

Maik Philipp | Simone Jambor-Fahlen (Hrsg.)

**Lesen: Prozess- und
Produktperspektiven
von der Wortebene
bis zu multiplen Texten**

BELTZ JUVENTA

Maik Philipp | Simone Jambor-Fahlen (Hrsg.)
Lesen: Prozess- und Produktperspektiven
von der Wortebene bis zu multiplen Texten

Maik Philipp | Simone Jambor-Fahlen (Hrsg.)

**Lesen: Prozess- und
Produktperspektiven
von der Wortebene
bis zu multiplen Texten**

BELTZ  **JUVENTA**

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Der Text dieser Publikation wird unter der Lizenz **Creative Commons Namensnennung – Nicht kommerziell – Keine Bearbeitungen 4.0 International (CC BY-NC-ND 4.0)** veröffentlicht.

Den vollständigen Lizenztext finden Sie unter: <https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode.de>. Verwertung, die den Rahmen der **CC BY-NC-ND 4.0 Lizenz** überschreitet, ist ohne Zustimmung des Verlags unzulässig. Das gilt insbesondere für die Bearbeitung und Übersetzungen des Werkes. Die in diesem Werk enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Quellenangabe/Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.



Dieses Buch ist erhältlich als:
ISBN 978-3-7799-6514-5 Print
ISBN 978-3-7799-5837-6 E-Book (PDF)

1. Auflage 2022

© 2022 Beltz Juventa
in der Verlagsgruppe Beltz · Weinheim Basel
Werderstraße 10, 69469 Weinheim
Einige Rechte vorbehalten

Herstellung: Ulrike Poppel
Satz: Helmut Rohde, Euskirchen
Druck und Bindung: Beltz Grafische Betriebe, Bad Langensalza
Printed in Germany

Weitere Informationen zu unseren Autor_innen und Titeln finden Sie unter: www.beltz.de

Inhalt

Vorwort <i>Maik Philipp & Simone Jambor-Fahlen</i>	7
Prozess- und Produktperspektiven des Lesens von der Wortebene bis zu multiplen Texten <i>Simone Jambor-Fahlen & Maik Philipp</i>	9
Methoden zur differenzierten Auswertung von Wortschatzkompetenzen und ihre Bedeutung für das frühe Lesen <i>Britta Juska-Bacher, Christoph Zangger & Martina Röthlisberger</i>	34
Die Wirksamkeit von Fördermaßnahmen zur Leseflüssigkeit in Abhängigkeit von Passung, Dosierung und Wiedergabetreue Herausforderungen für Forschung und Praxis <i>Mareike Ehlert & Elmar Souvignier</i>	56
Ein Lesestrategietraining im Distanzunterricht? Wie sich das veränderte Lernumfeld während der Corona-Pandemie auf die Entwicklung des Leseverstehens von Schülerinnen und Schülern auswirkt <i>Johannes Wild, Elisabeth Kraus, Maria Steinert, Sven Hilbert & Anita Schilcher</i>	77
Synthesen schreiben – eine textbasierte Längsschnittuntersuchung studentischer Lese- und Schreibperformanz <i>Alex Rickert & Maik Philipp</i>	100
Ist argumentatives materialgestütztes Schreiben dem informierenden überlegen? Eine Sekundäranalyse experimenteller Studien zu den Effekten der Zieltextsorte auf Verstehensleistungen <i>Maik Philipp</i>	128
Die Autorinnen und Autoren	155

Vorwort

Wenn von lesebezogenen Prozess- und Produktperspektiven ausgehend von der Wortebene bis zu multiplen Texten im Titel dieses Bandes die Rede ist, dann ist damit ein weit gefasstes Spektrum angesprochen. Dieses Spektrum umfasst nicht nur die klassischen Unterscheidungen dessen, worin sich kompetentes Lesen manifestiert, nämlich im Ergebnis oder auf dem Weg dorthin, also im Produkt oder im Prozess. Vielmehr ist mit diesem Kontinuum von einzelnen Wörtern bis hin zu mehreren Texten auch die Textmenge angesprochen, die wir im Alltag lesen. Zudem deutet sich darin ein Erwerbshorizont an, der vom Schriftspracherwerb bis zu den Anforderungen in der Hochschule reicht.

In diesem Band folgen wir diesem Kontinuum, indem wir sechs Beiträge zusammengestellt haben, die mehrheitlich aus dem Umfeld des 2020er Symposions Deutschdidaktik stammen. Diese Beiträge befassen sich mit dem Konstrukt Lesekompetenz, dessen Hierarchieebenen, seiner Förderung und seiner Entwicklung. Sie tun dies bei verschiedenen Altersgruppen aus empirischer und theoretischer Perspektive und sie werfen verschiedene Schlaglichter auf Prozesse und Produkte. So geben die Beiträge dieses Bandes Antwort auf Fragen wie:

- Welche grundsätzlich beschreibbaren Prozesse umfasst Lesekompetenz aus kognitiver Perspektive? (Kap. 1, Jambor-Fahlen & Philipp)
- In welchem Verhältnis steht die Entwicklung von Leseverstehen und Wortschatz? (Kap. 2, Juska-Bacher, Zangger & Röthlisberger)
- Wie müssen Leseflüssigkeitstrainings beschaffen sein, damit sie möglichst wirksam werden? (Kap. 3, Ehlert & Souvignier)
- Was lehrt eine pandemisch begründete Adaption über die essenziellen, kompetenzsteigernden Merkmale von Lesestrategietrainings? (Kap. 4, Wild, Kraus, Steinert, Hilbert & Schilcher)
- Welche Schwierigkeiten haben Studierende beim Verfassen von Diskurs-synthesen in puncto erbrachter Leseleistungen? (Kap. 5, Rickert & Philipp)
- Welche Effekte haben verschiedene Schreibaufträge auf Verstehensleistungen beim Schreiben über multiple Texte? (Kap. 6, Philipp)

Dass dieser Band in der vorliegenden Fassung entstehen konnte, verdanken wir nicht nur den Beiträgerinnen und Beiträgern, die sich an einen engen Zeitplan

gehalten haben. Auch dem Beltz-Juventa-Verlag danken wir, ebenso der Pädagogischen Hochschule Zürich für die Finanzierung der Open-Access-Publikation und Anna Gold für ihre Durchsicht der Beiträge.

Zürich und Köln, im Juni 2021

Maik Philipp und Simone Jambor-Fahlen

Prozess- und Produktperspektiven des Lesens von der Wortebene bis zu multiplen Texten

Simone Jambor-Fahlen & Maik Philipp

Zusammenfassung: Lesekompetenz ist ein mehrdimensionales Konstrukt, das im Wesentlichen hierarchieniedrige sowie hierarchiehohe Lesefähigkeiten umfasst. Im Beitrag wird zunächst eine Synopse aufeinander aufbauender Lesefähigkeiten bei der Lesekompetenz in empirischen (Bildungs-)Studien skizziert. Es wird deutlich, dass sie ein Kontinuum abbilden, das von elementaren Lesefähigkeiten über das sinnentnehmende Lesen bis hin zum überwiegend reflektierenden und bewertenden Lesen reicht. Zunehmende Kompetenz ist einerseits durch eine Steigerung der Textmenge gekennzeichnet und andererseits durch ein zunehmend evaluatives und interpretatives Lesen. Der Beitrag legt dabei seine Schwerpunkte auf die Beschreibung der Kontinuumsenden, indem er die Leseflüssigkeit auf der einen Seite und das Lesen und Verarbeiten multipler Texte auf der anderen Seite fokussiert. Vor diesem Hintergrund argumentiert der Beitrag abschließend offene Forschungsfragen, die trotz der tradierten Forschung in diesem Bereich noch immer Desiderate darstellen.

Schlüsselwörter: Leseverstehen, Leseflüssigkeit, Lesestrategien, Leseprozesse, Lesekompetenz

Abstract: Reading literacy is a multidimensional construct that essentially comprises lower-order as well as higher-order reading skills. In this paper, we first outline a synopsis of successive reading skills in reading literacy in empirical (educational) studies. It becomes obvious that they represent a continuum ranging from basic reading skills to meaning making to predominantly reflective and evaluative reading. Expanding competence is characterized by an increase in the amount of text and by increasingly evaluative and interpretative reading behavior. The chapter focuses on the description of the continuum ends by focusing on reading fluency on the one hand and reading and processing multiple texts on the other hand. Against this background, the paper concludes by arguing open research questions that still represent desiderata despite the research in this area.

Keywords: reading comprehension, reading fluency, reading strategies, reading processes, reading literacy

1 Einleitung

Lesen ist eine unverzichtbare Voraussetzung für die Teilhabe an Bildung, denn das Lesen ermöglicht durch die Nutzung von Texten fachübergreifend einen selbstständigen Wissenserwerb. Der Wissenserwerb ist gefährdet, wenn Leserinnen und Leser nicht sinnentnehmend lesen können: Texte, die nicht verstanden werden, bleiben bedeutungslos. Dass dieser Umstand nach wie vor relevant für das Bildungssystem ist, zeigen die in regelmäßigen Abständen durchgeführten nationalen wie internationalen Bildungsstudien. So verlässt beispielsweise in Deutschland jede sechste Schülerin und jeder sechste Schüler die Grundschule, ohne über ausreichende Lesefähigkeiten zu verfügen (Hußmann et al., 2017).

Die Forschung hat sich dementsprechend vermehrt der Frage nach einer wirkungsvollen Leseförderung zugewandt und dabei zunehmend die hierarchieniedrigen (Automatisierung des Wort- und Satzlesens) und die hierarchiehohen (strategiebezogenen) Lesefähigkeiten in den Blick genommen. Lesefördermaßnahmen wie Lautlese-Tandems oder Strategieprogramme gelten als gut erforscht und wirkungsvoll (Philipp, 2013) und die Deutschdidaktik läuft Gefahr, sich einem Grundsättigungsgefühl hinzugeben. Ist das Forschungsfeld «Lesekompetenz» gesättigt? Wissen wir nun grundsätzlich, was zu tun ist, und die Schulpraxis muss es nur noch umsetzen?

Im vorliegenden Beitrag werden wir das Forschungsfeld beleuchten und tradierte wie neuere Aspekte des mehrdimensionalen Konstrukts Lesekompetenz darstellen. Zunächst werden wir anhand einer Synopse von Modellierungen zur Lesekompetenz in Bildungsstudien zeigen, dass diese ein Spektrum abbilden, das von basalen Lesefähigkeiten über das sinnentnehmende Lesen bis hin zum Lesen multipler Texte mit komplexen Inferenzbildungen und Interpretationen reicht. Dabei wird jedoch deutlich, dass die basalen Lesefähigkeiten nur partiell – nämlich bei PISA – Eingang in das Konstrukt Lesekompetenz finden. Fokussiert wird in den Bildungsstudien vor allem das Verstehen von Satz- und Textbedeutungen. Um zu verdeutlichen, dass die Lesekompetenz keinen produktbezogenen Ist-Zustand abbildet, sondern ihr vielmehr eine Erwerbsfolge im Sinne zunehmender Kompetenz inhärent ist, legt der Beitrag einen Schwerpunkt auf den Beginn des Lesenlernens, hier vor allem auf die basalen Prozesse des Dekodierens und Automatisierens, um dann das Kontinuum weiter aufzuspannen und den Nutzen und den Einsatz von Lesestrategien bis hin zum Verstehen und Verarbeiten multipler Texte darzustellen. Dabei wird abschließend deutlich, dass die Forschung zur Lesekompetenz zwar

einerseits insbesondere im englischsprachigen, aber auch im deutschsprachigen Raum etabliert ist und wir bereits viele Antworten kennen, es aber gute Argumente dafür gibt, das Forschungsfeld noch nicht als ausreichend beforscht anzusehen, sondern neue Herausforderungen zu erkennen und ihnen pro aktiv zu begegnen.

2 Lesekompetenz – produktbezogene Modellierungen des Konstrukts in Bildungsstudien

2.1 Lesekompetenz: Definition und Logik der empirischen Kompetenzstufen

In regelmäßigen zeitlichen Abständen werden die sprachlichen, mathematischen und naturwissenschaftlichen Leistungen von Schülerinnen und Schülern in Bildungsstudien überprüft. So vergleichen etwa PISA (Programme for International Student Assessment, OECD, 2019) und IGLU (Internationale Grundschul-Lese-Untersuchung, Hußmann et al., 2017) die Kompetenzen von Schülerinnen und Schülern auf internationaler Ebene, während die IQB-Bildungstrends (Stanat et al., 2017; vormals Ländervergleiche) die Schulleistungen innerhalb Deutschlands miteinander vergleichen. Im Gegensatz zur PISA-Studie, die die Kompetenzen von 15-Jährigen untersucht, erfassen IGLU und der aktuelle IQB-Bildungstrend die Lese- und Schreibleistungen von Viertklässlern. Für die Sekundarstufe erfasste der IQB-Bildungstrend zuletzt 2015 die Lesekompetenz von Schülerinnen und Schülern der 9. Klassen (Stanat et al., 2016). Die zugrundeliegenden Definitionen der Lesekompetenz in den verschiedenen Bildungsstudien ähneln sich naturgemäß. *Lesekompetenz wird übergreifend als Spektrum von Fähigkeiten verstanden, die es den Leserinnen und Lesern ermöglichen, kontinuierliche sowie diskontinuierliche Texte zu verstehen, zu nutzen und über deren Inhalte zu reflektieren.*

Um die Leseleistungen der Schülerinnen und Schüler interpretieren und (international) vergleichbar machen zu können, werden die erreichten Punktwerte Kompetenzstufen zugeordnet. Kompetenzstufen bilden Intervalle auf einer kontinuierlichen metrischen Skala von Leistungspunkten ab. Diese Intervalle entsprechen jeweils einer definierten Spanne von Punkten. In PISA 2018 bspw. wird die Lesekompetenz auf acht Kompetenzstufen dargestellt, wobei jede Stufe etwa 80 Punkte umfasst (OECD, 2019). Jede Kompetenzstufe ist ausführlich beschrieben, sodass sie eine qualitative Interpretation von Punktwerten ermöglicht und so die Rückbindung an die realen Leistungen von Schülerinnen und Schülern gewährleistet.

Die Kompetenzstufen der o. g. Bildungsstudien sind im Vergleich mehr oder weniger ausdifferenziert. Inhaltlich ähneln sich die kategorialen Beschrei-

bungen, diese werden jedoch jeweils gröber (fünf Kompetenzstufen) oder feiner (acht Kompetenzstufen) gefasst. Die Kompetenzstufen beruhen auf empirisch gewonnenen Daten zu den Lösungshäufigkeiten der Aufgaben und einer theoretisch ausgearbeiteten graduellen Beschreibung von Fähigkeitsniveaus, die in den Bildungsstudien unterschiedlich generiert und modelliert werden. Bei PISA (OECD, 2019) orientiert sich das Kompetenzniveau der Schülerinnen und Schüler an den Aufgaben, die sie gelöst haben. Das heißt, dass in PISA zum einen der Schwierigkeitsgrad der Aufgaben und zum anderen auch das Kompetenzniveau der Leserinnen und Leser auf Basis von Item-Response-Theorien dargestellt werden. In den IQB-Bildungstrends werden anhand von empirisch gewonnenen Daten Skalen definiert, die in Anlehnung an die von der KMK verabschiedeten Bildungsstandards (KMK, 2004) die Kompetenzstufen bilden. Auch in der IGLU Studie von 2016 (Hußmann et al., 2017) werden die Kompetenzstufen aufgrund von Schülerfähigkeiten und Aufgabenschwierigkeiten gebildet: Die Leseleistungen werden hier auf einer Skala abgebildet, die in gleich große Intervalle eingeteilt wurde, sodass fünf Kompetenzstufen entstehen.

Eine Modellierung von Lesekompetenz aufgrund von Kompetenzstufenmodellen ist demnach in den Bildungsstudien das Ergebnis einer punktuellen Leistungserfassung und kann aufgrund der zyklischen Testung verschiedener, aber zum Zeitpunkt der Messung gleichaltriger Personen keine Leistungsentwicklung von Individuen oder Gruppen abbilden. Im Folgenden werden die in den Bildungsstudien dargelegten Kompetenzstufen näher erläutert, um anschließend ein mögliches gemeinsames Muster zu extrahieren.

