



**University of  
Zurich**<sup>UZH</sup>

**Zurich Open Repository and  
Archive**

University of Zurich  
University Library  
Strickhofstrasse 39  
CH-8057 Zurich  
[www.zora.uzh.ch](http://www.zora.uzh.ch)

---

Year: 2023

---

**Professionswissen von Studierenden und von Regel- und Förderlehrkräften zu  
Grundlagen des Leserwerbs und zur Diagnostik von Leseschwierigkeiten:  
Entwicklung eines Instruments und Herausforderungen**

Lietz, Meike ; Moser Opitz, Elisabeth ; Stöckli, Meret

DOI: <https://doi.org/10.2440/003-0001>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-239669>

Journal Article

Published Version



The following work is licensed under a Creative Commons: Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) License.

Originally published at:

Lietz, Meike; Moser Opitz, Elisabeth; Stöckli, Meret (2023). Professionswissen von Studierenden und von Regel- und Förderlehrkräften zu Grundlagen des Leserwerbs und zur Diagnostik von Leseschwierigkeiten: Entwicklung eines Instruments und Herausforderungen. *Empirische Sonderpädagogik*, 15(2):103-122.

DOI: <https://doi.org/10.2440/003-0001>

**Empirische Sonderpädagogik**, 2023.15:103-122

DOI <https://doi.org/10.2440/003-0001>

ISSN 1869-4845 (Print) · ISSN 1869-4934 (ebook)

# Professionswissen von Studierenden und von Regel- und Förderlehrkräften zu Grundlagen des Leseerwerbs und zur Diagnostik von Leseschwierigkeiten: Entwicklung eines Instruments und Herausforderungen

*Meike Lietz<sup>a</sup>, Elisabeth Moser Opitz<sup>b</sup> und Meret Stöckli<sup>b</sup>*

<sup>a</sup> Pädagogische Hochschule Bern

<sup>b</sup> Universität Zürich

## Zusammenfassung

Leseschwierigkeiten (LS) gehören zu den häufigsten Lernproblemen und es ist wichtig, dass betroffene Schüler:innen im Unterricht angemessen unterstützt werden. Das erfordert entsprechendes Professionswissen von den Lehrkräften und Instrumente, mit denen deren Wissen valide und reliabel erfasst werden kann. Bislang fehlen solche Instrumente. Im Kontext von inklusivem Unterricht stellt sich zudem die Herausforderung, Instrumente zu entwickeln, mit denen das Professionswissen von unterschiedlich ausgebildeten Lehrkräften (Regellehrkräfte, Förderlehrkräfte) erfasst werden kann. Im Artikel wird ein Instrument vorgestellt und hinsichtlich seiner psychometrischen Qualität analysiert, mit dem das Professionswissen zu Grundlagen des Leseerwerbs und zur Diagnostik von LS erfasst werden kann. In Studie 1 nahmen  $n = 190$  Studierende am Ende des Lehramtstudiums Grundschule,  $n = 139$  Studierende am Anfang des Aufbaustudiums Förderpädagogik und  $n = 114$  Studierende am Ende des Aufbaustudiums Förderpädagogik teil. In Studie 2 wurden  $n = 90$  amtierende Grundschullehrkräfte und  $n = 52$  amtierende Förderlehrkräfte mit einer gekürzten Fragebogenversion befragt. In Studie 1 ist es gelungen, ein Instrument mit 18 Items zu entwickeln, das sich für den Einsatz bei Studierenden der Förderpädagogik eignet, für die Studierenden des Lehramtes Grundschule jedoch nur eingeschränkt. In Studie 2 konnte nur für die Förderlehrkräfte eine Skala mit fünf Items und einer knapp akzeptablen Reliabilität konstruiert werden. Das eingesetzte Instrument scheint vor allem Wissen zu erfassen, das während bzw. kurz nach der Ausbildung vorhanden ist. Abschließend erfolgen Überlegungen zu weiteren notwendigen Forschungsbemühungen zur Erfassung des Professionswissens von Lehrkräften zum Thema LS.

**Schlagwörter:** Lehrkräfte, Professionswissen, Leseschwierigkeiten, Leseerwerb, Diagnostik

## The development of a tool for assessing how well in-service and student regular and special education teachers understand the acquisition of reading skills and are able to diagnose reading difficulties: Process and challenges

### Summary

Reading difficulties (RD) is a common learning problem and it is important that teachers have the skills to support affected students. Valid and reliable assessment instruments can ensure that teachers possess the requisite knowledge. But, there are none. For an instrument to be valid for inclusive education, it also needs to be able to assess the professional knowledge of both general education teachers (GET) and special education teachers (SET). This paper analyzes the psychometric quality of a questionnaire that has been designed to assess the professional knowledge of GET and SET on reading acquisition and the diagnosis of RD. Study 1 included  $n = 190$  student GET at the end of their training,  $n = 139$  student SET at the beginning of their training and  $n = 114$  student SET at the end of their training. In study 2,  $n = 90$  GET and  $n = 52$  SET completed an abbreviated version of the questionnaire. Study 1 resulted in the successful development of an instrument with 18 items suitable for assessing the professional knowledge of SET students, but the instrument had limited validity when assessing GET students. In study 2 a five-item scale with just about acceptable reliability could only be constructed for SET. The instrument seems to primarily capture knowledge related to the (SET) training courses. Finally, reflections on future research into measuring the professional knowledge of teachers on the topic of RD are discussed.

*Keywords:* teachers, professional knowledge, reading, reading disabilities, diagnostics

Leseschwierigkeiten (LS) gehören zu den häufigsten Lernproblemen in der Schule (Galuschka & Schulte-Körne, 2015). Diese wirken sich oft negativ auf die Schulleistungen in verschiedenen Fächern aus (Naumann et al., 2010) und können auch die Alltagsbewältigung sowie die berufliche Integration negativ beeinflussen (Bohnert-Kraus et al., 2021; Sturm, 2012). Es ist deshalb wichtig, dass die Probleme der Betroffenen frühzeitig erkannt und geeignete Fördermaßnahmen eingeleitet werden. Das erfordert von den Lehrkräften entsprechendes professionelles Wissen. In der Forschung zur professionellen Kompetenz von Lehrkräften wird davon ausgegangen, dass „erfolgreiches Unterrichten vor allem aufgrund einer gut vernetzten und umfangreichen domänenspezifischen Wissensbasis ermöglicht wird und dass diese Wissensbasis im Rahmen der strukturierten Lehreraus- und Weiterbildung vermittelbar ist“ (Baumert & Kunter, 2011, S. 30). Dies wird

durch empirische Ergebnisse auch für die Domäne Lesen bestätigt. Studien zeigen, dass die Genauigkeit, mit der Lehrkräfte die Leseleistungen ihrer Schüler:innen einschätzen, bedeutsam ist für deren Leistungszuwachs (Behrmann & Souvignier, 2013; Karing, Pfof & Artelt, 2011; Karing, Matthäi & Artelt, 2011; Rjosk et al., 2011). Von den Lehrkräften wird deshalb ein vertieftes Professionswissen zu Prozessen des Lesens und zur Diagnostik erwartet (Artelt, 2009; Behrmann & Souvignier, 2013). Das betrifft auch die Thematik der LS. Studien dazu liegen erst vereinzelt vor. Hintz und Grünke (2009) untersuchten, ob angehende Lehrkräfte (Sonder- und Regelpädagogik) die Wirkung von Interventionen zum Schriftspracherwerb bei lernschwachen Kindern einschätzen konnten. Schmidt und Schabmann (2016) analysierten, wie Referendar:innen ihr Wissen im Umgang mit einer Lese-Rechtschreibschwäche beurteilen. Hanke et al. (2019) erhoben schließlich

das Wissen von Lehramtstudierenden über basale Prozesse des Lesen- und Schreibens. Untersuchungen, die sich spezifisch mit dem professionellen Wissen von Lehrkräften zu LS befassen, sind somit erst zu Teilbereichen vorhanden. Zudem basieren die vorhandenen Ergebnisse häufig auf Selbsteinschätzungen. Instrumente, mit denen das professionelle Wissen von Lehrkräften zum Thema LS durch Fremdeinschätzung untersucht werden können, fehlen. Die vorliegende Studie leistet einen Beitrag zur Schließung dieser Forschungslücke, indem ein Befragungsinstrument zur Erfassung des professionellen Wissens zum Thema LS präsentiert und dessen psychometrische Qualität überprüft wird. Dabei interessiert insbesondere, ob das Wissen von Regellehrkräften (RLK) und Förderlehrkräften (FLK) mit demselben Instrument erhoben werden kann. Die Bedeutung der Beantwortung dieser Fragestellung ergibt sich aus dem Fakt, dass Schüler:innen mit LS inklusiv unterrichtet werden und – in Abhängigkeit von den örtlichen Rahmenbedingungen – auch von FLK unterstützt werden. RLK und FLK werden unterschiedlich ausgebildet, deshalb ist es wichtig zu überprüfen, ob mit einem Instrument in beiden Gruppen dasselbe Konstrukt erfasst werden kann. Solche messinvarianten Instrumente sind notwendig, wenn es darum geht, diese Gruppen von Lehrkräften zu vergleichen oder im inklusiven Unterricht den Einfluss des Professionswissens der beiden Lehrkrafttypen auf die Lernenden zu überprüfen.

Schließlich ist zu berücksichtigen, dass die vorhandenen Studien zum professionellen Wissen zum Thema Lesen und Schriftspracherwerb fast ausschließlich mit Studierenden durchgeführt wurden. Erkenntnisse zum Wissen zu LS von Lehrkräften, deren Ausbildung unter Umständen schon länger zurückliegt, fehlen. Mit Blick auf allenfalls notwendige Fortbildungen ist es wichtig, auch dazu Erkenntnisse zu gewinnen.

Ausgehend von diesen Forschungslücken werden zwei Studien präsentiert. In Studie 1 wird die Entwicklung und Validie-

rung eines Instruments zum Erfassen des professionellen Wissens von Studierenden (Regel- und Förderpädagogik) zum Thema LS beschrieben. Studie 2 hatte zum Ziel zu prüfen, ob das Professionswissen zu LS von amtierenden Lehrkräften (Regel- und Förderpädagogik) mit einer kürzeren bzw. angepassten Version des Instruments erfasst werden kann.

