

Diehl, Kirsten; Hartke, Bodo  
**Zur Reliabilität und Validität des formativen Bewertungssystems IEL-1.  
Inventar zur Erfassung der Lesekompetenz von Erstklässlern**  
*Empirische Sonderpädagogik 3 (2011) 2, S. 121-146*



Quellenangabe/ Reference:

Diehl, Kirsten; Hartke, Bodo: Zur Reliabilität und Validität des formativen Bewertungssystems IEL-1. Inventar zur Erfassung der Lesekompetenz von Erstklässlern - In: Empirische Sonderpädagogik 3 (2011) 2, S. 121-146 - URN: urn:nbn:de:0111-opus-93201 - DOI: 10.25656/01:9320

<https://nbn-resolving.org/urn:nbn:de:0111-opus-93201>

<https://doi.org/10.25656/01:9320>

in Kooperation mit / in cooperation with:

Pabst Science Publishers <https://www.psychologie-aktuell.com/journale/empirische-sonderpaedagogik.html>

#### Nutzungsbedingungen

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

#### Terms of use

We grant a non-exclusive, non-transferable, individual and limited right to using this document.

This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

#### Kontakt / Contact:

peDOCS  
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation  
Informationszentrum (IZ) Bildung  
E-Mail: [pedocs@dipf.de](mailto:pedocs@dipf.de)  
Internet: [www.pedocs.de](http://www.pedocs.de)

Mitglied der

  
Leibniz-Gemeinschaft

# EMPIRISCHE SONDERPÄDAGOGIK

ISSN 1869-4845

3. Jahrgang 2011 | Heft 2



*Manfred Hintermair, Andrea Schenk,  
Klaus Sarimski*

Exekutive Funktionen, kommunikative Kompetenz und Verhaltensauffälligkeiten bei hörgeschädigten Kindern.

Eine explorative Studie mit Schülern einer schulischen Einrichtung für Hörgeschädigte

*Alexandra Lenhard, Wolfgang Lenhard*

Computerbasierte Intelligenzförderung mit den "Denkspielen mit Elfe und Mathis" – Vorstellung und Evaluation eines Computerprogramms für Vor- und Grundschüler

*Kirsten Diehl, Bodo Hartke*

Zur Reliabilität und Validität des formativen Bewertungssystems IEL-1: Inventar zur Erfassung der Lesekompetenz von Erstklässlern

*Michael Grosche*

Effekte einer direkt-instruktiven Förderung der Lesegenauigkeit

*Kirsten Endrikat*

Wie integriert sind Schülerinnen und Schüler mit und ohne Schulabschluss? Soziale Desintegrationsbelastungen und deren mögliche Folgen



PABST SCIENCE PUBLISHERS

Empirische Sonderpädagogik, 2011, Nr. 2, S. 121-146

## Zur Reliabilität und Validität des formativen Bewertungssystems IEL-1: Inventar zur Erfassung der Lesekompetenz von Erstklässlern

*Kirsten Diehl, Bodo Hartke*

*Institut für Sonderpädagogische Entwicklungsförderung und Rehabilitation, Universität Rostock*

Eine sich zunehmend auf Prävention und Integration hin entwickelnde (Sonder-)Pädagogik benötigt Messverfahren, die fehlende Lernvoraussetzungen und im Lernprozess entstehende Lernlücken frühzeitig erfassen. Gelingt es zudem mittels eines Messverfahrens Einblick in den Lernverlauf zu bekommen, erhalten Lehrkräfte Informationen zu Lernfortschritten, stagnierenden oder ggf. auch rückläufigen Leistungsentwicklungen. Solche Daten sind eine gute Grundlage für einen adaptiven Unterricht und eine passende Förderung. Das Inventar zur Erfassung der Lesekompetenzen von Erstklässlern (IEL-1) wurde als pädagogisches Instrumentarium zur Leistungserfassung über drei Messzeitpunkte entwickelt. Sollte sich das Inventar als objektiv, reliabel und valide sowie geeignet für die Dokumentation von Lernfortschritten erweisen, liegt ein schulalltagstaugliches Verfahren zur formativen Evaluation des Erstleseunterrichts und präventiver Förderung innerhalb des Schriftspracherwerbs vor. Im vorliegenden Beitrag werden die Konzeption des IEL-1, das Design der Überprüfungsstudie mit  $N = 1649$  Erstklässlern und Ergebnisse zur Güte des Verfahrens vorgestellt.

Schlüsselwörter: Lesekompetenz, Lernverlaufsdiagnostik, niederländisches Leerlingsonderwijsvolgsysteem, Überwachung der Lernentwicklung, Dokumentation von Lernfortschritt, formative Bewertungssysteme, RTI (response-to-intervention)

### **Reliability and Validity of the Formative Evaluation System IEL-1: Inventory for the Assessment of the Reading Proficiency of First Graders**

Developing more and more towards prevention and integration, education needs innovative methods of measurement to recognize missing learning prerequisites and learning gaps which occurred in the process of learning at an early stage. If a measurement method could provide insight into the learning process, teachers could get feedback on advancements, stagnations or even regressive learning developments. Such data are a valid foundation for an adaptive teaching and advancement. The inventory for measuring the reading competence of first grade children (IEL-1) was developed as a pedagogic instrument for the facility recording over three measurement moments. Should the developed and verified inventory prove objective, reliable and valid as well as appropriate to show learning advancements, then it is a procedure for the formative evaluation of learning processes for successful preventive actions in the process of learning to read and write that is suitable for everyday school life. This paper presents the concept of the IEL-1, the design of the validation-study with  $N=1649$  first grade children and results concerning the quality of the procedure.

Key words: reading competence, curriculum-based measurement, Dutch Leerlingsonderwijsvolgsysteem, student monitoring progress, documentation of learning progress, formative evaluation system, RTI (response-to-intervention)

## Forschungsstand

Die Entwicklung und die Anwendung von formativen Messverfahren sind in den USA gegenwärtig hoch aktuelle Themen. Sowohl in der Praxis als auch in der Forschung fokussiert dieser Trend insbesondere auf den Bereich Lesen (Bamonto Graney & Shinn, 2005; Deno, 2003a,b; Fuchs, 2004; Fuchs, Fuchs, & Compton, 2004; Fuchs & Fuchs, 2005; Haager, Klingner & Vaughn, 2007; Stecker & Fuchs, 2000). Ein wesentlicher Teil der hierzu in den USA verfassten Studien zentriert sich um die Frage, ob ein formatives Evaluationsystem zur Verbesserung der Wirksamkeit pädagogischen Handelns bei Schülern mit Lernstörungen führt (Deno, 2003a). Scriven prägte (1972) den Begriff „formative Evaluation“. Bereits vor über 35 Jahren setzte er sich mit den verschiedenen Möglichkeiten der Einschätzung von pädagogischen Prozessen auseinander und führte unter anderem die Begriffe „formative“ und „summative“ Evaluation ein. Bei der Beschreibung der Begriffe bezieht er sich auf eine Unterscheidung von Cronbach: „Evaluation, used to improve the course while it is still fluid, contributes more to improvement of education than evaluation used to appraise a product already placed on the market“ (Cronbach zit. n. Scriven, 1972, S.43). Bezogen auf diese Unterscheidung nennt Scriven als ein charakteristisches Merkmal formativer Evaluation, dass während eines Entwicklungsprozesses eine laufende Überprüfung dieses Prozesses stattfindet, welche wiederum Einfluss auf die weitere Gestaltung des Prozesses habe. Unter summativer Evaluation versteht Scriven eine abschließende Wertung der Ergebnisse eines Prozesses (ebd.). Übertragen auf den Lernprozess von Kindern bedeutet dieses, dass bei der formativen Evaluation nicht das Endergebnis des Lernens abgeprüft wird, sondern der Lernprozess im Sinne des „response-to-intervention“ bzw. „response-to-instruction“-Paradigmas (RTI) kontinuierlich beobachtet und ausgewertet wird. Im Unterschied

zur summativen wird bei der formativen Evaluation „der LehrLernprozess in der Schule als von den Unterrichts- und Förderergebnissen abhängiger datenbasierter Problemlöseprozess verstanden“ (Walter, 2008a, S. 206). Es wird dementsprechend nicht nur der aktuelle Lernstand erfasst, sondern der Lernverlauf über einen längeren Zeitraum dokumentiert und evaluiert. In diesem Sinne sprechen einige Autoren anstelle von Lernprozessdiagnostik auch von Lernverlaufsdiagnostik (Strathmann & Klauer, 2008). *Formative* Bewertungssysteme dienen der Beantwortung der Fragen: Was hat der Schüler unter dem Einfluss des gegenwärtigen Unterrichts und von aktuell praktizierten Fördermaßnahmen bereits gelernt? Wo hat er noch Schwierigkeiten? *Formative* Bewertungssysteme werden parallel zum Lernprozess, also wiederholt durchgeführt. Sie sind Grundlage für Feedbacks, spezifische Fördermaßnahmen und die Verbesserung von Unterricht. Im Vergleich dazu unterstützen *summative* Bewertungssysteme dabei, vorhandenes Wissen beim Schüler zu ermitteln und dessen Leistung fair und präzise zu beurteilen. Sie kommen am Ende eines Lernprozesses zum Tragen und unterstützen die Zensurengebung sowie Selektions- und Platzierungsentscheidungen. *Formative* Bewertungen können den Lernerfolg noch beeinflussen, *summative* Bewertungen stellen ihn im Nachhinein fest (Stern, 2004).

*Formative* Bewertungssysteme sind ein wesentlicher Baustein innerhalb des Response-to-Intervention-Konzepts. Als ein aus Forschungsbemühungen in den USA resultierendes Rahmenkonzept integriert das Response-to-Intervention-Konzept *formative* Messverfahren und spezifische Interventionen innerhalb eines Mehrebenenpräventionssystems. Dieses zielt darauf ab, Schulleistungen zu verbessern und Schwierigkeiten im Verhalten der Kinder zu mindern. Untersuchungen (Compton, Fuchs, Fuchs & Bryant, 2006; Smith, Pepler & Rigby, 2004) haben gezeigt, dass selbst Kinder mit massiven Lernstörun-

gen von diesem gestuften System der Förderung profitieren. Mit dem RTI-Konzept werden den Schulen in den USA Informationen und Materialempfehlungen zu folgenden Schwerpunkten offeriert:

1. Identifikation von Kinder mit einem erhöhten Risiko für die Entwicklung von Schwierigkeiten im Lernen,
2. Überwachung des Lernfortschritts der Schüler (monitor student progress),
3. evidenzbasierte Interventionen (welche Verfahren in welcher Intensität?),
4. Identifikation von Kindern mit einer Lernbehinderung oder anderen Beeinträchtigungen (National Center of Response to Intervention, 2010).

Handlungsleitend bei der Entwicklung des hier vorgestellten Inventars zur Erfassung der Lesekompetenzen von Erstklässlern (IEL-1) waren Erkenntnisse, die sich auf die Bedeutung der Identifikation von Kindern mit einem erhöhten Risiko für die Entwicklung von Schwierigkeiten im Lernen und die Überwachung des Lernfortschritts der Schüler (monitor student progress) innerhalb des RTI-Konzeptes beziehen (Diehl & Hartke, 2007).

In den USA wird zur Überwachung des Lernfortschritts insbesondere auf curriculumbasierte Kurztests, so genannte CBMs (curriculum-based measurement) zurückgegriffen, die in relativ kurzen Abständen eingesetzt werden (vierwöchig, wöchentlich oder teilweise mehrmals die Woche). *Curriculum-based* hebt darauf ab, dass es in den Verfahren um schulische, lehrplanorientierte Aspekte der Entwicklung von Kindern geht. Die Inhalte der Messungen sind dem Unterrichtsmaterial entnommen und beziehen sich auf schulisch vermittelte Fertigkeiten, Wissenssysteme und Kompetenzen. Erste CBMs wurden innerhalb eines Forschungsprogramms an der Universität Minnesota (Deno, 1985) entwickelt. Ziel des Forschungsprogramms war es, ein praktikables, leicht durchführbares und wirkungsvolles formatives Bewertungssystem als Grundlage für Entscheidun-

gen zur besseren Abstimmung des Unterrichts auf Bedürfnisse der Kinder zu entwickeln. Neben CBMs werden in den Schulen in den USA vielfältige curriculumbasierte diagnostische Verfahren verwendet, die wiederholt eingesetzt werden, um Einblicke in Lernverläufe zu erzielen. Sie werden als curriculum based assessment (CBA) bezeichnet und dienen ebenfalls der Adaption von Unterricht und Förderung an die Lernausgangslage der Schüler. Auch in den Niederlanden wird an den Schulen bereits seit mehreren Jahren kontinuierlich der Lernverlauf der Schüler mit Hilfe eines Monitoringsystems, dem Leerlingonderwijsvolgsysteem (LOVS), erfasst und dokumentiert. Wesentlicher Unterschied zu den US-amerikanischen CBMs ist die Anzahl der Messungen, die sich in der Regel auf zwei Messungen pro Jahr beschränken, und der zeitliche Aufwand der Durchführung. Während die CBMs durchschnittlich ein bis drei Minuten in der Durchführung beanspruchen, dauern die niederländischen Verfahren bis zu zwei Schulstunden. Das LOVS zielt darauf ab, die Lernverläufe bei der Aneignung komplexer Lerninhalte über mehrere Schuljahre abzubilden. Im Gegensatz dazu zielen CBMs auf zeitnahe Abbildungen von Lernverläufen bezogen auf einzelne Fähigkeiten oder auch umfassende Kompetenzen ab (Diehl & Hartke, 2007). Im deutschsprachigen Raum finden sich seit geraumer Zeit Veröffentlichungen über Lernverlaufdiagnostik im Kontext formativer Bewertungssysteme (Diehl & Hartke, 2007; Hartmann, 2008; Klauer, 2006; Strathmann, Klauer & Greisbach, 2010; Strathmann & Klauer, 2010; Walter, 2008a,b; Walter, 2009a,b). Erste Forschungsergebnisse zur Güte von CBMs im deutschsprachigen Raum bestätigen, was die US-amerikanische Forschung seit mehr als 20 Jahren zeigt: Die Entwicklung von CBMs ist außerordentlich aufwendig, will man den hohen Standards, die mit diesen Verfahren verbunden sind, gerecht werden.

