in der Lesegenauigkeit (0,64–0,75) und in der Lesegeschwindigkeit (0,76–0,90) eine hohe Leistungsstabilität über die Zeit an. Die Autoren betonten hierbei, dass nur sehr wenige Wechsel zwischen Leistungsgruppen stattfanden (d. h. gute Leser blieben eher in der Gruppe der guten Leser und umgekehrt). Vor allem in der Lesegeschwindigkeit fächerten sich die Leistungsunterschiede zwischen guten und schwächeren Lesern auf. Diesen Matthäus-Effekt – d. h. den Effekt, dass sich die Variabilität in der Leseleistung über die Zeit hinweg vergrößert, während die individuelle Rangordnung der Leistung stabil bleibt – konnten Bast und Reitsma (1997, 1998) für das Leseverständnis von Grundschulkindern nicht zeigen. In ihren Untersuchungen fanden sie allerdings zunehmende interindividuelle Unterschiede im Dekodieren und Erkennen von Wörtern, die mit dem Leseverständnis eng in Beziehung standen. Kinder, die eine schnellere Entwicklung in ihren Worterkennungsfähigkeiten zeigten, erreichten später auch höhere Testwerte im Leseverständnis. Die Variabilität in der Worterkennungsfähigkeit nahm jedoch am Ende der 3. Klasse wieder ab, was Bast und Reitsma als ein entwicklungsbedingtes Limit in der Worterkennung interpretierten. Sie folgerten daraus, dass wenn ein bestimmtes Niveau in der Worterkennung erreicht wäre, eine Ausweitung individueller Unterschiede in anderen Teilfähigkeiten des Lesens und dem Leseverständnis sichtbar würde.

## 2.3 Einflüsse durch basale kognitive Prozesse

Basale Lesefähigkeiten und Arbeitsgedächtnisprozesse spielen vor allem beim Leseerwerb in der Grundschule eine wichtige Rolle (vgl. Bast und Reitsma 1997, 1998; Kendeou et al. 2012, 2014; Richter et al. 2012). Dementsprechend finden sich starke Zusammenhänge mit dem Leseverständnis (Kendeou et al. 2012; Oakhill et al. 2011; Richter et al. 2012; Tilstra et al. 2009). Beispielsweise zeigten Farnia

und Geva (2013) für kanadische SuS der 4. bis 6. Klasse, dass Worterkennung ein konsistenter Prädiktor für das Textverständnis der Kinder war. Im Gegensatz zu den Ergebnissen von Bast und Reitsma (1997, 1998) fanden sie jedoch für diese Schülergruppe keinen bedeutsamen Zusammenhang der Worterkennung mit dem Anstieg in latenten Wachstumskurven des Lesens. Farnia und Geva (2013) räumten dennoch ein, dass nicht nur Exaktheits-, sondern auch Flüssigkeitsmaße der Worterkennung berücksichtigt werden sollten.

Für Jugendliche wurden Zusammenhänge des Leseverständnisses mit basalen Lesefähigkeiten bisher seltener untersucht, da das Dekodieren und Erkennen von Wörtern „für das Problem des Textverständnisses als abgeschlossen, und zwar erfolgreich abgeschlossen“ (Groeben 1982, S. 19; vgl. auch Roeschl-Heils et al. 2003) gilt. Dennoch zeigen Studien, dass effiziente Worterkennungsfähigkeiten für norwegische Jugendliche eine wichtige Rolle beim Verstehen multipler Textdokumente spielen (Bråten et al. 2013) und dass Erwachsene mit gering aufgeprägten Lesekompetenzen langsamer und fehleranfälliger in der Identifikation von Wörtern sind (PIAAC 2012, Rammstedt et al. 2013). Wenn erwachsenen Leser über hochwertige mentale Wortrepräsentationen verfügen (Perfetti 2007), fällt es ihnen auch leichter, Wörter in einen Text und paraphrasierte Inhalte in ein Verstehensmodell zu integrieren. Leser, die Schwierigkeiten mit der Wort-Text-Integration haben, zeigen auch Defizite im Textverständnis (Perfetti 2007; Perfetti und Stafura 2014).

Das Arbeitsgedächtnis weist eine enge Beziehung mit dem Leseverständnis auf. Typischerweise berichten Studien Korrelationen um 0,50 (Dehn 2008). Defizite des Arbeitsgedächtnisses sind dementsprechend häufig mit Defiziten im Leseverständnis assoziiert (Kendeou et al. 2012; Oakhill et al. 2011; Seigneuric und Ehrlich 2005), sofern Leseverständnisschwierigkeiten nicht auf die phonologische Verarbeitung oder Wortdekodierung zurückgehen (Cain et al. 2004; Goff et al. 2005; Swanson und Berninger 1995). Ähnlich ausgeprägte Zusammenhangsmuster finden sich auch für erwachsene Leser (Hannon 2012). Obwohl das Arbeitsgedächtnis ein starker Prädiktor der Leseleistung ist, konnten Farnia und Geva (2013) keinen signifikanten Zusammenhang mit der Entwicklung des Textverständnisses in ihrer Leseentwicklungsuntersuchung feststellen. Jedoch ist einzuschränken, dass sie ein Kapazitätsmaß des Arbeitsgedächtnisses nutzten, das sich weniger zur Abbildung mentaler Informationsspeicherung und -verarbeitung eignet (vgl. Feldman Barrett et al. 2004; Engle und Conway 1998).

## 2.4 Die vorliegende Studie

Die vorliegende Arbeit untersucht unter kognitionspsychologischen Gesichtspunkten, welche Fähigkeiten und Teilprozesse des Lesens für das Leseverständnis und dessen Entwicklung notwendig sind. Es soll somit ein Beitrag zur Rolle kognitiver Basisfähigkeiten in der Leseentwicklung Heranwachsender geleistet und eventuelle Anknüpfungspunkte zur Förderung des Leseverständnisses aufgezeigt werden. Aufbauend auf dem zuvor dargestellten theoretischen Fundament wird als erste Forschungsfrage untersucht, ob sich die Lesekompetenz Jugendlicher durch Prozesse des Leseverstehens auf Wort- und Satzebene sowie durch Arbeitsgedächtnisprozesse vorhersagen lässt. Bisherige empirische Befunde sprechen dafür, dass kognitive

Basisfähigkeiten des Lesens und des Arbeitsgedächtnisses auch im Jugendalter prädiktiv für das Leseverständnis sind. Dementsprechend wurde erwartet, dass Jugendliche ein besseres Leseverständnis zeigen, je präziser sie Wörter erkennen (H1.1), je effizienter sie einen semantischen Kontext beurteilen (H1.2) und je flexibler sie ihre mentalen Arbeitsgedächtnisressourcen zum Behalten und Verarbeiten von Informationen nutzen können (H1.3).

Daran anknüpfend untersucht die zweite Fragestellung, ob SuS einen höheren Leistungszuwachs im Lesen innerhalb eines Schuljahres erfahren, wenn sie über ausgeprägte kognitive Basisfähigkeiten verfügen. Ausgehend von bisherigen Längsschnittuntersuchungen (z. B. Klicpera und Schabmann 1993; van Gelderen et al. 2007) kann erwartet werden, dass die SuS nach einem Jahr eine höhere mittlere Leseleistung zeigen (H2.1). Dabei ist zu erwarten, dass der mittlere Leistungszuwachs verhältnismäßig gering ausfällt (vgl. Bloom et al. 2008) und sich Leistungsunterschiede in Form einer größeren Variabilität auffächern (vgl. Boland 1993). Weiterhin wird erwartet, dass ein Leistungszuwachs in der Lesekompetenz durch die Effizienz von Fähigkeiten zur Worterkennung (H2.2) und semantischen Integration (H2.3) und durch das Arbeitsgedächtnis erklärt werden kann (H2.4).