## **Professionelles Wissen von Lehrkräften**

Zur Thematik der professionellen Kompetenz von Lehrkräften und deren Erfassung wurden im letzten Jahrzehnt zahlreiche Publikationen veröffentlicht. Dieser Diskurs kann im Rahmen dieses Artikels nicht umfassend aufgenommen werden und es werden nur diejenigen Grundlagen vorgestellt, die zum Verständnis des Instruments notwendig sind. Eine ausführliche Auseinandersetzung mit der Thematik bezogen auf das Thema LS findet sich bei Lietz (2021).

Professionelle Kompetenz im engeren Sinn schließt an den Begriff des Professionswissens an und umfasst theoretisches, deklaratives, prozedurales und konditionales Wissen, das „als notwendige Voraussetzung für professionelles Handeln im Unterricht zu verstehen [ist]“ (Krauss et al., 2017, S. 38). Professionelle Kompetenzen von Lehrkräften können somit allgemein als kognitive Fähigkeiten und Fertigkeiten bezeichnet werden, die zum Lösen bestimmter Probleme notwendig sind (König, 2015). Sie werden durch Wissen und Können definiert, werden im Folgenden als Professionswissen bezeichnet und beinhalten nach Seifert (2015) vier Eigenschaften:

1. Es handelt sich um eine Disposition, die nicht direkt beobachtbar ist und entsprechend operationalisiert werden muss.
2. Sie sind erlernbar und können durch pädagogische Maßnahmen beeinflusst werden.

3. Der Erwerb erfolgt in spezifischen Situationen und kommt in diesen zur Anwendung.
4. Die kognitiven Fähigkeiten haben einen Handlungsbezug und bilden sich im erfolgreichen Handeln in konkreten Situationen ab.

Professionswissen kann in verschiedenen Bereichen erfasst werden. Häufig werden in Anlehnung an Shulman (1986) drei Wissensbereiche unterschieden: Pädagogisches Wissen (z.B. Klassenführung, Wissen über Lernprozesse), Fachwissen (allgemeines Wissen einer Person zu einem Fachbereich) und fachdidaktisches Wissen (Wissen zu geeigneten Vorgehensweisen im Unterricht, Wissen zu Vorstellungen der Lernenden, usw.; z.B. Baumert & Kunter, 2011; Krauss et al., 2017). Das fachliche Wissen gilt dabei als Voraussetzung für das fachdidaktische Wissen.

Hinsichtlich von Lesen und LS stellt sich die Frage, ob und wie fachliches und fachdidaktisches Wissen voneinander abzugrenzen sind. Zudem gilt es zu überlegen, wie das diagnostische Wissen in das Shulman-Modell einzuordnen ist. Diagnostischen Kompetenzen wird in der Ausbildung von Lehrkräften eine hohe Bedeutung zugemessen (Kultusministerkonferenz, 2019). Nach Artelt (2009) ist es wichtig, das diagnostische Wissen domänenspezifisch zu konzeptualisieren. Sie schlägt vor, das diagnostische Wissen im Bereich Lesen dem fachdidaktischen Wissen zuzuordnen, da das diesbezügliche fachliche Wissen die Grundlage für das diagnostische Wissen sei. Auch Hanke et al. (2019) haben das diagnostische Wissen dem fachdidaktischen Wissen zugeordnet. Nicht diskutiert wird dabei die Frage nach dem Verhältnis von Wissen und Können. Wenn eine Lehrkraft beispielsweise weiß, dass die Lesegeschwindigkeit zentral ist, um das Leseverständnis zu fördern, heißt das noch nicht, dass sie Möglichkeiten zu deren Förderung kennt und anwendet. Blömeke (2011) und Bremerich-Vos et al. (2011) haben im Rah-

men der TEDS-Studien zum Umgang mit dieser Herausforderung und zur Erfassung der entsprechenden Wissensbestände vorgeschlagen, beim Fachwissen und beim fachdidaktischen Wissen verschiedene kognitive Anforderungen zu unterscheiden. Das Fachwissen wird mit Anforderungen im Sinn von Erinnern und Abrufen operationalisiert und rückt in die Nähe von deklarativem Wissen. Beim fachdidaktischen Wissen geht es um Verstehen/Analysieren/Anwenden, die Erfassung ist entsprechend an Problemstellungen gebunden und wird durch anwendungsbezogene Items erfasst (Bremerich-Vos et al., 2011).

Schließlich stellt sich die Frage, welche Inhalte zum Professionswissen zu LS gehören. Dazu ist der Forschungsstand zum Thema LS zu berücksichtigen.

## Leseschwierigkeiten

Für Kinder, die Schwierigkeiten beim Lesenlernen haben und unterdurchschnittliche Leseleistungen aufweisen, werden verschiedene Begriffe wie Lesestörung, Leseschwäche oder Legasthenie verwendet (Schulte-Körne, 2021). Ungeachtet des Definitionsdiskurses, der hier ausgeklammert wird, besteht weitgehend Einigkeit darüber, dass Kinder mit LS in ihrer Leseentwicklung dieselben Prozesse wie unauffällige Kinder durchlaufen und sich Unterschiede vor allem in einer zeitlichen Verschiebung und einem größeren Aufwand zum Erwerb von Lesekompetenzen zeigen (Scheerer-Neumann, 2018). Deshalb gehört zur Auseinandersetzung mit dem Thema LS auch der Prozess des Leseerwerbs.

Kinder mit LS fallen bereits vor dem eigentlichen Leseerwerb durch Defizite in der phonologischen Bewusstheit und der Benennungsgeschwindigkeit auf (Mayer, 2016), wobei besonders Schwierigkeiten mit der Benennungsgeschwindigkeit ein bedeutsamer Indikator für eine spätere LS sind (Gorecki & Landerl, 2015; Mayer, 2018). In der Anfangsphase des Leseerwerbs sind zudem

Probleme mit dem Aufbau der Phonem-Graphemkorrespondenz erste Anzeichen für eine LS (Klicpera et al., 2017; Schneider, 2017). Wichtig für den Leselernprozess ist weiter die Ausbildung des phonologischen Rekodierens (Fischer & Klicpera-Gasteiger, 2013). Kindern mit LS gelingt zwar das Lautieren von Buchstaben, aber nicht das Zusammenlauten bzw. Synthetisieren (Phonemsynthese) von Buchstabenfolgen (Klicpera et al., 2017). Weitere Probleme zeigen sich auf der Wortebene. Die Kinder können Wörter nicht oder nur eingeschränkt in orthographische Einheiten wie Silben, Morpheme und Signalgruppen segmentieren (Tacke, 2011). Das macht den Leseprozess nicht nur fehleranfälliger, sondern auch langsamer (Schründer-Lenzen, 2013). Zudem verzögert sich der Aufbau eines Sichtwortschatzes (Scheerer-Neumann, 2018), von Wörtern „die im mentalen orthographischen Lexikon gespeichert sind und holistisch, das heißt ohne den Umweg über die Analyse und Synthese von einzelnen Buchstaben, erkannt werden können“ (Nix, 2011, S. 77f.). Kinder mit und ohne LS unterscheiden sich weiter bezüglich der Leseflüssigkeit, insbesondere bezüglich der Dekodier- und somit der Lesegeschwindigkeit (Landerl & Wimmer, 2008). Rosebrock et al. (2011, S. 62) gehen davon aus, dass ab einem Richtwert von 100 Wörtern pro Minute und einer Genauigkeit von 95% korrekt gelesener Wörter ein „Fließtext auf Unabhängigkeitsniveau verstehend gelesen werden“ kann.

Die Probleme mit der Leseflüssigkeit sind mit Schwierigkeiten beim Leseverständnis assoziiert (Müller et al., 2013; Nix, 2011). Rosebrock und Nix (2012) empfehlen deshalb, bei leseschwachen Kindern vorrangig die Leseflüssigkeit durch geeignete Lautlese-Verfahren zu trainieren. Stilles Lesen ist dagegen nicht geeignet, u.a. wegen fehlender basaler Voraussetzungen und fehlender Lesestrategien zur Herstellung von Kohärenzen (Rosebrock & Nix, 2012).

Verschiedene Untersuchungen weisen schließlich darauf hin, dass Kindern mit

LS Lesestrategien fehlen oder dass diese ineffektiv angewendet werden (Scheerer-Neumann, 2018). Die genannten Schwierigkeiten auf Seiten der Lesenden werden in hohem Masse auch durch die textseitigen Anforderungen beeinflusst (Artelt, 2009). Diese können beispielsweise durch die Berücksichtigung von Merkmalen der Satz- und Textkomplexität kontrolliert werden. Ein einfaches Verfahren dazu stellt der Lesbarkeitsindex (Lix) dar, der auf den Textmerkmalen der durchschnittlichen Satzlänge und dem Prozentsatz der Wörter mit mehr als sechs Buchstaben beruht (Björns-son, 1968).

Ausgehend von den Überlegungen zum professionellen Wissen von Lehrkräften und den Erkenntnissen zum Thema LS stellt sich die Frage, wie das Professionswissen von Lehrkräften zu dieser Thematik modelliert werden kann.

### **Modellierung von Professionswissen im Bereich Lesen**

Zur Modellierung von Professionswissen LS gilt es zur reflektieren ob und wie die Unterscheidung zwischen verschiedenen Inhalts- und Wissensbereichen (Fachwissen, fachdidaktisches Wissen, diagnostisches Wissen) vorgenommen werden kann. Zudem müssen die kognitiven Anforderungen bestimmt werden, die erfasst werden sollen.

Bisher vorliegende Studien haben diese Überlegungen unterschiedlich berücksichtigt: In den Studien von Moats und Forman (2003) sowie von Schmid und Schabmann (2016) lag der Fokus auf inhaltlichen Dimensionen (z.B. phonologischen Bewusstheit, Wissen zum Lesen und zur Leseentwicklung) und somit auf Grundlagenwissen zum Leseerwerb. Es erfolgte keine Unterscheidung zwischen unterschiedlichen kognitiven Anforderungen oder unterschiedlichen Wissensdomänen. Rutsch (2016) fokussierte das fachdidaktische Wissen mit den inhaltlichen Facetten Förderwissen/ Intervention und Diagnose/Inhaltswissen.

Eine Unterscheidung zwischen fachlichem und fachdidaktischem Wissen wurde in der Studie von Phelps und Schilling (2004) vorgenommen. Dort konnten zwei fachliche Faktoren – Leseverständnis und Wortanalyse – sowie ein Faktor zum fachdidaktischen Wissen empirisch unterschieden werden.

