

## 4 Ein empirischer Vergleich von Konzepten der Lesekompetenz in PIRLS

*Ann Cathrice George, Alexander Robitzsch, Michael Krelle & Simone Breit*

*Im vorliegenden Artikel werden verschiedene Konzepte von Lesekompetenz aus fachwissenschaftlichen Annahmen, die der PIRLS-Studie zugrunde liegen, abgeleitet und anhand der PIRLS-2016-Daten von Österreich empirisch verglichen. Die betrachteten Konzepte unterscheiden sich durch die Anzahl und Anordnung der die Lesekompetenz beschreibenden Subkompetenzen. Zum Zweck der mehrdimensionalen statistischen Modellierung der unterschiedlichen Konzepte werden kognitiv-diagnostische Modelle (CDMs) eingesetzt. Im Modellvergleich wird ein achtdimensionales Modell vorgezogen, welches zwischen zwei Leseabsichten mit jeweils vier Verstehensprozessen unterscheidet. Implikationen für das Large-Scale Assessment und den Unterricht werden diskutiert.*

Large-Scale-Studien wie die Progress in International Reading Literacy Study (PIRLS; Mullis & Martin, 2015) informieren über Lesekompetenzen von Schülerinnen und Schülern aus einer Reihe von Teilnahmeländern. Die Daten sollen dabei helfen, die Bildungssysteme der teilnehmenden Länder zu vergleichen und zu evaluieren. Um genauere Aussagen zuzulassen, wird im Rahmen der nationalen (z. B. Wallner-Paschon, Itzlinger-Bruneforth & Schreiner, 2017; Hußmann et al., 2017) und internationalen Berichterstattung (Mullis, Martin, Foy & Hooper, 2017) nicht nur auf globale Lesekompetenz eingegangen, sondern auf mehrere untergeordnete Subkompetenzen. Diese Subkompetenzen wurden aus einem der PIRLS-Studie zugrundeliegenden fachwissenschaftlichen Modell zur Beschreibung von Lesekompetenz abgeleitet (Campbell, Kelly, Mullis, Martin & Sainsbury, 2001; Irwin, 1986), das auf Theorien zum Leseverstehen beruht. Die Anzahl der berichteten Subkompetenzen wurde allerdings über die Erhebungszyklen hinweg adaptiert: Während in PIRLS 2001 neben der globalen Lesekompetenz nur zwei untergeordnete Subkompetenzen dargestellt werden, finden sich in den Berichten ab 2006 vier Subkompetenzen.

Der Frage, welche Anzahl von fachwissenschaftlich abgeleiteten Subkompetenzen zu einer möglichst guten Beschreibung der Schülerantworten auf die PIRLS-Testitems beiträgt, wurde empirisch kaum nachgegangen. Durch eine solche Analyse könnte allerdings geklärt werden, ob die Daten für die Berichterstattung optimal genutzt werden: Einerseits könnte das Potenzial bestehen, Ergebnisse über mehr Subkompetenzen als aktuell angenommen zu berichten und damit den Informationsgehalt zu erhöhen. Andererseits könnte abgesichert werden, dass die Ergebnisse in den aktuell berichteten Subkompetenzen nicht redundant sind.

Mit Blick auf den aktuellen Forschungsstand kann auf eine Vielzahl empirischer Studien Bezug genommen werden, in denen Lesekompetenz durch zugrundeliegende Subkompetenzen beschrieben ist. Diese Studien teilen sich im Wesentlichen in zwei Richtungen auf. Der erste Teil der Studien untersucht, welche Subkompetenzen dem Leseprozess zugrunde liegen und welche statistischen Modelle sich eignen,

Lesekompetenz anhand der gefundenen Subkompetenzen zu beschreiben. Bei dem zweiten Teil der Studien handelt es sich um empirisch-methodische Studien mit dem Ziel der Validierung der Eindimensionalitätsannahme, die als Voraussetzung für die statistische Beschreibung von Lesekompetenz anhand von Item-Response-Modellen gilt (Jang & Roussos, 2007; Ponocny, Bauernfeind, Zins & Radinger, 2009; Seo, Chui & Roussos, 2008).

Die erstgenannten empirischen Studien legen diverse Subkompetenzen vor, die auf teils unterschiedliche Aspekte bezogen sind: So wurden beispielsweise mitunter die Länge der Texte (lang vs. kurz) oder das Format der Fragen (Multiple Choice vs. offen oder schlussfolgernd vs. bewertend) als Subkompetenzen betrachtet (Muijselaar et al., 2017). Zudem wurden Aspekte der Sprachkompetenz (u. a. Lexika, Syntax), Lesefertigkeiten bzw. Aspekte der Leseflüssigkeit (z. B. Dekodierungsgeschwindigkeit) modelliert (Foorman, Koon, Petscher, Mitchell & Truckenmiller, 2015; Kendeou, Papadopoulos & Spanoudis, 2012). Es werden eher weit gefasste Subkompetenzen wie auch eher „feinkörnige“ Subkompetenzen beschrieben. Zu den eher weit gefassten gehören Subkompetenzen wie beispielsweise das Verständnis von Wortbedeutungen, das Abgewinnen spezifischer Informationen, das Ziehen von Schlussfolgerungen bzw. die Verknüpfung von Informationen etc. (Chen & Chen, 2016; Chen & de la Torre, 2014; Lee & Sawaki, 2009; Li, Hunter & Lei, 2016; Ravand, 2016; von Davier, 2010). In Studien, in denen eher „feinkörnige“ Subkompetenzen modelliert sind, spielen dagegen Aspekte des (u. a. sprachlichen) Vorwissens oder des Arbeitsgedächtnisses etc. eine Rolle (Buck et al., 2014; Buck, Tatsuoka & Kostin, 1997; Jang, 2009; Jang, Dunlop, Wagner, Kim & Gu, 2013; Kim, 2014).

Bei der Ermittlung der Anzahl relevanter Subkompetenzen zur Beschreibung der Lesekompetenz geht keine der oben genannten Studien direkt auf die in PIRLS genutzten fachwissenschaftlichen Subkompetenzen ein. Eine Ausnahme bilden die Studien von Voss, Carstensen und Bos (2005) und – in einer Wiederholung dieser Studie – Bos, Valtin, Voss, Hornberg und Lankes (2007). Beide Studien beziehen

sich auf fachwissenschaftliche Modelle des Leseverstehens, wie sie u. a. in der Kognitionspsychologie diskutiert werden (vgl. u. a. Rickheit & Strohner, 1993). In der Studie von Voss et al. (2005) wird auf Basis der Antworten deutscher Schüler/innen zu den PIRLS 2001 Items ein Modell mit drei Subkompetenzen im empirischen Vergleich bevorzugt. Die Subkompetenzen bestehen aus den zwei Leseabsichten *Lesen als literarische Erfahrung* und *Lesen zum Informationsgewinn* und der dritten Subkompetenz *Bewerten und Kritisieren von Inhalt und Textelementen beim Lesen zur literarischen Erfahrung*. In der späteren Studie von Bos et al. (2007) wird dagegen, basierend auf den Antworten deutscher Schüler/innen zu den PIRLS-2006-Items, ein Modell mit vier Subkompetenzen gewählt. Die Subkompetenzen umfassen vier Verstehensprozesse: *Erkennen und Wiedergeben explizit angegebener Informationen*, *Ziehen einfacher Schlussfolgerungen*, *Interpretieren sowie Verknüpfen von Gedanken und Informationen* und *Bewerten und Kritisieren von Inhalt und Textelementen*. Beide Studien führen somit zu anderen Anzahlen von Subkompetenzen, als in den internationalen Berichten eingesetzt werden.

In einer Ausweitung der Studien von Voss et al. (2005) und Bos et al. (2007) werden im folgenden Artikel zunächst Konzepte von Lesekompetenz aus verschiedenen der PIRLS-Studie zugrundeliegenden, fachwissenschaftlichen Annahmen abgeleitet. Die Konzepte unterscheiden sich durch die Anzahl und den Zusammenhang zwischen den die Lesekompetenz beschreibenden Subkompetenzen. Anhand der Antworten österreichischer Schüler/innen auf die Items der PIRLS-2016-Studie werden die unterschiedlichen Konzepte anschließend statistisch modelliert und verglichen. Zum Zweck der statistischen Modellierung kommen kognitiv diagnostische Modelle (CDMs; DiBello, Roussos & Stout, 2007) zum Einsatz.

## Ableitung von Konzepten der Lesekompetenz

PIRLS misst die Lesekompetenz von Schülerinnen und Schülern in der vierten Jahrgangsstufe. Dabei wird basierend auf der Definition von Elley (1992) Lesekompetenz als Resultat einer längeren Entwicklungsphase definiert als

the ability to understand and use those written language forms required by society and/or valued by the individual. Young readers can construct meaning from variety of texts. They read to learn, to participate in communities of readers in school and everyday life, and for enjoyment. (Mullis & Martin, 2015)

Diese Definition spiegelt sich in der Operationalisierung der Lesekompetenz in PIRLS anhand zweier Aspekte wider: PIRLS unterscheidet mehrere (*Lese-*)*Verstehensprozesse*, um

zu messen, wie Schüler/innen den Texten Bedeutungen abgewinnen, und legt *Leseabsichten* fest, die mögliche Gründe benennen, aus denen Viertklässler lesen. Die geschilderten Aspekte bilden die Basis des fachwissenschaftlichen Lesekompetenzmodells, welches der PIRLS-Studie zugrunde liegt.