Mit den beiden hier kurz skizzierten Monitoringverfahren, den CBMs bzw. CBAs und

dem LOVS wurden Möglichkeiten von Lernverlaufsdagnostik beschrieben, die mehr oder minder zeitnah Veränderungen im Lernverlauf aufzeigen können. Abhängig von der Methode der Aufgabengestaltung und Zielstellung der Verfahren zur Lernverlaufsdagnostik können Lücken in der Entwicklung von Wissen und Fertigkeiten aufgezeigt werden. Forschungsergebnisse aus den USA (Fuchs & Fuchs, 1986) und Praxiserfahrungen in den Niederlanden sprechen dafür, derartige Verfahren auch für den deutschsprachigen Raum zu entwickeln. Gerade in Hinblick auf den Anfangsunterricht in Deutsch und Mathematik besteht ein deutlicher Handlungsbedarf. Studien konnten aufzeigen, dass sich insbesondere frühe Lernlücken gravierend auf den weiteren Lernerfolg in höheren Klassenstufen auswirken (Balzer, Fritz, Ricken & Jäger, 2007; Grube & Hasselhorn, 2006; Moser Opitz, 2007). Das IEL-1 wurde mit dem Anspruch entwickelt, eine präzise Beschreibung des Lesestands von Erstklässlern im Lesen zu drei Messzeitpunkten (Schuljahresmitte, Mitte zweites Halbjahr, Schuljahresende) zu ermöglichen. Die Messwiederholungen sollen einen Einblick in die Leistungsentwicklung bieten und bei ausbleibenden Lernerfolgen Veränderungen innerhalb der Förderung anregen.

## Das Inventar zur Erfassung der Lesekompetenzen (IEL-1)

Mit dem IEL-1 soll ein Spektrum an Lesefähigkeiten erfasst werden, wie es zum Ende des Schuljahres von Erstklässlern erwartet werden kann. Die wiederholten Messungen ermöglichen es, den Lernstand einzelner Lesefähigkeiten und genereller Lesekompetenz zum entsprechenden Messzeitpunkt zu beschreiben und ggf. Lernlücken aufzudecken. Diese Zielstellung erfordert ein dem Verfahren zugrunde gelegtes Modell des Schriftspracherwerbs sowie eine hohe Reliabilität und Validität der Messungen. Das Inventar

soll zum einen prüfen, ob aktuell in der Schule vermitteltes Wissen tatsächlich erworben wurde. Zum anderen werden innerhalb des Verfahrens curriculare Inhalte der gesamten ersten Klassenstufe berücksichtigt, um den Stand der Aneignung der in der Klassenstufe geforderten Kenntnisse und Fertigkeiten zu erfassen.

Die sieben Aufgabenbereiche des Inventars (Gruppentest s. Tabelle 1) und deren Items (insgesamt 135) basieren auf einer Analyse gängiger Fibeln und Arbeitsmaterialien sowie des pädagogisch-didaktischen und entwicklungs- und kognitionspsychologischen Wissensstandes zur Entwicklung der Lesefähigkeiten (Büttner, 2003a,b; Dehn & Sjölin, 1996; Groeben & Hurrelmann, 2006; Meiers, 1996; Klicpera & Gasteiger-Klicpera, 1995; Scheerer-Neumann, 1995). Der Zusatztest *Eine-Minute-Lese-Aufgabe* enthält 60 Wörter. Mit dieser zusätzlichen *Eine-Minute-Lese-Aufgabe* (Einzeltest) wird erfasst, wie viele Wörter ein Kind innerhalb einer Minute laut und korrekt vorlesen kann. Die Ergebnisse lauten Vorlesens unter Zeit, wie in der *Eine-Minute-Lese-Aufgabe*, korrelieren erfahrungsgemäß hoch mit den Ergebnissen komplexerer Lesetestungen (Deno, 2003a; Fuchs, 2004). Deshalb eignet sich die *Eine-Minute-Lese-Aufgabe*, wenn entsprechende Normen vorliegen, als wenig zeitaufwendiges Screeningverfahren zur Identifikation schwacher Leser. Liegen die Leseleistungen eines Kindes in der *Eine-Minute-Lese-Aufgabe* im unteren Quartil, ist dies ein Hinweis darauf, dass eine umfassendere und differenziertere Diagnostik des Lesestandes mit den sieben Subtests des IEL-1 angezeigt ist.

### **Aufgabenbereich 1(A): Buchstaben-Laut-Zuordnungen**

Untersuchungen im Bereich von Prävention und Frühförderung im Schriftspracherwerb (zusammenfassend Walter, 2001) zeigen die zentrale Rolle des Konzepts der phonologi-

Tab. 1: Aufgabenbereiche/Subtests des IEL-1

Aufgabenbereich (I-VII)	Subtest (A-G)	Anzahl der Items
I	Buchstaben-Laut-Zuordnungen (A)	43
II	Wörter in Silben segmentieren (B)	25
III	Silben zu Wörtern verbinden (C)	19
IV	Wortebene: Buchstaben analysieren und synthetisieren (D)	21
V	Wörter lesen und schreiben (E)	13
VI	Satzlesen (sinnverstehend) (F)	9
VII	Textlesen (sinnverstehend) (G)	5
Zusatztest	Eine-Minute-Lesen (Lesegeschwindigkeit und -genauigkeit)	60

schen Bewusstheit. Es konnte gezeigt werden, dass leseschwache Schüler besonders bei Aufgaben auffällig sind, die eine Analyse und/oder Synthese des Lautstroms gesprochener Sprache erfordern. Auch wenn einige wenige schrifterfahrene Kinder diese Leistung schon vor dem Beginn des bewussten Schriftspracherwerbs zeigen, differenziert sich die Fähigkeit der phonologischen Bewusstheit mit der Einführung der Buchstaben-Laut-Zuordnungen aus.

### **Aufgabenbereich II (B): Wörter in Silben segmentieren (Analyse)**

Im Aufgabenbereich II wird erfasst, inwieweit die Schüler Wörter in ihre einzelnen Silben segmentieren können. Bereits Vorschulkinder sind in der Lage, Wörter in ihre Silben zu zerlegen und zeigen somit eine Fähigkeit der phonologischen Bewusstheit im weiteren Sinn. Als prosodische Einheit (Betonungsmuster) ist ihnen die Silbe vertraut und kann entsprechend erkannt werden. Die Schwierigkeit bei diesem Aufgabenbereich besteht darin, dass die Schüler die Silben des Wortes

durch Silbenbögen genau kennzeichnen müssen. Das setzt voraus, dass sie die Binnenstrukturen der Silben kennen. Als Lesestrategie innerhalb der alphabetischen Phase erhöht das gekonnte silbenweise Lesen die Lesegeschwindigkeit.

### **Aufgabenbereich III (C): Silben zu Wörtern verbinden**

Bei diesem Aufgabenbereich wird erfasst, inwieweit die Schüler in vier unterschiedlichen Aufgabengruppen Silben zu Wörtern verbinden können. Der Schwierigkeitsgrad steigt von einer zur nächsten Aufgabengruppe. Das Segmentieren oberhalb der Buchstabenebene und unterhalb der Wortebene ist eine Lesestrategie, die es den Schülern ermöglicht, die Lesegeschwindigkeit zu erhöhen. Studien (Scheerer-Neumann, 1981; Walter, 1996) zeigen, dass gute Leser im Gegensatz zu schwächeren Lesern Gliederungsmöglichkeiten oberhalb der Buchstabenebene und unterhalb der Wortebene von sich aus beim Lesen nutzen. Leseschwache Kinder sind weniger gut in der Lage, die Binnenstruktur der Silbe zu erkennen. Das wiederum hindert sie,

die Einheit der Silbe leseunterstützend zu nutzen. Sie wählen die vermeintlich einfachere Strategie innerhalb der alphabetischen Phase, indem sie die Buchstaben eines Wortes in Laute umsetzen und diese sukzessive aneinanderreihen.

### **Aufgabenbereich IV (D): Buchstaben zu Wörtern verbinden**

Der Aufgabenbereich IV beinhaltet zwei Aufgabengruppen. Bei Aufgabengruppe 1 (Setze die fehlenden Buchstaben ein!) gilt es, zu vorgegebenen Wörtern mit Bildunterstützung lautgetreu die fehlenden Grapheme einzusetzen. Gesucht sind einfache Konsonanten und Diphthonge sowie schwierigere Buchstaben bzw. Buchstabenverbindungen wie <x>, <Z> und <Qu> an unterschiedlichen Positionen im Wort. Die Aufgabe gilt als korrekt gelöst, wenn die Schüler die fehlenden Laute lautgetreu bestimmen (für <Qu> kann <Kw> oder für <x> auch <ks> eingesetzt werden). Aufgabengruppe 2 (Setze die fehlenden Buchstaben ein!) unterscheidet sich dahingehend, dass die Wörter nicht bildlich dargeboten und somit nur durch korrektes phonologisches Rekodieren zu identifizieren sind. Bei den Wörtern <Banane> und <Telefon> sind zwei Grapheme einzutragen, allerdings handelt es sich jeweils um das gleiche (<e> bei <T \_ l \_ f o n>; <a> bei <B \_ n \_ n e>). Mit diesen Aufgaben wird erfasst, inwieweit die Schüler die Buchstaben-Laut-Zuordnungen, das Durchgliedern von Wörtern und das Erkennen von Buchstabenclustern, wie z. B. Signalgruppen, beherrschen.

### **Aufgabenbereich V (E): Wörter lesen und schreiben**

Im Aufgabenbereich V wird in unterschiedlichen Aufgabengruppen das sinnverstehende Lesen auf Wortebene erfasst. Aufgabenbereich V Wörter lesen, schreiben und identifi-

zieren beinhaltet drei Aufgabengruppen auf Wortebene. In der Aufgabengruppe 1 (Verbinde das Bild mit dem richtigen Wort!) müssen die Kinder anfänglich einfach strukturierte Wörter entsprechenden Bildern zuordnen. Auch die später komplexeren Wörter (viersilbig, Konsonantenhäufung am Wortanfang) innerhalb dieser Aufgabengruppe ähneln sich am Wortanfang und zwingen somit die Schüler, die Wörter über die ersten Buchstaben hinaus zu lesen. Aufgabengruppe 2 (Schreibe die Wörter!) prüft, ob die Schüler zu einem vorgegebenen Bild das passende Wort schreiben können. Die Ergebnisse lassen eine quantitative und qualitative Analyse zu. Eine qualitative Analyse ließe auf der Grundlage des Kompetenzentwicklungsmodells von Klicpera et al. (2003) Rückschlüsse auf die bevorzugte Lese- und Schreibstrategie zu. Die quantitative Analyse zeigt auf, wie viele Buchstaben die Schüler richtig verschriftet haben. Bis auf die letzten zwei Items handelt es sich um lautgetreu zu schreibende Wörter. Mit der Aufgabengruppe 3 (Welche Wörter findest du? Kreise die Wörter ein!) soll erfasst werden, inwieweit die Schüler in der Lage sind, einfache und komplexe Wörter simultan zu erfassen.