Ausgehend von diesen Vorarbeiten sowie den Überlegungen im Abschnitt «Pro-

fessionelles Wissen von Lehrkräften» wird Professionswissen zum Thema LS wie folgt konzeptualisiert: Erstens werden in Anlehnung an Artelt (2009) Grundlagenwissen und diagnostisches Wissen unterschieden (Tabelle 1).

Die im Abschnitt »Leseschwierigkeiten« dargestellten Erkenntnisse werden für das Grundlagenwissen mit vier Facetten sys-

**Tabelle 1**

*Modellierung des Professionswissens zu Grundlagen des Leseerwerbs und zu Leseschwierigkeiten*

Inhaltsbereich	Kognitive Anforderung	
	Fachwissen	Fachdidaktisches Wissen/ Können
	Kennen/Erinnern/Abrufen	Verstehen/Analysieren/Anwenden
<b>Facetten zum Grundlagenwissen</b>		
<i>Facette 1</i> Linguistische Merkmale deutsche Schriftsprache		
<i>Facette 2</i> Stufenmodelle Leseerwerb/ Leseentwicklung		
<i>Facette 3</i> Prozessmodelle Lesen und Determinanten Lesekompetenz		
<i>Facette 4</i> Didaktische Konzepte, Fördermaßnahmen Leseunterricht		
<b>Facetten zum diagnostischen Wissen</b>		
<i>Facette 1</i> Symptome Leseschwierigkeiten		
<i>Facette 2</i> Leseverläufe unauffällig und schwach lesenden Kinder; typische Schwierigkeiten		
<i>Facette 3</i> Diagnostische Verfahren und Aufgaben		

tematisiert: linguistische Merkmale der deutschen Schriftsprache; Stufenmodelle Leserwerb und Leseentwicklung; Prozessmodelle Lesen und Determinanten Lesekompetenz sowie didaktische Konzepte und Fördermaßnahmen Leseunterricht. Beim diagnostischen Wissen geht es darum, dass Lehrkräfte das Grundlagenwissen nutzen und Schwierigkeiten in verschiedenen Bereichen erkennen bzw. Produkte von Schüler:innen analysieren können. Dieses Wissen wurde mit drei Facetten beschrieben: Symptome Leseschwierigkeiten; Leseverläufe bei unauffällig und schwach lesenden Kindern bzw. typische Schwierigkeiten sowie diagnostische Verfahren und Aufgaben. Weiter werden in Anlehnung an Blömeke (2011) und Bremerisch-Vos et al. (2011) kognitive Anforderungen unterschieden: Fachwissen im Sinn von kennen, erinnern und abrufen sowie fachdidaktisches Wissen und Können als verstehen, analysieren und anwenden (Tabelle 1).

Zielsetzung der hier präsentierten Studie war, auf der Grundlage dieser Konzeption ein Befragungsinstrument zu entwickeln und folgende Fragen zu untersuchen:

### Studie 1

1. Lässt sich das fachliche und das fachdidaktische Professionswissen zu Grundlagen des Leserwerbs und zur Diagnostik von Leseschwierigkeiten von Studierenden der Regel- und der Förderpädagogik am Ende des Studiums mit demselben Instrument valide und reliabel erfassen?
2. Lassen sich die theoretisch angenommenen Dimensionen Grundlagenwissen /diagnostisches Wissen sowie die kognitiven Anforderungen (Fachwissen/fachdidaktisches Wissen) empirisch bestätigen?
3. Welche Unterschiede zeigen sich zwischen Studierenden der Regel- und Förderpädagogik?

Die meisten Studien zum Wissen von Lehrkräften im Bereich Lesen wurden mit Studierenden durchgeführt. Da es sein kann, dass gewisse Wissensbestände vor allem während bzw. kurz nach der Ausbildung vorhanden sind, wird eine zweite Studie mit ausgebildeten und erfahrenen Lehrkräften durchgeführt.

### Studie 2

1. Lässt sich das fachliche und fachdidaktische Professionswissen bezüglich Grundlagen des Leserwerbs und zur Diagnostik von Leseschwierigkeiten von ausgebildeten und erfahrenen Regel- und Förderlehrkräften mit abgeschlossener Ausbildung reliabel und valide erfassen?

### Methode

Zu den in Tabelle 1 dargestellten Inhalts- und Anforderungsbereichen wurden ausgehend vom im Abschnitt Leseschwierigkeiten zusammengefassten Erkenntnisstand Operationalisierungen vorgenommen, die zu einem Pool von 28 Items führten (Lietz, 2021). Mit diesen Items wurden mehrere Vorstudien durchgeführt (ebd.). Ausgehend von deren Ergebnissen wurden 18 Items für das finale Instrument für Studie 1, die im Folgenden präsentiert wird, ausgewählt (Fragebogen siehe elektronisches Supplement <https://doi.org/10.17605/OSF.IO/TQABN>). Die Items umfassten unterschiedliche Formate: Offene Kurzantworten, Single-Choice-Items, Multiple-Choice-Items, Richtig-Falsch-Aufgaben, Zuordnungsaufgaben sowie problembasierte Audio-Items (lesebezogene Kompetenzen von Kindern anhand von Audioaufnahmen einschätzen). Für Studie 2 wurden 13 Items verwendet (elektronisches Supplement). Die Begründung für deren Auswahl erfolgt bei der Beschreibung von Studie 2. Tabelle S1 im elektronischen Supplement zeigt die Zuordnung der Items zu den einzelnen Facetten bzw. Bereichen.



## Studie 1: Befragung von Studierenden

### Stichprobe Studierende (Stichprobe<sub>STUD</sub>)

Um die Eignung des Instruments für Studierende der Regel- und der Förderpädagogik zu überprüfen, wurde eine für die Zielsetzung ausreichend große Stichprobe ( $N = 507$ ) von Studierenden aus unterschiedlichen Studiengängen und mit unterschiedlichem Erfahrungshintergrund rekrutiert. Grundschullehrkräfte absolvieren in der Schweiz ein BA-Studium. Förderpädagogik ist ein Aufbaustudium, für die hier untersuchte Stichprobe galt in der Regel als Zulassungsbedingung ein BA-Abschluss für die Regelschule sowie mindestens zwei Jahre Berufserfahrung als Lehrkraft. Diese Ausgangslage führte zu drei Stichprobengruppen: Grundschullehrkräfte am Ende des Studiums, Förderlehrkräfte am Anfang des Aufbaustudiums (gleiche Ausbildung wie Grundschullehrkräfte, aber Berufserfahrung als Lehrkraft) sowie Förderlehrkräfte am Ende des Aufbaustudiums. Die Stichprobe setzte sich wie folgt zusammen:

- GRUNDSCHULE<sub>ABSCHLUSS</sub>,  $n = 190$ : Studierende Lehramt Grundschule am Ende des Studiums
- FÖRDER<sub>ANFANG</sub>,  $n = 139$ : Studierende am Anfang des MA-Aufbaustudiums für Förderlehrkräfte (BA als Regellehrkraft und Berufserfahrung)
- FÖRDER<sub>ABSCHLUSS</sub>,  $n = 114$ : Studierende Förderlehrkraft am Ende des Aufbaustudiums (BA als Regellehrkraft und Berufserfahrung)

Zudem wurde eine Gruppe von LAIEN als Kontrollgruppe gewonnen. Deren Daten wurden nicht für die Skalierung verwendet.

- Laien (LAIEN,  $n = 64$ ): Studierende Lehramt Grundschule zu Beginn der Ausbildung.

Die Proband:innen der Stichprobe GRUNDSCHULE stammten von zwei Pädagogischen Hochschulen (PH). In der Stichprobe FÖRDER waren Studierende von vier deutschsprachigen Ausbildungsinstitutionen für Förderlehrkräfte (PH und Universität, entspricht dem Lehramt Sonderpädagogik) beteiligt.

### Durchführung

Die Durchführung der Befragung erfolgte mittels Paper-Pencil-Verfahren unter standardisierten Anweisungen der Testleitenden. Die durchschnittliche Bearbeitungsdauer betrug 50 Minuten. Die Antworten der Items mit offenem Antwortformat (9, 10, 13, 16, 18) wurden anhand eines Manuals codiert (dichotom 0-1; Partial-Credit 0-1-2). Die Codierung erfolgte durch drei Personen. 20% des gesamten Materials wurde von allen Codierenden bearbeitet (Lombard et al., 2002). Die Übereinstimmung lag zwischen  $\kappa = .75$  und  $\kappa = .92$ . Gemäß Döring und Bortz (2016) sind diese Werte gut bis sehr gut.

## Studie 2: Befragung von amtierenden Lehrkräften

### Stichprobe Praxis (Stichprobe<sub>PRAXIS</sub>)

Die Stichprobe<sub>PRAXIS</sub> umfasst 142 Grundschul- und Förderlehrkräfte, die ihre Ausbildung abgeschlossen hatten und in einer inklusiven Klasse (Jahrgangstufe 2-4) Kinder mit LS unterrichteten. Befragt wurden Teams von Förder- und Regellehrkräften<sup>1</sup>, die sich freiwillig an einer größeren Studie beteiligten.

- GRUNDSCHULE<sub>PRAXIS</sub>,  $n = 90$ : Erfahrene Grundschullehrkräfte (Berufserfahrung  $M = 13.06$  Jahre,  $SD = 11.39$ ).
- FÖRDER<sub>PRAXIS</sub>,  $n = 52$ : Erfahrene Förderlehrkräfte (Berufserfahrung  $M = 12.27$  Jahre,  $SD = 6.76$ ).

<sup>1</sup> Zum Teil übernahmen Lehrkräfte ohne entsprechende Ausbildung die Aufgaben einer Förderlehrkraft. Diese Personen wurden der Stichprobe GRUNDSCHULE<sub>PRAXIS</sub> zugeordnet.

## Durchführung

In der Stichprobe<sub>PRAXIS</sub> musste aus forschungsökonomischen Gründen und damit die Lehrkräfte überhaupt zur Teilnahme bereit waren, eine gekürzte Version des Fragebogens aus Studie 1 eingesetzt werden. Ausgewählt wurden 13 Items (Elektronisches Supplement). 10 Items waren deckungsgleich mit dem Fragebogen aus Studie 1. Drei Items (19, 20, 21), die sich spezifisch auf die Förderung bei LS bezogen, wurden aus dem ursprünglichen Itempool von Lietz (2021) dazu genommen. Die Lehrkräfte füllten den Fragebogen online aus, das Itemformat entsprach der Paper-Pencil-Version. Die Bearbeitung dauerte durchschnittlich 32 Minuten.