## Ableitung von Konzepten aus dem fachwissenschaftlichen Modell der Lesekompetenz

Die im fachwissenschaftlichen Lesekompetenzmodell der PIRLS-Studie beschriebenen Verstehensprozesse erlauben Leserinnen und Lesern, Texten Bedeutungen abzugewinnen. Die Verstehensprozesse sind auf kognitionspsychologische Modelle zurückzuführen (Irwin, 1986; van Dijk & Kintsch, 1983). Es werden vier Verstehensprozesse unterschieden: (I) *Erkennen und Wiedergeben explizit angegebener Informationen*, (II) *Ziehen einfacher Schlussfolgerungen*, (III) *Interpretieren sowie Verknüpfen von Gedanken und Informationen* und (IV) das *Bewerten und Kritisieren von Inhalt und Textelementen* (Campbell et al., 2001). Die Verstehensprozesse sind jeweils durch verschiedene Kompetenzen der Schüler/innen charakterisiert. Einige Aspekte der Beschreibung der Verstehensprozesse können der rechten Spalte von Tabelle 2 entnommen werden (vgl. Wallner-Paschon & Itzlinger-Bruneforth, 2017; Wallner-Paschon et al., 2017). Die ersten beiden Verstehensprozesse (I) und (II) sind primär durch die Nutzung sogenannter „textimmanenter Informationen“ bestimmt und werden daher häufig zu einem übergeordneten Verstehensprozess (I/II) *Wiedergeben und Schlussfolgern* zusammengefasst. Die Prozesse (III) und (IV) basieren dagegen auf mehr oder weniger komplexem Vorwissen bzw. „externem Wissen“. Sie werden zum übergeordneten Verstehensprozess (III/IV) *Interpretieren, Verknüpfen und Bewerten* zusammengefasst. Dem Kompetenzmodell folgend, sind die vier Verstehensprozesse bzw. die zwei übergeordneten Verstehensprozesse als nicht hierarchisch zu verstehen. Es werden also explizit nicht die Fähigkeiten eines Verstehensprozesses als Voraussetzung für Fähigkeiten in einem anderen Verstehensprozess gesehen.

Zusätzlich zu den Verstehensprozessen unterscheidet PIRLS zwei Leseabsichten: das Lesen von Texten, um (a) *literarische Erfahrungen* zu machen und um (b) *Information zu gewinnen*. Im Lesekompetenzmodell wird nicht von einer Verbindung zwischen den Verstehensprozessen und den Leseabsichten ausgegangen. D. h., Schüler/innen, die beispielsweise den Verstehensprozess (I) beherrschen, tun dies unabhängig davon, ob sie vor dem Hintergrund literarischer Erfahrungen lesen oder ob sie lesen, um Informationen zu gewinnen.

Aus diesen fachwissenschaftlichen Annahmen zur Lesekompetenz in PIRLS können folglich zwei unterschiedliche Kon-

zepte von Lesekompetenz abgeleitet werden. Demnach kann die Beschreibung der Lesekompetenz erfolgen durch die

- vier Verstehensprozesse (I) bis (IV), die als nicht hierarchisch angesehen werden und unabhängig von der Leseabsicht sind, oder durch die
- zwei Leseabsichten (a) und (b).

### Ableitung von Konzepten aus der Itemkonstruktion

Die Leseintentionen und Verstehensprozesse werden im PIRLS-Test durch die Testitems verankert. Im Erstellungsprozess wird jedes Item so konstruiert, dass es genau eine Leseabsicht (a oder b) und genau einen Verstehensprozess (I bis IV) abdeckt. Somit lässt sich jedes Item genau einer Zelle in Tabelle 1 zuordnen. Ein Item, welches beispielsweise das *Erkennen und Wiedergeben explizit angegebener Informationen* (I) während des Lesens zur *literarischen Erfahrung* (a) testet, liegt diesem Schema zufolge in Zelle (Ia).

Die beiden Leseabsichten werden durch ihre Verbindung zu bestimmten Textgattungen bzw. -sorten operationalisiert: Während das Lesen zur literarischen Erfahrung anhand fiktionaler Texte stattfindet, ist das Lesen zum Informationsgewinn mit Sachtexten verbunden. Aus den Zellen von Tabelle 1 sind beispielhafte generische Fragestellungen zu jeder Kombination von Leseabsichten und Verstehensprozessen zu entnehmen (Bos et al., 2007, S. 87). Laut Voss et al. (2005, S. 21) können zu jeder Kombination von Leseabsicht und Verstehensprozess Items jeder Schwierigkeit erstellt werden. Aus dieser Sicht sind die Items im Mittel in keiner Zelle

schwieriger als in einer anderen; schließlich werden die Verstehensprozesse als nicht hierarchisch angenommen. Allerdings sind die Verstehensprozesse hier als abhängig von der Leseabsicht definiert. D. h., Schüler/innen, die beispielsweise Verstehensprozess (I) beim Lesen zur literarischen Erfahrung beherrschen, beherrschen den Verstehensprozess nicht unbedingt auch beim Lesen zum Informationsgewinn.

Aus den Annahmen zur Itemkonstruktion kann somit ein Konzept abgeleitet werden, in dem die Beschreibung der Lesekompetenz erfolgt durch

- acht Kombinationen (Ia bis IVb) aus jeweils vier Verstehensprozessen und zwei Leseabsichten. Dabei werden die Verstehensprozesse jeweils als nicht hierarchisch angesehen. Durch die Kombinationen sind die Fähigkeiten der Schüler/innen in den Verstehensprozessen von den Leseabsichten abhängig.

### Ableitung von Konzepten aus der statistischen Modellierung und Berichterstattung

Für jedes teilnehmende Land werden die Ergebnisse zur Lesekompetenz auf einer globalen Skala (L) und in vier untergeordneten Subkompetenzen berichtet. Bei den vier untergeordneten Subkompetenzen handelt es sich um die beiden oben genannten Leseabsichten *literarische Erfahrung* (a) und *Informationsgewinnung* (b) sowie die beiden übergeordneten Verstehensprozesse *Wiedergeben und Schlussfolgern* (I/II) und *Interpretieren, Verknüpfen und Bewerten* (III/IV). Die beiden Subkompetenzen (I/II) und (III/IV) wurden im internationalen Bericht von PIRLS 2001 nicht berichtet.

		Leseabsicht	
		(a) Literarische Erfahrung	(b) Informationsgewinn
(I/II) Wiedergeben und Schlussfolgern	(I) Erkennen und Wiedergeben explizit angegebener Informationen	(Ia) Was macht die Hauptfigur am Anfang der Geschichte?	(Ib) Was für Konsequenzen werden in diesem Text genannt?
	(II) Ziehen einfacher Schlussfolgerungen	(IIa) Wie haben die Handlungsträger die Probleme bewältigt?	(IIb) Warum hat die in diesem Artikel beschriebene Person ...?
(III/IV) Interpretieren, Verknüpfen und Bewerten	(III) Interpretieren sowie Verknüpfen von Gedanken und Informationen	(IIIa) Was würde in der Geschichte geschehen, wenn ... ?	(IIIb) Was würde nach diesen Informationen geschehen, wenn ...?
	(IV) Bewerten und Kritisieren von Inhalt und Textelementen	(IVa) Wie gelang dem Autor dieses überraschende Ende?	(IVb) Was kannst du mit den Informationen anfangen?

Tabelle 1: Beispielhafte generische Fragestellungen zur Operationalisierung der Leseabsichten (Spalten) und der Verstehensprozesse (Zeilen) in PIRLS (entnommen aus Bos, Valtin, Voss, Hornberg & Lankes, 2007, S. 87).

Um zu den Werten der Lesekompetenz in einer Subkompetenz zu gelangen, werden jeweils alle im Test eingesetzten Items skaliert, die während der Itemkonstruktion der jeweiligen Subkompetenz zugewiesen wurden. Zur Bestimmung des globalen Lesekompetenzwerts werden somit alle Items gemeinsam skaliert. Zur Bestimmung der Kompetenzwerte in der Subkompetenz Lesen zur *literarischen Erfahrung* (a) werden alle Items skaliert, die (a) zugeordnet werden, und zwar unabhängig davon, welchen Verstehensprozess sie abdecken. In die Berechnung von Subkompetenz (a) fließen also alle Items ein, die in Tabelle 1 den Zellen (Ia), (IIa), (IIIa) und (IVa) zuordenbar sind. Analog fließen beispielsweise in die Berechnung der Subkompetenz *Wiedergeben und Schlussfolgern* (I/II) alle Items ein, die der Subkompetenz (I/II) zugeordnet werden, unabhängig davon welche Leseabsicht sie überprüfen (Zellen Ia, Ib, IIa und IIb). Durch die so vorlie-

genden Berechnungen werden die Verstehensprozesse dort als unabhängig von den Leseabsichten aufgefasst.

Die im vorherigen Absatz beschriebenen Berechnungen dienen hauptsächlich der Ermittlung und dem Vergleich von Kompetenzmittelwerten. Um auch inhaltliche Aussagen über die Kompetenzen der Schüler/innen machen zu können, werden für die Skala der globalen Lesekompetenz Kompetenzstufen gebildet. Dazu wird die Skala in gleich große Abschnitte geteilt. Die Kompetenzen der Schüler/innen in den Abschnitten werden jeweils durch typische Aufgabenmerkmale charakterisiert (Mullis et al., 2017, S. 52 f.). Da die Abschnitte stufenweise aufeinander aufbauen, verfügen Schüler/innen einer Stufe jeweils auch über die Kompetenzen der niedrigeren Stufe(n). Die Beschreibungen der Kompetenzstufen sind in der linken Spalte von Tabelle 2 gegeben. Während