### **Aufgabenbereich VI (F): Satzlesen (sinnverstehend)**

Im Aufgabenbereich VI wird erfasst, inwieweit die Schüler Einzelinformationen auf der Basis von Wörtern in größeren Sinneinheiten erfassen können. Auf Grund einer zu hohen Belastung der Kapazität des Arbeitsgedächtnisses erschwert eine defizitäre Wortlesefähigkeit (langsames und fehlerhaftes Rekodieren und Dekodieren), Informationen aus Sätzen oder kleineren Texten sinnerfassend aufzunehmen und für den weiteren Leseprozess zu nutzen. Willson und Rupley (1997) konnten in Untersuchungen mit Schülern der ersten bis sechsten Klasse zeigen, dass in ersten Klassen sehr starke kausale Beziehungen zwi-



schen Wortlesen und Dekodierstrategien und dem Sinnverständnis zu beobachten sind. Diese Beziehung nimmt in den höheren Klassenstufen ab. Zunehmend beeinflussen sprachunspezifische Aspekte – wie Motivation, Strategiewissen – über die Art und Weise, wie man einen Text liest und allgemeines Vorwissen das Leseverständnis. Alle Aufgabengruppen des Aufgabenbereichs VI unterliegen einer zeitlichen Begrenzung. Insofern erweist sich das Lesetempo ausschlaggebend für das Lösen der Aufgaben.

### **Aufgabenbereich VII (G): Textlesen (sinnverstehend)**

Mit den Aufgabengruppen aus dem Aufgabenbereich VII soll erfasst werden, inwieweit die Schüler in der Lage sind, aus Texten mit und ohne Bildunterstützung Informationen sinnverstehend zu entnehmen. Erst nachdem die Schüler Wortleseprozesse relativ automatisch vollziehen, wird Arbeitsgedächtniskapazität für höhere Formen der Informationsverarbeitung frei. Zu ihnen zählen z. B. das Verknüpfen der Informationen mit Vorwissen und Weltwissen, Strategiewissen und Strategieanwendung. Buchstaben-Laut-Zuordnungen spielen in dieser Phase eine sekundäre Rolle, die Wörter werden größtenteils simultan erfasst. Weniger das einzelne Wort, sondern die Beziehungen der Wörter zueinander rücken beim sinnverstehenden Lesen in den Vordergrund. Ein Text ist in diesem Sinne auch nicht als Summe vieler Einzelwörter zu verstehen, sondern als Abbildung von Sprache mit unterschiedlichen funktionalen Systemen (Semantik, Morphologie, Syntax). Diese Aufgabe ist vermutlich nur von Schülern zu lösen, die fortgeschrittene Lesefertigkeiten besitzen.

### **Screening Eine-Minute-Lese-Aufgabe**

Lesen als komplexer Prozess umfasst unterschiedliche Teilfertigkeiten. Eine wesentliche Voraussetzung, um das eigentliche Ziel der Sinnerfassung zu erreichen, ist die technische Lesefertigkeit. In einigen Untersuchungen konnte aufgezeigt werden, dass die entscheidende Beeinträchtigung des lese-recht-schreibschwachen Kindes eine schwache Wortlesefähigkeit ist. Beeinträchtigt sinnerfassendes Textlesen ist hiernach eine Folge des langsamen und mühevollen Wortlesens. Dinges (2002) konnte mit einem Verfahren zur Erfassung der Lesefertigkeit in einer Minute zeigen, dass sich das „Eine-Minute-Lesen“ von einzelnen Wörtern zu mehreren Messzeitpunkten zur Dokumentation des Lernfortschritts sehr gut eignet. In Anlehnung an Dinges (2002) wurde die *Eine-Minute-Lese-Aufgabe* für Erstklässler entwickelt. Ebenso wie bei Dinges wurden die Wörter hierbei nach Schwierigkeitsgraden geordnet. Die *Eine-Minute-Lese-Aufgabe* hebt sich von den anderen Aufgabenbereichen ab, da sie in der Einzelsituation mit dem Schüler durchzuführen ist. Laut Aufgabenstellung sollen die Schüler innerhalb einer Minute so viele Wörter wie möglich laut vorlesen. Jeweils fünf Wörter sind in einer Folge mit annähernd gleichen Schwierigkeitsmerkmalen gebündelt. Bei den Wörtern handelt es sich um Nomen aus unterschiedlichen Inhaltsbereichen, lediglich der erste Block beinhaltet Präpositionen wie <mit>, <am>, das Kopula <und>, eine gebeugte Verbform <ist> und das Adverb <da>. Diese Wörter entstammen dem üblichen Sichtwortschatz für die erste Klasse. Die Lehrkraft protokolliert während des Vorlesens die Anzahl der richtig und falsch gelesenen Wörter, was Rückschlüsse auf die Lesegeschwindigkeit und die Lesegenauigkeit zulässt.

Das IEL-1 enthält Aufgaben, die unter unterschiedlichen Zeitvorgaben gelöst werden müssen. Es liegt für die Durchführung eine wortwörtliche Anleitung vor. Die Auswertung der Aufgabenbereiche erfolgt quantitativ mit-

tels der erreichten Punktzahl. Eine qualitative Auswertung kann individuell für jeden Schüler vorgenommen werden. Zwecks Klärung der Fragestellungen dieser Untersuchung wurden lediglich die auf den einzelnen Skalen erreichten Punkte ermittelt. Sie ergeben die Rohwerte für die statistische Analyse. Insgesamt werden für die Durchführung des IEL-1 zwei Unterrichtsstunden benötigt, die entweder auf einem Tag mit einer dazwischen liegenden Pause oder auch auf zwei aufeinanderfolgenden Tagen liegen sollten. Das IEL-1 liegt in einer Version vor. Testwiederholungseffekte sind nicht zu erwarten, da das Verfahren aus schulisch gängigen Aufgabenformaten besteht, vergleichbare „Item“ täglich im Erstleseunterricht vorkommen. Gerade die Fähigkeit zur Bewältigung schulischer Leseanforderungen wird mit Hilfe des IEL-1 geprüft.

## Fragestellung

Hauptziel der Erhebung mit dem Messverfahren IEL-1 ist die Überprüfung der mit dem Verfahren verbundenen Annahmen in der Praxis. Es soll geklärt werden, ob das entwickelte Messverfahren den gängigen testtheoretischen Gütekriterien (Objektivität, Reliabilität und Validität) genügt. Weiterhin gilt es zu prüfen, ob mit dem IEL-1 eine Schätzung des Lernzuwachses möglich ist. Folgende Fragestellungen sollen durch die Datenerhebung und -auswertung beantwortet werden.

- Können mit dem IEL-1 Leistungsunterschiede in der Lesekompetenz zu den unterschiedlichen Messzeitpunkten abgebildet werden?
- Wie zuverlässig misst das IEL-1 die Lesekompetenzen der Erstklässler (Reliabilität)?
- Welche Korrelationen zeigen sich mit anderen Testverfahren (WLLP) und dem Lehrerurteil (Validität)?
- Können mit Hilfe der wiederholten Messungen über drei Messzeitpunkte Leistungsunterschiede mit dem IEL-1 abgebil-

det werden? (1. Messzeitpunkt = Ende des ersten Schulhalbjahres, 2. Messzeitpunkt = Mitte des zweiten Schulhalbjahres, 3. Messzeitpunkt = Ende des Schuljahres)?

## Methode

### *Untersuchungsplan und verwendete Verfahren*

Die durchgeführte Untersuchung stellt einen geschlossenen Teilbereich der wissenschaftlichen Begleitung des Modellprojektes „Primarstufe“ dar. Eine genauere Beschreibung des Projektes findet sich bei Koch, Hartke, Blumenthal (2008). Der zeitliche Ablauf der eigenen Untersuchung wird in der nachfolgenden Abbildung 1 mitgeteilt.

Zum 1., 2. und 3. Messzeitpunkt wurden Daten über die Entwicklung der Lesekompetenz aller Erstklässler in Rostock und auf Rügen mit dem IEL-1 erhoben. In Mecklenburg-Vorpommern können Erstklässler in Grundschulklassen oder Diagnoseförderklassen (DFK) eingeschult werden (Koch, 2007). In den DFK haben Schüler, die vermutlich aufgrund einer stark verzögerten allgemeinen Entwicklung im Anfangsunterricht der Grundschule nicht erfolgreich lernen können, die Möglichkeit, die ersten beiden Grundschuljahre in drei Jahren zu absolvieren. Zusätzlich zum IEL-1 wurden andere Diagnoseverfahren zur Bestimmung der Lernausgangslage der Schüler in Grundschul- und Diagnoseförderklassen eingesetzt. Mit ihnen wurden u. a. Bereiche wie sprachliche Fähigkeiten, Konzentration, Motorik, Verhalten oder pränumerisches Wissen (Zahlbegriff) näher untersucht. Die Untersuchung wurde in Rostock von studentischen Hilfskräften und auf Rügen durch Sonderschullehrkräfte durchgeführt. Die Testleiter wurden vorab auf ihre Aufgabe vorbereitet. Die Testung fand im regulären Unterricht als Gruppentest statt. Relevant für die eigene Untersuchung zur Prüfung der Gütekriterien des IEL-1 sind der Ein-

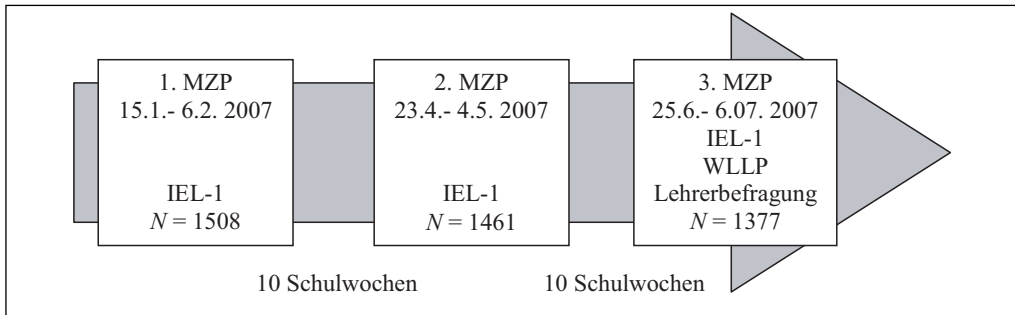


Abb. 1: Übersicht zur Datenerhebung

Erläuterung: M.Z.P. = Messzeitpunkt, N = Anzahl der einbezogenen Schüler, WLLP = Würzburger Leise Leseprobe

satz der Würzburger Leise Leseprobe von Küspert und Schneider (1998) und eine Lehrerbefragung zum dritten Messzeitpunkt. Alle Klassenlehrer der untersuchten Population wurden gebeten, in Form eines Fragebogens die Lesefertigkeit und das Leseverständnis der Schüler ihrer Klassen auf einer sechsstufigen Skala (1 = sehr gut, 2 = gut, 3 = befriedigend, 4 = ausreichend, 5 = mangelhaft, 6 = ungenügend) zu bewerten.

### Stichprobe

Von den insgesamt 1688 Erstklässlern bearbeiteten zum 1. Messzeitpunkt (M.Z.P.) 1508, zum zweiten 1461 und zum dritten M.Z.P. 1377 Schüler das IEL-1. Es gab Erstklässler, die zu keinem der drei M.Z.P. an der Studie teilnahmen, andere, die nur zu einem M.Z.P. mitarbeiteten und weitere, die an zwei bzw. allen drei M.Z.P. partizipierten. Auffallend ist die Reduzierung der Gesamtuntersuchungsgruppe vom M.Z.P. 1 bis zum M.Z.P. 3 um insgesamt 131 Schüler. Neben Gründen wie z. B. Krankheit, Umzug und entschuldigtes bzw. unentschuldigtes Fehlen sind die Ausfälle auf Schwierigkeiten bei der Datenerhebung zurückzuführen. Ein gerichteter Einfluss erscheint ausgeschlossen. Die Verteilung der Geschlechter ist nahezu ausgeglichen. Von 1168 Kindern liegen Daten zu allen M.Z.P. vor.

## Ergebnisse

### Verteilungsformen und Kennwerte der Verteilung

Bevor die Verteilungsformen der Gesamtpunkte des IEL-1 zu den Messzeitpunkten näher betrachtet wurden, wurde überprüft, inwieweit die IEL-Testitems in einem korrelativen Zusammenhang miteinander stehen. Die Ergebnisse dieser Analysen zeigen, dass sich das Verfahren durch eine hohe innere Konsistenz auszeichnet. Der Wert für Cronbachs Alpha ist mit  $r = .97$  sehr hoch. Die Subtests korrelieren zu den drei M.Z.P. mit dem Gesamtergebnis zwischen  $.75$  und  $.87$  (Pearson). Nur bei Subtest VII „Textlesen“ liegt die Korrelation mit dem Gesamttest mit  $.41$  (erster M.Z.P.) niedriger. Insofern erscheint es vertretbar, einen IEL-1-Gesamtwert als Indikator für Lesekompetenz zu bilden. Die *Eine-Minute-Lese-Aufgabe* korreliert mit dem IEL-1-Gesamttest zu den drei M.Z.P. zwischen  $r = .65$  und  $r = .74$ . ( $p < .01$ ). Im Folgenden werden zunächst die Verteilungsformen und Kennwerte der Verteilung für den Gesamttest vorgestellt. Angaben zu der Verteilung der Werte der Untertests finden sich bei Diehl (2009a).