2.2 Drei Beispiele und eine Synopse

2.2.1 IQB-Bildungstrends

In den IQB-Bildungstrends für die Primar- sowie für die Sekundarstufe (Stanat et al., 2016; Stanat et al., 2017) werden fünf Kompetenzstufen unterschieden. Während auf dem ersten Fähigkeitsniveau explizite Einzelinformationen, die sich an leicht auffindbaren Stellen im Text befinden, gewonnen werden können, ist die folgende Kompetenzstufe 2 dadurch gekennzeichnet, dass benachbarte Informationen miteinander verknüpft werden können, die weniger explizit sind. Auf Kompetenzstufe 2 sind die Texte – insbesondere in den Testungen für die Sekundarstufe – strukturell komplexer und länger. Auf diesem Fähigkeitsniveau sind die Schülerinnen und Schüler in der Lage, Informationen aufgrund von Schlussfolgerungen miteinander zu verknüpfen (Inferenzen, s. Teilkap. 3.2.3). So können die Leserinnen und Leser einfache kausale Zusammenhänge oder Motive von handelnden Personen identifizieren. Auf Kompetenzstufe 3 verfügen die Schülerinnen und Schüler über erweiterte Fähigkeiten im Lesen: Auf dieser Stufe können sie Informationen miteinander ver-

knüpfen, die über den Text verstreut sind, indem sie eigene Inferenzen bilden. Für die Sekundarstufe werden hier Aufgaben angeboten, die sich zum Teil auch auf diskontinuierliche Texte, die bspw. Diagramme enthalten, beziehen. Im Unterschied zu Kompetenzstufe 2 können die Schülerinnen und Schüler hier den Text weitestgehend als Ganzes erfassen und verstehen. Auf Kompetenzstufe 4 verstehen die Schülerinnen und Schüler Texte, die inhaltlich schwieriger und strukturell komplexer sind. Auf diesem Fähigkeitsniveau können Leserinnen und Leser wesentliche Aspekte eines Textes identifizieren, indem sie zentrale Aussagen aus Textpassagen miteinander in Beziehung setzen. Dabei sind zur Herstellung von Kohärenz eigene Schlüsse zu ziehen. Auf dieser Stufe können die Personen nicht nur Zusammenhänge erkennen, sondern darüber hinaus auch über die Textgestaltung reflektieren. Auf Kompetenzstufe 5 können die Schülerinnen und Schüler überwiegend reflektiert lesen, komplexe Schlussfolgerungen ziehen und unter Einbeziehung ihres eigenen Vorwissens zentrale Aussagen des Textes bewerten. Leserinnen und Leser auf diesem Fähigkeitsniveau können im Text formulierte Behauptungen reflektierend begründen. Schülerinnen und Schüler auf dieser Kompetenzstufe zeigen zudem ein umfassendes und detailliertes Textverständnis, das sie dazu befähigt, Aspekte des Textes zu interpretieren und eigene Bewertungsansätze zu formulieren.

2.2.2 IGLU

Auch IGLU 2016 definiert fünf Kompetenzstufen für die Beschreibung der Lesefähigkeiten von Kindern in der 4. Klasse. Anders als die IQB-Bildungstrends beginnt hier aber die Beschreibung der Fähigkeitsniveaus auf der Ebene des rudimentären Leseverständnisses. So wurden im Rahmen von IGLU 2016 (Deutschland) zur Erfassung der basalen Lesefähigkeiten zusätzlich zu den vorgegebenen Testaufgaben zwei Untertests des ELFE-Lesekompetenztest (ELFE II, Lenhard et al., 2017) eingesetzt. Der ELFE II dient der Erfassung der Leseflüssigkeit, der Lesegenauigkeit und des Leseverständnisses auf Wort-, Satz- und Textebene. Für IGLU 2016 wurden lediglich die Untertests für das Wort- und Satzlesen genutzt. So war es möglich, die Leseleistungen von Schülerinnen und Schülern zu erfassen, die im Rahmen des IGLU-Lesemodells nicht beschreibbar waren, da sie unterhalb der Schwelle von 390 bis 410 Punkten lagen. Die Definitionen der weiteren Kompetenzstufen ähneln denen der oben beschriebenen IQB-Bildungstrends.

2.2.3 PISA

Erstmals wurde im Jahr 2018 die PISA-Studie (in den meisten Ländern) digital durchgeführt, um dem sich verändernden Leseverhalten von Jugendlichen Rechnung zu tragen. Insbesondere die technischen Entwicklungen haben zu

neuen Lesegewohnheiten geführt, die auch das Lesen am Smartphone oder Bildschirm beinhalten. Die Aufgabenformate umfassen daher ein breites Spektrum an Texten: kontinuierliche Texte und diskontinuierliche Texte, die Tabellen, Grafiken und Abbildungen enthalten können, sowie multiple Texte, die auch Texte aus Internetforen inkludieren.

So wurden in PISA 2018 acht Kompetenzstufen unterschieden, wobei die erste Stufe in drei Substufen 1a, 1b und 1c aufgeteilt ist. Damit ist die Skalierung bei PISA differenzierter als in den oben beschriebenen Bildungsstudien. Im Jahr 2018 wurde erstmals die Leseflüssigkeit der Schülerinnen und Schüler evaluiert. Hierzu wurden Sätze vorgegeben, die inhaltlich richtig oder falsch waren und von den Schülerinnen und Schülern in einer definierten Zeit als richtig oder falsch bewertet werden mussten. Diese Leseleistung wird Stufe 1c zugeordnet; die Stufen 1b und 1a werden zunehmend schwieriger: Während Leserinnen und Leser auf Stufe 1b offensichtliche und explizite Informationen finden können, können sie auf Stufe 1a bereits mehrere benachbarte Informationen miteinander in Beziehung setzen und sogar das Hauptthema eines Textes oder die Absicht des Autors erkennen. Auf Stufe 2 können die Schülerinnen und Schüler durch Schlussfolgerungen Textzusammenhänge erkennen, auch wenn die Informationen nicht unmittelbar ersichtlich sind. Auf Stufe 3 ist es den Leserinnen und Lesern möglich, über Texte oder eine kleine Zusammenstellung von Texten zu reflektieren und Informationen zu interpretieren, die explizit auffindbar sind. In den weiteren Stufen 4 bis 6 werden die angebotenen Texte bzw. Textsammlungen zunehmend abstrakter. Jugendliche auf diesen Stufen können Informationen, die für die Aufgabenstellung relevant sind, gewichten und kritisch beurteilen. Die Leserinnen und Leser auf Stufe 6 sind in der Lage, lange und abstrakte Texte zu verstehen und tief eingebettete Informationen zu entnehmen. Texte dieser Stufe können widersprüchliche Perspektiven und Informationen enthalten, die von den Leserinnen und Lesern reflektiert und bewertet werden müssen, um zu einer richtigen Beantwortung der Fragestellungen zu gelangen.

2.2.4 *Synopse*

Den hier beschriebenen drei Modellierungen von Lesekompetenz ist gemeinsam, dass sie ein Spektrum abbilden, das von elementaren Lesefähigkeiten über das Verknüpfen von Informationen, also dem sinnentnehmenden Lesen bis hin zum überwiegend reflektierenden Lesen mit komplexen Inferenzbildungen und Interpretationen reicht. Übergreifendes Merkmal dieser Modellierungen von Lesekompetenz ist demnach, die propositionale Struktur von Texten besser zu erkennen und zu verarbeiten. Zunehmende Kompetenz ist dadurch gekennzeichnet, ein evaluatives, teils kritisches Lesen zu bemühen, um dadurch die aufgabenspezifischen Anforderungen zu erfüllen. Hinzu kommen Aspekte der

Leseflüssigkeit, die aber nur in bestimmten Modellen salient ist, sonst aber als Bestandteil vorausgesetzt wird. Zudem werden, wie in Abbildung 1 dargestellt, der Textumfang größer sowie die Struktur der Texte komplexer und die Inhalte abstrakter. Bei der Bedeutungserschließung entfernen sich die beschriebenen Kompetenzstufen immer weiter von der Textoberfläche, sodass sich das Lesekompetenzkontinuum von der Erfassung wörtlicher Informationen auf der Wort- und Satzebene bis hin zu umfangreichen, teils auch multimedialen Texten bewegt. In Abbildung 1 werden zentrale Aspekte der Kompetenzstufen in Verbindung mit Textlänge und Komplexität modellhaft zusammengeführt. Stufe A bildet das rudimentäre Leseverstehen ab, das bei PISA 2018 explizit als Leseflüssigkeit operationalisiert und in das Lesekompetenzmodell aufgenommen wurde. In IGLU 2016 wurde durch den Einsatz der oben beschriebenen zwei Untertests des ELFE-Verständnistests (ELFE II, Lenhard et al., 2017) das bisherige Lesekompetenzmodell um das rudimentäre Leseverstehen erweitert. Stufe B fokussiert auf das Verstehen von besonders leicht zu findenden expliziten Informationen. Auf Stufe C können explizite Informationen entnommen werden, die sich in benachbarten Sätzen befinden, während auf Stufe D die Informationen weniger explizit und zudem verstreut sind. Stufe E hingegen erfordert das höchste Maß an Inferenzbildung. Hier sind Informationen implizit vorhanden und können Aufgaben nur durch Reflektieren, Bewerten und Schlussfolgern gelöst werden.

Abbildung 1: Synopse der aufeinander aufbauenden Fähigkeiten bei der Lesekompetenz in empirischen Studien



(eigene Darstellung, * = nur bei den jüngsten IGLU- und PISA-Studien explizit enthalten)

3 Lesekompetenz – prozessbezogene Perspektiven auf zentrale Teilfähigkeiten

Die Lesekompetenz umfasst viele Aspekte des Lesens. Man kann den Begriff eher eng fassen und nur diejenigen Teilkompetenzen einbeziehen, die Prozesse des verstehenden Lesens miteinbeziehen (z. B. das Entwickeln mentaler Reprä-

sentationen im Ergebnis von Inferenzprozessen). Man kann Lesekompetenz auch weiter fassen – zumal unter der Erwerbsperspektive – und sich zusätzlich auf solche Aspekte des Lesens beziehen, bei denen es noch nicht um Sinnverstehen geht (z. B. das Dekodieren von Wörtern). Rosebrock (2012, S. 4) beschreibt die Lesekompetenz entsprechend als „ein ganzes Bündel von Teilfähigkeiten“. Diese Teilfähigkeiten sind auf verschiedenen Ebenen angesiedelt. Basale Leseleistungen wie die Automatisierung des Wort- und Satzlesens werden im Konstrukt der Leseflüssigkeit zusammengefasst und auch als hierarchieniedrige Leseprozesse bezeichnet.

So betonen Lenhard und Schneider (2006, S. 13), dass bei den hierarchieniedrigen Lesefähigkeiten Aspekte von zentraler Bedeutung sind, „die das Leseverstehen auf Wortebene beeinflussen. Dazu gehören der Wortschatz des Kindes, die Fähigkeit zum Dekodieren einzelner Wörter, die Erfassung der Wortbedeutung sowie das Wissen, wie diese Bedeutung durch den jeweiligen Kontext modifiziert wird“. Hier wird deutlich, dass das theoretische Konstrukt des Leseverständnisses auf hierarchieniedriger Ebene die visuelle Worterkennung als Voraussetzung für die Bedeutungserschließung einzelner Wörter sowie deren Vernetzung auf Satzebene einschließt. Die hierarchiehohen Leseprozesse umfassen das eigentliche Textverstehen und Durchdringen eines Textes. Neben der Bedeutungserschließung auf der Wort- und Satzebene müssen die Leserinnen und Leser die aufgenommenen Informationen mit ihrem Vorwissen verknüpfen, um den Text als Ganzes zu erfassen. Dabei generieren sie ein mentales Modell des Textes, in dem das bisher Gelesene repräsentiert und reflektiert ist (Richter & Schnotz, 2018).

Für das Lesen können also hierarchieniedrige und hierarchiehöhere Teilprozesse unterschieden werden: Zuerst wird durch Buchstaben- und Worterkennung das Wortmaterial erkannt (Dekodieren), das dann mit weiteren Wörtern in einem Satz in Beziehung gesetzt werden muss. Hierbei ist eine Analyse der syntaktischen Strukturen notwendig, um zu einer inneren Repräsentation der Satzbedeutung zu gelangen. Die gelesenen Sätze werden in einem weiteren Prozess fortlaufend verknüpft zu einem inneren Modell der Textbedeutung. Dieses wird eingebettet in das Weltwissen der Lesenden. Das Lesen ist damit ein „hochgradig aktiver Prozess der Auseinandersetzung mit den jeweiligen Inhalten“ (Lenhard & Schneider, 2006, S. 13). Die Leseprozesse auf hierarchieniedriger Ebene hingegen laufen automatisiert ab, d. h. praktisch ohne kognitive Kontrolle, wodurch Ressourcen für die hierarchiehöheren Prozesse des Textlesens frei bleiben.

Das in den Bildungsstudien modellierte Konstrukt der Lesekompetenz fokussiert auf das Verstehen von Satz- und Textbedeutungen. Zwar wird ein breites Spektrum von Teilkompetenzen durch die Kompetenzstufen abgebildet, jedoch wird deutlich, dass die hierarchieniedrigen Leseprozesse, wie das Dekodieren auf Wortebene oder die Leseflüssigkeit als Automatisierungsprozess nur

vereinzelte und vor allem eher als Randphänomen einbezogen werden. Betrachtet man Lesekompetenz jedoch nicht nur als theoretische Rekonstruktion eines empirischen Ist-Zustandes, der eine punktuelle Leistung verdeutlicht, sondern als Lern- und Entwicklungsprozess, muss notwendigerweise auch der Beginn des Leselernens betrachtet werden.

3.1 Leseflüssigkeit

3.1.1 Definition und Dimensionalität

Der Leseprozess beinhaltet den basalen Prozess des Dekodierens. Hierbei entschlüsseln Leserinnen und Leser die schriftlichen Zeichen, indem sie ihnen entsprechende konventionalisierte Bedeutungen zuordnen. Das Dekodieren wird durch Wiederholung automatisiert, um so das Arbeitsgedächtnis zu entlasten und dadurch wesentliche Informationen vorübergehend speichern und gleichzeitig verarbeiten zu können. So wird das flüssige Lesen erlangt. Dies ist zum Beispiel erforderlich, um einen Satz am Ende inhaltlich verstanden zu haben, ohne den Anfang bereits wieder vergessen zu haben. Damit bildet die Leseflüssigkeit eine notwendige Voraussetzung für die Bedeutungerschließung von Texten. Die Leseflüssigkeit gehört, wie oben dargestellt, zu den hierarchie-niedrigen Prozessen auf der Wort- bzw. Satzebene. Sie zählt zu den durch zahlreiche Forschungsergebnisse gut dokumentierten Komponenten des Lesens.

In der angloamerikanischen Forschungsgemeinschaft werden die hierarchieniedrigen Leseprozesse unter dem Begriff *Fluency* subsumiert und bereits seit 40 Jahren verstärkt in den Blick genommen. Forschungsergebnisse belegen, dass die Leseflüssigkeit stark mit dem Leseverständnis korreliert, also für das Verstehen eines Textes von zentraler Bedeutung ist (García & Cain, 2014; Wolters et al., in Druck). So ergeben die Studien, dass die Steigerung der Leseflüssigkeit eine Verbesserung des Leseverständnisses mit sich bringt, und zwar ohne dass zusätzlich hierarchiehöhere Verstehensprozesse gefördert werden müssen (National Reading Panel, 2000).

Leseflüssigkeit ist ein komplexes Konstrukt. Sie vereint in sich die genaue, automatisierte, schnelle, sinnkonstituierende Fähigkeit zu lautem und leisem Lesen, welche es ermöglicht, die Bedeutung des Gelesenen zu erschließen (Rosebrock & Nix, 2006, S. 94). Im Folgenden werden diese Einzelkomponenten kurz dargestellt:

- *Akkuratheit des Dekodierens*: Falsch erlesene Wörter führen zu Bedeutungsveränderungen (Pinnel et al., 1995). Nur das fehlerfreie Erlesen von Wörtern führt zur schnellen und sicheren Bedeutungerschließung, denn Schwächen auf dieser basalen Ebene führen notwendigerweise zu Einschränkungen im Textverständnis.

- *Automatisierung*: Das Lesen muss automatisiert werden, um möglichst wenige kognitive Ressourcen für den Dekodierprozess aufzuwenden. So ist gewährleistet, dass genügend Kapazitäten für Bedeutungerschließungen bzw. Schlussfolgerungen zur Verfügung stehen (Samuels, LaBerge & Bremer, 1978).
- *Geschwindigkeit*: Akkuratheit und Automatisierung bilden zusammen die Lesegeschwindigkeit. Langsame, stockende Leserinnen und Leser zeigen Schwächen im Leseverständnis, weil sie viel zu viele Informationen im Arbeitsgedächtnis speichern müssen (Berliner & Casanova, 1988; Brown & Hirst, 1983).
- *Prosodisches Lesen*: Leserinnen und Leser sollten im Stande sein, den Text in sinnkonstituierende Einheiten einzuteilen und dies durch ihre Betonung deutlich zu machen. Leserinnen und Leser, die einen Text zwar zügig und mit nur wenigen Fehlern, jedoch ohne jegliche Betonung lesen, verstehen den Text wahrscheinlich nicht umfassend (Schreiber, 1991). Neuere Befunde unterstreichen, dass es substanzielle positive Zusammenhänge zwischen Prosodie und Leseverstehen gibt (Wolters et al., in Druck).