Kompetenzstufen	Verstehensprozesse
<p><b>Stufe 4: sehr hohe Lesekompetenz</b></p> <ul style="list-style-type: none"> <li>■ Verknüpfen von Informationen (Ideen und Beweisen) über einen Text hinweg, um übergeordnete Themen zu erfassen (Erklärungen zu liefern);</li> <li>■ Interpretieren von vielschichtigen Informationen (Ereignissen und Handlungen);</li> <li>■ Interpretieren von bildlichen und inhaltlichen Merkmalen (Ereignissen und Handlungen), um die Autorenperspektive zu beleuchten (die Charakterentwicklung zu erklären).</li> </ul>	<p><b>Verstehensprozess IV: Bewerten und Kritisieren von Inhalt und Textelementen</b></p> <ul style="list-style-type: none"> <li>■ Beurteilung der Vollständigkeit und Klarheit der Informationen im Text;</li> <li>■ Die Wahrscheinlichkeit beurteilen, dass ein in der Geschichte beschriebenes Ereignis wirklich stattfinden könnte;</li> <li>■ Eine Sichtweise der Autorin/des Autors auf das zentrale Thema ermitteln</li> </ul>
<p><b>Stufe 3: hohe Lesekompetenz</b></p> <ul style="list-style-type: none"> <li>■ Auffinden und Unterscheiden wesentlicher Informationen (Handlungen, Details) die in umfangreiche Texte eingebettet sind;</li> <li>■ Ziehen von Schlussfolgerungen für Erklärungen von Sachverhalten (Beziehungen zwischen Absichten, Handlungen, Ereignissen und Gefühlen);</li> <li>■ Interpretieren und Verknüpfen von textlicher Information (Textereignissen);</li> <li>■ Bewerten und Verallgemeinern von Textelementen.</li> </ul>	<p><b>Verstehensprozess III: Interpretieren sowie Verknüpfen von Gedanken und Informationen</b></p> <ul style="list-style-type: none"> <li>■ Herauslesen der übergeordneten Botschaft eines Textes;</li> <li>■ Vergleichen und Gegenüberstellen von Textinformationen;</li> <li>■ Erwägung alternativer Verhaltensweisen von Charakteren.</li> </ul>
<p><b>Stufe 2: mittlere Lesekompetenz</b></p> <ul style="list-style-type: none"> <li>■ Auffinden und Wiedergeben von zwei oder drei Informationen im Text (Handlungen, Ereignisse, Gefühle);</li> <li>■ Ziehen einfacher Schlussfolgerungen;</li> <li>■ (Erstes) Interpretieren und Vernetzen von Informationen (Ereignisse, Gründe, Ursachen).</li> </ul>	<p><b>Verstehensprozess II: Ziehen einfacher Schlussfolgerungen</b></p> <ul style="list-style-type: none"> <li>■ Beschreibung der Beziehung zwischen zwei Charakteren;</li> <li>■ Schlussfolgern, dass ein Ereignis ein anderes verursacht hat;</li> <li>■ Generalisierungen im Text erkennen.</li> </ul>
<p><b>Stufe 1: niedrige Lesekompetenz</b></p> <ul style="list-style-type: none"> <li>■ Auffinden und Benennen explizit angegebener genannter Informationen;</li> <li>■ (Erstes) Ziehen einfacher Schlussfolgerungen;</li> <li>■ Vorläufiges Interpretieren von Leitgedanken und Ereignissen.</li> </ul>	<p><b>Verstehensprozess I: Erkennen und Wiedergeben explizit angegebener Informationen</b></p> <ul style="list-style-type: none"> <li>■ Nach expliziten Gedanken im Text suchen;</li> <li>■ Zeitliche oder örtliche Angaben erkennen;</li> <li>■ Nach Wortdefinitionen suchen.</li> </ul>
<p><b>Unter Stufe 1</b></p> <p>Keine Beschreibung</p>	Keine Verstehensprozesse

Tabelle 2: Beschreibung der Kompetenzstufen für die Leseabsichten Informationsgewinn und in Klammern literarische Erfahrung (linke Spalte) sowie Beschreibung der Verstehensprozesse aus dem Kompetenzmodell (rechte Spalte).

die Stufen getrennt für die beiden Leseabsichten beschrieben werden (Mullis et al., 2017; Wallner-Paschon et al., 2017), erfolgt die Berechnung der Stufen wie oben beschrieben für die Gesamtskala der Lesekompetenz.

Bei einem Vergleich der Beschreibungen der Kompetenzstufen und der Verstehensprozesse (vgl. Tabelle 2) fällt auf, dass die Kompetenzen der Schüler/innen auf den Stufen sehr ähnlich denen der jeweils korrespondierenden Verstehensprozesse beschrieben sind. So unterscheidet sich beispielsweise die Beschreibung von Stufe 2 nicht wesentlich von Verstehensprozess (II). Leitet man aus den ähnlichen Beschreibungen eine Analogie zwischen den Kompetenzstufen und den Verstehensprozessen ab, so entsteht die Annahme einer Hierarchie in den Verstehensprozessen. Diese macht Verstehensprozess (I) zum leichtesten und Verstehensprozess (IV) zum schwierigsten Prozess.

Aus der statistischen Modellierung und der Berichterstattung können also zwei weitere Konzepte abgeleitet werden. In diesen erfolgt die Beschreibung der Lesekompetenz durch die

- zwei übergeordneten Verstehensprozesse (I/II) und (III/IV), die nicht hierarchisch angesehen werden und unabhängig von der Leseabsicht sind und die
- vier Verstehensprozesse (I) bis (IV) bzw. die Kompetenzstufen, die hierarchisch angesehen werden und unabhängig von der Leseabsicht sind.

---

## Forschungsfragen

Insgesamt konnten wir aus den verschiedenen Quellen innerhalb des PIRLS Frameworks, d.h. dem fachwissenschaftlichen Lesekompetenzmodell, der Itementwicklung und der statistischen Modellierung bzw. Berichterstattung fünf alternative Konzepte zur Beschreibung der Lesekompetenz ableiten. Aus diesen Konzepten lassen sich die folgenden drei Forschungsfragen im Hinblick auf die Beschreibung der Lesekompetenz ableiten. Die Forschungsfragen bauen dabei in dem Sinne aufeinander auf, dass die nachstehenden Forschungsfragen abhängig von den empirischen Ergebnissen in den vorangehenden Forschungsfragen sind.

*Forschungsfrage 1:* Lassen sich die beiden Leseabsichten – (a) *literarische Erfahrungen* und (b) *Informationsgewinnung* – statistisch voneinander abgrenzen? Lassen sich die beiden übergeordneten Verstehensprozesse – *Wiedergeben und Schlussfolgern* (I/II) und *Interpretieren, Verknüpfen und Bewerten* (III/IV) – statistisch voneinander abgrenzen? Ist die Beschreibung der Lesekompetenz durch die Leseabsichten der Beschreibung der Lesekompetenz durch die Verstehensprozesse statistisch vorzuziehen?

*Forschungsfrage 2:* Lassen sich in jeder Leseabsicht die zwei übergeordneten Verstehensprozesse (I/II) und (III/IV) oder die vier Verstehensprozesse (I) bis (IV) statistisch voneinander abgrenzen? Unterstützen die Antworten der Schüler/innen dabei eine Hierarchieannahme zwischen den Verstehensprozessen (I) bis (IV) in den beiden Leseabsichten?

*Forschungsfrage 3:* Ist die gemeinsame Beschreibung der Lesekompetenz durch die beiden Leseabsichten (a) und (b) und durch die vier Verstehensprozesse (I) bis (IV) der alleinigen Beschreibung der Lesekompetenz durch die beiden Leseabsichten statistisch vorzuziehen?

---

## Methodik

Die verschiedenen Konzepte zur Lesekompetenz werden in der Reihenfolge der Forschungsfragen durch kognitiv diagnostische Modelle (CDMs; DiBello et al., 2007) statistisch modelliert und verglichen. Alle Modelle wurden in der statistischen Software R (R Core Team, 2018) mit dem Paket CDM (Robitzsch, Kiefer, George & Ünlü, 2018; George, Robitzsch, Kiefer, Groß & Ünlü, 2016) geschätzt.

### Kognitiv diagnostische Modelle

Die Klasse der CDMs erlaubt mehrdimensionale Kompetenzklassifikationen von Schülerinnen und Schülern. Die Dimensionen sind durch Subkompetenzen definiert, welche vorab von Didaktikerinnen und Didaktikern oder Fachexpertinnen und Fachexperten aus einer fachwissenschaftlichen Theorie (z. B. einem Kompetenzmodell) abgeleitet werden. Die Subkompetenzen werden im Kontext von CDMs häufig als Skills bezeichnet. Zur Lösung der Items benötigen die Schüler/innen einen oder mehrere der Skills. Die Expertinnen und Experten definieren, welche Skills dies sind und weisen sie den Items in einer sogenannten Q-Matrix zu (Tatsuoka, 1983).

Durch die Wahl eines spezifischen kognitiv diagnostischen Modells aus der Klasse der CDMs werden Annahmen über das Zusammenwirken der Skills untereinander und die gemeinsamen Einflüsse der Skills auf die Antwortwahrscheinlichkeiten zu den Items getroffen (für eine ausführliche Beschreibung zur Modellwahl siehe DiBello et al., 2007 oder Rupp, Templin & Henson, 2010). Bei den im Folgenden modellierten CDMs treten allerdings nur Q-Matrizen mit Einfachladungen auf, d. h., jedem Item ist jeweils nur ein Skill zugeordnet. In diesem Fall ist die Wahl des spezifischen CDMs einfacher, da keine Regeln für das Zusammenwirken von Skills definiert werden müssen und somit viele spezifische Modelle (bspw. Deterministic-Input-Noisy-“And“-Gate-Modell, Deterministic-Input-Noisy-“Or“-Gate-Modell, Generalized-Deterministic-Input-Noisy-“And“-Gate-Modell) übereinstimmen. Aus diesem Grund werden alle folgenden

Modelle durch Deterministic-Input-Noisy-“And“-Gate-Modelle (DINA; Haertel, 1989) beschrieben.

Die Hauptergebnisse aus einem CDM beschreiben, wie viel Prozent der Schüler/innen die jeweiligen Skills beherrschen und wie viel Prozent der Schüler/innen Kombinationen von bestimmten Skills (sog. Skillklassen) beherrschen. Da in CDMs zumeist davon ausgegangen wird, dass die Schüler/innen einen Skill beherrschen oder nicht beherrschen (für polytome Ausweitungen siehe z. B. Chen & de la Torre, 2013), ist bei  $K$  Skills eine Zuordnung der Schüler/innen in  $2^K$ -Klassen möglich. Im Falle von drei Skills wären somit  $2^3 = 8$  Skillklassen möglich, die als [000], [100], [010], [001], [110], [101], [011] und [111] notiert werden. Anhand statistischer Gütekriterien wie beispielsweise dem AIC, BIC oder CAIC können CDMs mit unterschiedlicher Anzahl von Skills oder unterschiedlichen Q-Matrizen verglichen werden. Basierend auf den im Vergleich jeweils niedrigsten Ausprägungen von AIC, BIC und CAIC wird das Modell gewählt, welches die Daten am besten beschreibt.

Durch spezielle Formen der Reduzierung der  $2^K$ -Klassen (d. h. durch Nullsetzen der Auftretenswahrscheinlichkeit bestimmter Skillklassen) können Annahmen über Strukturen zwischen den Skills wie beispielsweise Hierarchien modelliert werden (Leighton, Gierl & Hunka, 2004). Dazu werden die Auftretenswahrscheinlichkeiten aller Skillklassen, die der Modellannahme widersprechen (z. B. der Hierarchie), vor der Berechnung des Modells auf null gesetzt. Somit können bei der Schätzung des Modells keine Schüler/innen in diese Klassen gruppiert werden. Besteht beispielsweise im Falle von drei Skills die Annahme einer linearen Hierarchie zwischen den Skills, so werden die Auftretenswahrscheinlichkeiten der Skillklassen, die dieser Hierarchie widersprechen (d.h. [010], [001], [101], und [011]) auf null gesetzt.