Die Häufigkeitsverteilungen der richtigen Lösungen zu den drei Messzeitpunkten im IEL-1 werden in Abbildung 2, 3 und 4 dargestellt. Abbildung 2 – sie bezieht sich auf den

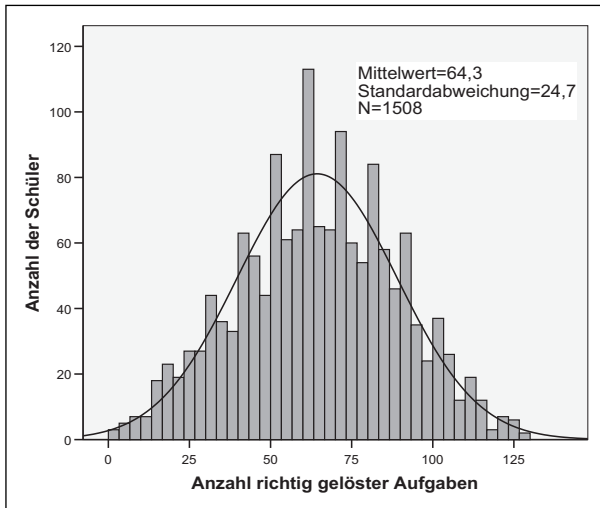


Abb. 2: Absolute Häufigkeiten der Anzahl richtig gelöster Items zum 1. MZP

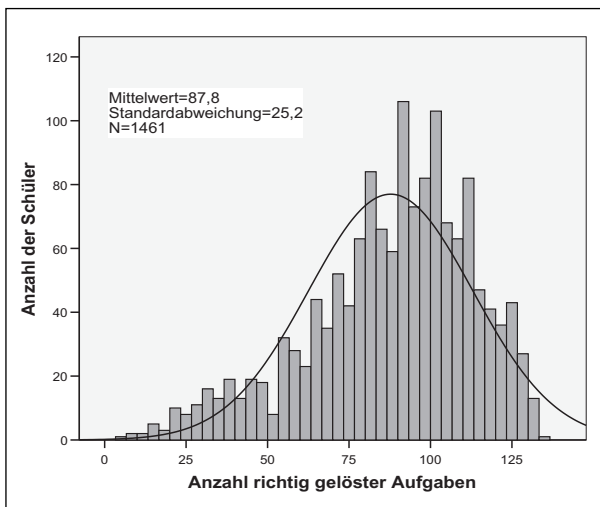


Abb. 3: Absolute Häufigkeiten der Anzahl richtig gelöster Items zum 2. MZP

ersten Messzeitpunkt – zeigt eine symmetrische, eingipflige Verteilungsform. Die Verteilung kann als glockenförmig und damit augenscheinlich normal verteilt angesehen werden, wobei die Ergebnisse des Kolmogorov-Smirnov-Anpassungstests diese Annahme für diese und weitere MZP negieren ( $p < .01$ , 2-seitig). Im Mittel wurden von 135 Aufgaben 64.3 Aufgaben zum 1. Messzeitpunkt bei einer Standardabweichung von  $SD = 24.7$  richtig gelöst. Auch zum 2. Messzeitpunkt (s.

Abb. 3) ist die Verteilung symmetrisch und eingipflig, wobei sich ein Deckeneffekt andeutet. Im Mittel wurden von den 135 Aufgaben 87.7 bei einer Standardabweichung von  $SD = 25.2$  richtig gelöst. Abbildung 4 zeigt die Verteilung der absoluten Häufigkeit der Anzahl richtig gelöster Aufgaben zum 3. Messzeitpunkt. Die Verteilung ist eingipflig und rechtssteil. Von den 135 Aufgaben wurden im Mittel 96.0 Aufgaben bei einer Standardabweichung von 25.0 richtig gelöst. Die

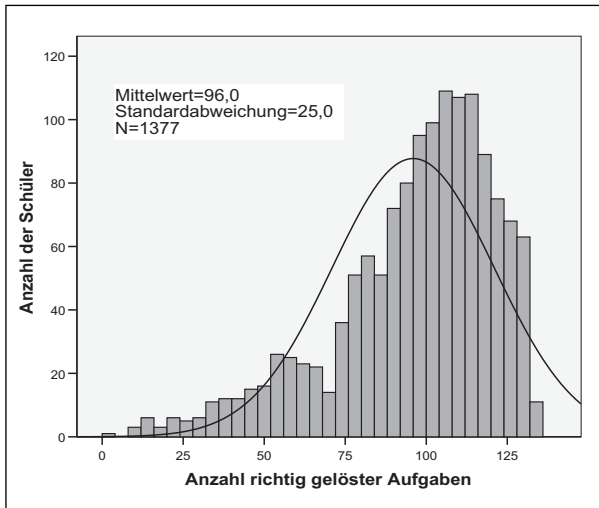


Abb. 4: Absolute Häufigkeiten der Anzahl richtig gelöster Items zum 3. MZP

Abbildung 4 verdeutlicht wie erwartet eine Verschiebung der Verteilung nach rechts. Die Verteilungskurve zeigt hingegen links einen leichten stetigen Anstieg.

In Tabelle 2 sind die Mittelwerte und Standardabweichungen für die Gesamtzahl richtig gelöster Aufgaben für die Gesamtuntersuchungsgruppe zu allen drei Messzeitpunkten aufgeführt. Es wird deutlich, dass die Standardabweichung bei steigender Anzahl richtig gelöster Aufgaben über die drei Messzeitpunkte etwa konstant bei 25 liegt. Auch wenn sich damit die Streuung über die Zeit der drei Messzeitpunkte im Verhältnis zum ansteigenden Mittelwert verringert, ist festzuhalten, dass der Wert der Standardabweichung zu den einzelnen Messzeitpunkten hoch ist, d. h. dass die Untersuchungsgruppe

hinsichtlich des gemessenen Merkmals heterogen zusammengesetzt ist.

Analog zum Gruppentest des IEL-1 werden für den Individualtest Eine-Minute-Lese-Aufgabe zunächst die Häufigkeitsverteilungen, Mittelwerte und die Standardabweichungen relevanter Kennwerte betrachtet. Die Abbildungen 5, 6 und 7 zeigen die Verteilungen der Anzahl richtig gelesener Wörter innerhalb von einer Minute zu den drei Messzeitpunkten.

Aus Abbildung 5 wird ersichtlich, dass die Verteilung der absoluten Häufigkeit richtig gelesener Wörter zum 1. Messzeitpunkt des Subtests Eine-Minute-Lesen einer Normalverteilung ähnlich sind. Die Kinder haben im Mittel 12.4 von 60 Wörtern bei einer Standardabweichung von  $SD = 8.2$  richtig gele-

Tab. 2: Mittelwerte und Standardabweichung der Rohwerte für die Gesamtuntersuchungsgruppe zu den drei Messzeitpunkten

	<b>N</b>	<b>M</b>	<b>SD</b>	<b>Min</b>	<b>Max</b>
1.MZP	1508	64.3	24.7	0	129
2.MZP	1461	87.8	25.2	6	135
3.MZP	1377	96.0	25.0	3	135

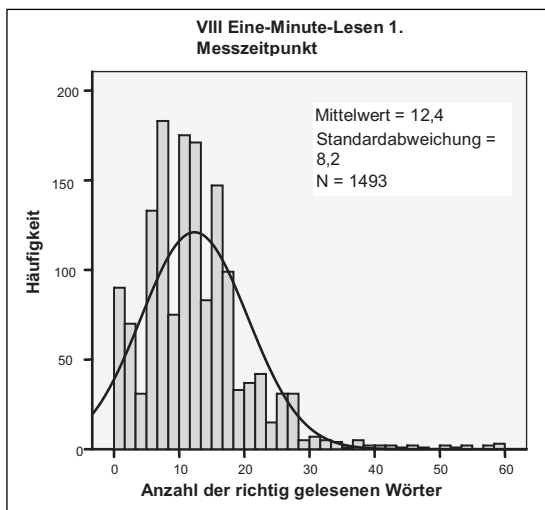


Abb. 5: Absolute Häufigkeiten der Anzahl richtig gelesener Wörter 1. Messzeitpunkt

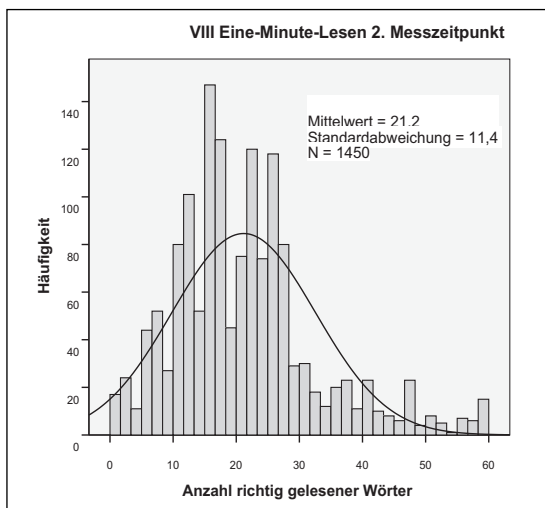


Abb. 6: Absolute Häufigkeiten der Anzahl richtig gelesener Wörter 2. Messzeitpunkt

sen. Die Verteilung zeigt am linken Verteilungsende einen leichten Bodeneffekt.

Der Abbildung 6 ist zu entnehmen, dass die absoluten Häufigkeiten richtig gelesener Wörter im Subtest Eine-Minute-Lesen zum 2. Messzeitpunkt glockenförmig verteilt sind. Die Kinder haben im Mittel 21,2 Wörter bei einer Standardabweichung von  $SD = 11,4$  richtig gelesen. Die Gruppe der untersuchten Kinder hat sich also im Vergleich zum 1. Messzeitpunkt deutlich in ihrer Lesefertigkeit

gesteigert. Die linkssteile Verteilung, wie sie sich beim 1. Messzeitpunkt zeigte, hebt sich zu Gunsten einer in etwa glockenförmigen Verteilung auf.

Die Abbildung 7 zeigt, dass sich die Verteilung der absoluten Häufigkeiten richtig gelesener Wörter des Subtests Eine-Minute-Lesen zum 3. Messzeitpunkt augenscheinlich als annähernd normal verteilt darstellt, wobei auch hier die Ergebnisse des Kolmogorov-Smirnov-Anpassungstests diese Annahme zu-

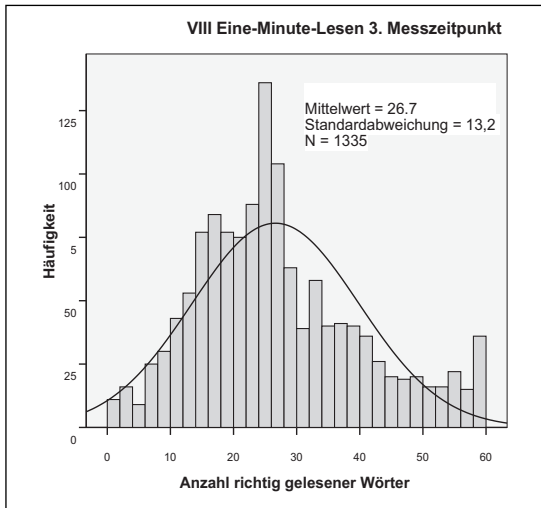


Abb. 7: Absolute Häufigkeiten der Anzahl richtig gelesener Wörter 3. Messzeitpunkt

rückweisen ( $p < .01$ , 2-seitig). Die Kinder lesen im Mittel 26.7 Wörter bei einer Standardabweichung von  $SD = 13.2$  in einer Minute korrekt. In der vorgegebenen Zeit wurden im Vergleich zum 1. Messzeitpunkt 14.2 Wörter und im Vergleich zum 2. Messzeitpunkt 5.5 Wörter mehr richtig gelesen. Auffallend ist zu allen drei Messzeitpunkten eine relativ hohe Streuung von  $SD = 8.2$  bis  $SD = 13.2$ . Die nachfolgende Tabelle 3 gibt die bereits genannten Mittel- und Streuungswerte der drei Messzeitpunkte wieder. Zusätzlich werden der Median, Angaben zu Minima und Maxi-

ma der richtig, falsch und gesamt gelesenen Wörter (Summe aus richtig und falsch gelesenen Wörtern) sowie deren Standardabweichung mitgeteilt.

Auffallend an den Werten in Tabelle 3 ist die relativ hohe Standardabweichung unter Berücksichtigung der zu jedem Messzeitpunkt im Mittel erreichten Rohwertpunkte und die breite Spannweite zwischen Minima und Maxima. Das bestätigt den Befund einer, hinsichtlich ihrer Leistungen, heterogenen Untersuchungsgruppe.