### 3 Methodisches Vorgehen

#### 3.1 Design und Durchführung

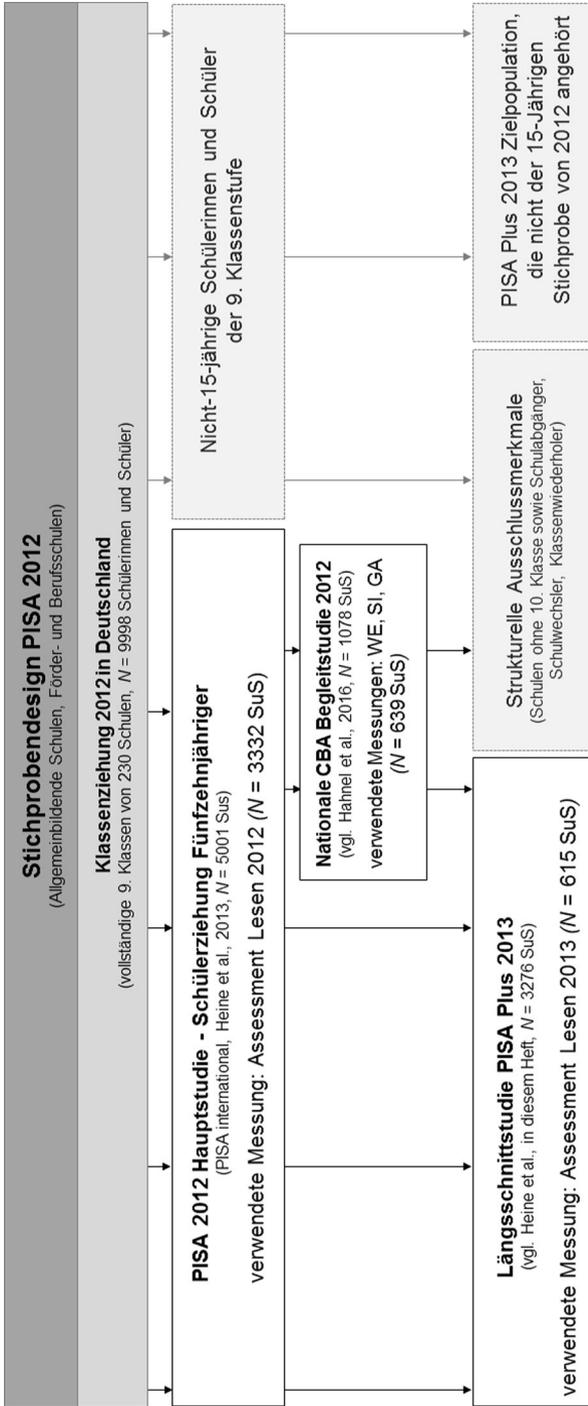
Die empirische Basis bildeten Teilstichproben aus Schülerinnen und Schülern (SuS), die zu verschiedenen Messzeitpunkten (1) am Lese-Assessment der Hauptstudie von PISA 2012, (2) an den Lese- und Arbeitsgedächtnistests der zusätzlichen CBA Begleitstudie sowie (3) am Lese-Assessment von PISA Plus teilnahmen (Abb. 1). Jede dieser Teilstudien umfasste verschiedene Versuchsbedingungen (sog. Booklets), denen die SuS zufällig zugeordnet wurden.

##### 3.1.1 PISA-Hauptstudie 2012

Lesen wurde in PISA 2012 als sogenannte *minor domain* behandelt. Das bedeutet, dass 9 von insgesamt 13 Booklets Leseaufgaben beinhalteten und somit etwa 2/3 aller an PISA teilnehmenden SuS eine von drei Aufgabensammlungen (sog. Cluster) zur Erfassung ihrer Lesekompetenz erhielten. Ein Cluster bestand aus vier bis fünf sog. Units (OECD 2013a). Eine Unit wiederum umfasste einen Text, an den sich drei bis fünf Aufgaben (sog. Items) anschlossen. Die drei verwendeten Lesecluster enthielten insgesamt 43 Items (Cluster 1: 4 Units mit 14 Items; Cluster 2: 4 Units mit 15 Items; Cluster 3: 5 Units mit 14 Items). Detaillierte Angaben zu Design, Durchführung und zur Stichprobe in PISA 2012 finden sich in Heine et al. (2013).

##### 3.1.2 CBA Begleitstudie

Die CBA Begleitstudie ergänzte in Deutschland die PISA-Erhebung von 2012 (vgl. Hahnel et al. 2016). Die teilnehmenden SuS, die zufällig aus der Stichprobe der



**Abb. 1** Stichprobendesign der drei Teilstudien (weiße Kästen) mit Angabe der gezogenen sowie der realisierten Stichprobe. (WE Worterkennung, SI semantische Integration, GA Gedächtnisaktualisierung)

**Tab. 1** Ergebnisse der gemeinsamen Skalierung des Leseverständnisses zum Messzeitpunkt 2012 und 2013 unter Berücksichtigung der Schulform ( $N = 3529$ )

Schulform	$N$	$\eta$	$E(\eta)$	$Var(\eta)$	$Cov(\eta_1, \eta_2)$	$Var(EAP)$	$Rel$
Nicht-Gymnasium	2095	Lesen 2012	0,00 (0,00)	1,16 (0,06)	0,90 (0,04)	0,88	0,76
		Lesen 2013	0,37 (0,12)	1,47 (0,16)			
Gymnasium	1237	Lesen 2012	1,27 (0,07)	0,57 (0,05)	0,89 (0,07)	0,32	0,56
		Lesen 2013	1,53 (0,14)	0,62 (0,10)			

$\eta$  latente Variable,  $Rel$  ist die marginale EAP-Reliabilität =  $Var(EAP)/Var(\eta)$

Hauptstudie 2012 ausgewählt wurden, absolvierten an einem zusätzlichen Testtag verschiedene computerbasierte Tests. Die Erhebungen fanden im Anschluss an die 2012er Haupterhebung statt. Insgesamt kamen Testinstrumente zum Einsatz, die computerbezogene Fertigkeiten, Lese- und Merkfähigkeiten sowie Einstellungen zu den Tests erfassten. Für die vorliegende Studie wurden die gesammelten Daten zweier Tests zur Erfassung basaler Lesefertigkeiten auf Wort- und auf Satzebene sowie eines Tests zur Messung der Fähigkeit zur Gedächtnisaktualisierung verwendet. Diese drei Verfahren wurden in 12 von insgesamt 16 Booklets eingesetzt, wobei sie den SuS in fester Reihenfolge vorlegt wurden (zunächst die Gedächtnisaktualisierungsaufgabe, anschließend die Tests zur Erfassung von Worterkennung und semantischer Integration). Alle Aufgaben wurden über Laptops administriert (Bildschirmauflösung: 1366 × 768px) und durch geschulte Testleiterinnen und Testleiter instruiert. Die Tests enthielten umfassende Anleitungen mit Übungsaufgaben.

### 3.1.3 PISA Plus 2013

Die Erhebungen zu PISA Plus fanden ein Jahr nach der PISA 2012 Hauptstudie statt. Lesekompetenz wurde mithilfe von zwei der in PISA 2012 verwendeten Lesecuster gemessen (Cluster 2 und 3). Diese beiden Lesecuster wurden auf 7 aus insgesamt 18 Booklets verteilt. Das entsprechende Booklet-Design kann Tab. 1 im Beitrag von Nagy et al. 2017a entnommen werden. Da die gezogene Stichprobe einer Zufallsstichprobe aus 2012 entspricht (Heine et al. 2013, 2017), bearbeiteten einige SuS identische Cluster zu 2012 und 2013. Cluster 2 wurde von 112 SuS zu beiden Messzeitpunkten bearbeitet; Cluster 3 von 117 SuS.

## 3.2 Verwendete Messungen und Auswertungsverfahren

### 3.2.1 Lesekompetenz zu 2012 und 2013

Die Texte der Lesekompetenz-Units wiesen unterschiedliche Formen (z. B. Aufsätze, Briefe, Tabellen, Listen) und Strukturen (z. B. Beschreibung, Erzählung, Argumentation) auf und deckten ein breites Situationsspektrum ab (z. B. Lesen zum Zweck des Informierens, Kommunizierens oder Lernens). Die zugehörigen Items erfrag-

ten explizite sowie implizite Informationen aus dem Text, konnten aber auch eine Reflexion über den Text oder dessen Bewertung erfordern. Die Antwortformate der Items umfassten Selektionsaufgaben (*multiple choice*) und offene Textantwortformate. Beispielaufgaben können unter <http://www.pisa.tum.de/beispielaufgaben/> eingesehen werden.

*Bestimmung der Testwerte.* Da ausschließlich die Gruppe der 15-Jährigen betrachtet wurde (Heine et al. 2017, Heine et al. 2013), war es notwendig, die Itemkalibrierung und die damit verbundene Schätzung der Lesekompetenzleistung zu 2012 und 2013 separat vorzunehmen. Die Skalierung wurde mit Mplus durchgeführt (Version 7; Muthén und Muthén, 1998–2012). Die betriebene Itemkalibrierung gleicht der Vorgehensweise von Nagy et al. 2017a. Das heißt, die kategorialen Ergebnisdaten aller SuS, die an den Lese-Assessments teilnahmen (Skalierungsstichproben:  $N_{2012} = 3332$ ,  $N_{2013} = 615$ ), wurden gemeinsam Rasch-skaliert, wobei die Itemschwierigkeiten identischer Items zwischen beiden Messzeitpunkten gleichgesetzt wurden. Verschiedene Positionen der Lesecluster in den Booklets wurden als Unterschiede in den Itemschwierigkeiten zusätzlich modelliert. Weiterhin ging die Schulform (nicht-gymnasiale Schulformen vs. Gymnasium) als Gruppierungsvariable in die Kalibrierung ein. Die hierarchische Datenstruktur, d. h. die Clusterung der Individualdaten in den Schulen, wurde in Mplus durch ein Verfahren berücksichtigt, dass für eine Unterschätzung der Standardfehler korrigiert (TYPE = COMPLEX). Die Metrik der spezifizierten latenten Kompetenzvariablen wurde in Referenz zur ersten Clusterposition des ersten Messzeitpunktes verankert, wobei die erste Clusterposition zur Modellidentifikation auf 0 fixiert wurde. Anschließend wurden *expected a posteriori* (EAP) Fähigkeitsschätzer bestimmt und als Testwerte der Lesekompetenz zu 2012 und 2013 verwendet, die aufgrund des Skalierungsvorgehens frei von Einflüssen des vorgelegten Leseclusters, der Clusterposition und des Booklets sind. Höhere Werte auf beiden Personenvariablen bilden ein besseres Leseverständnis in 2012 bzw. 2013 ab. Tab. 1 gibt die latenten Mittelwerte, Varianzen sowie die marginalen Reliabilitäten getrennt nach Schulform und Messzeitpunkt wieder. Ebenfalls dargestellt sind die latenten Korrelationen zwischen Lesen in 2012 und 2013 getrennt nach Schulform. Auffällig ist jedoch, dass die Reliabilität der EAP Testwerte zwar für Nicht-Gymnasiasten akzeptabel ist, für Gymnasiasten aber stark abnimmt.