3.1.2 *Das Konstrukt der Leseflüssigkeit und die Beiträge in diesem Band*

Die Fähigkeit zu flüssigem Lesen gilt als eine Voraussetzung für das verstehende Lesen. Für die Förderung der Leseflüssigkeit haben sich insbesondere Lautleseverfahren als wirksam erwiesen. Lautleseverfahren vereinen in sich unterschiedliche Aspekte der Förderung: Durch instruktionale Trainings wird vor allem die Automatisierung des Lesens erreicht. Beim wiederholenden Lesen lesen die Schülerinnen und Schüler einem Tutor (z. B. einem anderen Schüler oder einer anderen Schülerin) einen kurzen, für sie mittelschweren Text so lange immer wieder laut vor, bis sie eine zuvor festgelegte Flüssigkeit (definiert über Zeit) erreicht haben. Häufiges Wiederholen führt erwiesenermaßen zu einer besseren Verfügbarkeit von Kenntnissen und Fertigkeiten, denn auf diese Weise automatisiert sich der kognitive Zugriff auf das Gelernte (Hasselhorn & Gold, 2017). Dass die Förderung von Leseflüssigkeit erfolgreich ist und insbesondere durch Lautleseprogramme gute Effekte erzielt werden, ist eindeutig belegt (National Reading Panel, 2000).

Leseflüssigkeit, das zeigt sich an der Dimension Prosodie, trägt Aspekte des Verstehens in sich (Wolters et al., in Druck). Auf Wortebene müssen Personen automatisiert auf die Semantik von sprachlichen Einheiten zugreifen können, um eine angemessene Prosodie zu erzielen. Dafür sind sie auch auf ihren Wortschatz angewiesen, der seinerseits als wichtige Voraussetzung für das Leseverstehen gilt (Philipp, 2012). Studien zeigen, dass Wortschatzförderung das Leseverstehen erhöht (Elleman et al., 2009). Britta Juska-Bacher, Christoph Zangger und Martina Röthlisberger stellen in ihrem Beitrag in Kapitel 2 Ergebnisse ihrer

Studie zur Entwicklung von Wortschatz und Lesefähigkeiten bei Kindern der Primarstufe vor. Hierbei gehen sie der Frage nach, inwieweit Wortschatzkompetenz mit der Lesekompetenz zusammenhängt. Ihre Ergebnisse aus der Grundlagenforschung illustrieren, dass Wortschatz und Leseverstehen zu Beginn der Schullaufbahn in autoregressiven Kreuzpfadmodellen statistisch unverbunden sind, aber laut Fixed-Effects-Modell ein positiver Effekt vom Wortschatz auf das Leseverstehen besteht. Anders gesagt: Wer in der ersten Klasse einen sowohl qualitativ als auch quantitativ stärker ausgeprägten Wortschatz hatte, zeigte ein Jahr später bessere Leseleistungen.

Mareike Ehlert und Elmar Souvignier skizzieren in ihrem Beitrag zentrale Befunde zu wirksamen Fördermaßnahmen im Bereich der Leseflüssigkeit und stellen mit dem Programm „Der Lese-Sportler“ eine Möglichkeit zur Förderung in der Grundschule vor. Sie erläutern empirische Befunde, die zwar eine hohe Wirksamkeit gezielter Maßnahmen zur Förderung der Leseflüssigkeit belegen, aber auch zeigen, dass die Effekte in der schulischen Praxis niedriger ausfallen als in stärker kontrollierten Settings. Vor diesem Hintergrund diskutieren Mareike Ehlert und Elmar Souvignier drei Aspekte der Leseflüchtigkeitsförderung – Passung, Dosierung und Wiedergabetreue – und ihre Bedeutung für die optimale Entfaltung der Wirkung von Maßnahmen zur Förderung der Leseflüssigkeit.

3.2 Lesestrategien

Im vorangegangenen Abschnitt wurde Leseflüssigkeit als Teilaspekt der Lesekompetenz auf hierarchieniedrigem Niveau behandelt. Es wurde deutlich, dass das flüssige Lesen eine Voraussetzung für das Verstehen von Texten ist. Im Folgenden wird darauf aufbauend das Kontinuum der Lesekompetenz weiter aufgespannt: über den Nutzen und den Einsatz von Lesestrategien bis hin zum Verstehen und Verarbeiten multipler Texte.

3.2.1 Definition und Klassifikation

Die Trennlinie von hierarchieniedrigen und -hohen Prozessen lässt sich anhand mehrerer Unterscheidungsmerkmale ziehen. Erstens besteht eine Differenzierung von automatisiert ablaufenden, nicht steuerbaren Prozessen und solchen, die sich in Bezug auf Ziele prinzipiell steuern lassen – erstgenannte Prozesse bilden die hierarchieniedrigen Fähigkeiten, letztere die hierarchiehohen Fähigkeiten (Afflerbach et al., 2008). Zweitens ist auch die Textmenge und der sich daraus ergebende inhaltliche Verarbeitungsaufwand ein Distinktionsmerkmal: Bei hierarchieniedrigen Prozessen bilden Sätze in der Regel die Grundlage, während hierarchiehohe Prozesse sich auf längere Textteile, ganze Texte oder sogar multiple Texte beziehen (Lenhard, 2019; List & Alexander, 2019). Kenn-

zeichnend für die hierarchiehohe Prozesse ist damit, dass die lesende Person diverse steuerbare Prozesse auf eine solche Art orchestriert, dass textuelle Informationen und eigenes Wissen verstehensförderlich bzw. zielbezogen auf einer Objektebene interagieren lässt. Hinzu kommen weitere, über den eigentlichen mit der Sinnkonstruktion auf der Objektebene stattfindenden Überwachungs- und Regulationsprozesse, die auf einer Metaebene ablaufen und der Zielerreichung dienen (Bruin et al., 2020). Solche verstehensförderlichen zielbezogenen Aktivitäten, die zu weiten Teilen auf metakognitiven Wissensbeständen, also Wissen über Kognitionen, basieren, nennt man *Lesestrategien* (Philipp, 2015).

Die Lesestrategieforschung hat ihre Wurzeln zum Teil in Studien mit qualitativ gewonnenen Laut-Denk-Protokollen. Hierbei handelt es sich um Untersuchungen, in denen in aller Regel versierte Leserinnen und Leser lesebezogene Aufgaben erhielten und dabei parallel zur Aufgabenbearbeitung ihre Gedanken verbalisieren sollten. Diese aufgenommenen und transkribierten Gedankenströme bildeten als Laut-Denk-Protokolle dann den Gegenstand der Codierungen und Analysen. Solche Primärstudien wurden in Forschungsüberblicken zusammengefasst (Pressley & Afflerbach, 1995; Cho & Afflerbach, 2017; Israel, 2015), welche sich der Aufgabe angenommen haben, zum einen studienübergreifende Muster in Form von funktionsbezogenen Klassifikationen der Strategien (Pressley & Afflerbach, 1995; Cho & Afflerbach, 2017) vorzunehmen oder zum anderen entwicklungsbezogene Differenzen (Israel, 2015) herauszuarbeiten. Ein erstes wichtiges Ergebnis ist, dass gute Leserinnen und Leser variable Vorgehensweisen zeigen und dabei auf einen Verbund von Lesestrategien setzen, der je nach Stand im Kompetenzerwerb verschieden elaborierte Kombinationen von Strategien widerspiegelt. Ein zweites wichtiges Ergebnis lautet: Die einzelnen Strategien treten in der Regel im konzertierten, flexiblen und adaptiven Verbund auf und dienen dazu, aktuelle und übergeordnete Verstehensleistungen zu ermöglichen.

Strategien dienen beim Lesen unterschiedlichen Zwecken, erfüllen also unterschiedliche Funktionen. Die verschiedenen Funktionen der Lesestrategien für das Leseverstehen bilden den Grundstein für einen weithin etablierten Klassifikationsversuch, der seinerseits seinen Ursprung in Taxonomien von Lernstrategien hat (als übergeordnete, domänenunspezifischen Handlungspläne zur Unterstützung der inhaltlichen Aneignung und Speicherung von Lerninhalten; Weinstein & Mayer, 1986). Aus diesen Strategien sind zwei Gruppen, die ihrerseits noch weiter unterteilt werden können, besonders deutlich herausgearbeitet worden: kognitive Strategien und metakognitive Strategien. *Kognitive Strategien* dienen dazu, dass Leserinnen und Leser eine sinnvolle Interaktion zwischen text- und vorwissensbezogenen Prozessen vornehmen (sogenannte Objektebene), während die *metakognitiven Strategien* auf einer Metaebene diese interagierenden Prozesse initiieren, überwachen und regulieren – beides unter ef-

fektiver Nutzung der Arbeitsgedächtniskapazität (Bruin et al., 2020). Diese Gruppen von kognitionsbezogenen Strategien sind in Tabelle 1 dargestellt; es handelt sich hierbei dezidiert nur um einen Ausschnitt von Strategien (s. Philipp, 2015, für einen Überblick über weitere Lesestrategien, sowie Hattie & Donoghue, 2016, für einen allgemeinen Überblick über verschiedene Lesestrategien).

Tabelle 1: Selektiver Überblick über Lesestrategien

Gruppe von Strategien	Beschreibung	Beispiele
A) Kognitive Strategien		
<i>A1) Organisationsstrategien</i>	Inferenzbasierte Rekonstruktion der Struktur von Lerngegenständen	<ul style="list-style-type: none"> – Zusammenfassungen oder Synthesen schreiben – Hauptideen unterstreichen – Concept-Map oder ein anderes Schaubild erstellen
<i>A2) Elaborationsstrategien</i>	Aktives Verknüpfen von neuen Informationen mit Vorwissensbeständen, um gegebene Inhalte anzureichern (zu elaborieren)	<ul style="list-style-type: none"> – Textinhalte imaginieren – Unbekannte Wörter klären – Fragen an den Text formulieren
B) Metakognitive Strategien		
<i>B1) Planungsstrategien</i>	Vorbereitung der Leseaktivität durch Formieren von Lesezielen und Kohärenzstandards	<ul style="list-style-type: none"> – Entscheiden, welches Verstehensniveau benötigt wird (Kohärenzstandards festlegen) – Textteile portionieren und priorisieren
<i>B2) Überwachungsstrategien</i>	Überwachen des Fortschritts in Bezug auf die Zielerreichung	<ul style="list-style-type: none"> – Fragen an sich selbst formulieren, um das eigene Verstehen zu klären – Inhalte gezielt in eigenen Worten wiedergeben und erklären
<i>B3) Selbstregulationsstrategien</i>	Verändern und Optimieren des kognitiven und metakognitiven Strategieeinsatzes	<ul style="list-style-type: none"> – Verlangsamung des Lesetempos – Spezifische erneute Lektüren einzelner Teile mit teils geänderten Lesezielen – Auswahl einer neuen Herangehensweise

(Eigene Darstellung, basierend auf Friedrich & Mandl, 2006, und Pressley & Afflerbach, 1995)

Die Klassifikationen der Lesestrategien – auch jene aus Tabelle 1 – entstammen aus Studien, in denen häufig das möglichst umfassende Verstehen einzelner gedruckter (Sach-)Texte untersucht wurde. Damit geben sie Einblick in einen limitierten Ausschnitt dessen, was bei der Lesekompetenz gefordert ist. Nicht nur fehlen literarische Texte mit ihren sprachlich-ästhetischen Eigenheiten häufig, sondern es ergeben sich im sich wandelnden Lesealltag andere Herausforderungen (welche beispielsweise PISA 2018 schon explizit aufgegriffen hat,

s. Teilkap. 2.2.3). Darunter fallen die Lektüren elektronischer, multimodaler und multimedialer Texte ebenso wie das Lesen multipler Texte – unter Realbedingungen müssen sich Leserinnen und Leser diese Texte erst über Suchmaschinen suchen (Bråten et al., 2020). All dies setzt teils andere, teils neue Strategien voraus, um den lesebezogenen Anforderungen vollumfänglich gerecht zu werden (Cho & Afflerbach, 2017). Kurzum: Das Lesestrategiespektrum für die alltäglich nötige Lesekompetenz diversifiziert sich und erhöht damit die Anforderungen für alle Personen, die aktiv an der schriftbezogenen Kommunikation teilhaben wollen.

3.2.2 Die lesedidaktisch bedeutsame Frage nach der Vermittelbarkeit von Lesestrategien und die Beiträge in diesem Band

Lesestrategien gelten gemäß den Ausführungen aus dem Teilkapitel 3.2.1 als bedeutsam für das kompetente Lesen, und Entwicklungsmodelle betonen, dass kompetente Leserinnen und Leser über ein Inventar verschiedener Lesestrategien verfügen (Alexander, 2005). Auch andere theoretische Zugänge postulieren einen Zielhorizont des kompetenten Lesens, der auf einer Vielzahl von verfügbaren Lesestrategien verschiedener Funktionen basiert (Pressley & Afflerbach, 1995). Damit stellt sich die Frage, wie Personen solche Strategien in ihrem Verbund erwerben und erwerben sollen.

Als ein in der Forschung breit etablierter Zugang hat sich die explizite Strategievermittlung erwiesen, bei der Lernende von kompetenten anderen im Sinne des Beobachtungslernens die Strategieanwendung erklärt, vorgeführt und metakognitiv kommentiert stellvertretend erleben und analysieren können. Auf dieser Basis repräsentieren Lernende kognitiv zunächst die Anwendung der Strategien, ehe sie sie im Sinne eines allmählichen Übertrags der eigenverantwortlichen Strategieanwendung diese didaktisch portioniert selbst reproduzierend anwenden und sich sukzessive sowohl Automatisierungs- als auch metakognitive Evaluationsprozesse flankieren. Dies hat eine Doppelfunktion im Erwerb: Die eigentliche Ausführung der Strategie soll einerseits selbst weniger kognitive Belastung generieren und fehlerfreier erfolgen, andererseits soll das konditionale Strategiewissen konsolidiert werden (Philipp, 2015).

Trotz dieser aus verschiedenen Theoriesträngen plausibilisierbaren Abfolge von Vermittlungsschritten bzw. -phasen mit klar definierter Funktion im Gesamtzusammenhang besteht ein hartnäckiges Problem darin, dass diese explizite Vermittlung von kognitiven und metakognitiven Strategien im Unterricht nicht oder nur verkürzt vorkommt (z. B. für Befunde aus Deutschland: Schmitz, 2019). Möglicherweise hat dies damit zu tun, dass Strategievermittlung ein komplexes Unterfangen ist, das seitens der Lehrenden als Gelingensbedingung neben profundem lesedidaktischem Wissen auch eine Form der Unterrichtsentwicklung voraussetzt, damit Lernende umfassende Repräsentationen zum

Zweck und Ziel strategischen Lesens überhaupt erlangen (Duffy, 1993). Vor diesem Hintergrund stellt sich die Frage, was denn unverzichtbare, die Wirkung der Lesestrategievermittlung nicht nur in der Strategieanwendung, sondern auch für den sich im Leseverstehen niederschlagenden Erfolg beeinflussende Elemente sind. Solchen Fragen geht die Forschung inzwischen nach, wie es beispielsweise die Metaanalyse von Okkinga et al. (2018) illustriert.

Davon ungeachtet sind auch Primärstudien aus einer didaktischen Perspektive gefragt, die wirksamen Konstituenten einer Fördermaßnahme klar herauszuarbeiten. Eher unfreiwillig, nämlich forciert durch die Covid-19-Pandemie, ist dies zum Gegenstand der Analysen im Kapitel 4 von Johannes Wild, Elisabeth Kraus, Maria Steinert, Sven Hilbert und Anita Schilcher avanciert. Im Rahmen einer laufenden Interventionsstudie (FiLBY) mit Grundschulkindern der dritten Schulklasse konnten sie anhand von Leistungsentwicklungen im Leseverstehen und per Online-Fragebogen von den Lehrpersonen erfragten Unterrichtsmerkmalen auf breiter Basis zeigen, dass ein abrupter Wechsel auf nicht vom Programm so vorgesehene lesedidaktische Entscheidungen sogar nachteilig für die Entwicklung des Leseverstehens ist. Nur die Einführung der Sachtextstrategien durch die Lehrperson war hingegen vorteilhaft für einen Zuwachs im Leseverstehen. Solche Befunde sind wichtig, weil sie helfen, den Weg zu ebnen von evidenzbasierten Förderprogrammen in den Regelunterricht.