Tab. 3: Eine-Minute-Lesen. Richtig, falsch und gesamt gelesene Wörter: Median, Mittelwerte, Minimum/Maximum und Standardabweichung über die gesamte Stichprobe zu den drei Messzeitpunkten

	1. MZP N = 1493				2. MZP N = 1450				3. MZP N = 1335			
	Median	M	Min/Max	SD	Median	M	Min/Max	SD	Median	M	Min/Max	SD
richtig gelesene Wörter	11	12.4	0/60	8.2	20	21.2	0/60	11.4	25	26.7	0/60	13.2
falsch gelesene Wörter	2	2.4	0/18	2.4	2	2.5	0/32	2.5	2	3.0	0/21	2.7
gesamt gelesene Wörter	14	14.8	0/60	8.4	22	23.7	0/60	11.3	28	29.7	0/60	13.1

**Itemanalyse, innere Konsistenz, Testhalbierungsreliabilität und Retestrelia-  
bilität Gesamttest IEL-1  
und Zusatztest Eine-Minute-Lesen**

Im Folgenden werden die einzelnen Testaufgaben des IEL-1 mittels der Bestimmung der Schwierigkeitsgrade, der Trennschärfe sowie durch die Prüfung der inneren Konsistenz beschrieben. Zudem wird der Standardmessfehler des Verfahrens abgeschätzt. Eine Itemanalyse für den Zusatztest *Eine-Minute-Lese-Aufgabe* wurde aus forschungsökonomischen Gründen im Rahmen dieser Erhebung nicht durchgeführt. Um dennoch Angaben zur Reliabilität des Zusatztests geben zu können, wurden Korrelationen über die drei Messzeitpunkte berechnet.

In Tabelle 4 werden die genannten Kennwerte für den 1. Messzeitpunkt auf Subtestniveau und für das gesamte Messverfahren dargestellt.

Insgesamt ist anhand der Werte zur Itemschwierigkeit zum 1. MZP festzustellen, dass in jedem Subtest, mit Ausnahme des Subtests G, sehr leichte und sehr schwere Aufgaben enthalten sind. Der Gesamttest beinhaltet mittelschwere Aufgaben mit einem gemittelten Schwierigkeitsgrad von .41. Bei einem Alpha von .97 und einem Split-Half-Reliabilitätskoeffizienten von .88 für das gesamte Inventar kann das IEL-1 zum 1. Messzeitpunkt 1 als reliables Messverfahren eingeschätzt werden.

Die Tabelle 5 zeigt, dass sich die Werte erwartungsgemäß zum 2. Messzeitpunkt dahingehend verändern, dass die Aufgaben insgesamt für die Kinder leichter zu lösen waren. Der gemittelte Schwierigkeitsgrad für das gesamte Verfahren liegt nun bei  $P = .60$ . Die Reliabilität des IEL-1 zum 2. Messzeitpunkt ist bei einem  $\alpha = .97$  und einem Split-Half-Reliabilitätskoeffizienten von .88 als sehr hoch einzuschätzen. Subtest A beinhaltet die leichtes-

Tab. 4: Kennwerte für die Items des IEL-1, MZP 1

Items aus Subtest	Itemschwierigkeit Min-Max / M	Trennschärfe Min-Max / M	Schätzungen der Reliabilität	
			$\alpha$	Split-Half
Buchstaben-Laut-Zuordnung (Subtest A)	.03-.96 / .74	-.08-.56 / .40	.92	.85
Wörter in Silben segmentieren (Subtest B)	.01-.85 / .25	.17-.54 / .33	.84	.71
Silben zu Wörtern verbinden (Subtest C)	.06-.74 / .45	.29-.59 / .47	.88	.71
Buchstaben analysieren und synthetisieren (Subtest D)	.07-.80 / .33	.25-.62 / .47	.89	.86
Wörter lesen, schreiben und identifizieren (Subtest E)	.13-.82 / .49	.42-.65 / .54	.88	.81
Satzlesen (Subtest F)	.06-.82 / .47	.14-.53 / .41	.70	.65
Textlesen (Subtest G)	.11-.17 / .13	.20-.31 / .26	.65	.64
Gesamttest	.01-.96 / .41	-.08-.65 / .41	.97	.88

Erläuterung:  $\alpha$  = Cronbachs Alpha, Split-Half (Spearman-Brown-Koeffizient)



Tab. 5: Kennwerte für die Items des IEL-1, MZP 2

Items aus Subtest	Itemschwierigkeit	Trennschärfe	Schätzung der Reliabilität	
			$\alpha$	Split-Half
	Min-Max / M	Min-Max / M		
Buchstaben-Laut-Zuordnung (Subtest A)	.02-.98 / .84	-.16-.55 / .34	.89	.82
Wörter in Silben segmentieren (Subtest B)	.11-.93 / .40	.24-.58 / .39	.90	.83
Silben zu Wörtern verbinden (Subtest C)	.26-.89 / .67	.36-.59 / .50	.89	.75
Buchstaben analysieren und synthetisieren (Subtest D)	.03-.93 / .58	.39-.65 / .53	.91	.82
Wörter lesen, schreiben und identifizieren (Subtest E)	.42-.92 / .74	.30-.63 / .54	.88	.80
Satzlesen (Subtest F)	.31-.93 / .67	.03-.57 / .42	.73	.71..
Textlesen (Subtest G)	.28-.42 / .33	.33-.41 / .37	.69	.63
Gesamttest	.02-.98 / .60	-.16-.65 / .49	.97	.88

Erläuterung:  $\alpha$  = Cronbachs Alpha, Split-Half (Spearman-Brown-Koeffizient)

ten Aufgaben mit einem gemittelten Schwierigkeitsindex von .84. Demgegenüber beinhaltet Subtest G, mit einem gemittelten Schwierigkeitsindex von .33, die schwierigsten Aufgaben. Die innere Konsistenz ist bei allen Subtests, mit Ausnahme der Subtests F und G, als gut bis sehr gut zu bewerten. Die Subtests A, B und G zeigen recht niedrige Trennschärfekoeffizienten.

Der Tabelle 6 ist zu entnehmen, dass sich zum 3. Messzeitpunkt eine weitere Veränderung hinsichtlich der gemittelten Schwierigkeitsindizes der Subtests zeigt. Diese erhöhen sich in allen Subtests und folglich auch im Gesamtwert. Die Reliabilitäten der Subtests F und G sind mit einem  $\alpha = .73$  als hinreichend zuverlässig einzuschätzen. Alle anderen Subtests zeigen eine hohe Reliabilität mit Werten zwischen .88 - .92. Die Reliabilität für den Gesamttest ist zum 3. Messzeitpunkt mit einem  $\alpha = .97$  und einem gemittel-

ten Split-Half-Reliabilitätskoeffizienten von .87 als hoch einzuschätzen.

Die Tabelle 7 zeigt die Werte zur Retestreliabilität. Es wird deutlich, dass zwischen den Ergebnissen zu den drei Messzeitpunkten ( $p < 0,01$ ) signifikante Zusammenhänge bestehen. Auch wenn die Werte zur Retestreliabilität eher niedrig ausfallen, können sie in Anbetracht der Tatsache, dass zwischen den Messzeitpunkten *mehr als 10 Wochen* liegen, als erwartungsgemäß und damit zufriedenstellend eingeschätzt werden. Sie bieten einen weiteren Hinweis darauf, dass das IEL-1 als reliabel eingeschätzt werden kann.

Nachfolgend werden Angaben zum Standardmessfehler gemacht. Für den ersten Messzeitpunkt beträgt der Standardmessfehler  $Se = 8,4$ , zum zweiten Messzeitpunkt  $Se = 8,5$  und zum dritten Messzeitpunkt bei  $Se = 8,6$ .

Tab. 6: Kennwerte für die Items des IEL-1, MZP 3

Subtests	Itemschwierigkeit	Trennschärfe	Schätzung der Reliabilität	
	Min-Max / M	Min-Max / M	$\alpha$	Split-Half
Buchstaben-Laut-Zuordnung (Subtest A)	.01-.98 / .86	-.09-.54 / .37	.91	.84.
Wörter in Silben segmentieren (Subtest B)	.25-.96 / .46	.21-.58 / .41	.92	.85
Silben zu Wörtern verbinden (Subtest C)	.41-.93 / .74	.35-.62 / .51	.89	.79
Buchstaben analysieren und synthetisieren (Subtest D)	.38-.94 / .68	.40-.69 / .54	.92	.84
Wörter lesen, schreiben und identifizieren (Subtest E)	.53-.94 / .80	.30-.61 / .53	.88	.80
Satzlesen (Subtest F)	.46-.97 / .75	.01-.57 / .40	.73	.73
Textlesen (Subtest G)	.39-.55 / .46	.28-.44 / .37	.70	.64
Gesamttest	.01-.98 / .67	-.09-.69 / .45	.97	.87

Erläuterung:  $\alpha$  = Cronbachs Alpha, Split-Half (Spearman-Brown-Koeffizient)

Tab. 7: Retestreliabilität des IEL-1 Gesamttests über die drei Messzeitpunkte

	IEL Gesamt MZP 1	IEL Gesamt MZP 2	IEL Gesamt MZP 3
IEL Gesamt MZP 1			
IEL Gesamt MZP 2	.79**		
IEL Gesamt MZP 3	.74**	.83**	

Angaben zur Retestreliabilität des Zusatztests Eine-Minute-Lesen über die drei Messzeitpunkte werden in Tabelle 8 dargestellt.

Wie aus Tabelle 8 ersichtlich wird, bestehen zwischen den Ergebnissen zu den drei Messzeitpunkten signifikante Zusammenhänge ( $p < .01$ ). Auch hier sind die Werte zur Retestreliabilität aufgrund der relativ langen Zeiträume zwischen den Messzeitpunkten erwartungsgemäß eher niedrig und damit aber dennoch zufrieden stellend. Zwar liefern die Angaben zur Retestreliabilität nach über 10 Wochen nur Hinweise zur Einschätzung der Zuverlässigkeit der Messergebnisse

mit diesem Zusatztest, diese sprechen aber für eine ausreichende Reliabilität des Verfahrens.

### Ergebnisse zur Validität

Die kriteriumsbezogene Validität wurde durch Berechnungen der Korrelation zwischen den Testwerten des IEL-1 bzw. des Zusatztests Eine-Minute-Lesen und der WLLP (Küspert & Schneider, 1998) sowie dem Lehrerurteil zur Lesefertigkeit/Leseverständnis zum dritten Messzeitpunkt geprüft. Die so er-

Tab. 8: Retestreliabilität des Zusatztests Eine-Minute-Lesen (richtig gelesene Wörter) über die drei Messzeitpunkte

	Eine-Minute-Lesen MZP 1 N = 1493	Eine-Minute-Lesen MZP 2 N = 1329	Eine-Minute-Lesen MZP 3 N = 1210
Eine-Minute-Lesen MZP 1 N = 1493			
Eine-Minute-Lesen MZP 2 N = 1329	.76**		
Eine-Minute-Lesen MZP 3 N = 1210	.70**	.86**	

Erläuterung: \*\*  $p < .01$

mittelten mittleren (.62 bis .70) signifikanten ( $p < .01$ ) Zusammenhänge zwischen den unterschiedlichen Erhebungsverfahren weisen darauf hin, dass es sich bei dem IEL-1 um ein extern valides Verfahren handelt.

Die interne Validität des IEL-1 wurde durch einen Vergleich der Werte der Grundschulklassen mit den Werten der Diagnoseförderklassen (Extremgruppenvergleich) abgeschätzt. Die Entwicklung der Werte in Grundschulklassen im Vergleich zu den Werten in DFK über die drei Messzeitpunkte verläuft relativ parallel auf unterschiedlichen Leistungsniveaus. Insgesamt sprechen die Daten des Extremgruppenvergleichs für die Validität des Verfahrens. Tatsächliche Unterschiede in der Lesekompetenz von Schülern aus Grundschulklassen und DFK werden durch Messergebnisse mit dem IEL-1 abgebildet. Vergleicht man die Mittelwerte der Schüler aus DFK und aus Grundschulklassen mittels t-Test (zweiseitig,  $p < .05$ ) zeigen sich zu allen drei Messzeitpunkten signifikante Mittelwertsdifferenzen.

Um Aussagen zur prognostischen Validität des IEL-1 treffen zu können, wurden die Korrelationen zwischen dem Gesamtwert des IEL-1 zum 1. Messzeitpunkt und den

Werten der WLLP (Rohwertbasis) sowie dem Lehrerurteil zum 3. Messzeitpunkt ermittelt. Die bestehenden mittleren Korrelationen zwischen den Werten des IEL-1 zum 1. Messzeitpunkt und der WLLP ( $r = .58$ ) sowie dem Lehrerurteil/Lesefertigkeit ( $r = .56$ ) und dem Lehrerurteil/Leseverständnis ( $r = .62$ ) zum 3. Messzeitpunkt zeigen, dass zwischen den Verfahren ein Zusammenhang besteht. Zudem wurde, als weitere Möglichkeit, die prognostische Validität des Verfahrens einzuschätzen, eine Regressionsanalyse durchgeführt. Das Gütemaß für die Anpassung der Daten an die lineare Funktion steigerte sich vom 1. zum 3. MZP von  $R^2 = .55$  über  $R^2 = .68$  vom 2. auf den 3. Messzeitpunkt auf  $R^2 = .70$  unter Berücksichtigung sowohl des 1. und des 2. MZP. Daraus lässt sich schließen, dass die beste Vorhersage auf spätere Leistungen durch eine gemeinsame Betrachtung der Werte des 1. und 2. Messzeitpunktes erfolgt.