### 3.2.2 Basale Lesefertigkeiten auf Wort- und Satzebene

Die Fähigkeiten zur Worterkennung und zur semantischen Integration wurden anhand zweier Subskalen aus einem Inventar zur Erfassung hierarchieniedriger Lesefähigkeiten erhoben (Richter et al. 2012; Richter und Naumann 2009; Richter und van Holt 2005). Zur Abbildung von Worterkennung wurden 32 lexikalische Entscheidungsaufgaben verwendet (je 16 Wörter und Unwörter aus drei bis zehn Buchstaben bzw. ein bis drei Silben, z. B. Igel, Koveau); zur Abbildung semantischer Integration 24 Satzverifikationsaufgaben (je 12 wahre und falsche Sätze aus ein bis drei Propositionen, z. B. „Zucker ist süß“, „Ein Kaktus ist ein kleines, pelziges Tier“). In diesen Aufgaben wurden die SuS gebeten zu entscheiden, ob eine dargebotene Buchstabenkette ein Wort oder ein Unwort ergab bzw. ein dargebotener Satz semantisch wahr oder falsch war. Dabei sollten sie so schnell und genau wie

**Tab. 2** Stichprobengröße (*N*), Itemanzahl (*I*), mittlere Lösungsrate (*L*) sowie Mittelwert, Standardabweichung und Reliabilität der Indikatoren

Indikator	<i>N</i>	<i>I</i>	<i>L</i>	<i>M</i>	<i>SD</i>	$\alpha$
WE Zeit	618	16	0,90	1,28	0,19	0,87
SI Zeit	609	12	0,84	2,37	0,56	0,86
WE Driftrate	631	32	0,90	2,06	0,98	–
SI Driftrate	631	24	0,88	1,25	0,63	–

WE Worterkennung, SI semantische Integration

möglich arbeiten, wobei ihre Antworten (richtig vs. falsch gelöst) und ihre Reaktionszeiten erfasst wurden. Um Einflüsse extremer Reaktionszeiten zu verringern, wurde entsprechend der Empfehlung von Whelan (2008) eine untere Cut-off Grenze von 100 ms festgesetzt. Lange Reaktionszeiten wurden nicht berücksichtigt, wenn sie 3 Standardabweichungen über dem Itemmittelwert lagen (vgl. Richter et al. 2012). Für die Worterkennungsdaten ergab sich eine mittlere Rate fehlender Werte von 4,7 %; für die Daten zur semantischen Integration von 6,1 %. Als Indikatoren wurden die mittlere Reaktionszeit auf Wörter und semantisch wahre Sätze sowie Driftraten der Worterkennung und semantischen Integration gebildet.

**Indikator 1: Mittlere Reaktionszeit auf Wörter bzw. wahre Sätze.** Die mittlere Reaktionszeit wurde aus den Reaktionszeiten der SuS auf Wörter bzw. semantisch wahre Sätze ermittelt. Um einen speed-accuracy-tradeoff (vgl. Pachella and Pew 1968) – also das Phänomen, dass mit zunehmender Geschwindigkeit die Antwortgenauigkeit abnimmt – zu berücksichtigen, wurden bei der Indikatorenbildung ausschließlich die Reaktionszeiten korrekter Antworten verwendet (vgl. Vorgehen von van Gelderen et al. 2007). Die Lösungsraten der 16 Wörter des Worterkennungstests und 12 wahren Sätze der Satzverifikationsaufgabe sind in Tab. 2 dargestellt. Beantworteten die Probanden weniger als die Hälfte der Aufgaben richtig, wurden ihre Reaktionszeiten für die Indikatorenbildung nicht berücksichtigt. Hohe Geschwindigkeiten in der Identifikation von Wörtern zeigen an, dass die SuS existierende Worteinträge mental schnell abzurufen können, während hohe Geschwindigkeiten in der Verifikation von wahren Sätzen darauf hinweisen, dass sie die zugrundeliegende Semantik des Satzes schnell verarbeiten und validieren können.

**Indikator 2: Driftrate.** Es wurde ein weiterer Indikator der Worterkennung und semantischen Integration gebildet, um alle Antwort- und Reaktionszeitdaten (d. h. auch solche auf Unwörter bzw. semantisch falsche Aussagen) miteinzubeziehen und somit neben der Geschwindigkeit auch die Akkuratheit der Aufgabenlösungen zu berücksichtigen (vgl. Samuels und Flor 1997). Dazu wurden Driftraten aus sog. Diffusionsmodellen geschätzt (vgl. Ratcliff et al. 2004; Schroeder 2011). Diffusionsmodelle modellieren den Entscheidungsprozess in *forced-choice* Entscheidungsaufgaben. Sie gehen dabei von der Grundannahme aus, dass eine Entscheidung zwischen zwei Alternativen (z. B. ein Wort wird als Wort oder Unwort erkannt) dadurch entsteht, dass Informationen über den Stimulus mit der Zeit angehäuft werden. Die Informationsmenge akkumuliert dabei gegen ein Entscheidungskriterium zugunsten einer der beiden Antwortalternativen. Die mittlere Rate, mit der die Ent-

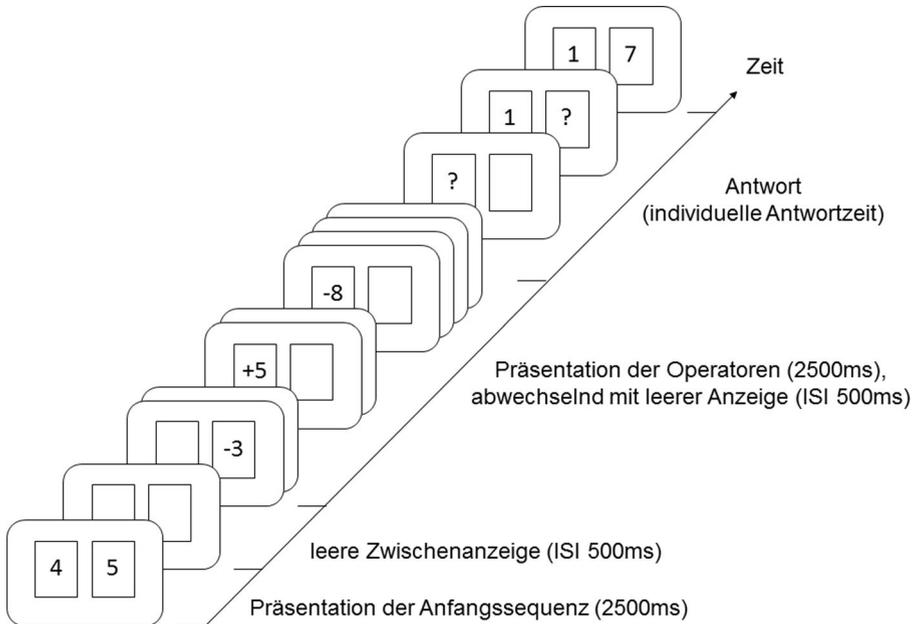
scheidung für eine Antwortalternative erfolgt, wird als Driftrate bezeichnet (Voss et al. 2004). Driftraten geben folglich die relative Informationsmenge an, die per Zeiteinheit aufgenommen wird. Dadurch werden sie als Effizienz interpretiert, mit der Informationen über den Stimulus verarbeitet werden und zu einer Entscheidung führen. Zur Schätzung der Parameter werden die empirischen Verteilungen der beobachteten Reaktionszeiten für richtige und falsche Antworten mit theoretisch erwarteten Verteilungen verglichen, die durch die Parameter des Diffusionsmodells (d. h. die Driftraten) bestimmt werden.

In der vorliegenden Studie entsprechen positive Driftraten der Tendenz, Stimuli korrekt zu identifizieren (z. B. Wörter werden als Wörter, Unwörter als Unwörter eingeschätzt), während negative Driftraten die Tendenz zu falschen Antworten widerspiegeln (z. B. Wörter werden als Unwörter, Unwörter als Wörter eingeschätzt). Höhere Driftraten sprechen somit für eine im Mittel schnellere und akkuratere Entscheidung. Zur Schätzung wurde die Software *fast-dm* (Voss und Voss 2007) verwendet. Lösungsraten sowie deskriptive Testwertstatistiken können Tab. 2 entnommen werden. Die Driftraten bilden die allgemeine Effizienz ab, mit der Buchstabenketten mit dem mentalen Lexikon abgeglichen werden bzw. mit der semantische Urteile gefällt werden.