3.2.3 Ein sich abzeichnendes neues Feld der Leseforschung (Lesen multipler Texte) und die Beiträge in diesem Band

Die Synopse der produktbezogenen Anforderungen von Leseverstehensleistungen, die im Teilkapitel 2.2.4 aus den Leseleistungsstudien extrahiert wurde, bildet einen Schwerpunkt ab, der auf das schlussfolgernde Kombinieren von Informationen aus distalen, sprich: weiter entfernten Textteilen und multiplen Texten fokussiert. Solche Schlussfolgerungen werden in der Leseforschung als (vorwissensbasierte) „Inferenzen“ bezeichnet und gelten als *conditio sine qua non* des Leseverstehens (Kendeou et al., 2016). Inferieren als Sammelbegriff bezeichnet – aus einer Prozessperspektive – die Fähigkeit einer lesenden Person, Informationen (aus dem Text und/oder dem Vorwissen) miteinander schlussfolgernd zu kombinieren, sodass Informationslücken in Texten gefüllt werden. Solche Inferenzen als Produkt werden nicht nur dazu benötigt, den Sinnzusammenhang in Textpassagen kohärent zu erkennen. Sie stehen aktuell auch deutlich erkennbar im Interesse der Forschung zu multiplen Texten. Das Lesen von mehreren Texten erfordert erhöhte Kohärenzleistungen, da die Materialität, Modalität und kommunikative Absicht verschiedener Texte nebst ihren textuellen Merkmalen (Textsorte, Lexik etc.) hochdivergierend sein kann. Entsprechend handelt es sich um schwierigkeitsgenerierende Merkmale, welche

eine erhöhte kognitive Belastung mit sich bringen (Ainsworth, 2018; Sweller, 2010). Hinzu kommt, dass intra- und intertextuelle Inferenzen gleichermaßen benötigt werden, damit die lesende Person mentale Modelle entwickeln kann. Sie muss also einzelne Texte ebenso verstehen wie die Gesamtheit der gelesenen Texte, und dafür braucht sie wiederum Strategien – viele fallen unter das Ziel, wie sich intertextuelle Kohärenz inferenzbasiert herstellen lässt, z. B. mit Querweisen in Notizen auf ähnliche thematische Aspekte in anderen gelesenen Texten (List et al., in Druck).

Die Forschung zum Lesen und Nutzen multipler Texte nimmt etwas auf, was andernorts bereits ein eigenes Feld der Leseforschung war: jene zu den *Aufgabeneffekten*. Die oftmals unhinterfragte Perspektive auf das Lesen bestand in der Vergangenheit darin, dass eine Person einen Text möglichst so versteht, dass ein adäquates mentales Modell das Resultat ist. Dieses Primat wurde von der aktuellen Leseforschung jedoch erschüttert, indem zunehmend die wertungsbezogene leserseitige Einschätzung und Verarbeitung von Informationen und Texten fokussiert wird. Es geht also um ein instrumentelles, ein wertendes Lesen, das beispielsweise darüber entscheidet, ob man Texte gänzlich verstanden oder singuläre Informationen aus ihnen extrahiert haben muss (Britt et al., 2018; Oudega & van den Broek, 2018). Dass und wie die Leserinnen und Leser sich dazu entscheiden, Informationen zu verwenden oder zu ignorieren, wird gegenwärtig verstärkt Gegenstand empirischer und theoretischer Arbeiten (McCrudden, 2018; McCrudden et al., 2011a) – nicht zuletzt, weil bereits die erfolgreiche Nutzung von Suchmaschinenergebnissen unter anderem auf dieser Fähigkeit der Relevanzzuweisung basiert.

Damit hängt zusammen, was sich ebenfalls deutlich in der Forschung herauschält: die Lesekompetenz multipler Texte und die Fähigkeit, das Leseverstehen in eigenen schriftlichen Texten zu dokumentieren. Tatsächlich zeigt sich in der Grundlagenforschung eine deutliche Tendenz, *Lesen und Schreiben zu integrieren* (Primor & Katzir, 2018). Das ergibt sich daraus, dass die Texte, die Personen über multiple Texte schreiben, als Produkt der aufgabenbezogenen Nutzung von multiplen Texten behandelt werden und besonders reichhaltige Analysegegenstände bilden. Sie eröffnen damit produktbezogene (und damit notwendigerweise eingeschränkte) Einblicke in Verstehensleistungen. Das bildet einerseits die sich wandelnden Kompetenzanforderungen ab, nach denen sich die Lesekompetenz auch daran bemisst, Verstandenes schriftlich adäquat weiter zu kommunizieren (Leu et al., 2017), was im deutschsprachigen Raum aktuell im Diskurs zum materialgestützten Schreiben behandelt wird. Andererseits weicht damit das Konstrukt Lesekompetenz auf: Es weist eine Schnittmenge mit der Schreibkompetenz auf. Dies zeigt sich in den Prozessen, die sich kaum noch eindeutig als nur lese- oder nur schreibbezogen entwirren lassen (Spivey, 1997).

Ein Nebeneffekt all dieser Entwicklungen ist, dass in der Leseforschung zunehmend auch *Leserinnen und Leser im Erwachsenenalter* stärker betrachtet werden. Zum einen bilden sie – zumal, wenn es sich um Studierende handelt – ein Convenience-Sample, das für Forschungsprozesse niederschwellig zu erreichen ist. Die Fokussierung auf Erwachsene – zumal solche aus Einrichtungen der Tertiärbildung – erlauben es auch, die wichtige Thematik der (normativen) Anschlüsse zwischen den Bildungsetappen genauer in den Blick zu nehmen. Dies zeichnet sich deutlich in der Logik der zyklischen Leseleistungsstudien ab, die nicht zufällig das Ende von Schulstufen fokussieren, aber noch selten Übergänge in andere Zweige des Bildungssystems in den Blick nehmen. Einen Sonderfall bilden die Lesefähigkeiten von Studierenden, die später Lehrpersonen werden und deren lesestrategisches Repertoire allein deshalb von Interesse sein sollte, weil es der Gegenstand späterer fachdidaktischer Kompetenz ist (Akyol & Ulusoy, 2010).

Diese Veränderungen in der Leseforschung spiegeln die Beiträge von Alex Rickert und Maik Philipp wider. Alex Rickerts und Maik Philipps Kapitel (Kapitel 5) geht beispielsweise darauf ein, welche Leistungen und welche Schwierigkeiten sich im Längsschnitt bei Lehramtsstudierenden aus der Schweiz zeigen. So divergieren die gezeigten Leistungen systematisch je nach kognitivem Anspruchsgrad und finden dort Grenzen, wo es um tiefe Verstehensleistungen mit hohem kognitiven Transformationsgrad geht. Markant sind zusätzlich die geringen Zusammenhänge zwischen den drei von ihnen untersuchten Prozessen des Lesens, die sie in den Texten aufwändig rekonstruiert haben. Mit ähnlicher Perspektive auf die Produkte des Leseverstehens widmet sich Maik Philipp den Aufgabeneffekten beim schriftlichen Synthetisieren anhand einer Sekundäranalyse von Experimentalstudien mit überwiegend erwachsenen Leserinnen und Lesern (Kapitel 6). Sein Hauptergebnis ist, dass verschiedene unabhängige Variablen – gerade in der Kombination – komplexe Befundmuster nach sich ziehen. Dennoch gibt es ein Muster: Je komplexer die nötigen Lese- und Schreibstrategien waren und je mehr Transformationen sie nach sich zogen, desto stärker fielen Vor- und Nachteile des argumentierenden bzw. informierenden Schreibens über mehrere Texte ins Gewicht.

4 Fazit: Warum sich mit den „Enden“ des gegenwärtigen Kompetenzspektrums befassen?

Wir haben in diesem Beitrag das Konstrukt Lesekompetenz behandelt und es aus produkt- und prozessbezogener Perspektive betrachtet, um daraufhin die fünf empirischen bzw. empiriebasierten Beiträge zu verorten. Diese Verortung erfolgte auf einem Kontinuum der Lesekompetenz, das sich von dem flüssigen Lesen (bzw. dessen Vorläufern) hin zum Leseverstehen und Nutzen multipler

Texte erstreckte (s. dazu die Synopse im Teilkapitel 2.2.4). Dieses Kontinuum ist zugegeben sehr weit, allerdings nicht neu (Perfetti, 1997). Es eröffnet den Blick auf Aufgaben und Felder der Lesedidaktik und der Leseforschung. In diesem Band werden insbesondere die Enden dieses Kontinuums, also die Leseflüssigkeit auf der einen und das Verstehen und Verarbeiten multipler Texte auf der anderen Seite, in den Blick genommen. Wir sammeln abschließend einige Argumente, warum wir die Beschäftigung mit diesen Kontinuumsenden für gewinnbringend halten.

4.1 Leseflüssigkeit

Wie oben beschrieben hat sich die Forschung zur Leseflüssigkeit im deutschsprachigen und insbesondere im angloamerikanischen Raum etabliert. Trotzdem ist dieses Forschungsfeld noch nicht vollständig erforscht worden. Dies illustrieren die folgenden drei Argumente, die dafür sprechen, sich weiterhin intensiv mit dieser Teilkompetenz zu befassen.

4.1.1 Der Leseflüssigkeit kommt als Voraussetzung für das Leseverstehen eine zentrale Bedeutung zu

Es ist evident, dass das schnelle, akkurate und automatisierte Lesen geringere kognitive Ressourcen benötigt und in der Folge kognitive Kapazitäten für höhere Verstehensleistungen frei werden. So belegen bereits 1974 die Arbeiten von LaBerge und Samuels die Relevanz der Automatisierung für das Arbeitsgedächtnis. Dass die Leseflüssigkeit als Prädiktor für das Leseverstehen gilt, zeigen z. B. die Arbeiten von Klicpera und Kollegen (1993). Langsame disfluente Leserinnen und Leser lesen stockend, konzentrieren sich auf das Dekodieren und können dem Gelesenen keinen Sinn entnehmen. Das Lernen in allen Fächern wird erschwert. Zudem ist die Schrift in unserer hoch literalen Gesellschaft omnipräsent. Das Beherrschen des Lesens ist deshalb unverzichtbare Voraussetzung für die Teilhabe an Bildung, Berufswelt und Gesellschaft. Das Lesen gilt daher zurecht als Schlüsselqualifikation.

4.1.2 Die Forschung zur Leseflüssigkeit weist noch blinde Flecken auf

Es gibt eine breite Studienlage zur Wirksamkeit von Förderprogrammen zur Steigerung der Leseflüssigkeit und ihrer Auswirkung auf das Textverstehen (z. B. Therrien, 2004; Chard et al, 2002). Eine differenzierte Forschung, die die moderierenden Faktoren beim Einsatz von Förderprogrammen in der Schulpraxis zum Gegenstand haben, muss jedoch noch geleistet werden (s. Kap. 3 in diesem Band). Neben der Frage nach Konzepttreue, Dosierung und Passung der Förderprogramme wissen wir z. B. noch wenig über mögliche Auswirkungen der Tandemkonstellationen bei kooperativen Lesesettings. Auch die einzel-

nen Teilaspekte des mehrdimensionalen Konstrukts Leseflüssigkeit finden unterschiedlich stark Beachtung. So findet das prosodische Lesen erst in jüngerer Zeit Eingang in die Leseforschung (Wolters et al., in Druck, Sappok et al., 2020).

4.1.3 Die Förderung der Leseflüssigkeit ist noch nicht ausreichend in der Schulpraxis etabliert

Neben der Erforschung von Lern- und Lehrgegenständen ist die Frage des Transfers von zentraler Bedeutung. Die Forschung kann und sollte einen Beitrag zu einer evidenzbasierten Bildungspolitik und -verwaltung sowie Bildungspraxis leisten und Impulse geben, die zu verbesserten (Lese-)Leistungen von Schülerinnen und Schülern führen sollen. So etabliert wie die Forschung zur Förderung der Leseflüssigkeit ist, so unbefriedigend ist bislang der Transfer evidenzbasierter Fördermethoden in die schulische Praxis (Philipp, 2014).

4.2 Lesestrategien (bei multiplen Texten)

Die Erforschung von Lesestrategien ist keineswegs neu, sondern im Gegenteil gibt es dazu eine reiche und reichhaltige Tradition (s. dazu aktuell z. B. Dinsmore et al., 2020, und Philipp, 2015). Wenn man sich spezifisch dem Lesen multipler (vor allem digitaler) Texte als einem historisch jungen Feld der Leseforschung zuwendet (Bräten et al., 2020; List, 2020), dann beginnt sich ein Forschungsfeld zu konturieren, welches vermutlich ein hohes Energetisierungspotenzial hat, indem es nach Grundlagen- und Interventionsforschung gleichermaßen verlangt, um neue Herausforderungen zu erkennen und proaktiv lesedidaktisch anzugehen. Drei Argumente sprechen dafür.

4.2.1 Die Prozesse des Lesens werden vielfältiger und komplexer

Wer mehr als einen Text liest und die Inhalte aufeinander beziehen will, benötigt dafür noch Prozesse auf der Ebene des einzelnen Textes, darunter jene hierarchieniedrigen des flüssigen Lesens. Aber es kommen Anforderungen hinzu (Philipp, 2018, 2020). Darunter fallen das Erkennen intertextueller Relationen, z. B. dass Textinhalte einander ergänzen oder widersprechen, dass ein Gegenstand als solcher divergierend dargestellt wird – oder sogar unvollständig («intertextuelles Integrieren»; List & Alexander, 2019). Weiterhin müssen Leserinnen und Leser darauf achten, dass Darstellungen tendenziös oder sogar sachlich unrichtig sein können, sie müssen also mit dem eigenen Weltwissen einen Abgleich vornehmen, um Falschaussagen zu detektieren und zurückzuweisen («epistemisches Validieren»; Richter & Maier, 2017). Insbesondere wenn das eigene Vorwissen nicht ausreicht, benötigen Leserinnen und Leser die «Sourcing» genannte Fähigkeit, die Metadaten – also Informationen zur Auto-

rin bzw. zum Autor oder zum Erscheinungskontext wie offiziellen Medien mit einer gewissen Reputation bzw. einem fragwürdigen Status – zum Evaluieren zu nutzen (Braasch & Scharrer, 2020). Dieses intertextuelle Integrieren und das Sourcing – und das sich dazwischen einfügende epistemische Validieren – sind Konstituenten des angemessenen Leseverstehens multipler Texte (Philipp, 2020) und bedürfen zukünftig erhöhter Aufmerksamkeit.

4.2.2 *Der Aufgabenbezug wird prominenter*

Ein weiterer Strang der Forschung, der sich deutlicher abzuzeichnen beginnt, ist, dass nicht mehr nur das Leseverstehen einzelner rein schriftsprachbasierter Texte beforscht wird. Vielmehr wird ein instrumentelles Lesen zunehmend akzentuiert, bei dem eine Aufgabenstellung und deren kognitive Repräsentation maßgeblich Prozesse und Produkte beeinflussen (Rouet et al., 2017). Der entscheidende Punkt ist hierbei, dass Zielvorstellungen zum Aufgabenprodukt früh und prägend auf die Vorgehensweisen der Leserinnen und Leser einwirken (List & Alexander, 2019). Solche Überlegungen sind stark gespeist aus der Forschung zu Relevanzeffekten, also zu der Forschung, die sich der Frage widmet, wie die Zuweisung von Eignung einzelner Informationen, Informationseinheiten und ganzer Texte für die Zielerreichung – ihrer Relevanz – günstige und ungünstige Auswirkungen hat (McCrudden & Schraw, 2007, s. einen Überblick bei Philipp, i. Vorb.). Solche Relevanzprozesse zu kennen, ist lesedidaktisch hochbedeutsam, weil gerade das Finden und Lesen multipler Texte im Internet eigene Anforderungen mit sich bringt, die ihrerseits mehr Selbstregulation erfordern (Wylie et al., 2018). Hierbei ist Relevanzzuweisung einerseits eine Entlastung, weil sie dabei hilft, die eigenen Ressourcen zielführend einzusetzen (McCrudden & Schraw, 2007), doch sie ist andererseits nicht voraussetzungslos, sondern im Gegenteil stark mit dem Vorwissen assoziiert – und mit dem Wissen, wie man Strategien sinnvoll für die Erreichung von Zielen nutzt, die unterschiedliche Verarbeitungstiefen und Transformationsgrade erfordern (Hattie & Donoghue, 2016). Dennoch klafft hier eine große Forschungslücke (McCrudden et al., 2011b).