Durch diese Strategie können CDMs auch in ihrer parametrischen Komplexität reduziert werden. Dazu werden Skillklassen mit niedrigen Auftretenswahrscheinlichkeiten nullgesetzt und somit nicht weiter geschätzt. Durch dieses Vorgehen wird die Anzahl der Modellparameter im Allgemeinen deutlich reduziert. Außerdem wird die Interpretation des Modells auf Skillklassen fokussiert, die in der Stichprobe hinreichend häufig und über zufällige statistische Variabilität hinausgehend auftreten. Eine Reduktion von Skillklassen kann außerdem dazu führen, dass sich die Skillklassen partiell hierarchisch oder vollständig hierarchisch ordnen lassen (Leighton et al., 2004).

### Vorgehen

Um die unterschiedlichen Konzepte der Lesekompetenz zu modellieren und damit die Forschungsfragen beantworten zu können, wählen wir ein adaptives Verfahren: Ausgehend von einem Basismodell werden nach und nach Modelle mit mehr Annahmen hinzugefügt (vgl. Abbildung 1). In jedem Schritt


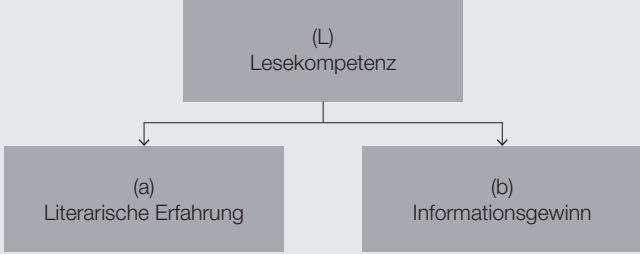
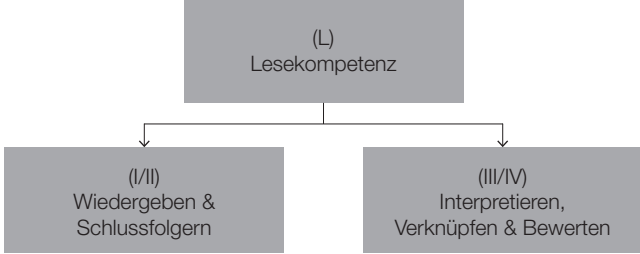
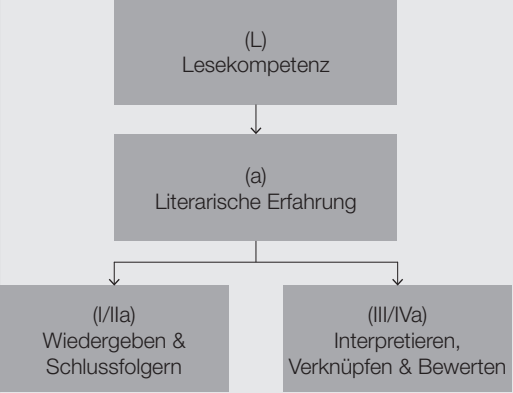
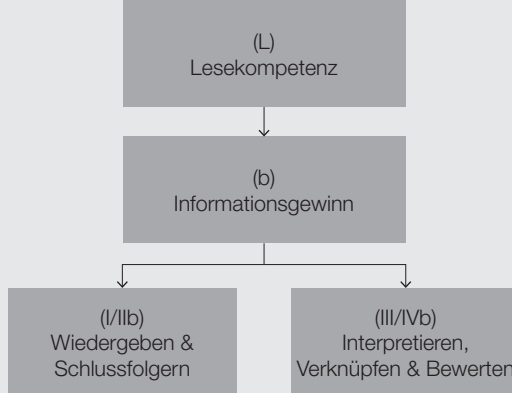
wird jeweils ein komplexeres Modell mit dem vorhergehenden anhand von Modellfitkriterien verglichen (siehe „Entscheidungsbaum“ in Abbildung 2). Dadurch wird schrittweise ein jeweils komplexeres Modell angepasst, ohne die Parameteranzahl unnötigerweise zu vergrößern.

Nach der Modellierung von Lesekompetenz als eindimensionales Konstrukt (Modell 0) wird zunächst geprüft (Forschungsfrage 1), ob die PIRLS Daten eine Trennung von Lesekompetenz in den beiden Leseabsichten (Modell 1) und in den beiden übergeordneten Verstehensprozessen (Modell 2) unterstützen. Es wird zusätzlich geprüft, ob die Beschreibung der Lesekompetenz durch die Leseabsichten der Beschreibung der Lesekompetenz durch die Verstehensprozesse statistisch vorzuziehen ist.

Sollten beide Modelle 1 und 2 einen besseren Fit aufweisen als das eindimensionale Modell 0, so wird für den nächsten Schritt der Modellierung Lesekompetenz anhand der beiden Skills (Leseabsichten oder Verstehensprozesse) beschrieben, die sich vorher als statistisch bedeutsamer zeigten. Für eine einfachere Darstellung in Abbildung 1 sei hier vorweggenommen, dass die Leseabsichten die Antworten österreichischer Schüler/innen auf die PIRLS-Items besser beschreiben konnten.

Im nächsten Schritt wird daher eine Aufgliederung des Lesekonstrukts zum einen in Lesen als literarische Erfahrung (a) und zum anderen in Lesen zum Informationsgewinn (b) angenommen. Es wird für jede der beiden Leseabsichten getrennt getestet (Forschungsfrage 2), ob eine weitere Trennung in die beiden übergeordneten Verstehensprozesse (Modell 3a für literarische Erfahrung und Modell 3b für Informationsgewinn) oder die vier Verstehensprozesse zusätzlichen Erklärungsgehalt liefert. Dabei werden die Verstehensprozesse unter der Annahme keiner Hierarchie (Modell 4a für literarische Erfahrung und Modell 4b für Informationsgewinn) und einer vorhandenen linearen Hierarchie geprüft (Modell 5a für literarische Erfahrung und Modell 5b für Informationsgewinn). Da die Modelle 4a und 4b unter Annahme nicht hierarchischer Verstehensprozesse parametrisch aufwändig sind, werden zur besseren Vergleichbarkeit dieser Modelle parametrisch reduzierte Varianten abgeleitet (Modelle 4a-r und 4b-r; siehe Abbildung 2).

Aus den Modellen 3a, 4a, 4a-r und 5a wird das Modell ausgewählt, welches die beste Beschreibung der Antworten österreichischer Schüler/innen auf die PIRLS-Items zur Leseabsicht literarische Erfahrung (a) liefert. Ebenso wird mit den Modellen 3b, 4b, 4b-r und 5b zur Beschreibung der Leseabsicht Informationsgewinn (b) vorgegangen. Die beiden gewählten Modelle werden abschließend zu einem Gesamtmodell 6 für beide Leseabsichten zusammengefügt und es wird geprüft (Forschungsfrage 3), ob die gemeinsame Beschreibung beider Leseabsichten mit den vorher ermittelten zugrundeliegenden Skills (übergeordnete Verstehensprozesse

Modell	Annahmen
0	<div style="text-align: center;">  </div> <p>Lesekompetenz ist ein eindimensionales Konstrukt, welches sich nicht weiter aufgliedern lässt. Alle denkbaren zugrundeliegenden Skills sind nicht empirisch unterscheidbar.</p>
1	<div style="text-align: center;">  </div> <p>Lesekompetenz lässt sich aufteilen in zwei unterschiedliche Leseabsichten.</p> <p>Schüler/innen, die eine hohe/niedrige Kompetenz in (a) haben, müssen also nicht zwangsweise auch eine hohe/niedrige Kompetenz in (b) haben oder umgekehrt. Die Werte können allerdings korreliert sein.</p>
2	<div style="text-align: center;">  </div> <p>Lesekompetenz lässt sich aufteilen in die zwei übergeordneten Verstehensprozesse <i>Wiedergeben und Schlussfolgern</i> (I/II) und <i>Interpretieren, Verknüpfen und Bewerten</i> (III/IV).</p> <p>Schüler/innen, die eine hohe/niedrige Kompetenz in (I/II) haben, müssen in dieser Perspektive nicht unbedingt auch eine hohe/niedrige Kompetenz in (III/IV) haben oder umgekehrt. Die Werte können allerdings korreliert sein.</p> <p>Die Kompetenzen der Schüler/innen in (I/II) und (III/IV) sind unabhängig von den Leseabsichten.</p>
3a 3b	<div style="display: flex; justify-content: space-around;"> <div style="text-align: center;">  </div> <div style="text-align: center;">  </div> </div> <p>Die beiden Leseabsichten (a) und (b) lassen sich jeweils in die übergeordneten Verstehensprozesse des <i>Wiedergebena und Schlussfolgerns</i> (I/II) und <i>Interpretierens, Verknüpfens und Bewertens</i> (III/IV) aufgliedern.</p> <p>Die beiden Leseabsichten mit ihren jeweiligen Verstehensprozessen sind unabhängig voneinander.</p>