Alle Antworten der Items mit offenem Antwortformat (9, 10, 18) wurden nach den Vorgaben von Lietz (2021) von zwei Personen codiert. Die Übereinstimmung lag zwischen  $\kappa = .67$  und  $\kappa = .98$ . Gemäß Döring und Bortz (2016) sind diese Werte gut bis sehr gut. Da sich bei einer ersten Reliabilitätsanalyse nach klassischer Testtheorie gezeigt hat, dass der Fragebogen<sub>PRAXIS</sub> keine zufriedenstellenden Reliabilität aufweist, wurden die Partial-Credit-Item (2, 4, 5, 9, 17, 18, 19, 20) in dichotome Items umcodiert (0/1 = 0; 2 = 1). Dies erfolgte mit dem Ziel, das professionelle Wissen eindeutiger erfassen zu können.

## Analysen

### Studie 1: Studierende

Mit den Daten der Stichprobe<sub>STUD</sub> wurden Item- und Skalenanalysen basierend auf der Item-Response-Theorie mit Conquest (Version 4) durchgeführt. Im Rasch-Modell werden Itemschwierigkeit und Personenfähigkeit auf einer gemeinsamen Skala abgebildet (Seifert, 2015). Mittels der gewichteten Abweichungsquadrate (Weighted Mean Square MNSQ) wird ermittelt, wie sich die erwarteten Item-Charakteristik Kurven (Item Characteristic Curves CC) und die auf der

Grundlage der erhobenen Daten basierenden Kurven zu einander verhalten bzw. wie stark sie voneinander abweichen (Bond & Fox, 2001). Für Multiple-Choice-Tests empfehlen Bond und Fox (2001) Grenzwerte von 0.80 bis 1.20. Wilson (2005) erachtet Werte zwischen 0.75 und 1.33 als akzeptabel.

Die Schätzung von Itemschwierigkeit und Personenfähigkeit erfolgte mittels des in ConQuest implementierten Marginal Maximum Likelihood Schätzverfahrens (MML) (Heine & Tarnai, 2015). Für die weiterführenden Gruppenvergleiche mittels einfaktorierlicher Varianzanalysen mit post-hoc Einzelvergleich (Games-Howell-Test für ungleiche Varianzen; Field, 2013) wurden die Personenfähigkeiten-Schätzer (Warm's Weighted Likelihood Estimates WLE) verwendet. Diese haben gegenüber von Rohwerten (Summenscores) den Vorteil, dass sie in Abhängigkeit von der Itemschwierigkeit und korrigiert um den Messfehler ermittelt werden (Lüdtke & Robitzsch, 2017). Entsprechend wird auch die WLE-Reliabilität berichtet. Die Werte können analog zu Cronbachs  $\alpha$  interpretiert werden. Aufgrund des explorativen Charakters der vorliegenden Studie ist eine Anpassung des Wertes auf .60 zulässig (Stancel-Piatak et al., 2013) bzw. in vergleichbaren Studien zur Erhebung des Professionswissens wird dieser Wert als akzeptabel bezeichnet (Bremerich-Vos et al., 2011).

Zur Überprüfung der Messinvarianz zwischen den Gruppen wurden differenzielle Itemfunktionsanalysen (DIF-Analysen) durchgeführt. Damit kann überprüft werden, ob die Items in Teilstichproben dasselbe Konstrukt messen (Wilbert & Linnemann, 2011). Dies erfolgt über die Berechnung der Abweichungen von der Null-Differenz ( $t > 1.96, p < .05$ ; Draba, 1977). Ist die Differenz von zwei Items  $\geq 0.43$  Logits, gilt die Differenz als moderat, ist sie  $\geq 0.64$  Logits ist sie moderat bis groß. Die statistische Absicherung erfolgt über die Abweichungen vom Ideal der Null-Differenz (Tristán, 2006). In den unterschiedlichen Stichproben zeigten sich zum Teil

signifikante DIF-Werte bezüglich mehrerer Items. Deren Ausschluss hätte zu massiven Einschränkungen der inhaltlichen Validität des Instruments geführt. Für die Mittelwertvergleiche wurde deshalb in Anlehnung an Schulte et al. (2013) ein interpretativer Umgang gewählt. Durch Itemsplitting (Ramseier, 2008) wurde die Itemschwierigkeit pro Stichprobe ermittelt. Anschließend wurden die Itemparameter auf die Skala der jeweils anderen Gruppe transformiert und die DIF-Werte wurden berechnet. Bei Items, die keine DIF aufwiesen, wurden für alle Skalen dieselbe Itemschwierigkeit verwendet. Für Items mit DIF wurden je eigene Itemschwierigkeiten geschätzt. Die Items tragen so je zur Schätzung der Personenfähigkeit bei, beeinflussen aber den jeweiligen Gruppenunterschied nicht (für Details siehe Lietz, 2021). Durch die sechs «Ankeritems» ohne DIF ist parallele Skalierung in den Gruppen gewährleistet. Anschließend wurde die Robustheit der messinvarianten Skalen durch einen Modellvergleich überprüft (Ramseier, 2008). D.h., die Skalen mit vollständig gesplitteten, teilweise gesplitteten und nicht gesplitteten Items wurden verglichen.

Zusätzlich wurde die Plausibilität der Distraktoren überprüft (Distraktorenanalyse). Gemäß Ingenkamp und Lissmann (2008) sollten Distraktoren von mindestens 5 % der Proband:innen gewählt werden, um als plausibel zu gelten.

Um zu prüfen, ob sich die theoretisch intendierten Subdimensionen empirisch unterscheiden lassen, wurden in Anlehnung an Ramseyer et al. (2008) ein Modelldifferenz-Test durchgeführt. Die Modelldifferenz wurde anhand der Informationsmasse Bayesian-Information-Criterion (BIC) und Akaike-Information-Criterion (AIC) beurteilt. Mittels Modelldifferenz-Test ( $\chi^2$ -Test) wurde überprüft, ob sich die Modelle signifikant voneinander unterscheiden (Ramseyer et al., 2008; Seifert, 2015). Ein signifikanter  $p$ -Wert bedeutet, dass das komplexere Modell zu bevorzugen ist (Grand et al., 2017).

## Studie 2: Stichprobe<sub>PRAXIS</sub>

In Studie 2 wurde mit den Daten der Stichprobe<sub>PRAXIS</sub> zunächst gemäß klassischer Testtheorie eine Item- und Skalenanalyse durchgeführt. Zusätzlich wurden analog zu Studie 1 die dabei entstandenen Skalen basierend auf der Item-Response-Theorie überprüft. Dazu wurde die Statistiksoftware R (Version 4.1.2) mit dem Package TAM verwendet.

## Ergebnisse

### Studie 1: Stichprobe<sub>STUD</sub>

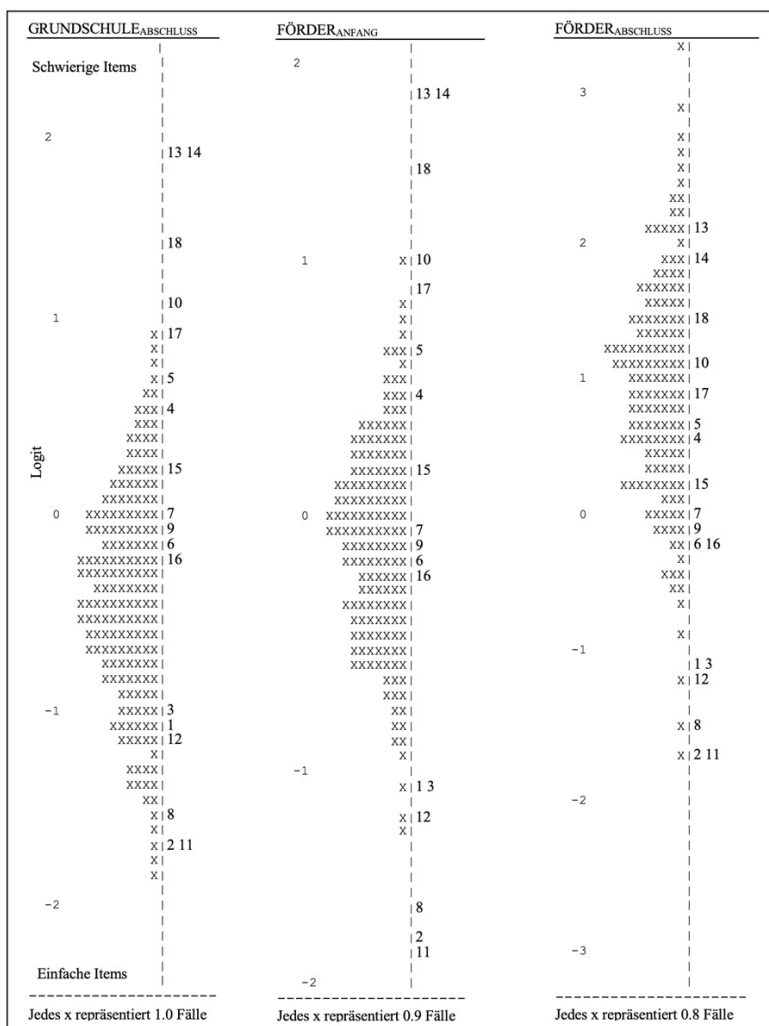
#### Item- und Skalenanalysen

In Tabelle 2 werden die Ergebnisse der Item- und Skalenanalysen des finalen Fragebogens für die Studierenden bezogen auf die Gesamtstichprobe und die Teilstichproben dargestellt. Die MNSQ-Infit-Werte liegen mit Ausnahme des Items 15 (Anlauttabelle) in der Gruppe FÖRDER<sub>ABSCHLUSS</sub> im akzeptablen Bereich (Wilson, 2005). Beim Outfit zeigt sich zusätzlich in derselben Stichprobe beim Item 11 (Schwierigkeiten der Einschätzung einer Leseaufgabe) ein Underfit. Der Wert von 1.36 liegt jedoch noch immer in einem als funktional geltenden Bereich, d.h. das Item liefert einen positiven Beitrag zur Messung der latenten Dimension, nur ist dieser Beitrag weniger groß als bei Items mit besserem Fit. Weil der Outfit sensibel reagiert auf Diskrepanzen zwischen erwarteter und effektiver Lösungswahrscheinlichkeit bezogen auf Aufgaben, deren Schwierigkeit weit vom Personenparameter entfernt liegt (z.B. einfache Aufgaben, die von kompetenten Testpersonen nicht gelöst werden konnten oder schwierige Aufgaben, die von wenig kompetenten Testpersonen gelöst werden konnten), wurde von Wright die Infit Statistik entwickelt. Sie nimmt eine Gewichtung nach Nähe zum Personenparameter bzw. Informationsgehalt vor. Aus diesem Grund kommt den Infit-Statistiken tendenziell mehr Bedeutung