4.2.3 *Lesen und Schreiben verschmelzen*

Mit dem Aufkommen des materialgestützten Schreibens als (wissenschaftspropädeutischem) Prüfungsformat ist die Hybridität von Lesen und Schreiben bei der Bearbeitung komplexer Aufgaben auch im deutschdidaktischen Diskurs – genauer: in der Schnittmenge des vor allem argumentativen Lesens und Schreibens – angelangt (Schüler, 2018). Ein besonderes Merkmal des materialgestützten Schreibens – also des Schreibens eines eigenen Sachtexts auf der Basis mehrerer zu nutzender Texte (bzw. «Materialien») – ist sein hoher kognitiver Anforderungsgrad: Personen müssen Informationen lesend und schreibend aus-

wählen, strukturieren und vorwissensbasiert im Sinne des Inferierens verknüpfen (Spivey, 1990). Für eine möglichst optimale Leistung ist die Hybridität nicht nur aus einer grundsätzlichen Logik des kombinierten (aufgabenbezogenen) Lesens und Schreibens entscheidend, sondern auch aus prozessualer Sicht: Lesen und Schreiben gehen ineinander über, es kommt zu einem teils schnellen Wechsel von Lese- und Schreibaktivitäten innerhalb eines komplexen Problemlöseprozesses, und es ist sogar für einen gelingenden eigenen Text notwendig, dass Personen selbstreguliert diese Hybridität im Prozess realisieren (Vandermeulen et al., 2020).

In den schlaglichtartigen Argumenten, warum es sich lohnt, sich weiterhin mit der Leseflüssigkeit sowie dem Lesen multipler Texte zu befassen, wird deutlich: Die Leseforschung hat längst nicht ausgedient, sondern im Gegenteil noch vieles zu leisten.

Literaturverzeichnis

- Afflerbach, P., Pearson, P. D. & Paris, S. G. (2008): Clarifying Differences between Reading Skills and Reading Strategies. *The Reading Teacher*, 61(5), 364–373. <https://doi.org/10.1598/RT.61.5.1>
- Ainsworth, S. E. (2018): Multi-Modal, Multi-Source Reading: A Multi-Representational Reader's Perspective. *Learning and Instruction*, 57, 71–75. <https://doi.org/10.1016/j.learninstruc.2018.01.014>
- Akyol, H. & Ulusoy, M. (2010): Pre-Service Teachers' Use of Reading Strategies in Their Own Readings and Future Classrooms. *Teaching and Teacher Education*, 26(4), 878–884. <https://doi.org/10.1016/j.tate.2009.10.026>
- Alexander, P. A. (2005): The Path to Competence: A Lifespan Developmental Perspective on Reading. *Journal of Literacy Research*, 37 (4), 413–436. https://doi.org/10.1207/s15548430jlr3704_1
- Berliner, D. & Casanova, U. (1988): How Do We Balance Test Anxiety and Achievement? *Instructor*, 97(8), 14–15.
- Braasch, J. L. G. & Scharrer, L. (2020): The Role of Cognitive Conflict in Understanding and Learning from Multiple Perspectives. In P. N. van Meter, A. List, D. Lombardi & P. Kendeou (Eds.), *Handbook of Learning from Multiple Representations and Perspectives* (pp. 205–222). New York: Routledge.
- Brañes, I., Braasch, J. L. G. & Salmerón, L. (2020): Reading Multiple and Non-Traditional Texts: New Opportunities and New Challenges. In E. B. Moje, P. Afflerbach, P. Enciso & N. K. Lesaux (Eds.), *Handbook of Reading Research: Volume V* (pp. 79–98). New York: Routledge.
- Britt, M. A., Rouet, J. F. & Durik, A. M. (2018): *Literacy beyond Text Comprehension: A Theory of Purposeful Reading*. New York: Routledge.
- Brown, P. & Hirst, S. B. (1983): *Writing Reading Courses: The Interrelationship of Theory and Practice*. *Language Teaching Projects for the Third World*. ELT Documents, 116. British Council English Teaching Information Centre.

- Bruin, A. B. H., Roelle, J., Carpenter, S. K., Baars, M. & EFG-MRE (2020): Synthesizing Cognitive Load and Self-Regulation Theory: A Theoretical Framework and Research Agenda. *Educational Psychology Review*, 32(4), 903–915. <https://doi.org/10.1007/s10648-020-09576-4>
- Chard, D. J., Vaughn, S. & Tyler, B. J. (2002): A Synthesis of Research on Effective Interventions for Building Reading Fluency with Elementary Students with Learning Disabilities. *Journal of Learning Disabilities*, 35(5), 386–406. <https://doi.org/10.1177/00222194020350050101>
- Cho, B. Y. & Afflerbach, P. (2017): An Evolving Perspective of Constructively Responsive Reading Comprehension Strategies in Multilayered Digital Text Environments. In S. E. Israel (Ed.), *Handbook of Research on Reading Comprehension* (2nd ed., pp. 109–134). New York: Routledge.
- Dinsmore, D. L., Fryer, L. K. & Parkinson, M. M. (Eds.) (2020): *Handbook of Strategies and Strategic Processing*. New York: Routledge.
- Duffy, G. G. (1993): Rethinking Strategy Instruction: Four Teachers' Development and Their Low Achievers' Understandings. *The Elementary School Journal*, 93(3), 231–247. <https://doi.org/10.1086/461724>
- Elleman, A. M., Lindo, E. J., Morphy, P. & Compton, D. L. (2009): The Impact of Vocabulary Instruction on Passage-Level Comprehension of School-Age Children: A Meta-Analysis. *Journal of Research on Educational Effectiveness*, 2(1), 1–44. <https://dx.doi.org/10.1080/19345740802539200>
- Friedrich, H. F. & Mandl, H. (2006): Lernstrategien: Zur Strukturierung des Forschungsfeldes. In H. Mandl & H. F. Friedrich (Hrsg.), *Handbuch Lernstrategien* (S. 1–23). Göttingen: Hogrefe.
- García, J. R. & Cain, K. (2014): Decoding and Reading Comprehension: A Meta-Analysis to Identify Which Reader and Assessment Characteristics Influence the Strength of the Relationship in English. *Review of Educational Research*, 84(1), 74–111. <https://doi.org/10.3102/0034654313499616>
- Hasselhorn, M. & Gold, A. (2017): *Pädagogische Psychologie: Erfolgreiches Lernen und Lehren*. Stuttgart: Kohlhammer.
- Hattie, J. A. C. & Donoghue, G. M. (2016): Learning Strategies: A Synthesis and Conceptual Model. *NPJ Science of Learning*, 1(1), 1–13. <https://doi.org/10.1038/npjscilearn.2016.13>
- Hußmann, A., Wendt, H., Bos, W., Bremerich-Vos, A., Kasper, D., Lankes, E. M., Mc Elvany, N., Stubbe, T. C. & Valtin, R. (2017): IGLU 2016. *Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster: Waxmann.
- Israel, S. E. (2015): *Verbal Protocols in Literacy Research: Nature of Global Reading Development*. New York: Routledge.
- Kendeou, P., McMaster, K. L. & Christ, T. J. (2016): Reading Comprehension: Core Components and Processes. *Policy Insights from the Behavioral and Brain Sciences*, 3(1), 62–69. <https://doi.org/10.1177/2372732215624707>
- Klicpera, C., Gasteiger-Klicpera, B. & Schabmann, A. (1993): *Lesen und Schreiben – Entwicklung und Schwierigkeiten: Die Wiener Längsschnittuntersuchungen über die Entwicklung, den Verlauf und die Ursachen von Lese- und Schreibschwierigkeiten in der Pflichtschulzeit*. Bern: Huber.
- KMK Kultusministerkonferenz (2004): *Bildungsstandards im Fach Deutsch für den Primarbereich*. München: Luchterhand.

- LaBerge, D. & Samuels, S. J. (1974): Toward a Theory of Automatic Information Processing in Reading. *Cognitive Psychology*, 6(2), 293–323. [https://doi.org/10.1016/0010-0285\(74\)90015-2](https://doi.org/10.1016/0010-0285(74)90015-2)
- Lenhard, W. & Schneider, W. (2006): *ELFE 1–6. Ein Leseverständnistest für Erst- bis Sechstklässler*. Göttingen: Hogrefe.
- Lenhard, W., Lenhard, A. & Schneider, W. (2017): *ELFE II – Ein Leseverständnistest für Erst- bis Siebtklässler*. Göttingen: Hogrefe.
- Lenhard, W. (2019): *Leseverständnis und Lesekompetenz: Grundlagen – Diagnostik – Förderung* (2., aktual. Auflage). Stuttgart: Kohlhammer.
- Leu, D. J., Kinzer, C. K., Coiro, J., Castek, J. & Henry, L. A. (2017): New Literacies: A Dual-Level Theory of the Changing Nature of Literacy, Instruction, and Assessment. *Journal of Education*, 197(2), 1–18. <https://doi.org/10.1177/002205741719700202>
- List, A. (2020): Six Questions Regarding Strategy Use When Learning from Multiple Texts. In D. L. Dinsmore, L. K. Fryer & M. M. Parkinson (Eds.), *Handbook of Strategies and Strategic Processing* (pp. 119–140). New York: Routledge.
- List, A. & Alexander, P. A. (2019): Toward an Integrated Framework of Multiple Text Use. *Educational Psychologist*, 54(1), 20–39. <https://doi.org/10.1080/00461520.2018.1505514>
- List, A., Du, H. & Lee, H. Y. (in Druck): How Do Students Integrate Multiple Texts? An Investigation of Top-Down Processing. *European Journal of Psychology of Education*. <https://doi.org/10.1007/s10212-020-00497-y>
- McCrudden, M. T. (2018): Text Relevance and Multiple Source Use. In J. L. G. Braasch, I. Bräten & M. T. McCrudden (Eds.), *Handbook of Multiple Source Use*, 168–183. New York: Routledge.
- McCrudden, M. T., Magliano, J. P. & Schraw, G. (2011a): Relevance in Text Comprehension. In M. T. McCrudden, J. P. Magliano & G. J. Schraw (Eds.), *Text Relevance and Learning from Text* (pp. 1–17). Charlotte: Information Age.
- McCrudden, M. T., Magliano, J. P. & Schraw, G. (2011b): Toward an Integrated View of Relevance in Text Comprehension. In M. T. McCrudden, J. P. Magliano & G. J. Schraw (Eds.), *Text Relevance and Learning from Text* (pp. 395–414). Charlotte: Information Age.
- McCrudden, M. T. & Schraw, G. (2007): Relevance and Goal-Focusing in Text Processing. *Educational Psychology Review*, 19(2), 113–139. <https://doi.org/10.1007/s10648-006-9010-7>
- National Reading Panel (US), National Institute of Child Health & Human Development (US) (2000): *Teaching Children to Read: An Evidence-Based Assessment of the Scientific Research Literature on Reading and its Implications for Reading Instruction: Reports of the Subgroups*. Washington: National Institute of Child Health and Human Development, National Institutes of Health.
- OECD (2019): *PISA 2018 Ergebnisse (Band I): Was Schülerinnen und Schüler wissen und können*. Paris: OECD.
- Okkinga, M., van Steensel, R., van Gelderen, A., van Schooten, E., Slegers, P. J. C. & Arends, L. R. (2018): Effectiveness of Reading-Strategy Interventions in Whole Classrooms: A Meta-Analysis. *Educational Psychology Review*, 30(4), 1215–1239. <https://doi.org/10.1007/s10648-018-9445-7>
- Oudega, M. & van den Broek, P. W. (2018): Standards of Coherence in Reading: Variations in Processing and Comprehension of Text. In K. Millis, D. L. Long, J. P. Magliano & K. Wiemer (Eds.), *Deep Comprehension: Multi-Disciplinary Approaches to Understanding, Enhancing, and Measuring Comprehension* (pp. 41–51). New York: Routledge.

- Perfetti, C. A. (1997): Sentences, Individual Differences, and Multiple Texts: Three Issues in Text Comprehension. *Discourse Processes*, 23(3), 337–355. <https://doi.org/10.1080/01638539709544996>
- Pinnell, G. S., Pikulski, J. J., Wixson, K. K., Campbell, J. R., Gough, P. B. & Beatty, A. S. (1995): *Listening to Children Read Aloud*. US Department of Education. Washington, DC: National Center for Education Statistics.
- Philipp, M. (2012): *Das vernachlässigte Füllhorn der Sprache: Einige Betrachtungen zum Zusammenhang von Wortschatz, Lesesozialisation und Textverstehen*. In: leseforum.ch. Online-Plattform für Literalität. www.forumlecture.ch – 1/2012. (Abruf am 30.06.2021).
- Philipp, M. (2013): *Lese- und Schreibunterricht*. Tübingen: Francke.
- Philipp, M. (2014): *Leseunterricht in der Grundschule – vom Ist-Zustand und vom Soll-Zustand. Was Beobachtungsstudien lehren*. In R. Valtin & I. Tarelli (Hrsg.), *Lesekompetenz nachhaltig stärken: Evidenzbasierte Maßnahmen und Programme* (S. 122–166). Berlin: Deutsche Gesellschaft für Lesen und Schreiben.
- Philipp, M. (2015): *Lesestrategien: Bedeutung, Formen und Vermittlung*. Weinheim: Beltz Juventa.
- Philipp, M. (2018): *Multiple Modelle des Leseverstehens multipler Texte: Eine Synopse aktueller kognitiver Modellierungen aus lesedidaktischer Perspektive*. In: leseforum.ch. Online-Plattform für Literalität. www.forumlecture.ch – 3/2018. (Abruf am 30.06.2021)
- Philipp, M. (2020): *Multiple Dokumente verstehen: Theoretische und empirische Perspektiven auf Prozesse und Produkte des Lesens mehrerer Dokumente*. Weinheim: Beltz Juventa.
- Philipp, M. (i. Vorb.): *Lesen digital: Komponenten und Prozesse einer sich wandelnden Kompetenz*. Weinheim: Beltz Juventa.
- Pressley, M. & Afflerbach, P. (1995): *Verbal Protocols of Reading: The Nature of Constructively Responsive Reading*. Hillsdale: Lawrence Erlbaum.
- Primor, L. & Katzir, T. (2018): Measuring Multiple Text Integration: A Review. *Frontiers in Psychology*, 9, Article 2294, 1–16. <https://doi.org/10.3389/fpsyg.2018.02294>
- Richter, T. & Maier, J. (2017): Comprehension of Multiple Documents with Conflicting Information: A Two-Step Model of Validation. *Educational Psychologist*, 52(3), 148–166. <https://doi.org/10.1080/00461520.2017.1322968>
- Richter, T. & Schnotz, W. (2018): Textverstehen. In D. H. Rost, J. R. Sparfeldt & S. R. Buch (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (5. Aufl., S. 826–837). Weinheim: Beltz Juventa.
- Rosebrock, C. & Nix, D. (2006): Forschungsüberblick: Leseflüssigkeit (Fluency) in der amerikanischen Leseforschung und -didaktik. *Didaktik Deutsch*, 20, 90–112.
- Rosebrock, C. (2012): *Was ist Lesekompetenz, und wie kann sie gefördert werden?* In: leseforum.ch. Online-Plattform für Literalität. www.forumlecture.ch – 3/2012. (Abruf am 11.06.21)
- Rouet, J. F., Britt, M. A. & Durik, A. M. (2017): RESOLV: Readers' Representation of Reading Contexts and Tasks. *Educational Psychologist*, 52(3), 200–215. <https://doi.org/10.1080/00461520.2017.1329015>
- Samuels, S. J., LaBerge, D. & Bremer, C. D. (1978): Units of Word Recognition: Evidence for Developmental Changes. *Journal of Verbal Learning and Verbal Behavior*, 17(6), 715–720. [https://doi.org/10.1016/S0022-5371\(78\)90433-4](https://doi.org/10.1016/S0022-5371(78)90433-4)