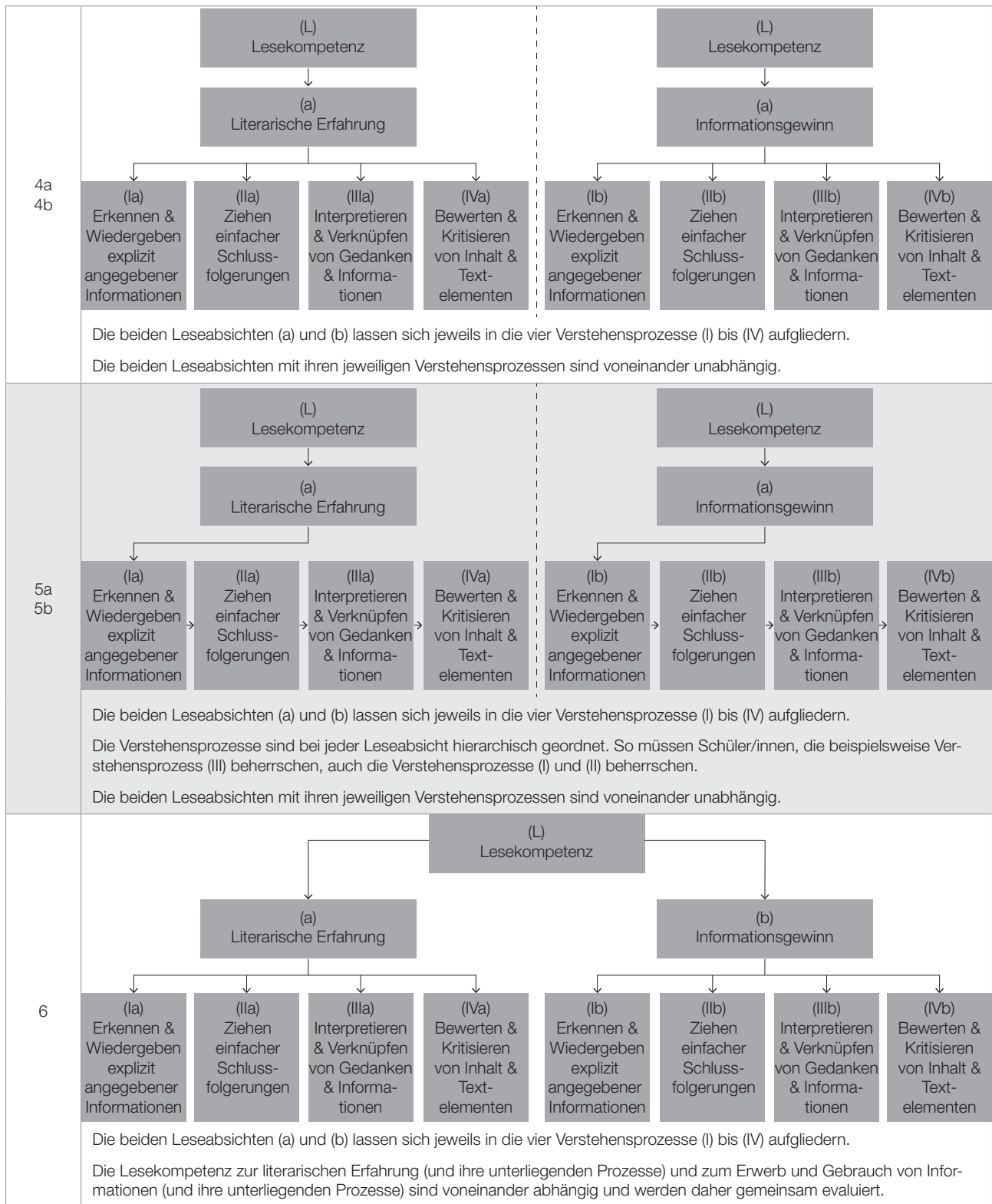


Abbildung 1: Modelle zur Evaluation der verschiedenen Lesekonzepte.



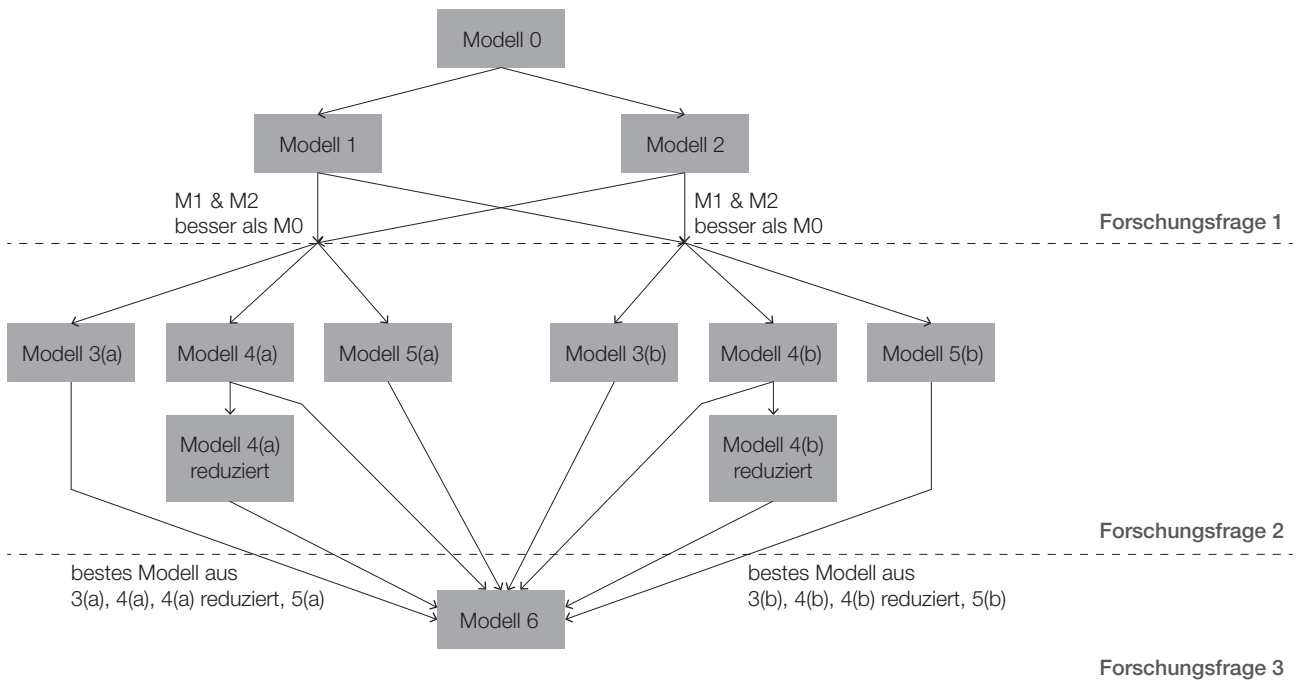


Abbildung 2: Abfolge der im Artikel durchgeführten Modellvergleiche. Die jeweiligen Beschreibungen der Modelle sind Abbildung 1 zu entnehmen.

oder Verstehensprozesse) zu einem höheren Informationsgewinn führt als die getrennte.

Für die Berechnung der DINA-Modelle wurden die Antwortdaten der Schüler/innen zu den Items dichotomisiert. Dabei wurden die jeweils höchste Antwortkategorie als richtig und

### Daten

Die in den Analysen genutzten Daten stammen aus der PIRLS-2016-Studie für Österreich. In Österreich nahmen 4360 Schüler/innen aus 150 Schulen und 257 Klassen an PIRLS teil. Die teilnehmenden Klassen wurden einem Stichprobenplan folgend zufällig ausgewählt (Martin, Mullis & Hooper, 2017). Die Stichprobe gilt als repräsentativ für Österreich (Wallner-Paschon et al., 2017, S. 24).

Um Lesekompetenzen in PIRLS zu messen, wurden in Österreich 181 Aufgaben bzw. Items zu 12 Texten administriert. Jede Schülerin/jeder Schüler erhielt ein Testheft, welches einen literarischen Text (zum Lesen als literarische Erfahrung) und einen Sach- und Gebrauchstext (zum Lesen als Informationsgewinn) enthielt. Zu den jeweiligen Texten wurden zwischen 12 und 20 Items vorgelegt. Die Aufteilung der Texte und der Items auf die Schüler/innen erfolgte nach einem Multi-Matrix-Design (Martin et al., 2017). Insgesamt bestanden die 181 Items aus 86 Multiple-Choice-Items und 95 offenen Items. Tabelle 3 enthält die Anzahl der Items, die jeweils einer Leseintention bzw. einem Verstehensprozess zugeordnet wurden.

		Verstehensprozess				
		(I)	(II)	(III)	(IV)	
Leseabsichten	(a) Literarische Erfahrung	24	24	31	14	93
	(b) Informationsgewinn	26	29	22	11	88
		50	53	53	25	181

Tabelle 3: Anzahl der Items in PIRLS 2016, die den beiden Leseintentionen und den vier Verstehensprozessen zugeordnet sind.

alle anderen Kategorien als falsch kodiert. Items, die nicht erreicht oder ausgelassen wurden, wurden ebenso als falsch kodiert. Alle Modelle wurden unter Berücksichtigung der Stichprobengewichte geschätzt.

Die Zuweisung der Items zu den jeweiligen Skills (Q-Matrix) ist dem PIRLS-Datenmaterial zu entnehmen („Item Information Table“; TIMSS and PIRLS International Study Center, 2018).

## Ergebnisse

Im Folgenden werden die Ergebnisse der Studie in der Reihenfolge der Forschungsfragen präsentiert.

*Forschungsfrage 1:* Nach der Modellierungsreihenfolge in Abbildung 2 wird zunächst das Modell, welches Lesekompetenz als eindimensionales Konstrukt (Modell 0) annimmt, mit den beiden Modellen verglichen, welche zum einen Lesekompetenz in die zwei Leseabsichten (a) und (b) aufgliedern (Modell 1) und zum anderen Lesekompetenz in zwei übergeordneten Verstehensprozessen (I/II) und (III/IV) auffassen (Modell 2). Aus Tabelle 4 geht hervor, dass sowohl das Modell, welches nach Leseabsichten unterscheidet, wie auch das Modell, welches auf den übergeordneten Verstehensprozessen aufbaut, in allen betrachteten Modellfitkriterien (AIC, BIC und CAIC) niedrigere Werte erzielt als das eindimensionale Modell. Somit sind die beiden zweidimensionalen Modelle dem eindimensionalen vorzuziehen. Darüber hinaus lassen sich die Modellgütekriterien so interpretieren, dass das Modell, welches Lesekompetenz anhand der beiden Leseabsichten beschreibt, dem Modell vorzuziehen ist, welches Lesekompetenz anhand der beiden übergeordneten Verstehensprozesse beschreibt. Daher wird bei der Modellierung von Lesekompetenz in der nächsten Forschungsfrage zwischen den Leseabsichten getrennt.

*Forschungsfrage 2:* Getrennt für die beiden Leseabsichten (a) und (b) werden nun jeweils vier Modelle evaluiert. Jede Leseabsicht (a) und (b) wird jeweils in die beiden übergeordneten Verstehensprozesse aufgliedert (Modell 3a und 3b), in die vier Verstehensprozesse (Modell 4a und 4b) und in die

vier Verstehensprozesse unter Annahme einer linearen Hierarchie in den Verstehensprozessen (Modell 5a und 5b). Zusätzlich wird zu den Modellen 4a und 4b eine parametrisch reduzierte Variante mit weniger Skillklassen geschätzt (Modell 4a-r und 4b-r).

Der obere Teil von Tabelle 5 zeigt für die Leseabsicht der literarischen Erfahrung (a), dass das Modell mit der Aufgliederung in vier Verstehensprozesse (Modell 4a) anhand der Gütekriterien im Vergleich zu dem Modell mit den zwei übergeordneten Verstehensaspekten (Modell 3a) als gleichwertig zu beurteilen ist. Das Modell mit den vier Verstehensprozessen (Modell 4a) ist dem Modell mit den als hierarchisch angenommenen vier Verstehensprozessen (Modell 5a) vorzuziehen. Das parametrisch reduzierte Modell unter Einbezug der vier Verstehensprozesse ohne Hierarchie (Modell 4a-r) liefert gemessen an den Gütekriterien die beste Anpassung an die Daten.