Die Korrelationen der inhaltlich und curricular begründeten Subskalen liegen (bis auf die Korrelation des Subtests VII (Textlesen) mit anderen Subtests zum ersten MZP) durchgängig im mittleren Bereich. Dies spricht neben inhaltlichen Überlegungen (je-

der Subtest prüft einen Lernzielbereich innerhalb des Leselehrganges) für eine Betrachtung der Ergebnisse im IEL-1 auf Subtestebene. Trotz der hohen Korrelationen der Subtests mit dem Gesamtwert messen die Aufgabenbereiche I bis VII unterschiedliche Teilfertigkeiten, die im Leselernprozess vermittelt werden. Das Verfahren erweist sich als curricular valide. Hierfür sprechen auch die berichteten Veränderungen in den Schwierigkeitsgraden der Aufgaben in den genannten Aufgabenbereichen. Zunächst schwierige Aufgaben nehmen im Schwierigkeitsgrad ab.

### **Ergebnisse zu Messungen von Lernstandsveränderungen**

Entsprechend dem Untersuchungsdesign wiederholen die Schüler im Verlauf des ersten Schuljahres die Bearbeitung des IEL 1 in identischer Form dreimal. Ohne der abschließenden Diskussion vorgreifen zu wollen, kann an dieser Stelle bereits festgehalten werden, dass die bisherigen Ergebnisse nahe legen, das IEL-1 als reliables und valides Verfahren zur Messung von Lesekompetenzen zu drei Messzeitpunkten anzusehen. Zudem differenziert es insbesondere im unteren Leistungsbereich, d. h. tatsächliche Leistungsunterschiede zwischen relativ schwachen Schülern werden zu allen drei Messzeitpunkten aufgezeigt. Notwendige Voraussetzungen für Aussagen zu Veränderungen des Lernstandes zu den drei Messzeitpunkten sind also gegeben. Durch den Vergleich der Mittelwerte und die Darstellung der Verteilung der absoluten Häufigkeiten der „richtig-Differenzen“

von Messzeitpunkt zu Messzeitpunkt sollen weitere Hinweise zur Einschätzung der Anwendungsmöglichkeit „Lernfortschrittsdokumentation“ gewonnen werden. In den folgenden Berechnungen wurden die Schüler (N = 1171) berücksichtigt, die an allen drei Durchführungen teilgenommen haben. Der Lernfortschritt im Lesen soll durch den Vergleich zwischen den Gesamtwerten, zusammengesetzt aus den Subtests A bis G (Gruppentestung) zu den drei Messzeitpunkten beschrieben werden. Die Tabelle 9 zeigt den Lernfortschritt der Gesamtgruppe zu den drei Messzeitpunkten. Sie basiert auf den Mittelwerten des Gesamtwertes des IEL-1.

Die konstante Standardabweichung von etwa  $SD = 24$  Rohwertpunkten über alle Messzeitpunkte führt dazu, dass die Varianz im Verhältnis zum ansteigenden Mittelwert von Durchführung zu Durchführung abnimmt. Die Gruppe weist insgesamt einen Lernfortschritt auf, sie wird hinsichtlich der Merkmalsausprägung homogener.

Die Mittelwertunterschiede sind signifikant (t-Test, zweiseitig,  $\alpha = .05$ ). Es kann davon ausgegangen werden, dass sich die Lesekompetenzen der Schüler vom 1. über den 2. bis zum 3. Messzeitpunkt stetig steigern.

Das Messverfahren IEL-1 ist aus sieben Aufgabenbereichen zusammengesetzt. Eine bereichsspezifische Auswertung der richtig gelösten Aufgaben bietet die Möglichkeit, Aussagen über den Lernverlauf eines Schülers innerhalb der einzelnen Bereiche zu treffen. In der nachfolgenden Abbildung 8 werden die Mittelwerte der richtig gelösten Aufgaben für jeden Aufgabenbereich pro Durchführung aufgezeigt.

Tab. 9: Vergleich absolute Gesamtmittelwerte zu den drei Messzeitpunkten

	<b>M</b>	<b>SD</b>	<b>N</b>
IEL-1 Gesamt MZP 1	66	24,1	1171
IEL-1 Gesamt MZP 2	88	24,4	1171
IEL-1 Gesamt MZP 3	96	24,3	1171

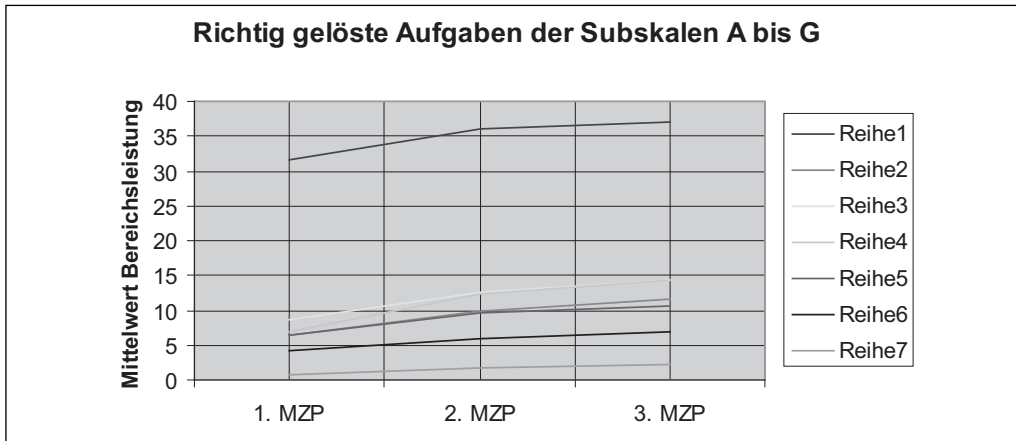


Abb. 8: Bereichsmittelwerte (Rohwertbasis) pro Durchführung

Reihe 1 = Aufgabenbereich A Buchstaben-Laut-Zuordnungen, Reihe 2 = Aufgabenbereich B Wörter in Silben segmentieren, Reihe 3 = Aufgabenbereich C Silben zu Wörtern verbinden, Reihe 4 = Aufgabenbereich D Buchstaben zu Wörtern verbinden, Reihe 5 = Aufgabenbereich E Wörter lesen und schreiben, Reihe 6 = Aufgabenbereich F Sätze, Reihe 7 = Aufgabenbereich G Texte

Abbildung 8 zeigt sehr deutlich, dass die Kinder im Mittel in jedem Bereich von Messzeitpunkt zu Messzeitpunkt mehr Aufgaben richtig lösen konnten. Demnach ist davon auszugehen, dass alle gewählten Aufgabenbereiche einen Beitrag zur Abbildung von Lernfortschritten durch das Gesamtergebnis leisten. Der Lernfortschritt je Bereich verläuft – ersichtlich durch die parallele Anordnung der Verbindungslinien – nahezu identisch. Die Kurven der Aufgabenbereiche F und G zeigen die geringsten Steigerungen. Generell ist zu beobachten, dass die Kurven aller Aufgabebereiche vom 1. zum 2. Messzeitpunkt steiler verlaufen und vom 2. zum 3. Messzeitpunkt nur geringfügig steigen.

Durch die Berechnung der Differenzen zwischen der Anzahl der richtig gelösten Aufgaben (richtig-Differenzen) im Einzelfall zwischen den Messzeitpunkten werden Lernstandsveränderungen sichtbar. Für die Gesamtgruppe wurde die Häufigkeitsverteilung dieser „richtig-Differenzen“ berechnet. Die Tabelle 10 stellt die Verteilung der Lernzuwächse zwischen den Messzeitpunkten dar.

### Diskussion der Ergebnisse

#### **Können mit dem IEL-1 Leistungsunterschiede in der Lesekompetenz zu den unterschiedlichen Messzeitpunkten abgebildet werden?**

Bevor auf diese Frage eingegangen wird, soll die Repräsentativität der Stichprobe diskutiert werden. Da sich die Stichprobe aus allen Erstklässlern der Hansestadt Rostock und der Insel Rügen zusammensetzt und sich die Ergebnisse der ländlichen und der städtischen Regionen nicht unterscheiden, ist davon auszugehen, dass die Häufigkeitsverteilungen der Ergebnisse und deren Kennwerte für Mecklenburg-Vorpommern repräsentativ sind. Auch ist auf Grund der Größe der Stichprobe und der Einbeziehung städtischer und ländlicher Gebiete zu vermuten, dass eine Verallgemeinerung der Werte auf andere Bundesländer zumindest hypothetisch zulässig ist. Zudem sprechen die zufrieden stellenden Werte hinsichtlich der externen Validität (gemessen mit der WLLP) für eine überregionale Aussagekraft von Messergebnissen mit

Tab. 10: Deskriptive Statistik Differenzen der Messzeitpunkte

	<b>N</b>	<b>M</b>	<b>SD</b>
IEL-1 Gesamt MZP 2 - MZP 1	1171	22,6	15,3
IEL-1 Gesamt MZP 3 - MZP 2	1171	8,0	14,1
IEL-1 Gesamt MZP 3 - MZP 1	1171	30,5	17,8

dem IEL-1. Demgegenüber zeigen die Ergebnisse der PISA-Studien, dass es zwischen den einzelnen Bundesländern auffallende Leistungsunterschiede in der Lesekompetenz der 15-jährigen gibt. Daher wird empfohlen, weitere Überprüfungen der Kennwerte des IEL-1 in anderen Bundesländern und hier insbesondere in Ländern mit einem höheren Anteil von Kindern nichtdeutscher Herkunft durchzuführen.

Die Verteilungsformen für den Gesamttest IEL-1 bestätigen die Annahme eingipfliger und tendenziell normal verteilter Werte. Die Verteilungen verschieben sich erwartungsgemäß zum 2. und 3. Messzeitpunkt auf der x-Achse leicht nach rechts. Am linken Rand der Verteilung (unterer Leistungsbereich) steigen die Häufigkeiten von nächsthöheren Werten kontinuierlich an. Die Kinder lösen von Messzeitpunkt zu Messzeitpunkt zunehmend mehr Aufgaben des IEL-1 richtig. Bei Mittelwerten und Standardabweichungen von  $M = 64$  und  $SD = 25.0$  zum 1. Messzeitpunkt, von  $M = 88$  und  $SD = 25.0$  zum 2. Messzeitpunkt und  $M = 96$  und  $SD = 25.0$  zum 3. Messzeitpunkt lassen die Ergebnisse zudem auf Grund einer relativ hohen Streuung der Werte auf Unterscheidungsmöglichkeiten mittels des IEL-1 zwischen unterschiedlich leistungsstarken Schülern schließen. Diese liegen, so wie es von der Konzeption her intendiert war, zu allen drei Messzeitpunkten im unteren Leistungsbereich vor. Die Konzeption des Inventars zielt auf eine objektive und valide Beurteilung früher Lesekompetenzen ab, um Rückstände im Leselernprozess frühzeitig zu erkennen und gezielt Fördermaßnahmen einleiten zu können. Aus son-

derpädagogischer Sicht ist es besonders wichtig, Kinder mit Lernschwierigkeiten hinsichtlich ihrer Leistungen präzise einschätzen zu können. Es liegen zum 1. Messzeitpunkt in etwa 16.5 % der Kinder eine und 1.7 % der Kinder zwei Standardabweichungen unter dem Durchschnitt. Beim 2. Messzeitpunkt sind es 15.8 % der Kinder, die eine und 5.5 % der Kinder, die zwei Standardabweichungen unter dem Durchschnitt liegen. Zum 3. Messzeitpunkt zeigen sich die Werte ähnlich wie zum 2. Messzeitpunkt. Diese Ergebnisse weisen darauf hin, dass ein doch erheblicher Teil der Kinder im Vergleich zur Gesamtuntersuchungsgruppe Schwierigkeiten im Leselernen hat und diese durch den IEL-1-Gesamtwert zu den verschiedenen Messzeitpunkten abgebildet werden können. Auf Grund der erwartungsgemäßen Häufigkeitsverteilungen des IEL-1 ist eine Differenzierung insbesondere im unteren Leistungsbereich zu allen drei Messzeitpunkten gewährleistet. Die Werte des 1. Messzeitpunktes lassen zudem eine gute Differenzierung im oberen Leistungsbereich zu. Das Gesamtergebnis des IEL-1 bietet den Lehrkräften also auch Hinweise darauf, welche Kinder mit sehr guten Ergebnissen im Leselernprozess zu diesem Zeitpunkt stärker zu fordern sind, um bei ihnen den Prozess des Leselernens angemessen voranzubringen. Ergebnisse einer kleineren Untersuchungsgruppe ( $N = 107$ ) zum Messzeitpunkt 2 (rekrutiert aus der Gesamtstichprobe), die das IEL-1 zum ersten Mal durchgeführt haben, erlauben, einen Testwiederholungseffekt zu vernachlässigen. Im Mittel liegt diese Gruppe mit 26 Rohwertpunkten lediglich um einen Rohwertpunkt

niedriger als die Gesamtstichprobe bei einer vergleichbaren Standardabweichung von  $SD = 27$  (Gesamtuntersuchungsgruppe  $SD = 25$ ). Das Ergebnis spricht gegen einen Übungseffekt durch ein wiederholtes Bearbeiten der Aufgaben des IEL-1, und für die Annahme, dass die unterrichtsnahe Art der IEL-Aufgaben und deren Menge zu aussagekräftigen Messergebnissen führen, unabhängig davon, ob das IEL-1 bereits vorher bearbeitet wurde. Dennoch empfiehlt es sich, diese Annahme mit einer unabhängigen Untersuchungsgruppe auch zum 3. Messzeitpunkt zu prüfen. Eine Untersuchung dazu mit einer vergleichbaren Untersuchungsgröße ( $N = 100$ ) findet derzeit statt. Ergebnisse sind im August dieses Jahres zu erwarten.