- Sappok, C., Linnemann, M. & Stephany, S. (2020): Leseflüssigkeit – Prosodie – Leseverstehen. Eine Longitudinalstudie zur Entwicklung der Leseflüssigkeit von Jahrgangsstufe 3 bis 7. In I. Rautenberg (Hrsg.), *Evidenzbasierte Forschung zum Schriftspracherwerb* (S. 175–209). Baltmannsweiler: Schneider Verlag Hohengehren.
- Schmitz, A. (2019): Reading Instruction in 5th Grade: Teachers' Perspectives on Promoting Self-Regulated Reading in Language and Content Area Teaching. *RISTAL: Research in Subject-Matter Teaching and Learning*, 2, 16–31. <https://doi.org/10.23770/rt1822>
- Schreiber, P. A. (1991): Understanding Prosody's Role in Reading Acquisition. *Theory into Practice*, 30(3), 158–164. <https://doi.org/10.1080/00405849109543496>
- Schüler, L. (2018): Wissenschaftlich argumentieren lernen durch materialgestütztes Schreiben. In S. Schmölzer-Eibinger, B. Bushati, C. Ebner & L. Niederdorfer (Hrsg.), *Wissenschaftliches Schreiben lehren und lernen: Diagnose und Förderung wissenschaftlicher Textkompetenz in Schule und Universität* (S. 147–169). Münster: Waxmann.
- Spivey, N. N. (1990): Transforming Texts: Constructive Processes in Reading and Writing. *Written Communication*, 7(2), 256–287. <https://doi.org/10.1177/0741088390007002004>
- Spivey, N. N. (1997): *The Constructivist Metaphor: Reading, Writing and the Making of Meaning*. San Diego: Academic Press.
- Stanat, P., Böhme, K., Schipolowski, S. & Haag, N. (Hrsg.) (2016): *IQB-Bildungstrend 2015: Sprachliche Kompetenzen am Ende der 9. Jahrgangsstufe im zweiten Ländervergleich*. Münster: Waxmann.
- Stanat, P., Schipolowski, S., Rjosk, C., Weirich, S. & Haag, N. (2017): *IQB-Bildungstrend 2016. Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im zweiten Ländervergleich*. Münster: Waxmann.
- Sweller, J. (2010): Element Interactivity and Intrinsic, Extraneous, and Germane Cognitive Load. *Educational Psychology Review*, 22(2), 123–138. <https://doi.org/10.1007/s10648-010-9128-5>
- Therrien, W. J. (2004): Fluency and Comprehension Gains as a Result of Repeated Reading: A Meta-Analysis. *Remedial and Special Education*, 25(4), 252–261. <https://doi.org/10.1177/07419325040250040801>
- Vandermeulen, N., van den Broek, B., van Steendam, E. & Rijlaarsdam, G. (2020): In Search of an Effective Source Use Pattern for Writing Argumentative and Informative Synthesis Texts. *Reading and Writing*, 33(2), 239–266. <https://doi.org/10.1007/s11445-019-09958-3>
- Weinstein, C. E. & Mayer, R. E. (1986): The Teaching of Learning Strategies. In M. C. Wittrock (Ed.), *Handbook of Research on Teaching* (3rd ed., pp. 315–327). New York: Macmillan.
- Wolters, A. P., Kim, Y. S. G. & Szura, J. W. (in Druck): Is Reading Prosody Related to Reading Comprehension? A Meta-Analysis. *Scientific Studies of Reading*. <https://doi.org/10.1080/10888438.2020.1850733>
- Wylie, J., Thomson, J., Leppanen, P. H. T., Ackerman, R., Kannianen, L. & Prieler, T. (2018): Cognitive Processes and Digital Reading. In M. Barzillai, J. Thomson, S. Schroeder & P. van den Broek (Eds.), *Learning to Read in a Digital World* (pp. 57–90). Amsterdam: John Benjamins.

Methoden zur differenzierten Auswertung von Wortschatzkompetenzen und ihre Bedeutung für das frühe Lesen

Britta Juska-Bacher, Christoph Zangger & Martina Röhlisberger

Zusammenfassung: In der vorliegenden Studie wurden die Zusammenhänge zwischen Leseverständnis und Wortschatzkompetenzen sowie deren Entwicklung bei Schulbeginn untersucht. In einem längsschnittlichen Design wurden dieselben Grundschul Kinder ($N = 313$) in der ersten Klasse (t1) und in der zweiten Klasse (t2) auf verstehendes Lesen von Wort, Satz und Text, die drei Wortschatzvariablen Wortschatzumfang, relationales und semantisches Wortwissen sowie die Vorläuferfertigkeiten phonologische Bewusstheit, schnelles Benennen und Rekodiergeschwindigkeit getestet. Das Cross-Lagged-Modell, mit dem die Daten längsschnittlich analysiert wurden, zeigte nach Korrektur der autoregressiven Effekte keinen signifikanten Einfluss des Wortschatzes zu t1 auf das Leseverständnis zu t2. Eine vertiefte Analyse der Beziehung zwischen Wortschatzvariablen und verstehendem Lesen offenbarte indes ein spezifisches Muster: Die Korrelationen zwischen Leseverständnis und Wortschatzvariablen waren zu t2 deutlich höher als zu t1. Zudem zeigten Varianzanalysen eine deutlich stärkere Entwicklung des Leseverständnisses im Vergleich zum Wortschatz. Im Fixed-Effects-Modell schließlich manifestierten sich diese Ergebnisse in Form eines hochsignifikanten Effekts der Wortschatzentwicklung auf das verstehende Lesen. Die komplexe Natur der Beziehung zwischen Wortschatz und Lesen erfordert ganz offensichtlich eine differenzierte Betrachtung der Wechselwirkungen.

Schlüsselwörter: Leseerwerb, Wortschatzentwicklung, Wortschatztests, statistische Analyse

Abstract: The present study investigated the relationships between early reading comprehension and vocabulary skills and their development. In a longitudinal design, the same primary school children ($N = 313$) were tested in first grade (t1) and in second grade (t2) for reading of words, sentences and texts, the three vocabulary variables vocabulary size, network and semantic word knowledge as well as the precursors phonological awareness, rapid naming and decoding speed. The cross-lagged model, with which the data were analysed longitudinally, showed no significant influence of

vocabulary at t1 on reading at t2 after correction for autoregressive effects. However, a more in-depth analysis of the relationship between vocabulary variables and reading revealed a specific pattern: the correlations between reading and vocabulary variables were significantly higher at t2 than at t1. In addition, variance analyses showed a significantly stronger development of reading compared to vocabulary. Finally, in a fixed-effects model, these results manifested themselves in the form of a highly significant effect of vocabulary development on comprehension reading. The complex nature of the relationship between vocabulary and reading thus clearly requires a differentiated view of the interactions.

Keywords: Reading acquisition, vocabulary development, vocabulary tests, statistical analysis

1 Einleitung

Lesen beinhaltet die Fähigkeit, geschriebene Symbole arbiträr Sprachlauten zuordnen zu können und diese zu größeren Einheiten zu synthetisieren. Die Lautfolgen als sprachliche Form werden wiederum arbiträren Bedeutungen zugeordnet, mittels derer sich der Inhalt der geschriebenen Sprache erschließt. In diesem Sinne ist Schriftsprache doppelt arbiträr, und Lesen muss als komplexes Konstrukt, das aus mehreren Teilprozessen besteht, betrachtet werden. Während die Teilprozesse beim erwachsenen, geübten Leser kaum mehr isoliert betrachtet werden können, spielen die Entschlüsselung der Form oder das Rekodieren und die Entschlüsselung des Inhalts oder das Dekodieren beim frühen Lesen von Kindern klar erkennbare, unterschiedliche Rollen. So wurden in der Leseforschung verschiedene Faktoren nachgewiesen, die das frühe Lesen entscheidend beeinflussen. Dabei dominieren auf dieser Entwicklungsstufe Studien, die auf Vorläuferfertigkeiten wie die phonologische Bewusstheit, das schnelle Benennen und die Rekodiergeschwindigkeit fokussieren, also die formale Seite des Lesens untersuchen. Der Wortschatz – der Abruf und Zuordnung der Wortbedeutung, also die inhaltliche Seite maßgeblich beeinflusst – hat beim beginnenden Lesen besonders in der deutschen Leseforschung bisher wenig Beachtung erhalten. Dies ist eine empfindliche Lücke, denn etwa ab dem dritten Schuljahr, wenn die Anforderungen an das Leseverständnis anspruchsvoller werden, wird Wortschatz ein immer wichtigerer Prädiktor für den Leseerfolg (z. B. Biemiller, 2012). Es drängt sich die Frage auf, welche Rolle Wortwissen beim Lesen in den ersten beiden Schuljahren spielt, wenn ein beträchtlicher Teil kognitiver Kapazitäten noch vom Re- und Dekodierungsprozess absorbiert ist und die Anforderungen an das Leseverstehen auf einer sehr basalen Ebene bleiben. Eben dieser Frage ist das Projekt „Die Entwicklung von Wortschatz und Lesen. Eine Untersuchung auf der Unterstufe“ (2017–2021, kurz:

EnWoLe, <http://p3.snf.ch/Project-173245>) gewidmet. In dieser Längsschnittstudie werden neben der Rekodierfähigkeit und den Vorläuferfertigkeiten verschiedene Wortschatzaspekte von über 350 Kindern vom 1. bis 3. Schuljahr untersucht, um das Zusammenwirken dieser Kompetenzbereiche und ihren Einfluss auf das Lesen zu ergründen. Der folgende Beitrag stellt die Auswertung der Studiendaten aus dem 1. und 2. Schuljahr vor.

Für die hier angestrebte differenzierte Auswertung des Zusammenhangs von Lesen und Wortschatz ist einerseits zentral, dass im Bereich des Wortschatzes nicht nur – wie dies häufig der Fall ist – exemplarisch der Umfang bekannter Wörter erfasst wird, sondern darüber hinaus auch die Tiefe semantischen Wissens sowie die Verbindungen zwischen Einträgen im mentalen Lexikon. Ziel dieses Vorgehens ist es, Hinweise darauf zu finden, welche Wortschatzkompetenzen bzw. welche Eigenschaften des mentalen Lexikons sich positiv auf das frühe Lesen auswirken. Um der Komplexität des Zusammenspiels der Kompetenzbereiche gerecht zu werden, werden andererseits neben dem Wortschatz die erwiesenermaßen gewichtigen Einflussvariablen phonologische Bewusstheit und schnelles Benennen sowie die Rekodiergeschwindigkeit als Indikator für die Leseflüssigkeit berücksichtigt.

Diese Komplexität bedingt überdies ein mehrschrittiges statistisches Vorgehen, das verdeutlicht, wie erst durch eine Methodenkombination quer- und längsschnittliche Zusammenhänge zwischen Lesen und Wortschatz verlässlich sichtbar gemacht werden können. In einem ersten Schritt zeigen die Ergebnisse von Korrelationen für die erste bzw. zweite Klasse zeitgleiche Zusammenhänge zwischen Lese-, Wortschatz- und Einflussvariablen auf. In einem zweiten Schritt wird mittels eines Cross-Lagged-Designs eine längsschnittliche Perspektive eingenommen. Hierbei werden alle Konstrukte zum zweiten auf alle Konstrukte zum ersten Erhebungszeitpunkt regressiert, was eine Einschätzung der Einflussrichtung unter Berücksichtigung autoregressiver Effekte erlaubt. Da diese Modelle jedoch auf starken Verteilungsannahmen beruhen und nicht ausgeschlossen werden kann, dass unbeobachtete Merkmale die Zusammenhänge beeinflussen, wird in einem dritten Schritt ein Fixed-Effects Panel-Modell gerechnet. Dieses erlaubt die kausale Rückführung der Entwicklung der Lesefähigkeiten auf die Veränderungen im Wortschatz, dem schnellen Benennen und die phonologische Bewusstheit.

2 Wortschatzkompetenzen und ihre Erhebung auf der Zielstufe

Das mentale Lexikon, d. h. derjenige Teil des Langzeitgedächtnisses, in dem das Wortwissen gespeichert ist, lässt sich einerseits durch die Zahl der gespeicherten Worteinträge, d. h. die Wortschatzbreite (vocabulary breadth; entspricht dem Wortschatzumfang), andererseits durch die Qualität des Wortwissens,

d. h. die Wortschatztiefe (vocabulary depth) charakterisieren (z. B. Anderson und Freebody, 1981; Ouellette, 2006; Tannenbaum et al., 2006; Vermeer, 2001).

Definition und Bedeutung des Wortschatzumfangs sind unumstritten. Ab dem Schulalter wird in der Regel ein Wortschatzauszug getestet, für den gemessen wird, ob die Zielwörter bekannt sind (dichotom bekannt vs. unbekannt). Für die Messung des rezeptiven Wortschatzes im Deutschen liegt für das Schuleingangsalter bspw. der aus dem Englischen übersetzte *Peabody Picture Vocabulary Test* für 3- bis 16-Jährige von Lenhard et al. (2015) vor.

Wortschatztiefe hingegen wird in der Literatur unterschiedlich ausgelegt (z. B. Cain & Oakhill, 2014; Read, 2004). Read (2004, S. 211–212) führt drei Möglichkeiten an: erstens die Präzision semantischen Wissens (precision of meaning), die Auskunft darüber gibt, wie differenziert semantische Repräsentationen zu Wörtern im mentalen Lexikon sind. Gemessen wird hier eine graduelle Abstufung, die von einer vagen Idee bis zur differenzierten Kenntnis der Bedeutungen verläuft (im Folgenden: „semantisches Wortwissen“). Weniger spezifisch ist die zweite Auslegung von Wortschatztiefe, nämlich das sog. umfassende Wortwissen (comprehensive word knowledge), das ebenfalls semantisches Wissen, darüber hinaus aber auch phonologisches, orthografisches und grammatisches Wissen zu Wörtern beinhaltet. Die dritte Auslegung schließlich umfasst Wissen über Verbindungen eines Wortes zu anderen Einträgen im mentalen Lexikon (network knowledge), z. B. zu Synonymen, Antonymen und Hyperonymen (Read, 2004), das im Folgenden als „relationales Wortwissen“ (nach Juska-Bacher & Jakob, 2014) bezeichnet wird. Im hier beschriebenen Projekt wird im Bereich der Wortschatztiefe semantisches und relationales Wortwissen betrachtet, formales Wortwissen hingegen bleibt – da auf der Schuleingangsstufe im Vergleich zu den späteren Primarschuljahren noch von geringerer Relevanz (zu grammatischem Wissen siehe Clahsen & Fleischauer, 2014 und Deacon & Kirby, 2004; orthografisches Wissen wird auf dieser Stufe gemäß aktuellem Lehrplan¹ in der Schweiz in den beiden ersten Klassen kaum vermittelt) – unberücksichtigt.

Während mit dem Wortschatz- und Wortfindungstest WWT 6-10 (Glück, 2011) im Deutschen ein valides Instrument für die Messung relationalen Wortwissens zur Verfügung steht, fehlt ein solches für das semantische Wortwissen. Bisher besteht nicht einmal Einigkeit darüber, wie dieses am besten zu erheben ist (Cain & Oakhill, 2014). Häufig wird es in Form von Definitionsaufgaben gemessen (weitere Möglichkeiten siehe Read 2004, S. 212–216), die anhand von Skalen bewertet werden (bspw. Ouellette, 2006 mit 0–3 Punkten).

1 Erziehungsdirektion des Kantons Bern (2016). Lehrplan 21. Kompetenzaufbau Deutsch. 1. Zyklus. Verfügbar unter: https://be.lehrplan.ch/lehrplan_printout.php?k=1&z=1&ekalias=0&fb_id=1&f_id=11.

Dabei werden dekontextualisierte, abstrakte Definitionen oft höher bepunktet als Beispiele und Angaben zur Funktion (Vermeer, 2001). Bei diesem Vorgehen besteht die Gefahr, dass nicht nur semantisches Wissen, sondern auch Weltwissen und die Form der Definition und damit die kognitiven Fähigkeiten des Kindes bewertet werden (Kurland & Snow, 1997; Juska-Bacher et al., 2021). Als Option steht im Deutschen wie in anderen Sprachen der Teiltest Wortaufgaben aus dem IQ-Test des *Hamburg-Wechsler-Intelligenztests für Kinder* (Petermann & Petermann, 2007) zur Verfügung, in dem Probanden Wörter mit steigendem Schwierigkeitsgrad definieren sollen. Allerdings ist dieser Test nicht für eine detaillierte Messung semantischen Wortwissens konzipiert, sondern für die Ermittlung eines Teilwertes des IQs der Testpersonen (zur Problematisierung siehe Kapitel Erhebungsinstrumente).