Der untere Teil von Tabelle 5 zeigt für die Leseabsicht des Informationsgewinns, dass das Modell mit der Aufgliederung in die vier Verstehensprozesse (Modell 4b) im Vergleich zu dem Modell mit den zwei übergeordneten Verstehensaspekten (Modell 3b) anhand der Gütekriterien vorzuziehen ist. Auch ist wiederum das Modell mit den vier Verstehensprozessen (Modell 4b) dem Modell mit den als hierarchisch angenommenen vier Verstehensprozessen (Modell 5b) vorzuziehen. Ebenso wie beim Lesen zur literarischen Erfahrung liefert auch beim Lesen zur Informationsgewinnung das parametrisch reduzierte Modell unter Einbezug der vier Verstehensprozesse ohne Hierarchie (Modell 4b-r) gemessen an den Gütekriterien die beste Anpassung an die Daten.

Zu einer detaillierteren Bewertung der Modelle bietet sich eine Analyse der Skillklassenverteilungen, d. h. der prozentualen Verteilungen der Schüler/innen in die Skillklassen, an. Die Modelle 4a, 4a-r und 5a sowie 4b, 4b-r und 5b beziehen jeweils die vier Verstehensprozesse zur Modellierung der Lesekompetenz in jeweils einer Leseabsicht ein. Die verschiedenen an die Modelle gestellten Annahmen spiegeln sich in den unterschiedlichen Skillklassen wider: Die Modelle 4a und 4b stellen keine Annahmen an die Verstehensprozesse und somit werden alle  $2^4 = 16$  möglichen Skillklassen geschätzt. Zur parametrischen Reduzierung werden von den 16 Skill-

Modell	Beschreibung	Nitems	Npars	loglike	AIC	BIC	CAIC
0	eindimensionales Konstrukt	181	363	-62.972,29	126.671	128.987	129.350
1	2 übergeordnete Verstehensprozesse	181	365	-62.875,88	126.482	128.811	129.176
2	2 Leseabsichten	181	365	-62.751,27	126.233	128.561	128.926

*Tabelle 4: Vergleich von Modellen zur Beschreibung von Lesekompetenz anhand der Modellgütekriterien AIC, BIC und CAIC. Zusätzlich sind zu jedem Modell die Anzahl der Items (Nitems), die Anzahl der Modellparameter (Npars) und die Loglikelihood (loglike) gegeben. Die Modellannahmen sind aus Abbildung 1 zu entnehmen.*

Modell	Beschreibung	Nitems	Npars	loglike	AIC	BIC	CAIC
<b>Lesen als literarische Erfahrung</b>							
3a	2 übergeordnete Verstehensprozesse	93	189	-31.891,02	64.349	65.366	65.555
4a	4 Verstehensprozesse	93	201	-31.843,53	64.290	65.371	65.572
4a-r	4 Verstehensprozesse; reduzierte Skillklassen	93	194	-31.843,31	64.269	65.312	65.507
5a	4 hierarchische Verstehensprozesse	93	190	-31.899,59	64.369	65.391	65.581
<b>Lesen zum Informationsgewinn</b>							
3b	2 übergeordnete Verstehensprozesse	88	179	-31.576,00	63.510	64.652	64.831
4b	4 Verstehensprozesse	88	191	-31.513,59	63.409	64.628	64.819
4b-r	4 Verstehensprozesse; reduzierte Skillklassen	88	185	-31.514,05	63.398	64.578	64.763
5b	4 hierarchische Verstehensprozesse	88	180	-31.594,31	63.549	64.697	64.877

*Tabelle 5: Modellvergleich von Modellen zur Beschreibung des Lesens als literarische Erfahrung (oben) und des Lesens zum Informationsgewinn (unten) anhand der Modellgütekriterien AIC, BIC und CAIC. Zusätzlich sind zu jedem Modell die Anzahl der Items (Nitems), die Anzahl der Modellparameter (Npars) und die Loglikelihood (loglike) gegeben.*

klassen all diejenigen Skillklassen, in die in den Modellen 4a und 4b weniger als 0,5 % der Schüler/innen klassifiziert wurden, als nicht relevant angesehen und somit nicht geschätzt. Es entstehen die Modelle 4a-r und 4b-r. In den Modellen

Skill-klasse	4a	4a-r	5a	4b	4b-r	5b
0000	29,2	29,2	32,6	31,6	31,6	41,2
1000	0,8	0,8	2,1	0,0	–	0,0
0100	1,1	1,3	–	0,6	0,7	–
1100	0,7	0,7	1,0	0,3	–	0,9
0010	1,2	1,3	–	2,4	2,5	–
1010	1,2	1,2	–	0,0	–	–
0110	2,6	2,5	–	4,6	4,6	–
1110	0,4	–	1,2	7,0	7,1	14,0
0001	0,0	–	–	3,4	3,3	–
1001	0,0	–	–	0,0	–	–
0101	0,3	–	–	1,6	1,5	–
1101	0,0	–	–	0,0	–	–
0011	0,1	–	–	1,3	1,3	–
1011	0,1	–	–	0,0	–	–
0111	5,3	5,6	–	7,9	8,0	–
1111	57,0	57,3	63,1	9,5	39,5	43,9

*Tabelle 6: Prozentsatz an Schülerinnen und Schülern in den Skillklassen aufgliedert nach Modellen.*

5a und 5b werden, ausgehend von allen 16 Skillklassen, nur die Skillklassen geschätzt, die die Annahme der linearen Hierarchie zwischen den Skills widerspiegeln. In Tabelle 6 ist für jedes Modell eingetragen, wie viel Prozent der Schüler/innen jeweils in die einzelnen Skillklassen klassifiziert wurden. Skillklassen, die nach den oben erläuterten Annahmen nicht geschätzt wurden sind, mit einem „–“ gekennzeichnet.

Die erste Spalte von Tabelle 6 enthält die Skillklassenverteilung für Modell 4a beim Lesen zur literarischen Erfahrung, in dem vier nicht hierarchische Verstehensprozesse beschrieben werden. Tabelle 6 zeigt zunächst, dass 13,8 % der Schüler/innen nicht in die Skillklassen [0000] oder [1111] fallen. Dies zeigt, dass es Schüler/innen gibt, die nur eine bestimmte Teilmenge aller Skills beherrschen, die zum Lesen als literarische Erfahrung notwendig sind. Vergleichen wir nun die Skillklassenverteilung aus Modell 4a mit der Skillklassenverteilung in Modell 5a, in dem eine lineare Hierarchie zwischen den Verstehensprozessen postuliert wird. Es fällt auf, dass in Modell 4a 11,9 % der Schüler/innen in die Skillklassen [0100], [0010], [1010], [0110], [0001], [1001], [0101], [1101], [0011], [1011] und [0111] eingeordnet werden, in denen die Beherrschung der Verstehensprozesse von der in Modell 5a angenommenen linearen Hierarchie abweicht. Beispielsweise beherrschen nach Modell 4a knapp 5,3 % der Schüler/innen [0111] also alle Verstehensprozesse bis auf das Erkennen und Wiedergeben explizit angegebener Informationen. Der Prozentsatz an Schülerinnen und Schülern, die nicht in die hierarchischen Klassen fallen, erklärt, warum Modell 5a mit Hierarchie-Annahme zwischen den Verstehensprozessen die Daten weniger gut be-

Modell	Beschreibung	Nitems	Npars	loglike	AIC	BIC	CAIC
0	eindimensionales Konstrukt	181	363	-62.972,29	126.671	128.987	129.350
1	2 übergeordnete Verstehensprozesse	181	365	-62.875,88	126.482	128.811	129.176
2	2 Leseintentionen	181	365	-62.751,27	126.232	128.561	128.926
6	2 Leseintentionen mit jeweils 4 Verstehensprozessen, reduziert	181	361	-62.486,93	125.696	127.999	128.360

*Tabelle 7: Vergleich von Modellen zur Beschreibung von Lesekompetenz anhand der Modellgütekriterien AIC, BIC und CAIC. Zusätzlich sind zu jedem Modell die Anzahl der Items (Nitems), die Anzahl der Modellparameter (Npars) und die Loglikelihood (loglike) gegeben. Die Modellannahmen sind aus Abbildung 1 zu entnehmen.*

schreibt. Werden ausgehend von allen 16 Skillklassen nun nur die Skillklassen geschätzt, die in Modell 4a mit mehr als 0,5 % der Schüler/innen besetzt sind, entsteht die Verteilung in Modell 4a-r. Die im reduzierten Modell geschätzten Skillklassen weichen von den Skillklassen ab, die in Modell 5a unter Annahme der Hierarchie geschätzt wurden. Auch dies kann als Erklärung gewertet werden, warum das Modell mit der Hierarchie die Daten weniger gut beschreibt. Daher verwerfen wir die Annahme der Hierarchie in den Verstehensprozessen für das Lesen zur literarischen Erfahrung und behalten die in Modell 4a-reduziert geschätzten Skillklassen und deren Annahmen bei.

Bei einer analogen Analyse der Skillklassenverteilungen für das Lesen zum Informationsgewinn zeigt Tabelle 6 zunächst, dass in Modell 4b unter Annahme nicht hierarchischer Skillklassen 29,1 % der Schüler/innen nicht in die Skillklassen [0000] oder [1111] fallen. Somit beherrschen beim Lesen zum Informationsgewinn mehr Schüler/innen nur eine bestimmte Teilmenge aller notwendigen Skills als beim Lesen zur literarischen Erfahrung (hier sind es 13,8 %). Vergleichen wir nun wieder die Skillklassen in Modell 4b mit den unter Annahme der linearen Hierarchie in den Verstehensprozessen geschätzten Skillklassen in Modell 5b. In Modell 4b werden 21,8 % der Schüler/innen in Skillklassen eingeteilt, die von der linearen Hierarchie abweichen. Der prozentuale Anteil der Schüler/innen, die in nicht hierarchische Klassen fallen ist beim Lesen zum Informationsgewinn deutlich größer als beim Lesen zur literarischen Erfahrung (hier sind es 11,9 %). Dies kann als Anzeichen gewertet werden, dass die Verstehensprozesse beim Lesen zum Informationsgewinn mehr von der in Modell 5b postulierten linearen Hierarchie abweichen als beim Lesen zur literarischen Erfahrung. Auch weichen die nach Reduzierung geschätzten Skillklassen in Modell 4b-r deutlich von denen ab, die in Modell 5b unter

Annahme der Hierarchie geschätzt wurden. Daher verwerfen wir auch für das Lesen zum Informationsgewinn die Annahme der Hierarchie in den Verstehensprozessen und behalten die in Modell 4b-reduziert geschätzten Skillklassen und ihre Annahmen bei.