**Tabelle 2**  
Deskriptive Angaben auf der Grundlage des Rasch-Modells

Substichproben	n	M <sub>WLE</sub> (SD)	Varianz	MNSQ Infit	MNSQ Outfit	WLE-Rel.
GRUNDSCH <sub>ABSCHLUSS</sub>	190	-0.45 (.73)	.18	0.91-1.04	0.76-1.11	.61
FÖRDER <sub>ANFANG</sub>	139	-0.08 (.69)	.11	0.94-1.07	0.84-1.09	.53
FÖRDER <sub>ABSCHLUSS</sub>	114	0.93 (.95)	.42	0.88-1.33	0.60-1.36	.77
LAIEN <sup>a</sup>	64	-0.87 (.64)				
Gesamt	443	0.12 (0.75)		0.93-1.12		.78

<sup>a</sup> Die Daten der Laienstichprobe werden nur für den Gruppenvergleich genutzt und es wurden keine Skalenanalyse vorgenommen.



**Abbildung 1**  
Itemschwierigkeit der Substichproben GRUNDSCHULE<sub>ABSCHLUSS</sub>, FÖRDER<sub>ANFANG</sub> UND FÖRDER<sub>ABSCHLUSS</sub> im Vergleich

zu als Outfit-Statistiken (vgl. Linacre 2002). Die WLE-Reliabilität für die Stichprobe FÖRDER<sub>ABSCHLUSS</sub> ist befriedigend, für die Stichprobe GRUNDSCHULE<sub>ABSCHLUSS</sub> ist sie niedrig und für die Gruppe FÖRDER<sub>ANFANG</sub> ungenügend. Der Wert von .61 der Stichprobe GRUNDSCHULE<sub>ABSCHLUSS</sub> kann gemäß Stancel-Piatak et al. (2013) aufgrund des explorativen Charakters der Studie akzeptiert werden.

In Abbildung 1 ist die Itemschwierigkeit für die Substichproben in der Personen-Item-Map dargestellt. Das Instrument differenziert im unteren und oberen Leistungsbereich, wobei im mittleren Bereich Items fehlen. In der Stichprobe FÖRDER<sub>ABSCHLUSS</sub> ist das Fähigkeitsspektrum am besten abge-

deckt. In den Substichproben FÖRDER<sub>ANFANG</sub> und GRUNDSCHULE<sub>ABSCHLUSS</sub> ist das Fähigkeitsspektrum durch die Items nicht befriedigend erfasst.

Die Distraktorenanalyse, in der auch die Antworten der Laienstichprobe einbezogen war, ergab, dass einige Distraktoren (Elektronisches Supplement: 2 abd) von weniger als 5 % der Proband:innen gewählt wurden und angepasst werden müssten (Lietz, 2021). Das Item 2 wurde in die definitive Skala aufgenommen, weil es sich um ein einfaches „warm-up“-Item handelt. Drei Distraktoren (1f „übertroffen“, 7d, 8d) wurden nie gewählt und könnten bei einem weiteren Einsatz des Instruments weggelassen werden (ebd.).

**Tabelle 3**

Differenzielle Itemfunktion in den drei Substichproben im paarweisen Vergleich

Item Nr.	GRUNDSCHULE <sub>ABSCHLUSS</sub> / FÖRDER <sub>ANFANG</sub>		FÖRDER <sub>ANFANG</sub> / FÖRDER <sub>ABSCHLUSS</sub>		GRUNDSCHULE <sub>ABSCHLUSS</sub> / FÖRDER <sub>ABSCHLUSS</sub>	
	Betrag Differenz der Logit-Werte	<i>t</i>	Betrag Differenz der Logit-Werte	<i>t</i>	Betrag Differenz der Logit-Werte	<i>t</i>
1	0.26	1.37	0.17	0.60	0.43	1.57
2	0.60 <i>moderat</i>	<b>2.04*</b>	0.41	0.99	0.20	0.54
3	0.48 <i>moderat</i>	1.89	0.76 <i>hoch</i>	<b>2.40*</b>	0.28	0.99
4	0.06	0.31	0.28	1.42	0.23	1.15
5	0.43 <i>moderat</i>	<b>2.28*</b>	0.62 <i>moderat</i>	<b>3.09**</b>	0.17	0.90
6	0.53 <i>moderat</i>	<b>2.32*</b>	0.31	1.14	0.83 <i>hoch</i>	<b>3.28**</b>
7	0.09	0.39	0.35	1.26	0.26	0.97
8	0.04	0.16	0.18	0.46	0.14	0.37
9	0.25	1.72	0.39	1.95	0.14	0.72
10	0.23	0.81	0.98 <i>hoch</i>	<b>3.31**</b>	0.75 <i>hoch</i>	<b>2.69**</b>
11	0.15	0.59	0.19	0.34	0.34	0.66
12	0.65 <i>hoch</i>	<b>2.62**</b>	0.32	<b>1.05</b>	0.97 <i>hoch</i>	<b>3.32**</b>
13	0.79 <i>hoch</i>	<b>2.26*</b>	1.15 <i>hoch</i>	<b>3.50**</b>	0.37	1.99
14	1.22 <i>hoch</i>	<b>2.19*</b>	2.55 <i>hoch</i>	<b>4.87**</b>	1.31 <i>hoch</i>	<b>3.87**</b>
15	0.12	0.71	0.64 <i>moderat</i>	<b>3.09**</b>	0.52 <i>moderat</i>	<b>2.63*</b>
16	0.76 <i>hoch</i>	<b>3.22**</b>	0.95 <i>hoch</i>	<b>3.44**</b>	0.19	0.75
17	0.20	1.09	0.38	1.80	0.58 <i>moderat</i>	<b>2.95**</b>
18	0.21	0.57	0.78 <i>hoch</i>	<b>3.03**</b>	0.99 <i>hoch</i>	<b>3.19**</b>

\* $p < .05$ . \*\* $p < .01$ .

Die DIF-Analysen weisen mit Ausnahme von sechs Items eine Differenz von  $\geq 0.43$  Logits und/oder eine signifikante Abweichung von  $t > 1.96$  auf (Tabelle 3).

Die größten Abweichungen zeigten sich im Vergleich zwischen den Gruppen FÖRDER<sub>ANFANG</sub> versus FÖRDER<sub>ABSCHLUSS</sub> und GRUNDSCHULE<sub>ABSCHLUSS</sub> versus FÖRDER<sub>ABSCHLUSS</sub>. Keine DIF zeigte sich bei den einfachsten Items (1, 2, 8, 11) sowie bei drei Items mit mittlerem Schwierigkeitsgrad (4, 6, 9).

Der Modellvergleich mit drei durch Itemsplitting gebildeten Skalen ergab, dass sich die Mittelwerte der drei Substichproben in allen drei Skalen (vollständig gesplittete, teilweise gesplittete und nicht gesplittete Items) in derselben Richtung unterscheiden. Das kann als Hinweis auf eine ausreichende Robustheit der Basis-Skala mit nicht gesplitteten Items interpretiert werden. Allerdings muss berücksichtigt werden, dass die finale Skala stichprobenabhängig unterschiedliches misst (siehe Limitationen).

**Dimensionalität**

Zuerst wurde überprüft, ob sich die theoretisch angenommenen Dimensionen Grundlagenwissen versus diagnostisches Wissen empirisch unterscheiden lassen. Diese zweifaktorielle Struktur ließ sich empirisch nicht bestätigen. Die höheren AIC- und BIC-Werte des zweidimensionalen Modells zei-

gen, dass die Abweichung von den Daten bei diesem Modell größer ist als beim eindimensionalen Modell (Tabelle 4). Der  $\chi^2$ -Test ( $\chi^2[2] = 2.55, p > .05$ ) bestätigt das Ergebnis. Die Korrelation der beiden latenten Dimensionen ist zudem mit .95 sehr hoch.

Auch die kognitiven Dimensionen Fachwissen versus fachdidaktisches Wissen ließen sich empirisch nicht trennen. Sowohl die AIC und die BIC-Werte (Tabelle 5) als auch der Modellvergleich ( $\chi^2 [2] = 0.37, p > .05$ ) zeigten, dass das zweidimensionale Modell keine bedeutsame Verbesserung gegenüber dem reduzierten Modell bringt.

**Gruppenvergleich**

Abschließend interessierte, ob sich die drei Substichproben und die Kontraststichprobe LAIEN in ihrem professionellen Wissen unterscheiden. Da mit Ausnahme der LAIEN in den Stichproben keine Normalverteilung vorlag und auch Unterschiede in den Varianzen vorhanden waren ( $p < .01$ ), wurde die ANOVA mit dem robusten Verfahren Welch-Test angewendet. Es ergab sich ein signifikanter und substantieller Haupteffekt für den Faktor Gruppenzugehörigkeit ( $F(3, 221.47) = 92.751, p < .01$ ). Die Gruppenzugehörigkeit erklärt unter Einbezug der Laien 39.9 % der Wissensunterschiede. Die Hauptaussage veränderte sich nicht, auch wenn die Laien nicht berücksichtigt wurden ( $F(2, 253.31) = 90.12, p < .01$ ). Die Gruppenzugehörigkeit

**Tabelle 4**  
*Modellvergleich Grundlagenwissen und diagnostisches Wissen vs. eindimensionales Modell*

Modell	Modelldifferenz	Parameter	n	AIC	BIC
1-dimensional	10 917.999	27	443	10 972	10 989
2-dimensional	10 915.445	29	443	10 973	10 992

**Tabelle 5**  
*Modellvergleich Fachwissen und fachdidaktisches Können vs. eindimensionales Modell*

Modell	Devianz	Parameter	n	AIC	BIC
1-dimensional	10 917.999	43	443	10 972	10 989
2-dimensional	10 917.626	45	443	10 976	10 994

erklärt hier 35.2 % der Wissensunterschiede zwischen den Substichproben.