Da in der internationalen Leseforschung immer wieder Korrelationen zwischen Wortwissen und Lesen aufgezeigt wurden, die mit zunehmendem Alter der Probanden höher wurden (vgl. z. B. Tannenbaum et al., 2006), wird der Einbezug des Wortschatzes bereits auf der Primarstufe dringlich gefordert (z. B. Ouellette, 2006). Dabei wird angeregt, seine Bedeutung möglichst differenziert zu untersuchen, d. h. neben dem Wortschatzumfang Kriterien wie semantisches und relationales Wissen einzubeziehen (u. a. Tannenbaum et al., 2006). In dieser Studie wurde für Drittklässler gezeigt, dass Wortschatzumfang und -tiefe deutlich unterscheidbare Dimensionen von Wortschatz sind, auch wenn sie stark korrelieren (Tannenbaum et al., 2006).

3 Weitere für das frühe Lesen relevante Faktoren

Unbestritten in ihrer Relevanz für das frühe Lesen ist eine Reihe weiterer Faktoren, die deutlich besser erforscht ist als der Wortschatz. Dazu zählen die Vorläuferfertigkeiten phonologische Bewusstheit und schnelles Benennen (rapid automatized naming oder RAN) sowie die Rekodiergeschwindigkeit als Indikator für die Leseflüssigkeit. Besonders am Anfang des Leserwerbs spielt die phonologische Bewusstheit eine zentrale Rolle, wenn die Kinder beim rekodierenden Lesen auf dem indirekten, phonologischen Weg die Wörter erlesen (zum Zwei-Wege-Modell siehe Castles & Coltheart, 2004). Ihre Bedeutung wurde für eine Reihe von Sprachen eindeutig nachgewiesen, für das Deutsche z. B. von Ennemoser et al. (2012), Fricke et al. (2016) und Juska-Bacher et al. (2015). Deutlich weniger Beachtung erhielt bisher die Zugriffsgeschwindigkeit auf die Wortbedeutung (das schnelle Benennen), die beim Dekodieren eine wichtige Rolle spielt. Da ein schneller lexikalischer Zugriff beim Lesen eine kognitive Entlastung darstellt, wird in der internationalen Forschung die Berücksichtigung dieses Faktors postuliert (z. B. Schmitt, 2010). Zur Messung des schnellen Benennens erhalten die Probanden die Aufgabe, eine Reihe gleichzei-

tig präsentierter Stimuli (z. B. Farben) möglichst schnell „vorzulesen“. Den wichtigen Einfluss des schnellen Benennens konnten im Deutschen bspw. Studien von Ennemoser et al. (2012), Fricke et al. (2016) und Landerl et al. (2019, sprachvergleichend) nachweisen.

Die Rekodiergeschwindigkeit setzt ebenso wie die phonologische Bewusstheit am Rekodierprozess, an der technischen Seite des Lesens an. Gemessen wird die Flüssigkeit durch das Lesen von Pseudowörtern, bei denen der Zugriff auf eine Wortbedeutung (Dekodieren) ausgespart wird. Der deutliche Einfluss der Rekodiergeschwindigkeit auf das verstehende Lesen (im Deutschen bspw. belegt von Landerl & Wimmer, 2008; Mayer, 2018), verwundert wenig, stellt doch dieser Prozessschritt des Rekodierens besonders am Anfang der Schulzeit einen beträchtlichen Anteil der basalen Lesekompetenzen auf der Wort-, Satz- und Textebene dar. Flüssiges Lesen entlastet den Leseprozess und setzt damit Ressourcen für das Verstehen frei (De Jong & Van der Leij, 2002; Kim et al., 2011).

4 Ausgangshypothesen der Studie EnWoLe

Die Ziele des hier präsentierten Studienausschnitts sind, die Auswirkungen eines gefestigten Wortschatzes als Ganzes sowie der Teilbereiche Wortschatzumfang, relationales und semantisches Wortwissen auf das verstehende Lesen sowohl zeitgleich als auch longitudinal zu untersuchen. Um diesen Einfluss klar von den technischen Lesefähigkeiten abgrenzen zu können, müssen auch die Vorläuferfertigkeiten und die Rekodierfähigkeit im statistischen Modell berücksichtigt werden. Während für Ersteres korrelative Analysen verwendet werden, wird Zweiteres mittels längsschnittlicher Kausalanalysen untersucht. Ein signifikanter Effekt des Wortschatzes aus dem ersten Schuljahr auf das Lesen im zweiten Schuljahr war gemäß den obenstehenden Befunden nicht per se zu erwarten, da zu diesem frühen Zeitpunkt der rekodierende Prozess und dessen Vorläuferfertigkeit, die phonologische Bewusstheit, eine dominante Rolle beim Lesen spielen und die autoregressiven Effekte des Lesens den Wortschatzeinfluss überlagern dürften. Dennoch möchten wir im Folgenden der Frage nachgehen, ob Wortschatzkompetenzen nicht doch einen bedeutsamen Einfluss auch bereits auf das frühe Lesen haben, insbesondere wenn als abhängige Variable verstehendes Lesen und damit der dekodierende Teilprozess erhoben wird. Die daraus abzuleitenden Hypothesen bilden den Ausgangspunkt der Untersuchung:

- 1) Werden die beiden Zeitpunkte isoliert betrachtet, ist der Einfluss des Wortschatzes auf das verstehende Lesen signifikant. (Querschnittuntersuchung)
- 2) Weil die Dekodierfähigkeit gegenüber dem Rekodieren mit steigenden Lesekompetenzen an Bedeutung gewinnt, ist in der zweiten Klasse der Zusammenhang zwischen Wortschatz und verstehendem Lesen größer als in der ersten.
- 3) Es gibt einen Effekt des Wortschatzes im ersten Schuljahr auf das verstehende Lesen im zweiten Schuljahr. (Längsschnittuntersuchung)
- 4) Dieser Effekt wird möglicherweise nicht direkt sichtbar. Der Nachweis eines solchen Effektes bedingt die Anwendung eines geeigneten Datenmodells, das die Zunahme der Korrelationen des Wortschatzes mit dem verstehenden Lesen zulässt.

5 Material und Methoden

5.1 Stichprobe

In der Langzeitstudie EnWoLe wurden im Zeitraum von Mitte März bis Ende Mai 2018 mit 348 Erstklässlern und im selben Zeitraum im Jahr 2019 mit 361 Zweitklässlern verschiedene Tests zu Wortschatz- und Lesekompetenzen sowie Einflussvariablen durchgeführt. Für die hier vorgestellten Analysen wurden ausschließlich Datensätze von Kindern berücksichtigt, von denen vollständige Daten beider Erhebungszeitpunkte vorlagen ($N = 313$ aus 39 Klassen, davon 179 Mädchen und 134 Jungen mit einem Durchschnittsalter von 7;5 Jahren zu t_1 und 8;5 Jahren zu t_2). 27 der Klassen waren aus dem städtischen Bereich, 12 aus dem ländlichen. Alle Kinder sprachen Deutsch als Erstsprache und wiesen keine diagnostizierten Sprachentwicklungsbeeinträchtigungen auf.

5.2 Erhebungsinstrumente

Lesekompetenzen wurden anhand des standardisierten und im Deutschen gängigen Lesetests ELFE II (Lenhard et al., 2018) erhoben, der das Leseverständnis von Einzelwörtern, Sätzen und kurzen Texten prüft.

Im Bereich des Wortschatzumfangs kam der aus dem Englischen übersetzte PPVT-4 (Lenhard et al., 2015) zum Einsatz. Abgesehen von den Herausforderungen, die die Übersetzung eines Testinstruments in sprachlicher wie konzeptioneller Hinsicht mit sich bringt (siehe z. B. Peña, 2007), erforderte die Verwendung dieses Tests in der Schweiz einige Anpassungen. Dies betraf den Ersatz der bundesdeutschen durch Schweizer Varianten, u. a. „Schornstein“ durch „Kamin“ und „Umschlag“ durch „Kuvert“. Aus Zeitgründen wurde mit einer Testhalbierung gearbeitet (nur ungerade Items).

Zur Erfassung des semantischen Wortwissens wurde der Untertest 6 Wortschatztest des HAWIK-IV (Petermann & Petermann, 2007) verwendet. Dieser Test kam bereits in verschiedenen anderen Sprachen zur Messung des semantischen Wortwissens zum Einsatz (z. B. im Englischen bei Cain & Oakhill, 2014, im Niederländischen bei Swart et al., 2017, und im Norwegischen bei Rydland et al., 2012). In EnWoLe wurde auch dieser Test halbiert (Wortaufgaben 7–35, nur die 15 ungeraden Items). Da das dreistufige Bewertungssystem dieses IQ-Tests mit 0 bis 2 Punkten² pro Antwort nicht auf einer linguistisch überzeugenden semantischen Analyse der einzelnen Wörter beruht und nur den Mindestdifferenzierungsgrad für eine Beurteilung von semantischem Wortwissen bietet, wurden in EnWoLe nach einer semantischen Analyse eigene Bewertungskriterien erarbeitet und die Punkteskala auf 0 bis 3 Punkte erhöht (siehe Juska-Bacher et al., in Druck; ebenfalls mit einer vierstufigen Skala hat bspw. Ouellette, 2006 gearbeitet). Die transkribierten Antworten der Kinder wurden in einem ersten Schritt von jeweils zwei unabhängigen Raterinnen beurteilt. Mit Cohens Kappa $\kappa = 0.79$, $SD = 0.1$, Spannweite 0.6 – 1 war die Interrater-Reliabilität zufriedenstellend. Bei Abweichungen in der Bewertung folgte ein zweiter Schritt, in dem eine dritte Raterin die Definitionen unabhängig beurteilte.

Das relationale Wortwissen schließlich wurde mit der Kurzfassung 2 des WWT 6-10 expressiv (Glück, 2011; Items T16 bis 55 für die Zielgruppe von 7 bis 9 Jahren) bestimmt. In diesem Test geht es u. a. um die Benennung von Antonymen und Hyperonymen. Da am Projekt nur erstsprachige Kinder ohne diagnostizierte Sprachentwicklungsverzögerungen beteiligt waren, wurde auf Abrufhilfen verzichtet.

Als Kontextvariablen wurden die Rekodiergeschwindigkeit mit einer Aufgabe zum Lesen von Pseudowörtern mit dem SLRT-II (Moll & Landerl, 2014) sowie die Vorläuferfertigkeiten phonologische Bewusstheit mit dem BAKO 1-4 (Stock et al., 2017; Vokalerersetzung, Restwortbestimmung und Phonemvertauschung) und schnelles Benennen mit dem TEPHOBE (Mayer, 2016; Teilstests Farben, Buchstaben und Zahlen) erhoben.

Die verwendeten Testinstrumente sind zusammenfassend in Tabelle 1 dargestellt.

2 0 Punkte werden gemäß Testmanual für eine offensichtlich falsche Antwort vergeben, 2 Punkte gibt es für ein treffendes Synonym, die Angabe des hauptsächlichen Verwendungszwecks oder ein Hyperonym. Die übrigen Antworten, die zwar korrekt, aber nicht ausreichend sind, werden mit einem Punkt bewertet (Petermann & Petermann, 2007, S. 206–207). Diese Vorgaben legen relativ großes Gewicht auf eine dekontextualisierte Form der Definition. Die Verwendung von Synonymen und Hyperonymen binden stark den Bereich des relationalen Wortwissens ein, was im Zusammenhang mit der Beurteilung kognitiver Fähigkeiten des Kindes durchaus legitim ist, bei einer Untersuchung seines semantischen Wortwissens aber möglichst weitgehend reduziert werden sollte.

Tabelle 1: Untersuchte Variablen und die verwendeten Testinstrumente

Variable	Testinstrument
Lesen	
Wortlesen	Lenhard et al. (2018): ELFE II
Satzlesen	Lenhard et al. (2018): ELFE II
Textlesen	Lenhard et al. (2018): ELFE II
Wortschatz	
Wortschatzumfang	Lenhard et al. (2015): PPVT (Testhalbierung)
Semantisches Wortwissen	Petermann & Petermann (2007): WISC-IV (Untertest 6 Wortschatztest mit Testhalbierung)
Relationales Wortwissen	Glück (2011): WWT (Kurzfassung 2 expressiv)
Kontextvariablen	
Pseudowörterlesen	Moll & Landerl (2014): SLRT-II
Phonologische Bewusstheit	Stock et al. (2017): BAKO (Teilttests Vokalersetzung, Restwortbestimmung und Phonemvertauschung)
Schnelles Benennen	Mayer (2016): Tephobe (Teilttests Farben, Buchstaben und Zahlen)

Zu beiden Zeitpunkten wurden mit denselben Kindern dieselben Tests durchgeführt. Für eine detailliertere Beschreibung der Tests vgl. Juska-Bacher et al. (in Druck). In der ersten Klasse wurden die Daten ausschließlich mit Papierversionen in Einzeltestungen erhoben, in der zweiten Klasse bildete der ELFE-Lesetest (Lenhard et al., 2018) eine Ausnahme, er wurde – da die Kinder nun in der Lage waren, leise zu lesen – in Gruppen von bis zu sieben Kindern durchgeführt.

5.3 Auswertungsmethoden

Um der komplexen Natur der Wortschatzkompetenzen und des Leseverständnisses sowie deren Entwicklung und gegenseitigem Einfluss gerecht zu werden, ist die Wahl der passenden statistischen Modelle zur Auswertung und ihre Kombination elementar. Während das Wesen der Wortschatz- und Lesekompetenzen an sich deskriptiv ist und durch Mittelwerte abgebildet wird, sind die simultanen Einflüsse der vorliegenden Daten korrelativ und werden entsprechend durch korrelative Analysen bewertet. Die Frage nach der Kausalität im Längsschnitt wiederum muss mittels eines linearen Panel-Modells beantwortet werden. In Längsschnittstudien mit latenten Konstrukten wird diese Situation typischerweise durch sogenannte Cross-Lagged-Designs abgebildet (Mayerl & Andersen, 2019; Reinders, 2006). Die Methoden, um ein Cross-Lagged-Design zu untersuchen, sind grundsätzlich Partialkorrelationen, Regressionsanalysen oder ein Strukturgleichungsmodell. In solchen Modellen werden alle Konstrukte zu t2 auf alle Konstrukte zu t1 regressiert. Einschränkend wirkt sich

dabei eine allfällige Kollinearität aus, die eine Interpretation der relativen Veränderung über die Zeit erschwert, was hier der Fall ist. Deshalb wird im Folgenden ein Fixed-Effects-Modell verwendet, das die Veränderung in den Konstrukten frei von unbeobachteter zeitlich konstanter Heterogenität in Verbindung setzt (Allison, 2009).

6 Ergebnisse

6.1 Deskriptive Statistik der Einzelvariablen

Tabelle 2 enthält die Mittelwerte der einzelnen im Projekt erhobenen Variablen. In allen Tests war die Verbesserung der Kinderleistungen zwischen t1 und t2 hochsignifikant. In der Restwortbestimmung trat zu t2 ein Deckeneffekt auf, der aber in keiner anderen Variable feststellbar war.

Tabelle 2: Mittelwerte (Rohwerte) der Einzelvariablen

	1. Klasse				2. Klasse			
	Min	Max	Mittelwert	SD	Min	Max	Mittelwert	SD
Wortlesen	2	56	20.1	8.0	11	65	33.5	9.8
Satzlesen	0	27	5.2	4.1	2	36	12.9	6.0
Textlesen	0	16	2.8	2.8	1	24	8.4	4.7
Pseudowortlesen	4	52	20.1	7.7	7	71	30.2	9.3
Wortschatzumfang	90	197	142.7	19.3	100	211	155.5	19.4
Relationales Wortwissen	1	36	18.2	6.6	4	37	24.2	6.3
Semantisches Wortwissen	2	29	14.6	4.8	0	32	20.2	5.7
RAN Farben ^a	0.34	1.39	0.79	0.2	0.46	2.14	0.91	0.22
RAN Buchstaben ^a	0.57	2.78	1.29	0.33	0.69	2.78	1.64	0.37
RAN Zahlen ^a	0.61	2.63	1.25	0.33	0.54	2.78	1.57	0.36
PHB Vokalersetzung	0	12	4.5	3.3	0	12	6.7	3.3
PHB Restwortbestimmung	0	7	4.4	2.1	0	7	5.5	1.7
PHB Phonemvertauschung	0	11	4.3	3.6	0	11	7.2	2.9

N = 313; ^a Der Rohwert ist in Anzahl korrekt benannte Items pro Sekunde angegeben.