### 3.2.3 Gedächtnisaktualisierung

Zur Erfassung der Fähigkeit zur Gedächtnisaktualisierung wurden 21 numerische Updating-Aufgaben (vgl. Oberauer und Kliegl 2006) verwendet. Wilhelm et al. (2013) zeigten, dass sich Updating-Aufgaben auf latenter Ebene – unabhängig von verbalem, numerischem oder figuralem Aufgabenmaterial – sehr gut eignen um Arbeitsgedächtnisprozessen zu repräsentieren. Das numerische Aufgabenmaterial wurde gewählt um auszuschließen, dass gefundene Effekte auf eine materialspezifische Komponente zurückgehen, wobei die Verarbeitung von numerischem und verbalen Aufgabenmaterial eine enge Beziehung aufweist (Oakhill et al. 2011).

Zur Aufgabenbearbeitung wurden die SuS gebeten, eine zwei- bis vierstellige Zahlenfolge zu memorieren und diese durch mentale Addition oder Subtraktion zu verändern (vgl. Abb. 2). Pro Ziffernstelle wurden je zwei Additions- oder Subtraktionsoperationen (–8 bis +8) vorgegeben. Für Aufgaben mit zwei Zifferstellen wurden somit vier, für drei Stellen sechs und für vier Stellen acht Operatoren präsentiert. Nach der Präsentation wurden die SuS aufgefordert, ihr Endergebnis per Keyboard einzugeben. Korrekturen waren möglich. Der Stimulus und die Operatoren wurden für jeweils 2500 ms angezeigt und durch Interstimulusintervalle von 500 ms unterbrochen. Alle Zwischen- und Endergebnisse enthielten pro Stelle nie mehrstellige oder negative Zahlenwerte. Die dichotomen Ergebnisdaten wurden Rasch-skaliert ( $N = 639$ ; mittlere Lösungsrate 0,37; 0,08–0,71 pro Item). Als Personenvariablen wurden EAP-Schätzer verwendet (EAP Reliabilität = 0,88). Höhere Werte bilden ein effizienteres Überwachen und Aktualisieren von Arbeitsgedächtnisinhalten ab.



**Abb. 2** Schematischer Ablauf einer Updating-Aufgabe mit zwei Stellen. Die dargebotenen Operatoren werden zu den Anfangsziffern an den entsprechenden Stellen hinzuaddiert bzw. subtrahiert. Pro Stelle werden jeweils zwei Operatoren mit einem Interstimulusintervall (ISI) vorgegeben

### 3.3 Stichprobe

Während zur Ermittlung der Testwerte alle verfügbaren Daten dienen, wurden für die Untersuchung der aufgestellten Forschungsfragen Teilstichproben der SuS betrachtet. Für die erste Forschungsfrage über die Vorhersagbarkeit des Leseverständnisses durch basale Leseprozesse auf Wort- und Satzebene sowie durch das Arbeitsgedächtnis wurde die Teilstichprobe herangezogen, die am Lese-Assessment von PISA 2012 und an der CBA Begleitstudie teilnahm. Somit wurden die Daten von 429 Jugendlichen (52,9 % männlich, 47,1 % weiblich) von 75 Schulen (nicht gymnasiale Schulformen 64,8 %, Gymnasium 35,2 %) analysiert, die zum Zeitpunkt der ersten Erhebung 15 Jahre alt ( $M = 15,83$ ,  $SD = 0,29$ ) waren.

Zur Untersuchung der zweiten Forschungsfrage über die Vorhersagbarkeit der Veränderung des Leseverständnisses wurden die Daten der SuS betrachtet, die an allen drei Messzeitpunkten teilnahmen. Insgesamt konnten somit die Daten von 53 Jugendlichen (58,5 % männlich, 41,5 % weiblich; Alter:  $M = 15,72$ ,  $SD = 0,27$ ) aus 32 Schulen (nicht gymnasiale Schulformen 50,9 %, Gymnasium 49,1 %) verwendet werden, wobei zu beachten ist, dass aufgrund struktureller Merkmale (z. B. Schulen ohne 10. Klasse; vgl. Heine et al. 2017) SuS von Gymnasien häufiger vertreten sind.

**Tab. 3** Deskriptive Statistiken und Interkorrelationen des Leseverständnisses zu 2012 und der kognitiven Basisfähigkeiten ( $N = 429$ )

		<i>M</i>	<i>SD</i>	1	2	3	4	5					
1	Lesen 2012	0,42	1,06	–	–	–	–	–	–	–	–	–	
2	WE Zeit	1,28	0,18	–0,18	***	–	–	–	–	–	–	–	
3	WE Drift	2,04	1,00	0,38	***	–0,28	***	–	–	–	–	–	
4	SI Zeit	2,37	0,56	–0,24	***	0,40	***	–0,28	***	–	–	–	
5	SI Drift	1,26	0,60	0,38	***	–0,11	*	0,31	***	–0,02	–	–	
6	GA	0,00	1,74	0,55	***	–0,17	***	0,30	***	–0,16	**	0,33	***

WE Worterkennung, SI semantische Integration, GA Gedächtnisaktualisierung

\*  $p < 0,05$ , \*\*  $p < 0,01$ , \*\*\*  $p < 0,001$

**Tab. 4** Deskriptive Statistiken und Interkorrelationen des Leseverständnisses zu 2012, 2013 und der kognitiven Basisfähigkeiten ( $N = 53$ )

		<i>M</i>	<i>SD</i>	1	2	3	4	5	6						
1	Lesen 2012	0,61	0,93	–	–	–	–	–	–	–	–	–			
2	Lesen 2013	0,95	0,94	0,97	***	–	–	–	–	–	–	–			
3	WE Zeit	1,30	0,20	–0,41	**	–0,38	**	–	–	–	–	–			
4	WE Drift	2,25	0,97	0,15	–	0,13	–	–0,51	***	–	–	–			
5	SI Zeit	2,39	0,54	–0,45	***	–0,42	**	0,66	***	–0,32	*	–			
6	SI Drift	1,35	0,53	0,31	*	0,31	*	–0,19	–	–0,03	–	–0,21	–		
7	GA	0,11	1,59	0,56	***	0,54	***	–0,14	–	0,19	–	–0,28	*	0,16	–

WE Worterkennung, SI semantische Integration, GA Gedächtnisaktualisierung

\*  $p < 0,05$ , \*\*  $p < 0,01$ , \*\*\*  $p < 0,001$

## 4 Ergebnisse

Zunächst wurden deskriptive Stichprobenkennwerte des Leseverständnisses zu beiden Messzeitpunkten und der kognitiven Basisfähigkeiten bestimmt. In Tab. 3 und 4 sind die Mittelwerte und Standardabweichungen der Testwerte sowie ihre Interkorrelationen dargestellt. Tab. 3 enthält die deskriptiven Kennwerte der Teilstichprobe, die im Rahmen der ersten Fragestellung betrachtet wurde; Tab. 4 die zur Untersuchung der zweiten Fragestellung. Vergleicht man die deskriptiven Kennwerte beider Teilstichproben fällt auf, dass die Mittelwerte im Lesen und in den kognitiven Basisfähigkeiten in der zweiten Teilstichprobe tendenziell höher bzw. die zugehörigen Standardabweichungen niedriger ausfallen. Ein Vergleich mit dem bundesweiten Durchschnitt im Lesen in 2012 (in Kompetenzpunkten:  $M = 508$ ,  $SD = 90$ ; vgl. Prenzel et al. 2013) deutete ebenso darauf hin, dass sich die Leseleistung der SuS in

**Tab. 5** Lineare Regression des Leseverständnisses zu 2012 sowie der Differenzvariable auf die kognitiven Basisfähigkeiten

Prädiktor	Lesen 2012 ( $N = 429$ )		Differenzwerte ( $N = 53$ )		
	$\beta$ ( $SE$ )	$p$	$\beta$ ( $SE$ )	$p$	
WE Zeit	0,01 (0,04)	0,369	–	–0,03 (0,20)	0,449
WE Drift	0,17 (0,04)	<0,001	***	–0,06 (0,16)	0,347
SI Zeit	–0,13 (0,04)	0,002	**	0,13 (0,19)	0,250
SI Drift	0,18 (0,04)	<0,001	***	0,04 (0,14)	0,284
GA	0,42 (0,04)	<0,001	***	–0,04 (0,15)	0,406
$R^2$	0,39			0,02	

WE Worterkennung, SI semantische Integration, GA Gedächtnisaktualisierung

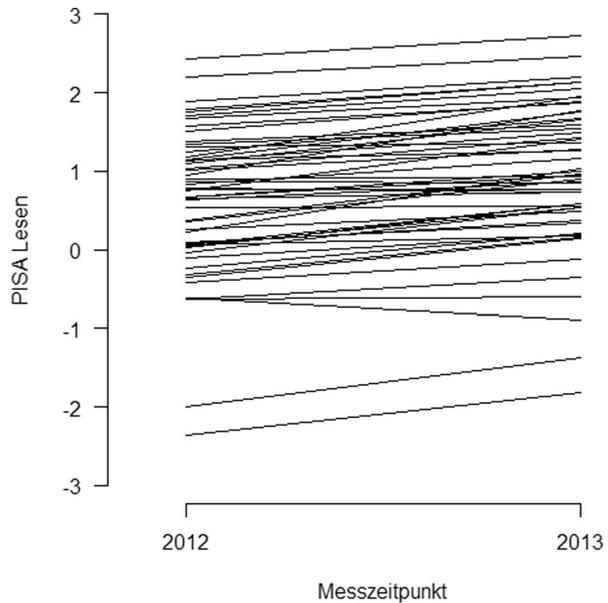
\*  $p < 0,05$ , \*\*  $p < 0,01$ , \*\*\*  $p < 0,001$

der ersten Teilstichprobe nicht erheblich vom Bundesdurchschnitt unterschied ( $M = 506$ ,  $SD = 85$ ); die Leistung in der zweiten Stichprobe fiel aber tendenziell höher aus ( $M = 513$ ,  $SD = 79$ ).