Der Einzelvergleich der vier Gruppen zeigte, dass die Substichprobe LAIEN das geringste Wissen aufwies, gefolgt von der Substichprobe GRUNDSCHULE<sub>ABSCHLUSS</sub> und der Substichprobe FÖRDER<sub>ANFANG</sub>. Die Gruppe FÖRDER<sub>ABSCHLUSS</sub> verfügte über das höchste Professionswissen. Im Einzelvergleich der Substichproben (inklusive LAIEN) wurden folgende Effektstärken ermittelt: Laien vs. GRUNDSCHULE<sub>ABSCHLUSS</sub>  $d = 0.60$ , GRUNDSCHULE<sub>ABSCHLUSS</sub> vs. FÖRDER<sub>ANFANG</sub>  $d = 0.55$  und FÖRDER<sub>ANFANG</sub> vs. FÖRDER<sub>ABSCHLUSS</sub>  $d = 1.29$ .

### Studie 2: Stichprobe<sub>PRAXIS</sub>

In der STICHPROBE<sub>PRAXIS</sub> der Studie 2 zeigte sich für die Skala mit 13 Items (Elektronisches Supplement) unter Verwendung der Partial-Credit-Codierung eine ungenügende Reliabilität (FÖRDER<sub>PRAXIS</sub> Cronbachs Alpha = .50; GRUNDSCHULE<sub>PRAXIS</sub> Cronbachs Alpha = .49). Wie im Abschnitt Methoden beschrieben, wurde auch eine dichotome Codierung vorgenommen (0/1 = 0; 2 = 1). Für die Stichprobe FÖRDER<sub>PRAXIS</sub> verbesserte sich dadurch die Reliabilität und ließ sich durch den Ausschluss von Items auf .60 erhöhen. Übrig blieben die folgenden fünf Items: Lesekompetenzen 2. Klasse (19), Lesesaufgabe Ebene Strategie (4), Zusammenhang Leseverständnis - Leseflüssigkeit (6), 1-Minute Leseprobe (9) und Phonemanalyse (17). Die MNSQ-Werte lagen zwischen 0.92 und 1.02 und erfüllen die Anforderungen. Die WLE-Reliabilität (.33) war deutlich niedriger als Cronbachs Alpha (.60). Dies kann daran liegen, dass bei der Berechnung der WLE-Reliabilität Extremwerte durch korrigierte Werte ersetzt werden, die nicht-linear – abhängig von der Distanz zum Zentrum – verschoben werden (Linacre, o.J.).

In der Stichprobe GRUNDSCHULE<sub>PRAXIS</sub> führte die Umcodierung nicht zum Erfolg (Cronbachs  $\alpha = .53$ ). Deshalb wurde auf die Durchführung von Rasch-Analysen und DIF-Analysen verzichtet.

## Diskussion

Die beiden Studien verfolgten das Ziel, ein Instrument zu entwickeln und zu evaluieren, mit dem das Professionswissen zu LS von Studierenden der Regel- und Förderpädagogik sowie von erfahrenen Regel- und Förderlehrkräften valide und reliabel erfasst werden kann. In der Stichprobe der Studierenden ließ sich die theoretisch angenommene Dimensionalität (Grundlagenwissen/diagnostisches Wissen sowie Fachwissen/fachdidaktisches Wissen) empirisch nicht bestätigen. Die Korrelation der latenten Dimensionen weist darauf hin, dass es sich beim Professionswissen zu LS um ein dimensionales Konstrukt handelt, zu dem Wissen aus verschiedenen Bereichen gehört. Es könnte sein, dass die Unterscheidung zwischen Fachwissen und fachdidaktischem Wissen im Gegensatz zu anderen Fächern im Bereich Lesen nicht möglich war, weil sich Lesen strukturell von anderen Fächern unterscheidet (Phelps & Schilling, 2004). Zudem ist davon auszugehen, dass Grundlagenwissen eine Voraussetzung für diagnostisches Wissen ist (Artelt, 2009). Nicht ausgeschlossen werden kann – und dies ist eine Limitation der Studie – dass die Modellierung der Wissensbereiche durch die Items nicht gelungen ist (Lietz, 2021).

In der Studie 1 mit der Stichprobe der Studierenden war es möglich, ein Instrument mit 18 Items zu entwickeln, dass die psychometrischen Anforderungen nicht optimal, jedoch mit akzeptablen Werten erfüllt. In der Gruppe FÖRDER<sub>ABSCHLUSS</sub> streuen die Items über das gesamte Schwierigkeitspektrum. Das heißt, dass das Instrument spezifisches Wissen, das in der Ausbildung für Förderlehrkräfte erworben wird, erfasst und sich auch vom Wissen von Laien unterscheidet.

Für die Lehrkräfte ohne förderpädagogische Ausbildung ist die Entwicklung des Instruments nicht befriedigend gelungen und die Reliabilitätswerte sind für die Stichproben GRUNDSCHULE<sub>ABSCHLUSS</sub> und FÖRDER<sub>ANFANG</sub> niedrig. Auch weisen die DIF-

Analysen darauf hin, dass das Instrument bei Personen mit und ohne förderpädagogischem Studium Unterschiedliches misst. DIF zeigte sich vor allem bei Items, die Grundlagenwissen zu LS überprüften (z.B. Lix, Benennungsgeschwindigkeit). Bei mehreren Items zum diagnostischen Wissen, bei denen es um die Analyse bzw. Einschätzung von Schwierigkeiten der Schüler:innen ging, gab es keine DIF. Lehrkräfte ohne förderpädagogische Ausbildung scheinen somit vor allem über anwendungsbezogenes diagnostisches Wissen zu verfügen.

Die Schwierigkeiten, das Professionswissen von Grundschullehrkräften zum Thema LS zu erfassen, weisen auch darauf hin, dass sie nur teilweise über dieses verfügen (Karing, Matthäi & Artelt, 2011; Rjosk et al., 2011; Schmidt & Schabmann, 2010). Nun könnte argumentiert werden, dass dies aufgrund ihrer Ausbildungen erklärt werden kann und die Unterschiede zwischen Regel- und Förderlehrkräfte zu erwarten sind. Bei dieser Argumentation bleibt aber unberücksichtigt, dass in der Regel die Regellehrkräfte für die Förderung bei LS zuständig sind und es häufig vorkommt, dass wenig oder keine Unterstützung durch Förderlehrkräfte zur Verfügung steht. Das heißt, dass das fehlende Wissen sich auch auf die Qualität des Unterrichts und der Leseförderung auswirken kann.

In der Studie 2, in der Lehrkräfte befragt wurden, deren Ausbildung schon länger zurückliegt, musste aus forschungsökonomischen Gründen ein reduzierter Fragebogen eingesetzt werden. Werden die Daten mittels klassischer Testtheorie analysiert, ergibt sich für die Förderlehrkräfte für die Skala mit fünf Items ein Wert von Cronbachs Alpha = .60, der niedrig ist, jedoch dem Richtwert in der TEDS-LT Studien entspricht (Bremerich-Vos et al., 2011). Allerdings erfassen die verbliebenen Items nur einen eingeschränkten Bereich von professionellem Wissen zu LS. Zum einen blieben Items in der Skala, die Wissen zum Schriftspracherwerb erfassen (Leseaufgabe Ebene Strategie [4], Phonemanalyse [17]); zum anderen Items, die

in verbreiteten Förderprogrammen und Diagnoseempfehlungen vorkommen (Lesekompetenzen 2. Klasse [19], 1-Minute Leseprobe [9], Zusammenhang Leseverständnis - Leseflüssigkeit [6]). Das heißt, es könnte sein, dass den Lehrkräften diejenigen Wissensbestände präsent sind, denen sie in ihrem Alltag begegnen, beispielsweise im Manual von Förderprogrammen.

Einerseits kann somit festgestellt werden, dass die Förderlehrkräfte über zentrale Wissens Elemente verfügen, andererseits handelt es sich um eingeschränktes basales Wissen. Die Inhaltsvalidität der reduzierten Skala muss somit in Frage gestellt werden. Das Ergebnis könnte allerdings – und das ist eine weitere Limitation der Studie – auch durch die kleine Anzahl von Teilnehmenden und die Kürze der Skala bedingt sein.

In der Stichprobe GRUNDSCHULE<sub>PRAXIS</sub> hat die reduzierte Skala gar nicht funktioniert. Das heißt zusammenfassend, dass sich das Instrument für die Stichprobe<sub>PRAXIS</sub> nicht bewährt hat.

## Limitationen und Fazit

Auf mehrere Limitationen (knapp akzeptable Reliabilitätswerte, kleine Stichprobe<sub>PRAXIS</sub>, Einsatz eines gekürzten Instruments in der Stichprobe<sub>PRAXIS</sub>) wurde schon hingewiesen. Zudem ist die Messinvarianz hinsichtlich der verschiedenen Gruppen von Lehrkräften nicht gegeben. Da das Ziel der Untersuchung darin besteht, Stichproben mit unterschiedlichem berufsbiografischem Hintergrund und ausbildungsbezogenem Wissen zu vergleichen, wurde hier anstelle des Ausschlusses der betreffenden Items ein interpretativer Umgang gewählt. Die Item-Differenzen wurden gemäß dem Vorschlag von Schulte et al. (2013) als wertvolle Informationsquelle für weiterführende inhaltliche Analysen verwendet.

Nicht untersucht wurde, ob und wie sich Dispositionen der Lehrkräfte (Motivation, Überzeugungen usw.) oder organisatorische Bedingungen (z.B. Ausbildungsort) auf die Ergebnisse auswirken (Blömeke et al., 2013).



Obwohl eine Distraktorenanalyse nur bei wenigen Items auf die Anwendung von Rasterstrategien oder auf eine Plausibilitätsauswahl von Antworten hinweist, kann nicht ausgeschlossen werden, dass die Ergebnisse davon beeinflusst wurden. Es könnte zudem auch sein, dass die Ergebnisse durch die große Varianz der Itemformate beeinflusst worden sind. Dies – und das ist eine weitere Limitation – wurde nicht systematisch überprüft. Allerdings differiert die Schwierigkeit der Items sehr stark innerhalb eines Aufgabenformats. Die Multiple Choice Items 2 und 3 sind beispielsweise viel einfacher als die Items 5 und 6, die dasselbe Format haben (elektronisches Supplement und Abbildung 1). Es könnte sein, dass eher die Vielfalt der Itemformate, deren Auswahl stark auch mit dem Anliegen zusammenhing, das Professionswissen möglichst handlungsnah zu erfassen, eine hohe Flexibilität der Proband:innen erforderte und die Ergebnisse beeinflusste. In weiteren Untersuchungen müsste überprüft werden, ob sich das Instrument mit einem einheitlicheren Itemformat verbessern ließe.