Mit dem Zusatztest Eine-Minute-Lesen sind zum 1. und 2. Messzeitpunkt gute Differenzierungen im oberen Leistungsbereich möglich, zum 3. Messzeitpunkt lassen die Werte Aussagen sowohl zu Leistungsunterschieden im unteren als auch oberen Bereich zu. Die Werte zeigen, dass zum 1. Messzeitpunkt 12.8 % der Kinder eine Standardabweichung unter dem Durchschnitt liegen und 12.4 % der Kinder eine Standardabweichung sowie 3.1 % der Kinder zwei Standardabweichungen über dem Leistungsdurchschnitt liegen. Zum 2. Messzeitpunkt sind es 14.2 % der Kinder, die eine Standardabweichung unter dem Durchschnitt und 13.2 %, die eine sowie 5.7 %, die zwei Standardabweichungen über dem Durchschnitt liegen. Zum 3. Messzeitpunkt liegen 17 % eine und 0.8 % der Kinder zwei Standardabweichungen unter dem Durchschnitt. 15.1 % der Kinder liegen bei der 3. Messung eine und 5.5 % der Kinder zwei Standardabweichungen über dem Durchschnitt.

### **Wie reliabel misst das IEL-1 die Lesekompetenzen der Erstklässler?**

Die ermittelten Reliabilitätswerte zur inneren Konsistenz (Cronbachs-Alpha) der einzelnen Subtests sind sehr unterschiedlich. Sie schwanken zwischen  $\alpha = .65$  und  $\alpha = .92$ .

Der Gesamttest weist eine hohe innere Konsistenz auf ( $\alpha = .97$ ). Während die Reliabilität des Gesamttests als sehr gut einzuschätzen ist, erweisen sich die Subtests VI und VII als nicht zufriedenstellend reliabel, hingegen die Subtests II, III, IV, V als zufriedenstellend messgenau. Die Werte zur Retestreliabilität sind trotz der relativ langen Zeitspanne (10 Schulwochen) zwischen den Messungen als zufrieden stellend einzuschätzen ( $r_{tt} = .74$  bis  $r_{tt} = .83$ ). Vergleicht man die Werte zur Retestreliabilität mit anderen standardisierten Messverfahren, bei denen die Zeitspanne zwischen den Messwiederholungen annähernd gleich ist (z. B. der WLLP Retestmessung nach vier Monaten  $r_{tt} = .75$ ), weisen die Werte der Retestreliabilität des IEL-1 auf ein als zuverlässig einzuschätzendes Messverfahren hin. Zieht man alle ermittelten Kennwerte zur Reliabilität in Betracht, erweist sich das IEL-1 als zuverlässig. Der Wert zur inneren Konsistenz (Cronbachs-Alpha) auf Gesamtwertebene mit einem Wert von  $.97$  gilt als sehr hoch. Ebenso untermauern die Split-Half-Werte von  $.87$  bis  $.88$  für den Gesamttest IEL-1 zu den drei Messzeitpunkten und die Retestreliabilitätswerte die Zuverlässigkeit des Inventars.

### **Wie valide misst das IEL-1 die Lesekompetenz der Erstklässler?**

Die Validität des IEL-1 begründet sich zunächst durch die spezifische Auswahl der Items. In die Aufgabengestaltung des IEL-1 wurden solche Items einbezogen, die aus gängigen Arbeitsheften und Fibeln der Klasse 1 bekannt sind, und es ist somit davon auszugehen, dass sie Lesefertigkeiten überprüfen. Demnach kann das IEL-1 als inhaltswalide in Bezug auf das zu messende Merkmal Lesekompetenz eingeschätzt werden. Die Items des IEL-1 sind zu inhaltlich begründeten Subtests gebündelt. Die Subtests sollen dabei unterschiedliche Teilfertigkeiten, also mögliche Merkmalsfacetten von Lesekompetenz, abbilden. Die Items sind so angelegt, dass sie basierend auf der Theorie des Kompetenzmo-

dells von Klicpera et al. (2003) einzelne Fertigkeiten, von denen angenommen wird, dass sie zum erfolgreichen Lesen führen, in einem Subtest bündeln. Die durchgeführten Zusammenhangsanalysen bestätigen im Wesentlichen, dass mit dem IEL-1 hauptsächlich eine Fähigkeit gemessen wird, aber auch dass die inhaltlich begründete Zusammensetzung der Subtests vertretbar ist. Die Subtests zeigen mit Ausnahme des Aufgabenbereichs G (Textlesen) untereinander sowie zur Gesamtskala wie erwartet mittlere Korrelationen. Dass der Subtest G eher niedrige Korrelationen mit den weiteren Subtests zeigt, könnte daran liegen, dass die Aufgabenstellung für die Kinder zum 1. und 2. Messzeitpunkt zu schwer war. Es handelt sich um den letzten Aufgabenbereich im Test, möglicherweise beeinflussen Faktoren wie Konzentration und Aufmerksamkeitsspanne die Ergebnisse. Zudem könnte die relativ hohe Ratewahrscheinlichkeit (multiple-choice mit drei Auswahlmöglichkeiten) die Ergebnisse verfälschen.

Dass es sich bei der gemessenen Fähigkeit um Lesekompetenz handelt, bestätigt auch die mittlere Korrelation zwischen dem IEL-1 und der WLLP von Küspert und Schneider (1998) sowie dem Lehrerurteil. Das IEL-1 erweist sich hiernach als extern valide. Es mag sich die Frage stellen, warum ein recht zeitintensives Verfahren entwickelt wurde, wenn das ökonomischere Verfahren, die WLLP, zu ähnlichen Ergebnissen kommt. Die WLLP ist ein Speed-Test und für die Überprüfung der Lesefertigkeit konzipiert. Die Daten der WLLP geben keinerlei Hinweis darauf, welche Teilfertigkeit möglicherweise unzureichend ausgebildet ist. Die Stärken der WLLP liegen in der sehr ökonomischen Durchführung und Auswertung. Das Verfahren führt effizient zu einer validen Einschätzung der Lesefertigkeit. Mit dem IEL-1 sollen neben dem Erfassen der Lesefertigkeit auf Wortebene auch andere Teilfertigkeiten wie z. B. die phonematische Differenzierungsfähigkeit, das Silbensegmentieren und die Synthese von Silben zu Wörtern sowie das Satz- und

Textlesen erfasst werden. Darüber hinaus galt es, die Leistungsstände in den genannten Teilbereichen über drei Messzeitpunkte abzubilden. Die konzeptionellen Unterschiede der WLLP und des IEL-1 könnten als Erklärung für die ‚nur‘ mittleren Korrelationen zwischen den Tests herangezogen werden.

Zwischen der WLLP und dem Zusatztest Eine-Minute-Lesen besteht ebenfalls eine mittlere Korrelation. Auf Grund der Konzeption beider Verfahren, die Lesefertigkeit über die Geschwindigkeit zu messen, könnte eine noch höhere Korrelation vermutet werden. Eine Erklärung für die dennoch nur mittleren Zusammenhänge wird in der unterschiedlichen Auswahl und Reihenfolge der Items gesehen. Bei dem Zusatztest Eine-Minute-Lesen wurden die Items nach linguistischen Merkmalen ausgewählt und entsprechend einer Schwierigkeitssteigerung angeordnet. Das Wortmaterial der WLLP stammt größtenteils aus dem bayerischen Grundwortschatz. Die Reihenfolge der Items wurde per Zufall erstellt. Zudem ist zu beachten, dass die Aufgabenstellung bei der WLLP, neben der Anforderung, ein Wort zu lesen, auch eine optische Diskriminationsleistung beinhaltet. Diese zusätzliche Anforderung kann ein weiterer Grund für unterschiedliche Leistungen in der WLLP und der Eine-Minute-Lesen-Aufgabe sein. Die Korrelationen des Zusatztests Eine-Minute-Lesen mit dem Lehrerurteil liegen ebenfalls im mittleren Bereich und sind als zufrieden stellender Hinweis auf die Validität des Verfahrens (hier: externe Validität) anzusehen.

### **Können mit Hilfe der wiederholten Messungen über drei Messzeitpunkte Veränderungen des Lernstandes mit dem IEL-1 abgebildet werden?**

Die Häufigkeitsverteilungen zu den drei Messzeitpunkten, ihre Kennwerte sowie die deutlichen Mittelwertunterschiede zwischen den Messzeitpunkten legen die Vermutung nahe, dass mit den IEL-1-Ergebnissen zu mehreren Messzeitpunkten Hinweise auf eine



Abschätzung der Lernfortschritte gegeben sind. Methodenkritisch betrachtet, kann allerdings nur von Annäherungen an eine Messung des Lernfortschritts mittels des IEL-1 ausgegangen werden. Auch wenn die Voraussetzungen für eine valide Erfassung der Ist-Stände gewährleistet sind, ist das Verfahren als solches, auf Grund messtheoretischer Einwände, nicht als ausreichend reliabel für Veränderungs-messungen einzuschätzen (Rost, 2004, S. 272). Es können somit lediglich Aussagen darüber getroffen werden, wie die Leistung des einzelnen Schülers oder einer Schülergruppe im Vergleich zur Eichstichprobe zu einem bestimmten Zeitpunkt ausfällt. Zudem kann im Einzelfall das Antwortverhalten des Kindes zu den einzelnen Messzeitpunkten qualitativ analysiert und verglichen werden. Hierbei können quantitative Aspekte berücksichtigt werden. Durch Berechnungen der „richtig Differenzen“ können Leistungsveränderungen im Lesen eingeschätzt werden. Einschränkung ist anzumerken, dass es sich hierbei lediglich um Annäherungen zur Einschätzung von Leistungsveränderungen handelt. Präzise Aussagen über Lernfortschritte würden eine messtheoretische Fundierung des IEL-1 auf Basis von Rasch-Modellen voraussetzen (Rost, 2004, S. 272). Es kann nicht davon ausgegangen werden, dass es sich bei dem IEL-1 um ein valides Verfahren zur Lernfortschrittsmessung handelt. Mit dem IEL-1 können *Lernstandsveränderungen* über drei Messzeitpunkte valide dokumentiert werden.

## Fazit

Zusammenfassend kann das IEL-1 als objektiv sowie ausreichend reliabel und valide eingeschätzt werden. Mit dem IEL-1 ist es demnach möglich, valide Aussagen über die Lesekompetenz von Erstklässlern zu drei Messzeitpunkten zu treffen. Insbesondere im unteren Leistungsbereich differenziert das IEL-1 zu allen drei Messzeitpunkten. In Relation zu einer Eichstichprobe kann der Lehrer mit

dem IEL-1 erstmals zum Ende des ersten Schulhalbjahres Kinder mit geringen Leseleistungen identifizieren. Durch die wiederholten Ist-Stand-Messungen und deren Vergleich mit einer Eichstichprobe gewinnt der Anwender des IEL-1 einen ersten Hinweis darauf, ob Lernfortschritte bei den Kindern auftreten. Als theoriegeleitetes und curricularbasiertes Messverfahren bietet es den Lehrkräften wichtige Informationen über den Leistungsstand in unterschiedlichen Teilfertigkeiten des Leselernprozesses zu mehreren Messzeitpunkten. Im Kontext von Förderdiagnostik sollten nach einer weiteren Prüfung der Generalisierbarkeit der hier ermittelten Daten Forschungsbemühungen zur Entwicklung und Evaluation von Fördermaterialien ansetzen, die auf durch das IEL-1 erfasste Teilfertigkeiten abheben. So ein „*whole-in-one-Paket*“ (Hartke, Diehl & Vrban, 2008) könnte maßgeblich dazu beitragen, Kinder mit Schwierigkeiten im Lesenlernen angemessen und sinnvoll zu unterstützen. Als Teil eines Multiple Assessments innerhalb des RTI-Ansatzes, verstanden als Mehrebenen-Präventions-Programm (Diehl & Hartke, im Druck), kann es sowohl wertvolle Daten zur formativen Evaluation bisheriger Förderung, zur Bestimmung der Präventionsebene, auf dem das jeweilige Kind gefördert werden sollte, und zur Indikation evidenzbasierter Fördermaßnahmen bieten.