Zur Beantwortung der ersten Fragestellung über die Vorhersagbarkeit der Lesekompetenz von Jugendlichen durch die kognitiven Basisfähigkeiten (H1.1–H1.3) wurde ein lineares Regressionsmodell mit der Lesekompetenz von 2012 als abhängige Variable geprüft. Die Regression auf die fünf Indikatoren kognitiver Basisfähigkeiten ergab, dass alle Indikatoren bis auf die Zeit, mit der Wörter korrekt erkannt werden, prädiktiv waren ( $R^2 = 0,39$ ; rechtsseitige Signifikanztests; Spalte Lesen 2012 in Tab. 5). Das heißt, dass das Leseverständnis der SuS positiv damit assoziiert war, wie effizient Buchstabenketten mit mentalen Wortrepräsentationen verglichen, semantische Urteile gefällt sowie Gedächtnisinhalte überwacht und verändert werden konnten. Die Zeit zur Identifikation wahrer Sätze stand negativ mit dem Leseverständnis in Verbindung: Texte wurde umso besser verstanden, je mehr Zeit, die SuS zur Verifikation wahrer Aussagen benötigten.

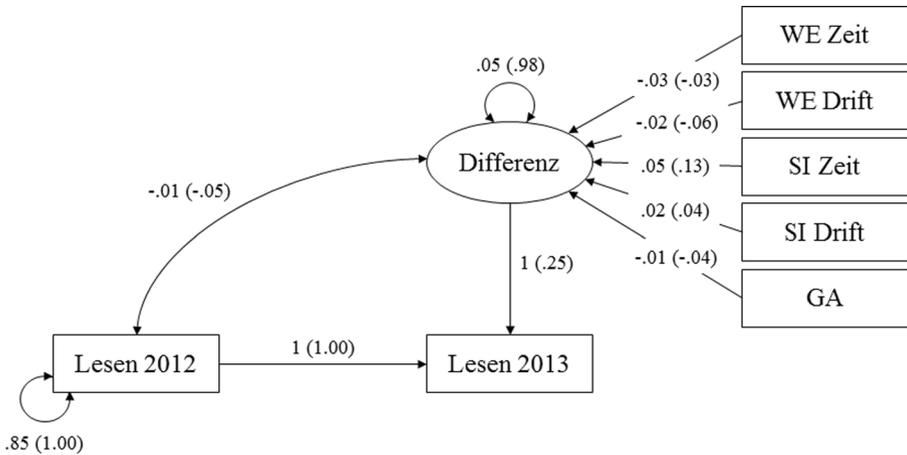
Um die zweite Fragestellung über die Vorhersagbarkeit der Leistungsveränderung im Lesen nach einem Jahr zu überprüfen, wurden zunächst die Leseleistungen zu 2012 und 2013 sowie ihre Differenz betrachtet. Das Leseverständnis in 2012 korrelierte sehr stark mit dem Leseverständnis in 2013 (Tab. 4). Abb. 3 zeigt, dass die Entwicklung des Leseverständnisses individuell unterschiedlich verlief: Die meisten Jugendlichen erfuhren einen Zuwachs, während die Leistung anderer stabil blieb oder sogar abnahm. Die Rangfolge der SuS blieb augenscheinlich erhalten. Damit spricht die hohe Autokorrelation dafür, dass das Leseverständnis nach Herauspartialisierung von Effekten des Leseclusters, der Position im Testablauf und der Schulzugehörigkeit über die Zeit hinweg sehr stabil blieb (d. h. gute Leser schnitten im Vergleich zu anderen Lesern weiterhin gut ab bzw. schwache Leser blieben im Vergleich schwach). Aus psychometrischer Sicht kann dieser Befund zudem als Hinweis auf eine hohe Messstabilität des PISA Lesetests gewertet werden. Die Differenz aus beiden Leseleistungen zeigte einen mittleren Leistungszuwachs von 0,33 ( $SD = 0,23$ ), was einem Effekt von 0,36 im Zuwachs des Leseverständnisses entspricht (vgl. Bloom et al. 2008). Ferner deutet die Richtung des Zusammenhangs zwischen der Lesekompetenz 2012 und den Differenzwerten zwar darauf hin, dass schwächere Leser einen stärkeren Leistungszuwachs zeigten; diese wurde statistisch allerdings

**Abb. 3** Veränderung der geschätzten Lesefähigkeit von 2012 zu 2013 ( $N = 55$ )



nicht signifikant,  $r = -0,11$ ,  $p = 0,419$ . Mit den kognitiven Basisfähigkeiten wiesen beide Messungen des Leseverständnisses ähnliche Zusammenhänge auf (Tab. 4).

Zur inferenzstatistischen Überprüfung der Veränderung der Lesekompetenz wurde ein Differenzwertmodell mit manifesten Variablen spezifiziert. Dazu wurde das in 2013 gemessene Leseverständnis auf das Leseverständnis in 2012 mit einem auf 1 fixierten Regressionsgewicht zurückgeführt. Die Leistungsveränderung wurde als Differenz zum Leseverständnis in 2013 definiert und technisch als latente Differenz modelliert, wobei die Korrelation des Leseverständnisses von 2012 mit der Differenzvariablen zugelassen wurde. Zur Identifikation der Differenzvariablen wurde die Faktorladung auf 1 und die Residualvarianz des Leseverständnisses in 2013 auf 0 fixiert. Im Vergleich zu 2013 verbesserte sich die Leseleistung der Jugendlichen im Mittel signifikant ( $M_{\text{latente Differenz}} = 0,33$ ;  $p < 0,001$ ), wobei sich die Leistungszuwächse interindividuell signifikant unterschieden ( $SD_{\text{latente Differenz}} = 0,05$ ,  $p < 0,001$ ). Zur Überprüfung, ob interindividuelle Unterschiede im Leistungszuwachs durch die kognitiven Basisfähigkeiten erklärt werden können (H2.2–H2.4), wurden die Differenzwerte auf die kognitiven Basisfähigkeiten zurückgeführt. Abb. 4 zeigt die Erweiterung des spezifizierten Pfadmodells. Die kognitiven Basisfähigkeiten leisten keinen bedeutsamen Beitrag zur Erklärung der individuellen Leistungszuwächse ( $R^2 = 0,02$ ; rechtsseitige Signifikanztests; Spalte Differenzwerte in Tab. 2). Aufgrund der geringen Stichprobengröße wurden zusätzlich Monte Carlo Poweranalysen mit Mplus angestellt ( $B = 1000$  Replikationen), die auf eine geringe Teststärke hinwiesen. Die einzelnen Regressionsparameter konnten nur mit einer geringen Wahrscheinlichkeit (0,08–0,13) über die Replikationen hinweg wiedergefunden werden.



**Abb. 4** Ergebnisse des Differenzwertmodells mit Regression der Differenzwerte auf die kognitiven Basisfähigkeiten. Korrelationen der Prädiktoren untereinander wurden im Modell zugelassen. Unstandardisierte Regressionskoeffizienten; dahinter *in Klammern* die standardisierten Koeffizienten. (*WE* Worterkennung, *SI* semantische Integration, *GA* Gedächtnisaktualisierung)