Insgesamt zeigt sich, dass es in den vorliegenden Studien nur teilweise gelungen ist, Instrumente zu entwickeln, die das Professionswissen von Lehrkräften zu Grundlagenwissen und diagnostischem Wissen bei LS bei verschiedenen Gruppen von Studierenden und Lehrkräften reliabel und valide erfassen. Solche Instrumente wären jedoch notwendig, wenn es um die Beantwortung der Frage geht, welches professionelle Wissen und damit verbunden, welche Ausbildung Lehrkräfte brauchen, um Schüler:innen mit LS zu unterstützen.

Weitere Anstrengungen zur Entwicklung von geeigneten Instrumenten sind somit notwendig. Diese könnten in einem ersten Schritt darin bestehen, ein Instrument zu entwickeln, das auf Grundschullehrkräfte ausgerichtet ist. Wichtig dabei ist jedoch, dass die Erhebungen nicht nur mit Studierenden, sondern auch mit Lehrkräften, die ihre Ausbildung schon länger abgeschlossen haben, durchgeführt werden. Zudem

stellt sich die Frage, ob andere Erhebungsformen – z.B. Interviews mit den Lehrkräften zu konkreten Diagnose- und Fördersituationen – eingesetzt werden könnten, um professionelles Wissen zum Thema LS zu erheben.

## Literatur

- Artelt, C. (2009). Diagnostische Urteile von Lehrkräften im Bereich der Lesekompetenz. In A. Bertschi-Kaufmann & C. Rosebrock (Hrsg.), *Literalität. Bildungsaufgabe und Forschungsfeld* (S. 125–136). Juventa.
- Baumert, J. & Kunter, M. (2011). Das Kompetenzmodell von COACTIV. In J. Baumert, M. Kunter, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (S. 29–53). Waxmann.
- Behrmann, L. & Souvignier, E. (2013). The relation between teachers' diagnostic sensitivity, their instructional activities, and their students' achievement gains in reading. *Zeitschrift für pädagogische Psychologie*, 27(4), 283–293. <https://doi.org/10.1024/1010-0652/a000112>
- Björnsson, C. H. (1968). *Läsbarhet*. Liber
- Blömeke, S. (2011). Teacher Education and Development Study: Learning to Teach (TEDS-LT) - Erfassung von Lehrerkompetenzen in gering strukturierten Domänen. In S. Blömeke, A. Bremerich-Vos, H. Haudeck, G. Kaiser, G. Nold, K. Schwippert & H. Willenberg (Hrsg.), *Kompetenzen von Lehramtsstudierenden in gering strukturierten Domänen. Erste Ergebnisse aus TEDS-LT* (S. 7–24). Waxmann.
- Blömeke, S., Buchholtz, C. & Bremerich-Vos, A. (2013). Zusammenhang institutioneller Merkmale im dem Wissenserwerb im Lehramtsstudium. In S. Blömeke, A. Bremerich-Vos, G. Kaiser, G. Nold, H. Haudeck, J.-U. Kessler & K. Schwippert (Hrsg.), *Professionelle Kompetenzen im Studienverlauf. Weitere Ergebnisse zur Deutsch-, Englisch- und Mathematiklehrausbildung aus TEDS-LT* (S. 167–187). Waxmann.

- Bohnert-Kraus, M., Kempe Preti, S., Löffler, C., Haid, A., Reising, L., & Feil, S. (2021). Zum Zusammenhang zwischen laut- sowie meta- und schriftsprachlichen Fähigkeiten bei Erwachsenen mit Illetrismus. *Lernen und Lernstörungen*, 10, 217–230. <https://doi.org/10.1024/2235-0977/a000351>
- Bond, T. G. & Fox, C. M. (2001). *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*. Lawrence Erlbaum.
- Bremerich-Vos, A., Dämmer, J., Willenberg, H. & Schwippert, K. (2011). Professionelles Wissen von Studierenden des Lehramts Deutsch. In S. Blömeke, A. Bremerich-Vos, H. Haudeck, G. Kaiser, G. Nold, K. Schwippert & H. Willenberg (Hrsg.), *Kompetenzen von Lehramtsstudierenden in gering strukturierten Domänen. Erste Ergebnisse aus TEDS-LT* (S. 47–99). Waxmann.
- Döring, N. & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften* (5. Aufl.). Springer.
- Draba, R. E. (1977). *The identification and interpretation of item bias. Educational Statistics Laboratory - Memo 25*. Verfügbar unter: [www.rasch.org/memo25.htm](http://www.rasch.org/memo25.htm)
- Field, A. (2013). *Discovering statistics using IBM SPSS Statistics* (4. Ed.). SAGE.
- Fischer, U. & Klicpera-Gasteiger, B. (2013). Prävention von Leseschwierigkeiten. Diagnose und Förderung im Anfangsunterricht. *Didaktik Deutsch. Halbjahresschrift für die Didaktik der deutschen Sprache und Literatur*, 19(35), 63–81. <https://doi.org/10.25656/01:17162>
- Galuschka, K. & Schulte-Körne, G. (2015). Evidenzbasierte Interventionsansätze und forschungsbasierte Programme zur Förderung der Leseleistung bei Kindern und Jugendlichen mit Lesestörung – Ein systematischer Review. *Zeitschrift für Erziehungswissenschaft*, 18, 473–487. <https://doi.org/10.1007/s11618-015-0650-6>
- Gorecki, B. & Landerl, K. (2015). Phonologische Bewusstheit. Ist die phonologische Bewusstheit ein Prädiktor für die Leseleistung? *Zeitschrift für Entwicklungspsychologie und pädagogische Psychologie*, 47(3), 139–146. <https://doi.org/10.1026/0049-8637/a000135>
- Grand, A., Dittrich, R. & Hatzinger, R. (2017). Präferenzmodell in der Praxis. Analyse von Paarvergleichen, Likert Items und Ranking mit R-prémod. *Facultas*.
- Hanke, P., König, J., Jäger-Biela, D., Pohl, T., Schabmann, A., Becker-Mrotzek, M., Träuble, B. & Schmitt, R. (2019). Professionelles Wissen von Lehramtsstudierenden zum basalen Lesen- und Schreibenlernen – ein interdisziplinäres Projekt. In C. Donie, F. Foerster, M. Obermayr, A. Deckwerth, G. Kammermeyer, G. Lenske, M. Leuchter & A. Wildemann (Hrsg.), *Grundschulpädagogik zwischen Wissenschaft und Transfer* (S.52–58). Springer VS.
- Heine, J.-H. & Tamai, C. (2015). Pairwise Rasch model item parameter recovery under sparse data conditions. *Psychological Test and Assessment Modeling*, 57(1), 3–36.
- Hintz, A.-M. & Grünke, M. (2009). Einschätzungen von angehenden Lehrkräften für Sonder- und allgemeine Schulen zur Wirksamkeit von Interventionen für den Schriftspracherwerb bei lernschwachen Kindern. *Empirische Sonderpädagogik*, 1(1), 45–61. <https://doi.org/10.25656/01:9463>
- Ingenkamp, K. & Lissmann, U. (2008). *Lehrbuch der pädagogischen Diagnostik* (6. Aufl.). Beltz.
- Karing, C., Matthäi, J. & Artelt, C. (2011). Genauigkeit von Lehrerurteilen über die Lesekompetenz ihrer Schülerinnen und Schüler in der Sekundarstufe I – Eine Frage der Spezifität? *Zeitschrift für pädagogische Psychologie*, 25(3), 159–172. <https://doi.org/10.1024/1010-0652/a000041>
- Karing, C., Pfost, M. & Artelt, C. (2011). Hängt die diagnostische Kompetenz von Sekundarstufenlehrkräften mit der Entwicklung der Lesekompetenz und der mathematischen Kompetenz ihrer Schülerinnen und Schüler zusammen? *Journal for educational research online*, 3(2), 119–147. <https://doi.org/10.25656/01:5626>
- Klicpera, C., Schabmann, A., Gasteiger-Klicpera, B. & Schmidt, B. (2017). *Legasthenie – LRS: Modelle, Diagnose, Therapie und Förderung* (5. überarbeitete und erweiterte Aufl.). Reinhardt UTB.
- König, J. (2015). Kontextualisierte Erfassung von Lehrerkompetenz. Einführung in den Thementeil. *Zeitschrift für Pädagogik*, 3(61), 305–309.
- Krauss, S., Lindl, A., Schilcher, A. & Tepner, O. (2017). Das Forschungsprojekt FALKO – ein einleitender Überblick. In S. Krauss, A. Lindl, A. Schilcher, M. Fricke, A. Göhring, B. Hofmann, P. Kirchhoff & R. Mulder (Hrsg.), *FALKO: Fachspezifische Lehrkompetenzen. Konzeption von Professionswissenstests in den Fächern Deutsch, Englisch, Latein, Physik, Musik, Evangelische Religion und Pädagogik* (S. 9–65). Waxmann.