*Forschungsfrage 3:* In einem letzten Schritt werden die beiden reduzierten Modelle des Lesens zur literarischen Erfahrung (Modell 4a-r) und des Lesens zum Informationsgewinn (Modell 4b-r) wieder zu einem Modell zur Beschreibung von Lesekompetenz (zu beiden Leseabsichten) zusammengeführt. Im entstehenden achtdimensionalen Modell (vier nicht hierarchische Verstehensprozesse für je zwei Leseabsichten; Modell 6) werden alle Skillklassen geschätzt, die sich aus der Kombination der beiden reduzierten Modelle ergeben<sup>1</sup>. Tabelle 7 zeigt, dass, bewertet anhand der Gütekriterien, das so reduzierte achtdimensionale Modell das Konstrukt Lesen besser beschreibt als das eindimensionale und die beiden zweidimensionalen Modelle (siehe Kapitel 3 in diesem Band, S. 47 f.).

## Diskussion und Implikationen

In diesem Beitrag wurden kognitiv diagnostische Modelle, die verschiedene fachwissenschaftliche Annahmen an das Konstrukt Lesen widerspiegeln, an die PIRLS-2016-Daten für österreichische Schüler/innen angepasst. Statistische Gütekriterien bevorzugen ein achtdimensionales Modell, in dem die beiden Leseabsichten gemeinsam sowie vier nicht hierarchische Verstehensprozesse in jeder Leseabsicht in reduzierter Form modelliert werden. Das gewählte achtdimensionale Modell spiegelt die fachwissenschaftlichen Annahmen aus der Itemerstellung bei der PIRLS-Studie wider.

<sup>1</sup> Bei den entstehenden achtdimensionalen Skillklassen beschreiben die ersten vier Einträge die vier Verstehensprozesse beim Lesen zur literarischen Erfahrung, während die letzten vier Einträge die 4 Verstehensprozesse zum Lesen zum Informationsgewinn kennzeichnen. In den ersten vier Einträgen werden alle Skillklassen übernommen, die in Modell 4a reduziert geschätzt wurden (bspw. [0000XXXX]). Diese werden dann in den letzten vier Einträgen mit allen Skillklassen kombiniert, die in Modell 4b reduziert geschätzt wurden (bspw. [00000000], [00000100], [00001100]).

Die Modellwahl zeigt für österreichische Schüler/innen: Die Kompetenzen der Schüler/innen sind abhängig davon, ob die Schüler/innen zur literarischen Erfahrung oder zum Informationsgewinn lesen. Die Kompetenzen der Schüler/innen in beiden Leseabsichten beeinflussen sich jedoch. Bei beiden Leseabsichten kann die Annahme der linearen Hierarchie zwischen den Verstehensprozessen verworfen werden. Bei einer genaueren Analyse durch die Skillklassen der verschiedenen Modelle zeigen sich allerdings hohe Anteile an Schülerinnen und Schülern, die in die hierarchischen Skillklassen klassifiziert werden. Der Anteil der Schüler/innen, die nicht in die hierarchische Klassen eingeordnet werden können, ist aber hinreichend, um – bewertet an relativen Fitindices – der Annahme der Hierarchie zu widersprechen. Die Abweichung von der Annahme der Hierarchie in den Verstehensprozessen ist, gemessen an der prozentualen Verteilung der Schüler/innen in die Skillklassen, beim Lesen zum Informationsgewinn stärker als beim Lesen als literarische Erfahrung.

Bedingt durch die Ergebnisse der Modellwahl lassen sich unter anderem die folgenden Aspekte diskutieren: Bei der PIRLS-Berichterstattung wäre es zumindest in Österreich möglich, mehr Skalen als bisher üblich zurückzumelden. Würden die Ergebnisse der vorliegenden Analysen auf formative Testsettings übertragen werden, die auf einem zu PIRLS analogen fachwissenschaftlichen Modell von Lesekompetenz aufbauen, könnte der Informationszugewinn breiter genutzt werden. In Tests zu formativen Zwecken würde die Untersuchung möglichst vieler Subkompetenzen auch von Vorteil sein, um individuelle Lernstände zu berichten und daraus fachdidaktische Konsequenzen abzuleiten. In solchen Kontexten wird hierarchischen Modellen angesichts der Fülle an Leistungen von Schülerinnen und Schülern eine nicht immer angemessene Informationsverdichtung nachgesagt (Bremrich-Vos & Böhme, 2009, S. 236).

Es liegt auf der Hand, dass mit den vorliegenden Ergebnissen solche fachwissenschaftlichen Theorien zur Lesekompetenz gestützt werden, in denen die Prozesse als nicht hierarchisch angesehen und zugleich Leseabsichten unterstellt werden (vgl. Christmann, 2010). Weiters können die Ergebnisse wichtige Impulse für die aktuelle Diskussion um den Zusammenhang von Subkompetenzen, Leseabsichten und Förderung geben: Schließlich werden hier seit geraumer Zeit „Varianten des Literacy-Konzepts und ihre Bedeutung für die Deutschdidaktik“ diskutiert (vgl. z. B. Wieler, 2013). Dabei ist zu bedenken, dass die meisten fachdidaktischen

Modelle theoretische Modelle sind, die sich aber auf die Ergebnisse dieser Untersuchung beziehen lassen (z. B. die kognitive Ebene in den Modellen von Hurrelmann, 2002; Rosebrock & Nix, 2008).

In der Folge könnten Bezüge zu solchen fachdidaktischen Modellen der Lesekompetenz hergestellt werden, die Subkompetenzen mit konkreten Fördermaßen in Verbindung bringen (u. a. Rosebrock, Nix, Rieckmann & Gold, 2011; Hoppe, 2012; Gold, Behrendt, Lauer-Schmaltz & Rosebrock, 2013; Schulte & Souvignier, 2013; Munser-Kiefer, 2014). Da sich das Konzept der „literarischen Kompetenz“ in Ansätzen mit der Leseabsicht „literarische Erfahrungen“ bei PIRLS vergleichen lässt, werden solche Modelle gestützt, die zwischen „literarischer Kompetenz“ und „Lesekompetenz“ unterscheiden. Darauf basierend sollte der Literaturunterricht mit speziellen Angeboten ausgestattet werden (vgl. Frederking, Meier, Stanat & Dickhäuser, 2008).

Fachdidaktisch wäre vor diesem Hintergrund auch zu überlegen, ob die Teilpopulation der Schüler/innen, die in nicht hierarchischen Skillklassen klassifiziert wurde, anders Lesen lernt als die Schüler/innen in den hierarchisch geordneten Klassen. In diesem Fall könnten ebenfalls Schlüsse über unterschiedliche didaktische Fördermaßnahmen gezogen werden, die das Ziel haben, den Gesamtschnitt der Lesekompetenz in Österreich weiter anzuheben. Eine solche Analyse könnte auch durch Fragen im PIRLS-Kontextfragebogen unterstützt werden: So könnte zum einen überprüft werden, ob zwischen den Teilpopulationen der Schüler/innen in und außerhalb hierarchischer Klassen Unterschiede in der Erstsprache bestehen. Zum anderen könnte aber auch die Bewertung des Leseunterrichts durch die Schüler/innen nach Teilpopulationen differenziert werden (bspw. anhand von Fragen wie: „Meine Lehrerin möchte, dass wir im Unterricht über das reden, was wir gelesen haben.“). Und schließlich ließen sich aus dem hier vorgelegten Modell auch Schlüsse auf die Gestaltung von differenzierten Aufgaben bzw. Fragen zu Lesetexten im Unterricht ziehen.

In einer weiteren vertiefenden Analyse könnten außerdem die Items analysiert werden, die von den beiden Teilpopulationen zumeist gelöst werden. Bei einer solchen Analyse könnte erforscht werden, ob bestimmte schwierigkeitsgenerierende Merkmale einer Teilpopulation schwerer fallen als der anderen (siehe auch Illetschko und George in diesem Band), und daraus könnten fachdidaktische Schlüsse gezogen werden.

---

## Literatur

- Bos, W., Valtin, R., Voss, A., Hornberg, S. & Lankes, E.-M. (2007). Konzepte der Lesekompetenz in IGLU 2006. In W. Bos, S. Hornberg, K.-H. Arnold, G. Faus, L. Fried, E.-M. Lankes, K. Schwippert & R. Valtin (Hrsg.), *IGLU 2006. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 81–104). Münster: Waxmann.
- Bremerich-Vos, A. & Böhme, K. (2009). Lesekompetenzdiagnostik – die Entwicklung eines Kompetenzstufenmodells für den Bereich Lesen. In D. Granzer, O. Köller & A. Bremerich-Vos (Hrsg.), *Bildungsstandards Deutsch und Mathematik: Leistungsmessung in der Grundschule* (S. 219–249). Weinheim: Beltz.
- Buck, G., Tatsuoka, K. & Kostin, I. (1997). The subskills of reading: Rule-space analysis of a multiple choice test of second language reading comprehension. *Language Learning*, 47 (3), 423–466. doi:10.1111/0023-8333.00016
- Buck, G., VanEssen, T., Tatsuoka, K., Kostin, I., Lutz, D. & Phelps, M. (2014). *Development, selection, and validation of a set of cognitive and linguistic attributes for the SAT I verbal: Critical reading section*. Princeton, NJ: Educational Testing Services.
- Campbell, J. R., Kelly, D. L., Mullis, I. V. S., Martin, M. O. & Sainsbury, M. (2001). *Framework and specifications for PIRLS assessment 2001*. Chestnut Hill, MA.: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.
- Chen, H. & Chen, J. (2016). Retrofitting non-cognitive-diagnostic reading assessment under the generalized DINA framework. *Language Assessment Quarterly*, 13 (3), 218–230. doi:10.1080/15434303.2016.1210610
- Chen, J. & de la Torre, J. (2013). A general cognitive diagnosis model for expert-defined polytomous attributes. *Applied Psychological Measurement*, 37, 419–437. doi:10.1177/0146621613479818
- Chen, J. & de la Torre, J. (2014). A Procedure for diagnostically modeling extant large-scale assessment data: The case of the programme for international student assessment in reading. *Psychology*, 5 (18), 1967–1978. doi:10.4236/psych.2014.518200
- Christmann, U. (2010). Lesepsychologie. In M. Kämper-van den Boogaart & K. H. Spinner (Hrsg.), *Lese- und Literaturunterricht I* (S. 148–200). Baltmannsweiler: Schneider-Verlag Hohengehren.
- DiBello, L., Roussos, L. & Stout, W. (2007). Review of cognitively diagnostic assessment and a summary of psychometric models. In C. R. Rao & S. Sinharay (Hrsg.), *Handbook of Statistics, Volume 26, Psychometrics* (pp. 979–1030). Amsterdam: Elsevier. doi:10.1016/S0169-7161(06)26031-0
- Elley, W. B. (1992). *How in the world do students read?* The Hague, Netherlands: International Association for the Evaluation of Educational Achievement (IEA).
- Frederking, V., Meier, C., Stanat, P. & Dickhäuser, O. (2008). Ein Modell literarästhetischer Urteilskompetenz. *Didaktik Deutsch*, 25, 11–31.
- Foorman, B. R., Koon, S., Petscher, Y., Mitchell, A. & Truckenmiller, A. (2015). Examining general and specific factors in the dimensionality of oral language and reading in 4th–10th grades. *Journal of Educational Psychology*, 107, 884–899.
- George, A. C., Robitzsch, A., Kiefer, T., Groß, J. & Ünlü, A. (2016). The R package CDM for cognitive diagnosis modeling. *Journal of Statistical Software*, 74 (2), 1–24. doi:10.18637/jss.v074.i02
- Gold, A., Behrendt, S., Lauer-Schmaltz, M. & Rosebrock, C. (2013). Förderung der Leseflüssigkeit in dritten Grundschulklassen. In C. Rosebrock & A. Bertschi-Kaufmann (Hrsg.), *Literalität erfassen: bildungspolitisch, kulturell, individuell* (S. 203–218). Weinheim: Beltz Juventa.
- Haertel, E. H. (1989). Using restricted latent class models to map the skill structure of achievement items. *Journal of Educational Measurement*, 26, 301–323. doi:10.1111/j.1745-3984.1989.tb00336.x