## 5 Diskussion

Die vorliegende Arbeit befasst sich mit dem Leseverständnis 15-jähriger Schülerinnen und Schüler (SuS) und dessen Zusammenhang mit basalen Lesefähigkeiten auf Wort- und Satzebene und dem Arbeitsgedächtnis. Dabei wurde den Forschungsfragen nachgegangen, ob sich das Leseverständnis Jugendlicher durch diese Basisfähigkeiten noch vorhersagen lässt und welche Rolle sie in der Leseentwicklung innerhalb eines Jahres einnehmen. Die zur Beantwortung der ersten Forschungsfrage aufgestellten Hypothesen konnten insgesamt bestätigt werden. Das Leseverständnis wurde auf Wortebene zwar nicht durch die Zeit, die zum Erlesen bekannter Wörter benötigt wird, aber durch die Effizienz vorhergesagt, mit der Personen Buchstabenketten mit ihrem mentalen Lexikon abgleichen können (H1.1). Daneben waren beide Indikatoren der Satzebene prädiktiv für das Leseverständnis: Sowohl die Zeit, mit der semantisch wahre Aussagen erkannt werden, als auch die Effizienz, mit der SuS semantische Entscheidungen treffen, waren prädiktiv (H1.2). Zuletzt sagte auch die Fähigkeit im Überwachen und Aktualisieren von Arbeitsgedächtnisinhalten die Leseverständnisleistung vorher (H1.3). Die Hypothesen der zweiten Forschungsfrage konnten hingegen nur teilweise bestätigt werden. Die erwartete Verbesserung im Leseverständnis nach einem Jahr wurde in den Daten gefunden (H2.1). Die Leseleistung legte um 0,36 Standardabweichungen im Mittel zu, wobei sich die Varianz zwischen den SuS signifikant auffächerte. Der gefundene Zuwachs wurde allerdings weder durch Fähigkeiten zur Worterkennung (H2.2), semantischen Integration (H2.3) noch durch das Arbeitsgedächtnis vorhergesagt (H2.4). Zusammengefasst zeigen die Ergebnisse, dass selbst im Jugendalter kognitive Basisfähigkeiten des Lesens und des Arbeitsgedächtnisses stark mit dem Verständnis in Verbindung stehen, das während des Lesens konstruiert wird. Die Entwicklung der Lesekompetenz nach

einem Schuljahr wird durch diese Basisfähigkeiten jedoch nicht beeinflusst, was im Folgenden näher beleuchtet wird.

### 5.1 Vorhersagbarkeit des Leseverständnisses 15-Jähriger

Auf der Wortebene entsprechen die Ergebnisse den zugrundeliegenden Annahmen über die lexikalische Qualität mentaler Wortrepräsentationen (Perfetti 2007) und die Automatisierbarkeit von Prozessen des Lesens (Samuels und Flor 1997; Walczyk 2000). Im Sinne der Annahmen deutet die korrekte und schnelle Identifikation von Wörtern auf den automatisierten Abruf hochwertiger Wortrepräsentationen hin, die das Leseverstehen konstituieren (Perfetti und Stafura 2014). Empirisch zeigten die SuS auch ein besseres Leseverständnis, wenn sie Buchstabenketten nicht nur korrekt, sondern auch schnell als Wörter oder Unwörter identifizieren konnten. Die Zeit, mit der Wörter korrekt als solche bestimmt wurden, war hingegen nicht prädiktiv für das Leseverständnis. Möglicherweise ist die Zeit zur Erkennung eines Wortes für das Niveau, auf dem in der 9. Klasse gelesen wird, nicht mehr bedingend und andere Leseteilprozesse treten verstärkt in den Vordergrund (vgl. Bast und Reitsma 1997, 1998). Technisch gesehen kann die fehlende Bedeutsamkeit auch auf eine lineare Beziehung zwischen den Regressionsprädiktoren zurückgehen. Die Reaktionszeit auf Wörter war moderat mit der Reaktionszeit auf semantisch wahre Aussagen korreliert, die das Leseverständnis signifikant vorhersagte. Es kann vermutet werden, dass beide Zeitmaße eine allgemeine Verarbeitungsgeschwindigkeit von geschriebenem Material abbilden, wodurch ein Vorhersagebeitrag im Leseverständnis nur durch ein Maß beobachtbar wäre, da geteilte Varianzanteile in einer multiplen Regression entfallen.

Bezüglich der Lesefähigkeitsindikatoren auf der Satzebene war neben der Reaktionszeit auf wahre Aussagen auch die Driftrate semantischer Integration für das Leseverständnis prädiktiv. Die SuS zeigten ein besseres Leseverständnis, wenn sie in der Lage waren, Satzkontexte in effizienter Weise semantisch zu beurteilen. Dass semantische Integration auch unabhängig von Prozessen der Worterkennung das Leseverständnis vorhersagte, stimmt mit theoretischen Überlegungen zur Wort-Text-Integration überein (Perfetti und Stafura 2014). Worterkennung ist ein bedeutender Teilaspekt semantischer Integration, da die Bedeutungen erkannter Wörter die Grundlage zur Integration eines kohärenten Satzsinns bilden (Richter et al. 2012; Richter und Naumann 2009). Stärker aber als auf der Wortebene wird das Hintergrundwissen einer Person zur semantischen Integration einbezogen (Perfetti und Stafura 2014). Interessanterweise stand die Reaktionszeit auf semantisch wahre Aussagen sogar stärker mit der Driftrate der Worterkennung als der zur semantischen Integration in Verbindung. Da hochwertige mentale Wortrepräsentationen fest mit ihren Wortbedeutungen assoziiert sein sollten (Perfetti 2007), würden sie bereits einen bestimmten semantischen Kontext aktivieren, wodurch Leser schneller in der Lage wären eine Aussage zu verifizieren. Wenn aber semantisch falsche Aussagen beurteilt werden müssen, wäre eine systematischere Überprüfung des individuellen Wissens notwendig.

Die Ergebnislage schließt keineswegs aus, dass andere Leseteilprozesse – wie beispielsweise die Kohärenzbildung zwischen Sätzen, Paragraphen oder ganzen Texten

– eine bedeutsame Rolle bei der Konstruktion eines mentalen Verstehensmodells spielen (z. B. Kintsch 1998; Richter et al. 2012). Sie verdeutlichen vielmehr, dass Defizite im Leseverständnis jugendlicher Leser mit Defiziten in basalen Lesefähigkeiten und mangelnder Routine in Verbindung stehen (Bråten et al. 2013; Farnia und Geva 2013; Schroeder 2011), die bis ins Erwachsenenalter weiterbestehen können (Rammstedt et al. 2013). Perfetti (2007) führt aus, dass mentale Wortrepräsentationen durch häufige Wortexposition innerhalb verschiedener semantischer Kontexte stabilisiert werden können. Zukünftige Forschung könnte hier ansetzen und prüfen, ob Zusammenhänge zwischen Wortidentifikations- und Integrationsprozessen und dem Leseverständnis durch Lesegewohnheiten oder bestimmte Lernstrategien (z. B. das wiederholte Lesen von Texten) vermittelt werden.

Hervorzuheben ist auch, dass das Arbeitsgedächtnis einen von den basalen Lesefähigkeiten unabhängigen Erklärungsbeitrag am Leseverständnis leistete. Das Leseverständnis der SuS wurde dabei in ähnlicher Weise wie in Studien mit anderen Altersgruppen vorhergesagt (vgl. Cain et al. 2004; Dehn 2008; Hannon 2012; Oakhill et al. 2011). Das kontinuierliche Überwachen und Aktualisieren von Arbeitsgedächtnisinhalten stellte folglich eine wichtige Ressource beim Lesen dar und steht im Einklang mit theoretischen Vorstellungen des Arbeitsgedächtnisses als kognitives System, in dem Textinformationen behalten und gezielt verarbeitet werden (Feldman Barrett et al. 2004; Engle und Conway 1998; Oberauer und Kliegl 2010). Für zukünftige Forschung stellt sich die Frage, ob leseschwache Schülerinnen und Schülern mit eingeschränkten Fähigkeiten zur Gedächtnisaktualisierung durch den Einsatz kognitiver und metakognitiver Lesestrategien stärker profitieren würden als vergleichsweise Leser mit hoch ausgeprägten Gedächtnisfähigkeiten.

## 5.2 Entwicklung des Leseverständnisses 15-Jähriger

Bezogen auf die untersuchte Leistungsveränderung fand sich die vermutete Entwicklung im Leseverständnis der SuS innerhalb eines Jahres. Im Vergleich zu den Befunden von Bloom et al. (2008), die einen mittleren Zuwachs von 0,19 unter Verwendung mehrerer Lesetests fanden, nahm das Leseverständnis mit einem Effekt von 0,36 zwar verhältnismäßig stark zu, stellt in Referenz zur Leseentwicklung in anderen Klassenstufen wie beispielsweise in der Grundschule immer noch eine geringe Veränderung dar (vgl. Farnia und Geva 2013; Tilstra et al. 2009). Bemerkenswert ist auch die hohe Autokorrelation des Leseverständnisses zwischen der 9. und 10. Klasse, aus der Effekte der Lesecluster, ihrer Positionen und der Schulform herauspartialisiert wurden. Lesefähigkeiten weisen im Allgemeinen eine hohe Merkmalsstabilität auf (Boland 1993; Klicpera und Schabmann 1993). Zusammen mit der signifikant größeren Varianz im Leseverständnis der SuS der 10. Klasse deutet sich damit ein Matthäus-Effekt an, d. h., gute Leser erfahren einen höheren Zuwachs in ihrem Leseverständnis als schwächere Leser. Bast und Reitsma (1997, 1998) fanden zwar keinen Matthäus-Effekt für das Leseverständnis, schlossen für ältere SuS aber entwicklungsbedingte Veränderungen im Beziehungsmuster von Lesefähigkeiten mit anderen kognitiven Variablen über die Zeit und somit eine spätere Auffächerung der Variabilität im Lesen nicht aus.