## Literatur

- Balzer, L., Fritz, A., Ricken, G. & Jäger, R.S. (2007). Der Rechenschwäche auf der Spur – eine Re-Analyse von Mathematik-Leistungsdaten eines kompletten Schülerjahrgangs der achten Klassenstufe in Rheinland-Pfalz. *Psychologie in Erziehung und Unterricht*, 54, 177-190.
- Bamonto Graney, S. & Shinn, M. (2005). Effects of Reading Curriculum-Based Measurement (R-CBM) Teacher Feedback in General Education Classrooms. *School Psychology Review*, 34, 184-201.

- Büttner, G. (2003a). Gedächtnis: Neurowissenschaftliche Grundlagen – Störungen und Therapie – Aspekte der Entwicklung – Funktionsweise des Arbeitsgedächtnisses. *Sprache – Stimme – Gehör*, 27, 1-2.
- Büttner, G. (2003b). Gedächtnisentwicklung im Kindes- und Jugendalter. *Sprache – Stimme – Gehör*, 27, 24-30.
- Compton, D.L., Fuchs, D., Fuchs, L.S. & Bryant, J.D. (2006). Selecting at-risk readers in first grade for early intervention: A two-year longitudinal study of decision rules and procedures. *Journal of Educational Psychology*, 98, 394-409.
- Dehn, M. & Sjölin, A. (1996). Frühes Lesen und Schreiben. In H. Günther & O. Ludwig (Hrsg.), *Schrift und Schriftlichkeit. Handbücher zur Sprach- und Kommunikationswissenschaft* (Band 10, Halbband 2, S. 1141-1152). Berlin: de Gruyter.
- Deno, S.L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children*, 52, 219-232.
- Deno, S.L. (2003a). Curriculum-Based Measures: Development and Perspectives. *Assessment for Effective Intervention*, 28, 3-12.
- Deno, S.L. (2003b). Developments in Curriculum-Based Measurement. *The Journal of Special Education*, 37, 184-192.
- Diehl, K. & Hartke, B. (2007). Curriculumnahe Lernfortschrittsmessungen. *Sonderpädagogik*, 37, 195-211.
- Diehl, K. & Hartke, B. (im Druck). *Schulische Prävention im Bereich Lernen*. Stuttgart: Kohlhammer.
- Diehl, K. (2009a). *Schriftspracherwerb und Lernfortschrittsdokumentation*. Saarbrücken: svh.
- Diehl, K. (2009b). Das Inventar zur Erfassung der Lesekompetenzen von Erstklässlern (IEL-1) – theoretische Grundlagen sowie empirische Befunde zur Validität. *Die Sprachheilarbeit. Fachzeitschrift für Sprachheilpädagogik und Sprachtherapie*, 54, 87-107.
- Dinges, E. (2002). Systematische Beurteilung und Förderung schulischer Leistungen. *Hornburg: Persen*.
- Fuchs, D., Fuchs, L.S. & Compton, D.L. (2004). Identifying reading disabilities by responsiveness-to-instruction: specifying measures and criteria. *Learning Disabilities Quarterly*, 27, 216-227.
- Fuchs, L.S. (2004). The Past, Present, and Future of Curriculum-Based Measurement Research. *School Psychology Review*, 33, 188-192.
- Fuchs, D. & Fuchs, L.S. (2005). Peer-Assisting Learning Strategies: Promoting Word Recognition, Fluency, and Reading Comprehension in Young Children. *The Journal of Special Education*, 39, 34-44.
- Fuchs, L. S. & Fuchs, D. (1986). Effects of systematic formative evaluation: A meta-analysis. *Exceptional Children*, 53 (3), 199-208.
- Groeben, N. & Hurrelmann, B. (2006). *Empirische Unterrichtsforschung in der Literatur- und Lesedidaktik. Ein Weiterbildungsprogramm*. Weinheim: Juventa.
- Grube, D. & Hasselhorn, M. (2006). Längsschnittliche Analysen zur Leseschreib- und Mathematikleistung im Grundschulalter: zur Rolle von Vorwissen, Intelligenz, phonologischem Arbeitsgedächtnis und phonologischer Bewusstheit. In I. Hosenfeld & F.-W. Schrader (Hrsg.), *Schulische Leistung. Grundlagen, Bedingungen, Perspektiven* (S.87-105). Münster: Waxmann.
- Haager, D., Klingner, J. K. & Vaughn, S. (Eds.) (2007). *Validated reading practices for three tiers of intervention*. Baltimore, MD: Brookes.
- Hartke, B., Diehl, K. & Vrban, R. (2008). Planungshilfen zur schulischen Prävention – Früherkennung und Intervention bei Lern- und Verhaltensproblemen. In J. Borchert, B. Hartke & P. Jogschies (Hrsg.), *Frühe Förderung entwicklungsauffälliger Kinder und Jugendlicher* (S. 218-234). Stuttgart: Kohlhammer.
- Hartmann, E. (2008). Konzeption und Diagnostik von schriftsprachlichen Lernstörungen im Responsiveness-to-Intervention-Modell: eine kritische Würdigung. In: *VHN* 77, 123-137.
- Klauer, K.J. (2006). Erfassung des Lernfortschritts durch curriculum-basierte Messung. *Heilpädagogische Forschung*, 32, 16-26.
- Klicpera, C. & Gasteiger-Klicpera, B. (1995). *Psychologie der Lese- und Schreibschwierigkeiten*. Beltz: Weinheim.
- Klicpera, C., Schabmann, A. & Gasteiger-Klicpera, B. (2003). *Legasthenie. Modelle, Diagnostik*.

- se, Therapie und Förderung. München: Reinhardt.
- Koch, K. (2007). Die wissenschaftliche Begleitung des Projektes Primarstufe in Mecklenburg-Vorpommern – Bilanz, ausgewählte Ergebnisse und Zukunft der Studie. In K. Diehl & K. Koch (Hrsg.), Neues aus Forschung und Lehre. 2. Fachtagung des Instituts für Sonderpädagogische Entwicklungsförderung und Rehabilitation (ISER) (S. 34-51). Universität Rostock.
- Koch, K., Hartke, B. & Blumenthal, Y. (2008). Die Lernausgangslage von Kindern mit besonderem Förderbedarf in Grundschulklassen 1 und Diagnoseförderklassen. Rostock: Universität, Institut für Sonderpädagogische Entwicklungsförderung und Rehabilitation.
- Küspert, P. & Schneider, W. (1998). Würzburger Leise Leseprobe (WLLP). Ein Gruppentest für die Grundschule. Göttingen: Hogrefe.
- Lenhard, W. & Schneider, W. (2006). Elfe 1-6. Ein Leseverständnistest für Erst- bis Sechsklässler. Göttingen: Hogrefe.
- Meiers, K. (1996). Aspekte und Probleme des Leseunterrichts: Erstlesen. In H. Günther & O. Ludwig (Hrsg.), Schrift und Schriftlichkeit. Handbücher zur Sprach- und Kommunikationswissenschaft (Band 10, Halbband 2, S. 1217-1224). Berlin: de.
- Moser Opitz, E. (2007). Rechenschwäche/Dyskalkulie. Theoretische Klärungen und empirische Studien an betroffenen Schülerinnen und Schülern. Bern: Verlag Paul Haupt.
- National Center on Response to Intervention. [www.rti4success.org](http://www.rti4success.org). Zugriff am 14.12.2010.
- Rost, J. (2004). Lehrbuch Testtheorie – Testkonstruktion (2. überarbeitete und erweiterte Auflage). Bern: Hans Huber.
- Scheerer-Neumann, G. (1981). Prozessanalyse der Leseschwäche. In R. Valtin, U. Jung & G. Scheerer-Neumann (Hrsg.), Legasthenie in Wissenschaft und Unterricht (S. 183-209). Darmstadt: Wissenschaftliche Buchgesellschaft.
- Scheerer-Neumann, G. (1995). Lesenlernen: Entwicklungsprozesse und Probleme (Potsdamer Studien zur Grundschulforschung (Heft 4). Potsdam: Universität.
- Scriven, M. (1972). The Methodology of Evaluation, M. Scriven, In R.W. Tyler, R.M. Gagné & M. Scriven (Hrsg.), Perspectives of Curriculum Evaluation (5. Aufl., S. 39-83). USA: Rand McNally & Company.
- Smith, P.K., Pepler, D. & Rigby, K. (2004). Bullying in schools: How successful can interventions be? Cambridge: Cambridge University Press.
- Stecker, P. M. & Fuchs, L. S. (2000). Effecting superior outcomes using curriculum-based measurement: The importance of individual progress monitoring. Learning Disabilities Research and Practice, 15, 128-143.
- Stern, T. (2004). Neue Wege zu einer förderlichen Leistungsbewertung. Eine neue Lernkultur braucht eine neue Prüfungskultur. Lernende Schule, 7 (28), 16-21.
- Strathmann, A. & Klauer, K. J. (2008). Diagnostik des Lernverlaufs. Eine Pilotstudie am Beispiel der Entwicklung der Rechtschreibkompetenz. Sonderpädagogik, 38, 5-25.
- Strathmann, A. & Klauer, K. J. (2010). Lernverlaufsdiagnostik: Ein Ansatz zur längerfristigen Lernfortschrittsmessung. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 42, 111-122.
- Strathmann, A., Klauer, K. J. & Greisbach, M. (2010). Lernverlaufsdiagnostik. Dargestellt am Beispiel der Rechtschreibkompetenz in der Grundschule. Empirische Sonderpädagogik, 2, 64-77.
- Walter, J. (1996). Förderung bei Lese- und Rechtschreibschwäche. Göttingen: Hogrefe.
- Walter, J. (2001). Förderung bei Lese- und Rechtschreibschwäche (2. unveränderte Auflage). Göttingen: Hogrefe.
- Walter, J. (2008a). Adaptiver Unterricht erneut betrachtet: Über die Notwendigkeit systematischer formativer Evaluation von Lehr- und Lernprozessen und die daraus resultierende Diagnostik und Neudefinition von Lernstörungen nach dem RTI-Paradigma. Zeitschrift für Heilpädagogik, 59, 202-215.
- Walter, J. (2008b). Curriculumbasiertes Messen (CBM) als lernprozessbegleitende Diagnostik: Erste deutschsprachige Ergebnisse zur Validität, Reliabilität und Veränderungssensibilität eines robusten Indikators zur Lernfortschrittsmessung beim Lesen. Heilpädagogische Forschung, 34, 62-79.

- Walter, J. (2009a). Theorie und Praxis Curriculumbasierter Messens (CBM) in Unterricht und Förderung. *Zeitschrift für Heilpädagogik*, 60, 162-170.
- Walter, J. (2009b). Eignet sich die Messtechnik „MAZE“ zur Erfassung von Lesekompetenzen als lernprozessbegleitende Diagnostik? *Heilpädagogische Forschung*, 35, 62-75.
- Willson, V.L. & Rupley, W.H. (1997). A Structural Equation Model for Reading Comprehension Based on Background, Phonemic, and Strategy Knowledge. *Scientific Studies of Reading*, 1, 45-63.

### **Anschriften der Autoren:**

*DR. KIRSTEN DIEHL*  
*Institut für sonderpädagogische*  
*Entwicklungsförderung und Rehabilitation*  
*August-Bebel-Str. 28*  
*18051 Rostock*  
*kirsten.diehl@uni-rostock.de*

*PROF. DR. BODO HARTKE*  
*Institut für sonderpädagogische*  
*Entwicklungsförderung und Rehabilitation*  
*August-Bebel-Str. 28*  
*18051 Rostock*  
*bodo.hartke@uni-rostock.de*

## Modernes Antiquariat



**50-90%** Preisvorteil für Bücher aus:

Belletristik, Mathematik, Medizin, Musik, Philosophie, Politik, Psychologie, Recht, Religion, Soziologie, Wirtschaft und Zeitgeschichte.

Bücher zum Teil Raritäten in bibliophiler Ausstattung.

➔ *Mehr als 200 Bücher neu eingestellt!*

Versandkostenfrei bei Bestellwert über 20,- Euro, bei geringerem Bestellwert Versandkostenpauschale von 2,- Euro.

[www.modernes-antiquariat.net](http://www.modernes-antiquariat.net)