- Hoppe, I. (2012). In *Lesewelten hineinwachsen – Leseförderung in der flexiblen Schulanfangsphase*. Ludwigsfelde-Struveshof: Landesinstitut für Schule und Medien Berlin-Brandenburg (LISUM). Verfügbar unter: [https://bildungsserver.berlin-brandenburg.de/fileadmin/bbb/schule/grundschulportal/publikationen\\_grundschule/Lesewelten.pdf](https://bildungsserver.berlin-brandenburg.de/fileadmin/bbb/schule/grundschulportal/publikationen_grundschule/Lesewelten.pdf)
- Hurrelmann, B. (2002). Leseleistung – Lesekompetenz. *Praxis Deutsch*, 176, 6–18.
- Hußmann, A., Wendt, H., Bos, W., Bremerich-Vos, A., Kasper, D., Lankes, E.-M. et. al. (Hrsg.). (2017). *IGLU 2016. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster: Waxmann.
- Irwin, J. W. (1986). *Teaching reading comprehension process*. Englewood Cliffs: Prentice-Hall Inc.
- Jang, E. E. (2009). Cognitive diagnostic assessment of L2 reading comprehension ability: Validity arguments for applying Fusion Model to LanguEdge assessment. *Language Testing*, 26 (1), 31–73. doi:10.1177/0265532208097336
- Jang, E. E., Dunlop, M., Wagner, M., Kim, Y. H. & Gu, Z. (2013). Elementary school ELLs' reading skill profiles using cognitive diagnosis modeling: Roles of length of residence and home language environment. *Language Learning*, 63 (3), 400–436. doi:10.1111/lang.12016
- Jang, E. E. & Roussos, L. (2007). An Investigation into dimensionality of TOEFL using conditional covariance-based non-parametric approach. *Journal of Educational Measurement*, 44 (1), 1–21. doi:10.1111/j.1745-3984.2007.00024.x
- Kendeou, P., Papadopoulou, T. C. & Spanoudis, G. (2012). Processing demands of reading comprehension tests in young readers. *Learning and Instruction*, 22, 354–367. doi:10.1016/j.learninstruc.2012.02.001
- Kim, A. Y. (2014). Exploring ways to provide diagnostic feedback with an ESL placement test: Cognitive diagnostic assessment of L2 reading ability. *Language Testing*, 32 (2), 227–258. doi:10.1177/0265532214558457
- Lee, Y.-W. & Sawaki, Y. (2009). Application of three cognitive diagnosis models to ESL reading and listening assessments. *Language Assessment Quarterly*, 6 (3), 239–263. doi:10.1080/15434300903079562
- Leighton, J. P., Gierl, M. J. & Hunka, S. M. (2004). The attribute hierarchy method for cognitive assessment: A variation on Tatsuoka's rule space approach. *Journal of Educational Measurement*, 41, 205–237. doi:10.1111/j.1745-3984.2004.tb01163.x
- Li, H., Hunter, C. V. & Lei, P.-W. (2016). The selection of cognitive diagnostic models for a reading comprehension test. *Language Testing*, 33 (3), 391–409. doi:10.1177/0265532215590848
- Martin, M. O., Mullis, I. V. S. & Hooper, M. (Hrsg.). (2017). *Methods and procedures in PIRLS 2016*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.
- Muijselaar, M. M. L., Swart, N. M., Steenbeek-Planting, E. G., Droop, M., Verhoeven, L. & de Jong, P. F. (2017). The dimensions of reading comprehension in Dutch children: Is differentiation by text and question type necessary? *Journal of Educational Psychology*, 109 (1), 70–83. doi:10.1037/edu0000120
- Mullis, I. V. S. & Martin, M. O. (Hrsg.). (2015). *PIRLS 2016. Assessment framework. 2nd ed.* Chestnut Hill, U.S.: TIMSS & PIRLS. International Study Center, Lynch School of Education, Boston College.
- Mullis, I. V. S., Martin, M. O., Foy, P. & Hooper, M. (Hrsg.). (2017). *PIRLS 2016. International results in reading*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.
- Munser-Kiefer, M. (2014). *Leseförderung im Leseteam in der Grundschule. Eine Interventionsstudie zur Förderung von basaler Lesefertigkeit und (meta-)kognitiven Lesestrategien* (Empirische Erziehungswissenschaften, Bd. 40). Münster: Waxmann.

- Ponocny, I., Bauernfeind, U., Zins, A. & Radinger, R. (2009). Die Dimensionalität der PIRLS-Skalen. In B. Suchan, C. Wallner-Paschon & C. Schreiner (Hrsg.), *PIRLS 2006. Die Lesekompetenz am Ende der Volksschule. Österreichischer Expertenbericht* (S. 243–251). Graz: Leykam.
- R Core Team. (2018). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.
- Ravand, H. (2016). Application of a cognitive diagnostic model of a high-stakes reading comprehension test. *Journal of Psychoeducational Assessment*, 34 (8), 782–799. doi:10.1177/0734282915623053
- Rickheit, G. & Strohner, H. (1993). *Grundlagen kognitiver Sprachverarbeitung. Modelle, Methoden, Ergebnisse*. Tübingen/Basel: Francke.
- Robitzsch, A., Kiefer, T., George, A. C. & Ünlü, A. (2018). *CDM: Cognitive diagnosis modeling. R package version 6.5-12*. Verfügbar unter: <https://CRAN.R-project.org/package=CDM>.
- Rosebrock, C. & Nix, D. (2008). *Grundlagen der Lesedidaktik und der systematischen schulischen Leseförderung*. Baltmannsweiler: Schneider Verlag Hohengehren.
- Rosebrock, C., Nix, D., Rieckmann, C. & Gold, A. (2011). *Leseflüssigkeit fördern: Lautleseverfahren für die Primar- und Sekundarstufe [CD]*. Seelze: Klett Kallmeyer.
- Rupp, A. A., Templin, J. L. & Henson, R. A. (2010). *Diagnostic measurement. Theory, methods, and applications*. New York: Guilford Press.
- Schulte, E. & Souvignier, E. (2013). Der Lese-Sportler. Ein Programm für individuelle Leseförderung. *Die Grundschulzeitschrift*, 27 (7), 58–61.
- Seo, M., Chiu, C. & Roussos, L. (2008, September). *Evaluating the dimensionality of the 2001 PIRLS reading assessment: An application of DIMTEST with DESM and CFA*. Paper presented at the 3rd IEA International Research Conference, Taipei, Chinese Taipei.
- Tatsuoka, K. K. (1983). Rule space: An approach for dealing with misconceptions based on item response theory. *Journal of Educational Measurement*, 20, 345–354. doi:10.1111/j.1745-3984.1983.tb00212.x
- TIMSS and PIRLS International Study Center. (2018). *PIRLS 2016. User guide for the international database*. Retrieved from: <https://timssandpirls.bc.edu/pirls2016/international-database/index.html>
- van Dijk, T. & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.
- von Davier, M. (2010). A general diagnostic model applied to language testing data. *British Journal of Mathematical and Statistical Psychology*, 61 (2), 287–307. doi:10.1348/000711007X193957
- Voss, A., Carstensen, C. H. & Bos, W. (2005). Textgattungen und Verstehensaspekte: Analyse von Leseverständnis aus den Daten der IGLU-Studie. In W. Bos, E. M. Lankes, M. Prenzel, K. Schwippert, R. Valtin & G. Walther (Hrsg.), *IGLU: Vertiefende Analysen zu Leseverständnis, Rahmenbedingungen und Zusatzstudien* (S. 1–33). Münster: Waxmann.
- Wallner-Paschon, C. & Itzlinger-Bruneforth, U. (Hrsg.). (2017). *PIRLS 2016. Technischer Bericht*. Salzburg: Bundesinstitut für Bildungsforschung, Entwicklung und Evaluation des österreichischen Schulwesens (BIFIE).
- Wallner-Paschon, C., Itzlinger-Bruneforth, U. & Schreiner, C. (Hrsg.). (2017). *PIRLS 2016. Die Lesekompetenz am Ende der Volksschule. Erste Ergebnisse*. Graz: Leykam.
- Wieler, P. (2013). Varianten des Literacy-Konzepts und ihre Bedeutung für die Deutschdidaktik. In U. Abraham, A. Bremerich-Vos, V. Frederking & P. Wieler (Hrsg.), *Deutschdidaktik und Deutschunterricht nach PISA*. (S. 47–68). Freiburg im Breisgau: Fillibach.