- Kultusministerkonferenz. (2019). Standards Lehrerbildung: Bildungswissenschaften. *Beschluss der Kultusministerkonferenz vom 16.12.2004 i. d. F. vom 16.05.2019* <https://www.kmk.org/themen/allgemeinbildende-schulen/lehrkraefte/lehrerbildung.html>
- Landerl, K. & Wimmer, H. (2008). Development of Word Reading Fluency and Spelling in a Consistent Orthography: An 8-Year Follow-Up. *Journal of Educational Psychology, 100*(1), 150-161. <https://doi.org/10.1037/0022-0663.100.1.150>
- Lietz, M. (2021). *Professionswissen von Lehrpersonen zu Grundlagen des Leserwerbs und Diagnostik bei Leseschwierigkeiten – Entwicklung und Validierung eines Befragungsinstrumentes für ausgebildete und angehende Lehrpersonen und Schulische Heilpädagoginnen und Heilpädagogen. Dissertation*. Universität Zürich, Philosophische Fakultät. <https://doi.org/10.5167/uzh-204969>
- Linacre, J. M. (o.J.). *Reliability and separation of measures*. <https://www.winsteps.com/winman/reliability.htm>
- Linacre, J.M. (2002) What do Infit and Outfit, Mean-square and Standardized mean? *Rasch Measurement Transactions, 16*(2).
- Lombard, M., Snyder-Duch, J. & Bracken, C. (2002). Content Analysis in Mass Communication. Assessment and Reporting of Intercoder Reliability. *Human Communication Research, 28*(4), 587–604. <https://doi.org/10.1111/j.1468-2958.2002.tb00826.x>
- Lüdtke, O. & Robitzsch, A. (2017). Eine Einführung in die Plausible-Values-Technik für die psychologische Forschung. *Diagnostica, 63*(3), 193–205. <https://doi.org/10.1026/0012-1924/a000175>
- Mayer, A. (2018). Benennungsgeschwindigkeit und Lesen. *Forschung Sprache, 6*(1), 20–42.
- Mayer, A. (2016). *Lese-Rechtschreibstörungen (LRS)*. Reinhardt UTB.
- Moats, L. C. & Foorman, B. R. (2003). Measuring Teachers' Content Knowledge of Language and Reading. *Annals of Dyslexia, 53*(1), 23–45. <https://doi.org/10.1007/s11881-003-0003-7>
- Müller, B., Krizan, A., Hecht, T., Richter, T. & Ennemoser, M. (2013). Leseflüssigkeit im Grundschulalter: Entwicklungsverlauf und Effekte systematischer Leseförderung. *Lernen und Lernstörungen, 2*(3), 131–146. <https://econtent.hogrefe.com/doi/pdf/10.1024/2235-0977/a000039>
- 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). Waxmann.
- Nix, D. (2011). *Förderung der Leseflüssigkeit: Theoretische Fundierung und empirische Überprüfung eines kooperativen Lautlese Verfahrens im Deutschunterricht*. Juventa.
- Phelps, G. & Schilling, G. S. (2004). Developing Measures of Content Knowledge for Teaching Reading. *Elementary School Journal, 105*(1), 31–48. <https://doi.org/10.1086/428764>
- Ramseier, E. (2008). Validation of competence models for developing education standards: Methodological choices and their consequences. *Mesure et évaluation en éducation, 31*(2), 35–53. <https://doi.org/10.7202/1025006ar>
- Ramseier, E., Moser, U., Moreau, J. & Antonietti, J.-P. (2008). *HarmoS – Entwicklung von Bildungsstandards. Schlussbericht der Methodologiegruppe*. Schweizerische Konferenz der kantonalen Erziehungsdirektoren (EDK).
- Rjosk, C., McElvany, N., Anders, Y. & Becker, M. (2011). Diagnostische Fähigkeiten von Lehrkräften bei der Einschätzung der basalen Lesefähigkeit ihrer Schülerinnen und Schüler. *Psychologie in Erziehung und Unterricht, 58*(2), 92–105. <http://dx.doi.org/10.2378/peu2011.art04d>
- Rosebrock, C. & Nix, D. (2012). *Grundlagen der Lesedidaktik und der systematischen schulischen Leseförderung* (5. Aufl.). Schneider-Verlag Hohengehren.
- Rosebrock, C., Nix, D., Rieckmann, C. & Gold, A. (2011). *Leseflüssigkeit fördern. Lautleseverfahren für die Primar- und Sekundarstufe*. Kallmeyer.
- Rutsch, J. (2016). *Entwicklung und Validierung eines Vignettentests zur Erfassung des fachdidaktischen Wissens im Leseunterricht bei angehenden Lehrkräften. Dissertation Pädagogische Hochschule Heidelberg*. [https://opus.ph-heidelberg.de/frontdoor/deliver/index/docId/215/file/Dissertation\\_Rutsch\\_Veroeffentlichung.pdf](https://opus.ph-heidelberg.de/frontdoor/deliver/index/docId/215/file/Dissertation_Rutsch_Veroeffentlichung.pdf)
- Scheerer-Neumann, G. (2018). *Lese-Rechtschreib-Schwäche und Legasthenie: Grundlagen, Diagnostik und Förderung* (2. Aufl.). Kolhammer.

- Schmidt, B. M. & Schabmann, A. (2010). „Es ist vorübergehend!“ Lehrereinschätzungen über mögliche Lese-Rechtschreibprobleme – eine klassifikatorische Analyse. *Heilpädagogische Forschung*, 36(3), 106–115.
- Schmidt, B. M. & Schabmann, A. (2016). Wissen und Kompetenzwahrnehmung von Referendaren im Umgang mit LRS. *Heilpädagogische Forschung*, 42(2), 96–108.
- Schneider, W. (2017). *Lesen und Schreiben lernen. Wie erobern Kinder die Schriftsprache?* Springer.
- Schründer-Lenzen, A. (2013). *Schriftspracherwerb* (4. Aufl.). Springer VS.
- Schulte, K., Nonte, S. & Schwippert, K. (2013). Die Überprüfung von Messinvarianz in international vergleichenden Schulleistungsstudien am Beispiel der Studie PIRLS. *Zeitschrift für Bildungsforschung*, 3(99), 99–118. <https://doi.org/10.1007/s35834-013-0062-8>
- Schulte-Körne, G. (2021). Verpasste Chancen: Die neuen diagnostischen Leitlinien zur Lese-, Rechtschreib- und Rechenstörung der ICD-11. *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie*, 49(6), 463–467. <https://doi.org/10.1024/1422-4917/a000791>
- Seifert, A. (2015). Kompetenzforschung in den Fachdidaktiken auf der Grundlage von IRT-Modellen. In U. Riegel, S. Schubert, G. Siebert-Ott & K. Macha (Hrsg.), *Kompetenzmodellierung und Kompetenzmessung in den Fachdidaktiken* (S. 131–161). Waxmann.
- Shulman, L. S. (1986). Those Who Understand: Knowledge Growth in Teaching. *Educational Researcher*, 15(2), 4–14.
- Stancel-Piatak, A., Buchholtz, C. & Schwippert, K. (2013). Anlage und Design der Studie TEDS-LT. In S. Blömeke, A. Bremerich-Vos, G. Kaiser, G. Nold, H. Haudeck, J.-U. Kessler & K. Schwippert (Hrsg.), *Professionelle Kompetenzen im Studienverlauf. Weitere Ergebnisse zur Deutsch-, Englisch- und Mathematiklehrausbildung aus TEDS-LT* (S. 25–46). Waxmann.
- Sturm, A. (2012). Wenn Lesen und Schreiben alles andere als selbstverständlich sind. *Schweizerische Zeitschrift für Heilpädagogik*, 18(5), 5–10.
- Tacke, G. (2011). Ein umfassendes Konzept zur schulischen und häuslichen Lese-Rechtschreibförderung von Klasse 1 bis in die Sekundarstufe. In G. Schulte-Körne (Hrsg.), *Legasthenie und Dyskalkulie: Stärken erkennen – Stärken fördern* (S. 135–163). Winkler.
- Tristán, A. (2006). An Adjustment for Sample Size in DIF Analysis. *Rasch Measurement Transactions*, 20(3), 88–92
- Wilbert, J. & Linnemann, M. (2011). Kriterien zur Analyse eines Tests zur Lernverlaufsdiagnostik. *Empirische Sonderpädagogik*, 3(3), 225–242.
- Wilson, M. (2005). *Constructing measures: An item response modeling approach*. N.J. Lawrence Erlbaum Associates

## Autorinnen- und Autorenhinweis

 Meike Lietz

<https://orcid.org/0000-0002-6522-4094>

 Elisabeth Moser Opitz

<https://orcid.org/0000-0002-5243-4770>

 Meret Stöckli

<https://orcid.org/0000-0003-1518-5960>

Die Autorinnen haben zu gleichen Teilen zur Veröffentlichung beigetragen.

*Korrespondenzadresse:*

Elisabeth Moser Opitz  
 Institut für Erziehungswissenschaft  
 Universität Zürich  
 Freiestrasse 36  
 CH-8032 Zürich  
[elisabeth.moseropitz@uzh.ch](mailto:elisabeth.moseropitz@uzh.ch)

Erstmals eingereicht: 19.07.2022

Überarbeitung eingereicht: 16.01.2023

Angenommen: 20.07.2023

Offene Daten	Die Daten von Studie 1 stehen nicht zur Verfügung. Die Daten von Studie 2 werden nach Projektabschluss zur Verfügung gestellt werden, in welcher Form ist im Moment noch nicht klar.
Offener Code	Keine Angabe
Offene Materialien	<p>Ausführliche Informationen zum Fragebogen sind in folgender Publikation verfügbar:</p> <p>Lietz, M. (2021). Professionswissen von Lehrpersonen zu Grundlagen des Leseerwerbs und Diagnostik bei Leseschwierigkeiten – Entwicklung und Validierung eines Befragungsinstruments für ausgebildete und angehende Lehrpersonen und Schulische Heilpädagoginnen und Heilpädagogen. Dissertation. Universität Zürich, Philosophische Fakultät. <a href="https://doi.org/10.5167/uzh-204969">https://doi.org/10.5167/uzh-204969</a></p> <p>Das Auswertungsmanual kann bezogen werden bei <a href="mailto:Meike.Lietz@phbern.ch">Meike.Lietz@phbern.ch</a></p>
Präregistrierung	Nein
Votum Ethikkommission	Die Ethikkommission der Philosophischen Fakultät der Universität Zürich hat die Studie am 18. August 2018 bewilligt. Für Studie 1 musste keine Bewilligung eingeholt werden.
Finanzielle und weitere sachliche Unterstützung	<p>Studie 1 wurde von der Pädagogischen Hochschule Bern gefördert (Fördernummer 14 s 0002 02).</p> <p>Studie 2 wurde vom Schweizer Nationalfonds zur Förderung der wissenschaftlichen Forschung gefördert (Fördernummer 100019-207807)</p>
Autorenschaft	Meike Lietz hat Studie 1 geplant und durchgeführt und Rückmeldung zum Manuskript gegeben. Elisabeth Moser Opitz hat Studie 2 geplant und das Manuskript verfasst. Meret Stöckli hat den Fragebogen für Studie 2 angepasst, die Daten analysiert und Rückmeldung zum Manuskript gegeben.