Unerwarteter Weise wurde der Leistungszuwachs im Leseverständnis durch keine der kognitiven Basisfähigkeiten erklärt. Die Entwicklung im Lesen geht demnach nicht auf individuelle kognitive Basisfähigkeiten zurück, sondern erfolgt durch andere Einflüsse – beispielsweise durch Unterricht oder den Einsatz von Lesestrategien. Die Belastbarkeit dieser Ergebnisse und ihre Interpretation sind jedoch stark eingeschränkt, was im nächsten Abschnitt näher erörtert wird.

### 5.3 Limitationen

Einschränkend müssen für die angestellte Studie einige Restriktionen festgehalten werden. Diese Restriktionen betreffen (1) die Stichprobenzusammensetzung der zweiten Teilstichprobe und eine damit einhergehende Restriktion des Fähigkeitspektrums, (2) ihre geringe Stichprobengröße und resultierende Einschränkungen der Teststärke des geschätzten Differenzwertmodells und (3) die Schätzung der Driftraten zur Abbildung basaler Lesefähigkeiten.

Erstens, im Gegensatz zur ersten zeichnete sich die zweite Teilstichprobe durch einen höheren Anteil an Gymnasiasten aus. Dieses Oversampling ist strukturellen Merkmalen von Schule – beispielsweise durch den Wegfall der 10. Klasse in Hauptschulen – geschuldet. Der Dropout wurde in der Leseleistung sichtbar, die im Gegensatz zur Leseleistung der ersten Teilstichprobe im Mittel erhöht war und eine geringere Variabilität aufwies. Es ist daher davon auszugehen, dass in der zweiten Teilstichprobe nicht das gesamte Fähigkeitspektrum der zu 2012 untersuchten 15-jährigen SuS abgedeckt wurde. Gute und schwache Leser weisen unterschiedliche Entwicklungspotentiale in der Entwicklung ihres Leseverständnisses auf (vgl. Bast und Reitsma 1997, 1998). Generalisierte Aussagen über Zusammenhänge des Leseverständnisses mit basalen kognitiven Fähigkeiten könnten daher verzerrt sein und nur für leistungsstarke SuS gelten. Hinzu kommt, dass die Reliabilität der Lesetestwerte, die als Varianzanteil der geschätzten Testwerte an der latenten Fähigkeitsvarianz operationalisiert wurde, insbesondere für SuS des Gymnasiums gering ausfiel. Das bedeutet, dass Testwertschätzungen innerhalb der Gruppe der Gymnasialschülerinnen und -schüler nur mit einer größeren Unsicherheit bestimmt werden können. Die niedrigere Reliabilität ist aber insofern plausibel, als das der PISA Lesetest für die Erfassung eines breiten Fähigkeitspektrums konzipiert wurde (OECD 2013a), wodurch die Fähigkeitsmessungen innerhalb der Randbereiche des Spektrums weniger genau differenzieren. Für künftige Forschungsarbeiten besteht folglich die Notwendigkeit, unterschiedliche Leistungsniveaus von SuS gezielt zu betrachten, die mit unterschiedlichen entwicklungsbedingten Anforderungen assoziiert sein könnten (vgl. Bast und Reitsma 1998).

Zweitens, die zweite Teilstichprobe kann zwar als eine Zufallsstichprobe der PISA 2012 Hauptstichprobe angesehen werden, jedoch nahmen aufgrund des Bookletdesigns und des Stichprobenausfalls nur wenige SuS zu allen drei Messzeitpunkten teil. Die angestellten statistischen Analysen beruhen somit auf einer geringen Probandenanzahl. Es wurde zwar auf den Einsatz latenter, messfehlerbereinigter Veränderungsmodelle, die eine hohe Fallzahl benötigen (vgl. MacCallum et al. 1996), verzichtet. Dennoch zeigen die durchgeführten Monte Carlo Poweranalysen für die im Differenzwertmodell spezifizierten Regressionskoeffizienten nur eine geringe

Teststärke auf. Dadurch wurden potentiell bestehende Zusammenhänge der kognitiven Basisfähigkeiten mit der Entwicklung des Leseverständnisses möglicherweise nicht aufgedeckt. Die Belastbarkeit der Ergebnisse aus dem Differenzwertemodell ist zwar somit eingeschränkt, dennoch sind die Ergebnisse hinsichtlich der Stabilität des Leseverständnisses und des Zusammenhangsmusters mit den kognitiven Basisfähigkeiten über die Zeit aufschlussreich.

Drittens, der Entscheidungsprozess, der durch Driftraten abgebildet wird, kann für verschiedene Aufgabenmerkmale variieren. Es kann daher sinnvoll sein, Driftraten – wie zur Abbildung basaler Lesefähigkeiten auf Wort- und Satzebene – nach verschiedenen Aufgabenmerkmalen zu differenzieren und separat zu schätzen (vgl. Ratcliff et al. 2004; Schroeder 2011). Für Worterkennung können solche Merkmale die Unterscheidung Wort vs. Unwort oder auch die Silbenanzahl betreffen; für semantische Integration der Wahrheitsgehalt einer Aussage oder die Anzahl der Propositionen in einem Satz. Um stabile Schätzer zu erhalten wird eine hinreichend große Aufgabenanzahl pro Merkmal benötigt. Für zukünftige Studien wäre es daher wünschenswert, Driftraten differenzierter zu betrachten um somit stärker differenzierte diagnostische Informationen zu erhalten.

Insgesamt kann aus der vorliegenden Arbeit festgehalten werden, dass kognitive Basisfähigkeiten des Lesens und des Arbeitsgedächtnisses mit dem Leseverständnis auch noch im Jugendalter zusammenhängen. Das Leseverständnis der untersuchten Jugendlichen entwickelte sich in diesem Zeitraum jedoch unabhängig von der Ausprägung der kognitiven Basisfähigkeiten. Generalisierte Aussagen über die individuelle Veränderung des Leseverständnisses im Verlauf eines Schuljahres sind aufgrund der Stichprobengröße und Zusammensetzung nicht möglich. Dennoch wird deutlich, dass ineffiziente kognitive Basisfähigkeiten symptomatisch für Defizite im Leseverständnis sind und auch noch im Jugendalter Entwicklungspotential aufweisen.

**Danksagung** Wir danken zwei anonymen Gutachtern für ihre Anmerkungen und Einblicke.

**Förderung** Die Studie wurde durch das Zentrum für internationale Vergleichsstudien (ZIB) unterstützt und finanziert.

## Literatur

- Arrington, C. N., Kulesz, P. A., Francis, D. J., Fletcher, J. M., & Barnes, M. A. (2014). The contribution of attentional control and working memory to reading comprehension and decoding. *Scientific Studies of Reading, 18*, 325–346.
- Artelt, C., Schiefele, U., & Schneider, W. (2001). Predictors of reading literacy. *European Journal of Psychology of Education, 16*(3), 363–383.
- Bast, J., & Reitsma, P. (1997). Matthew Effects in reading: A comparison of latent growth curve models and simplex models with structured means. *Multivariate Behavioral Research, 32*, 135–167.
- Bast, J., & Reitsma, P. (1998). Analyzing the development of individual differences in terms of Matthew Effects in reading: results from a Dutch longitudinal study. *Developmental Psychology, 34*, 1373–1399.
- Bloom, H. S., Hill, C. J., Black, A. R., & Lipsey, M. W. (2008). Performance trajectories and performance gaps as achievement effect-size benchmarks for educational interventions. *Journal of Research on Educational Effectiveness, 1*(4), 289–328.
- Boland, T. (1993). The importance of being literate: reading development in primary school and its consequences for the school career in secondary education. *European Journal of Psychology of Education, 8*(3), 289–305.

- Bråten, I., Ferguson, L. E., Anmarkrud, Ø., & Strømsø, H. I. (2013). Prediction of learning and comprehension when adolescents read multiple texts: the roles of word-level processing, strategic approach, and reading motivation. *Reading and Writing, 26*(3), 321–348.
- Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology, 96*(1), 31–42.
- Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of verbal learning and verbal behavior, 19*, 450–466.
- Dehn, M. J. (2008). *Working memory and academic learning: assessment and intervention*. Hoboken NJ: John Wiley & Sons Inc..
- Engle, R. W., & Conway, A. R. A. (1998). Working memory and comprehension. In R. H. Logie & K. J. Gilhooly (Hrsg.), *Working memory and thinking* (S. 67–91). Hove: Psychology Press.
- Farnia, F., & Geva, E. (2013). Growth and predictors of change in English language learners' reading comprehension. *Journal of Research in Reading, 36*(4), 389–421.
- Feldman Barrett, L., Tugade, M. M., & Engle, R. W. (2004). Individual differences in working memory capacity and dual-process theories of the mind. *Psychological Bulletin, 130*(4), 553–573.
- van Gelderen, A., Schoonen, R., Stoel, R. D., de Glopper, K., & Hulstijn, J. (2007). Development of adolescent reading comprehension in language 1 and language 2: a longitudinal analysis of constituent components. *Journal of Educational Psychology, 99*(3), 477–491.
- Goff, D. A., Pratt, C., & Ong, B. (2005). The relations between children's reading comprehension, working memory, language skills and components of reading decoding in a normal sample. *Reading and Writing, 18*, 583–616.
- Groeben, N. (1982). *Leserpsychologie: Textverständnis – Textverständlichkeit*. Münster: Aschendorff.
- Hahnel, C., Goldhammer, F., Naumann, J., & Kröhne, U. (2016). Effects of linear reading, basic computer skills, evaluating online information, and navigation on reading digital text. *Computers in Human Behavior, 55*, 486–500.
- Hannon, B. (2012). Understanding the relative contributions of lower-level word processes, higher-level processes, and working memory to reading comprehension performance in proficient adult readers. *Reading Research Quarterly, 47*(2), 125–152.
- Heine, J.-H., Sälzer, C., Borchert, L., Sibberns, H., & Mang, J. (2013). Technische Grundlagen des fünften internationalen Vergleichs. In M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *PISA 2012. Fortschritte und Herausforderungen in Deutschland* (S. 309–343). Münster: Waxmann.
- Heine, J. H., Nagy, G., Meinck, S., Zühlke, O., Mang, J. (2017). Empirische Grundlage, Stichprobenausfall und Adjustierung im PISA-Längsschnitt 2012-2013 *Zeitschrift für Erziehungswissenschaft*. doi:10.1007/s11618-017-0756-0.
- Kendeou, P., Papadopoulos, T. C., & Spanoudis, G. (2012). Processing demands of reading comprehension tests in young readers. *Learning and Instruction, 22*(5), 354–367.
- Kendeou, P., van den Broek, P., Helder, A., & Karlsson, J. (2014). A cognitive view of reading comprehension: implications for reading difficulties. *Learning Disabilities Research & Practice, 29*(1), 10–16.
- Kintsch, W. (1998). *Comprehension: a paradigm for cognition*. Cambridge: Cambridge University Press.
- Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological review, 85*(5), 363–394.
- Klicpera, C., & Schabmann, A. (1993). Do German-speaking children have a chance to overcome reading and spelling difficulties? A longitudinal survey from the second until the eighth grade. *European Journal of Psychology of Education, 8*(3), 307–323.
- Lesaux, N. K., Rupp, A. A., & Siegel, L. S. (2007). Growth in reading skills of children from diverse linguistic backgrounds: Findings from a 5-year longitudinal study. *Journal of Educational Psychology, 99*(4), 821–834.
- MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological methods, 1*(2), 130.
- Muthén, L. K., & Muthén, B. O. (1998–2012). *Mplus User's Guide* (7. Aufl.). Los Angeles CA: Muthén & Muthén.
- Nagy, G., Lüdtke, O., Köller, O., Heine, J. H. (2017a). IRT-Skalierung der Tests im PISA-Längsschnitt 2012/2013: Auswirkungen von Testkontexteffekten auf die Zuwachsschätzung. *Zeitschrift für Erziehungswissenschaft*. doi:10.1007/s11618-017-0749-z.
- Nagy, G., Retelsdorf, J., Goldhammer, F., Schiepe-Tiska, A., Lüdtke, O. (2017b). Veränderungen der Lesekompetenz von der 9. zur 10. Klasse: Differenzielle Entwicklungen in Abhängigkeit der Schulform, des Geschlechts und des soziodemografischen Hintergrunds? *Zeitschrift für Erziehungswissenschaft*. doi:10.1007/s11618-017-0747-1.

- Naumann, J., Artelt, C., Schneider, W., & Stanat, P. (2010). Lesekompetenz von PISA 2000 bis PISA 2009. In E. Klieme, C. Artelt, J. Hartig, N. Jude, O. Köller, M. Prenzel, W. Schneider & P. Stanat (Hrsg.), *PISA 2009. Bilanz nach einem Jahrzehnt* (S. 23–71). Münster: Waxmann.
- Oakhill, J., Yuill, N., & Garnham, A. (2011). The differential relations between verbal, numerical and spatial working memory abilities and children's reading comprehension. *International Electronic Journal of Elementary Education*, 4(1), 83–106.
- Oberauer, K., & Kliegl, R. (2006). A formal model of capacity limits in working memory. *Journal of Memory and Language*, 55(4), 601–626.
- Oberauer, K., & Kliegl, R. (2010). Interferenz im Arbeitsgedächtnis: Ein formales Modell. *Psychologische Rundschau*, 61(1), 33–42.
- Organisation for Economic Co-operation and Development (OECD) (2013a). *PISA 2012 assessment and analytical framework*. Paris: OECD Publishing.
- Organisation for Economic Co-operation and Development (OECD) (2013b). *OECD skills outlook 2013: first results from the survey of adult skills*. Paris: OECD Publishing.
- Pachella, R. G. & Pew, R. W. (1968). Speed-Accuracy Tradeoff in Reaction Time: Effect of Discrete Criterion Times. *Journal of Experimental Psychology* 76(1), 19–24.
- Parrila, R., Aunola, K., Leskinen, E., Nurmi, J.-E., & Kirby, J. R. (2005). Development of individual differences in reading: results from longitudinal studies in English and Finnish. *Journal of Educational Psychology*, 97(3), 299–319.
- Perfetti, C. (2007). Reading ability: lexical quality to comprehension. *Scientific Studies of Reading*, 11(4), 357–383.
- Perfetti, C., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading*, 18(1), 22–37.
- Prenzel, M., Baumert, J., Blum, W., Lehmann, R. H., Leutner, D., Neubrand, M. et al. (Hrsg.). (2006). *PISA 2003: Untersuchungen zur Kompetenzentwicklung im Verlauf eines Schuljahres*. Münster: Waxmann.
- Prenzel, M., Sälzer, C., Klieme, E., & Köller, O. (2013). *PISA 2012. Fortschritte und Herausforderungen in Deutschland*. Münster: Waxmann.
- Rammstedt, B., Ackermann, D., & GESIS – Leibniz-Institut für Sozialwissenschaften (2013). *Grundlegende Kompetenzen Erwachsener im internationalen Vergleich: Ergebnisse von PIAAC 2012*. Münster: Waxmann.
- Ratcliff, R., Gomez, P., & McKoon, G. (2004). A diffusion model account of the lexical decision task. *Psychological Review*, 111, 159–182.
- Richter, T., & Naumann, J. (2009). Was misst der ELVES-Subtest Satzverifikation? Analysen von Mess- und Itemeigenschaften mit hierarchisch-linearen Modellen. In W. Lenhard & W. Schneider (Hrsg.), *Diagnostik und Förderung des Leseverständnisses* (Bd. 7, S. 131–149). Göttingen: Hogrefe.
- Richter, T., Isberner, M.-B., Naumann, J., & Kutzner, Y. (2012). Prozessbezogene Diagnostik von Lesefähigkeiten bei Grundschulkindern. *Zeitschrift für Pädagogische Psychologie*, 26(4), 313–331.
- Richter, T., & van Holt, N. (2005). ELVES: Ein computergestütztes Diagnostikum zur Erfassung der Effizienz von Teilprozessen des Leseverstehens. *Diagnostica*, 51(4), 169–182.
- Roeschl-Heils, A., Schneider, W., & van Kraayenoord, C. E. (2003). Reading, metacognition and motivation: A follow-up study of German students in grades 7 and 8. *European Journal of Psychology of Education*, 18(1), 75–86.
- Rost, D. H., & Schilling, S. R. (2006). Leseverständnis. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (S. 450–460). Weinheim: Beltz.
- Samuels, S. J., & Flor, R. F. (1997). The importance of automaticity for developing expertise in reading. *Reading & Writing Quarterly*, 13, 107–121.
- Schroeder, S. (2011). What readers have and do: effects of students' verbal ability and reading time components on comprehension with and without text availability. *Journal of Educational Psychology*, 103(4), 877–896.
- Seigneuric, A., & Ehrlich, M.-F. (2005). Contribution of working memory capacity to children's reading comprehension: a longitudinal investigation. *Reading and Writing*, 18, 617–656.
- Swanson, H. L., & Berninger, V. (1995). The role of working memory in skilled and less skilled readers' comprehension. *Intelligence*, 21(1), 83–108.
- Tilstra, J., McMaster, K., Van den Broek, P., & Kendeou, P. (2009). Simple but complex: components of the simple view of reading across grade levels. *Journal of Research in Reading*, 32(4), 383–401.
- Voss, A., Rothermund, K., & Voss, J. (2004). Interpreting the parameters of the diffusion model: an empirical validation. *Memory & Cognition*, 32(7), 1206–1220.
- Voss, A., & Voss, J. (2007). Fast-dm: A free program for efficient diffusion model analysis. *Behavior Research Methods*, 39(4), 767–775.

- Walczyk, J. J. (2000). The interplay between automatic and control processes in reading. *Reading Research Quarterly*, 35(4), 554–566.
- Whelan, R. (2008). Effective analysis of reaction time data. *The Psychological Record*, 58(3), 475.
- Wilhelm, O., Hildebrandt, A., & Oberauer, K. (2013). What is working memory capacity, and how can we measure it? *Frontiers in Psychology*, 4. doi:[10.3389/fpsyg.2013.00433](https://doi.org/10.3389/fpsyg.2013.00433).