Tabelle 3: Korrelationen aller Variablen über beide Zeitpunkte

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	
1 t1 Wortlesen	-																									
2 t1 Satzlesen	.782**	-																								
3 t1 Textlesen	.675**	.785**	-																							
4 t1 Pseudowortlesen	.678**	.748**	.689**	-																						
5 t1 Vokalersatz	.294**	.372**	.333**	.276**	-																					
6 t1 Restwortbest.	.213**	.294**	.223**	.205**	.499**	-																				
7 t1 Phonemtausch	.367**	.411**	.355**	.308**	.447**	.546**	-																			
8 t1 RAN Farben	.277**	.225**	.214**	.271**	.157**	.184**	.232**	-																		
9 t1 RAN Buchst.	.406**	.354**	.434**	.474**	.222**	.167**	.338**	.309**	-																	
10 t1 RAN Zahlen	.347**	.259**	.299**	.374**	.146**	.126*	.272**	.513**	.624**	-																
11 t1 Wortschatzumf.	.01	.139*	.129*	.113*	.10	.04	.05	.06	-.03	-.06	-.04	-.06	-.04	-.06	-.04	-.06	-.05	-.06	-.06	-.06	-.06	-.06	-.06	-.06	-.06	-.06
12 t1 Ret. Wortwissen	.144*	.260**	.223**	.182**	.289**	.187**	.203**	.07	-.05	-.06	.404**															
13 t1 Sem. Wortwissen	.120*	.185**	.189**	.181**	.10	.00	.03	.05	-.02	.02	.356**	.425**														
14 t2 Wortlesen	.654**	.679**	.589**	.581**	.320**	.264**	.305**	.492**	.323**	.303**	.10	.174**	.156**													
15 t2 Satzlesen	.661**	.732**	.663**	.620**	.379**	.293**	.410**	.225**	.352**	.285**	.141*	.232**	.196**	.845**												
16 t2 Textlesen	.638**	.739**	.698**	.604**	.379**	.312**	.407**	.174**	.359**	.269**	.152**	.322**	.171**	.714**	.840**											
17 t2 Pseudowortlesen	.529**	.569**	.549**	.724**	.269**	.190**	.226**	.156**	.347**	.212**	.07	.114*	.138*	.637**	.695**	.648**										
18 t2 Vokalersatz	.233**	.264**	.199**	.163**	.484**	.422**	.355**	.115*	.04	-.01	-.01	.292**	.09	.276**	.332**	.294**	.237**									
19 t2 Restwortbest.	.205**	.218**	.188**	.191**	.407**	.470**	.359**	.152**	.155**	.11	.02	.134*	.00	.246**	.287**	.244**	.217**	.430**								
20 t2 Phonemtausch	.253**	.317**	.284**	.289**	.415**	.410**	.387**	.170**	.212**	.114*	.03	.168**	.03	.373**	.443**	.404**	.393**	.508**	.526**							
21 t2 RAN Farben	.211**	.189**	.198**	.252**	.131*	.05	.130*	.587**	.304**	.424**	.04	-.01	.06	.214**	.221**	.137**	.213**	-.10	.08	.169**						
22 t2 RAN Buchst.	.379**	.340**	.330**	.440**	.145*	.08	.157**	.257**	.431**	.395**	.04	-.02	.08	.446**	.402**	.376**	.451**	.131*	.06	.184**	.368**					
23 t2 RAN Zahlen	.350**	.319**	.297**	.401**	.131*	.10	.193**	.485**	.433**	.633**	-.01	-.04	.113*	.371**	.325**	.286**	.348**	.161**	.08	.142*	.554**	.572**				
24 t2 Wortschatzumf.	.08	.186**	.131*	.09	.181**	.04	.10	.02	-.05	-.07	.436**	.405**	.346**	.198**	.283**	.242**	.10	.185**	.09	.08	.03	.09	-.06			
25 t2 Ret. Wortwissen	.128*	.255**	.240**	.172**	.398**	.262**	.217**	.06	-.03	-.04	.442**	.612**	.402**	.227**	.341**	.357**	.174**	.372**	.290**	.255**	.03	.06	-.01	.542**		
26 t2 Sem. Wortwissen	.170**	.182**	.184**	.149**	.162**	.117*	.122*	.167**	.11	.08	.270**	.377**	.507**	.212**	.276**	.268**	.191**	.209**	.10	.11	.147**	.149**	.199**	.361**	.441**	

Anmerkung: N = 313; ** Die Korrelation ist signifikant auf dem 0.01-Level (2-seitig); * Die Korrelation ist signifikant auf dem 0.05-Level (2-seitig)

Zu beiden Zeitpunkten zeigten sich relativ konsistente Zusammenhänge bei den Items desselben Konstrukts. Tabelle 3 gibt die bivariaten Pearson-Korrelationskoeffizienten zwischen den Variablen wieder.

Während die Korrelationen zwischen den Lese- und den Wortschatzvariablen zu t1 schwach ausfielen und kaum 0.2 überstiegen, nahm die Stärke der Zusammenhänge über die Zeit zu: Zu t2 lagen – mit der Ausnahme des Wortschatzumfangs – die Korrelationen über 0.2 und waren oft höher als 0.3. Zudem nahmen auch die zu t1 kaum vorhandenen Korrelationen der Wortschatzgrößen mit dem RAN und der phonologischen Bewusstheit zu. Dieses Muster ist für das semantische Wortwissen am ausgeprägtesten.

Die hohen Korrelationen der Variablen innerhalb desselben Konstrukts zu beiden Zeitpunkten wiesen darauf hin, dass diese mittels einer konfirmatorischen Faktorenanalyse für längsschnittliche Daten zusammengefasst werden dürfen (Little, 2013). Da die Normalverteilung teilweise nicht gegeben war, wurde ein robustes Maximum Likelihood Verfahren verwendet (Gana & Broc, 2019) Eine Korrelation zwischen den jeweiligen Konstrukten zu t1 und t2 wurde zugelassen, um die zeitliche Stabilität abzubilden. Tabelle 4 gibt einige Maßzahlen zu den vier latenten Konstrukten wieder.

Tabelle 4: Maßzahlen zu den latenten Konstrukten

	Lesen	Phonologische Bewusstheit	RAN	Wortschatz
CFI	0.98	0.98	0.89	0.93
SRMR	0.03	0.03	0.08	0.05
R ² -Items	[0.67–0.92]	[0.46–0.58]	[0.35–0.80]	[0.32–0.66]
Alpha (t1, t2)	0.89, 0.92	0.75, 0.73	0.74, 0.75	0.66, 0.70
Korrelation t1-t2	0.83	0.84	0.79	0.99

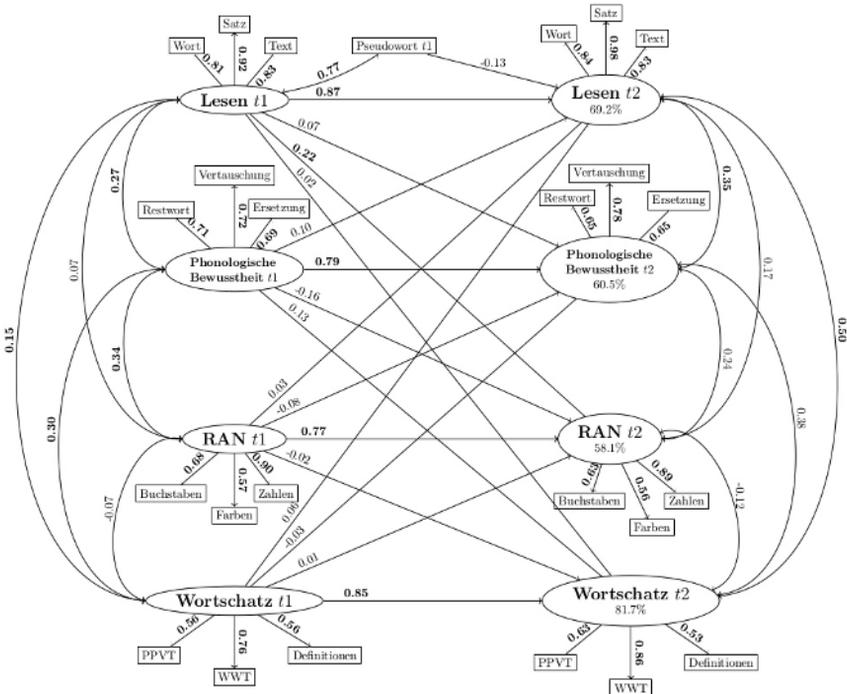
Das Lesen wurde durch die drei Variablen Wort-, Satz- und Textlesen vergleichsweise gut abgebildet und wies sehr gute Werte beim Modellfit (CFI > 0.95, SRMR < 0.05), der Dimensionalität (Cronbachs Alpha > 0.8) und der erklärten Varianz in den einzelnen Variablen auf (hohe R²-Werte für alle Items zu beiden Zeitpunkten). Dies gilt auch für die phonologische Bewusstheit und mit Einschränkungen für das schnelle Benennen. Für den Wortschatz fiel das Ergebnis hinsichtlich der Dimensionalität und der erklärten Varianz bescheiden aus. Wortschatzumfang, relationales und semantisches Wortwissen schienen nur mäßig auf ein gemeinsames Konstrukt „Wortschatz“ zurückzuführen zu sein. Das latente Konstrukt „Wortschatz“ vermochte insbesondere zu t1 nur rund ein Drittel der Varianz in Wortschatzumfang und semantischem Wortwissen zu erklären. Auf Grund der ansonsten akzeptablen Modellgüte und den

theoretischen Überlegungen wurden die drei Items weiter für die Konstruktion des Wortschatzes verwendet.

6.2 Zusammenhang zwischen Wortschatz und Lesen im Längsschnitt

Zur Klärung der Frage, wie die Entwicklung der Lesefähigkeiten mit dem Wortschatz zusammenhängt, wurden Cross-Lagged-Modelle verwendet, in welchen zusätzlich die beiden weiteren Konstrukte, RAN und phonologische Bewusstheit, kontrolliert wurden (Mayerl & Andersen, 2019). In solchen Modellen werden alle Konstrukte zu t2 auf alle Konstrukte zu t1 regressiert. Zudem wird das Lesen zu t2 auch auf das Pseudowortlesen zu t1 zurückgeführt. Erwartungsgemäß lud das Pseudowortlesen deutlich niedriger auf einem gemeinsamen Lesefaktor als die anderen Lese-Items, weshalb es als unabhängiges manifestes Merkmal mit in das Modell mitaufgenommen wurde. Abbildung 1 zeigt die entsprechenden Resultate, hervorgehobene Koeffizienten sind mindestens auf dem 5%-Signifikanzniveau signifikant, alle Koeffizienten sind standardisiert.

Abbildung 1: Cross-Lagged-Modell



CI robustler RMSEA: [0.055, 0.070]; robuster CFI: 0.933

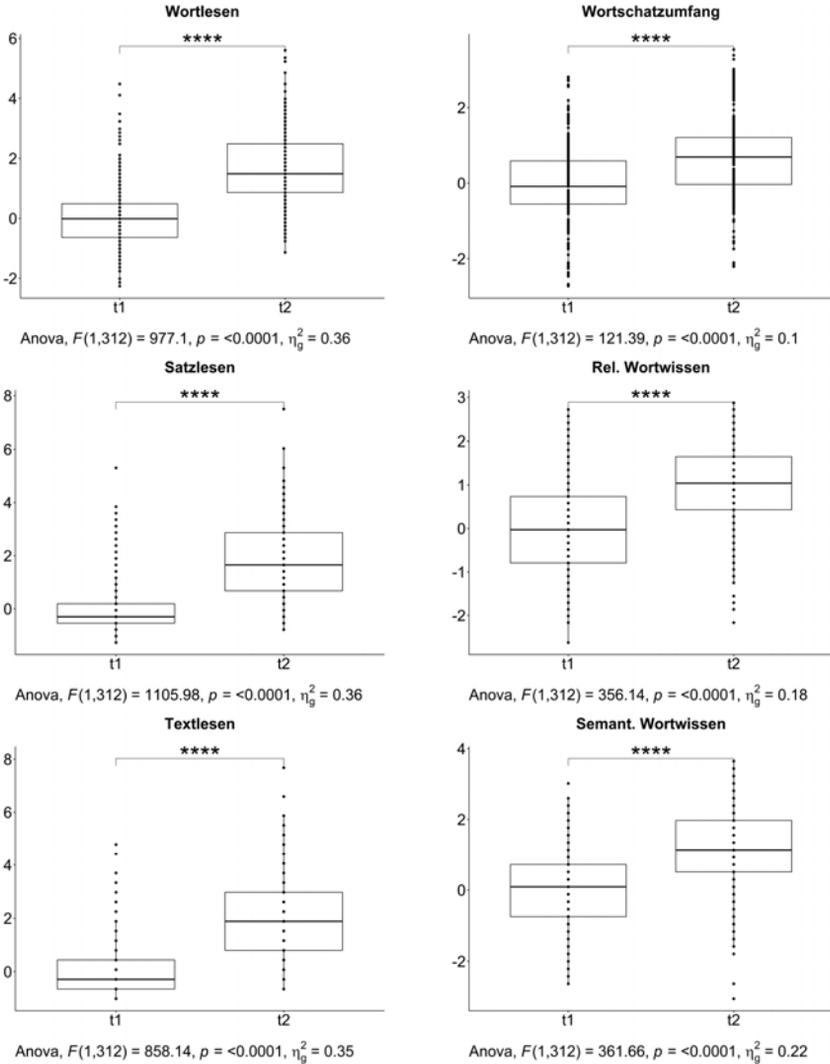
Die Abbildung verdeutlicht, dass die einzelnen Konstrukte zeitlich stabil sind, was sich in vergleichbaren standardisierten Ladungen zu t1 und t2 und starken autoregressiven Pfaden äussert: Alle Konstrukte zu t1 hatten einen hochsignifikanten Einfluss auf die entsprechenden Konstrukte zu t2. Vor dem Hintergrund dieser autoregressiven Einflüsse manifestierte sich ein einziger signifikanter Cross-Lagged-Effekt, der RAN zu t2 wurde weiter durch die Lesefähigkeiten zu t1 determiniert.

Dass keine weiteren Effekte der Konstrukte zu t1 auf die Lesefähigkeiten zu t2 unter Kontrolle der Lesefähigkeiten zu t1 bestanden, vermag auf den ersten Blick zu überraschen. Indes zeigt sich einerseits, dass insbesondere beim Lesen und beim Wortschatz zu t2 große Varianzanteile durch die jeweiligen Konstrukte zu t1 erklärt werden. Die entsprechenden Anteile belaufen sich auf 69.2 Prozent beim Lesen und auf 81.7 Prozent beim Wortschatz. Andererseits finden sich zwischen dem Lesen und dem Wortschatz wie auch mit der phonologischen Bewusstheit signifikante Kovarianzen zu beiden Messzeitpunkten (nicht gezeigt in Abbildung 1).

Die Zusammenhänge zwischen Wortschatz und Lesen, die über die gezeigten Korrelationen hinausgehen, werden veranschaulicht durch die parallele Entwicklung der Lese- und Wortschatzkompetenzen. Der vergleichbare, positive Koeffizient bedeutet, dass beide Konstrukte einen ähnlichen Zuwachs zwischen t1 und t2 zu verzeichnen haben.

Abbildung 2 gibt die paarweisen Vergleiche entsprechender Varianzanalysen wieder und zeigt den Anstieg in den Lese- wie auch den Wortschatzvariablen zwischen t1 und t2 deutlich, alle Differenzen waren hochsignifikant.

Abbildung 2: ANOVA standardisierte Lese- und Wortschatzvariablen



Anmerkung: Die Variablen wurden basierend auf den Werten zu t1 mit einem Mittelwert von 0 und einer Standardabweichung von 1 standardisiert. Eine Streuung in den negativen Bereich ist deshalb möglich. Die standardisierten Werte bilden die Skalen der y-Achsen in den Einzelgrafiken.

Im Mittel erhöht sich dabei die Lesekompetenz um 1.87 Punkte bei einer mittleren Standardabweichung von 1.47, während der Wortschatz lediglich 0.92 Punkte dazugewinnt, dies bei einer mittleren Standardabweichung von 1.26. Die entsprechenden Mittelwerte zu t2 sind 0.66 ($SD = 1$) für den Wortschatzumfang, 0.92 ($SD = 0.96$) für das relationale und 1.16 ($SD = 1.81$) für das semantische Wortwissen, bzw. 1.70 ($SD = 1.23$), 1.89 ($SD = 1.46$) und 2.00 ($SD =$

1.71) für das Wort-, Satz- und Textlesen. Zusammengenommen mit den steigenden Korrelationen (Abbildung 1) übersteuern diese Effekte einen Einfluss des Wortschatzes t1 auf das Lesen t2.

Die Folgen dieser Konstellation werden sichtbar, wenn in einem Fixed-Effects-Modell in Tabelle 5 simultane Effekte und die Veränderung in den Konstrukten längsschnittlich und frei von unbeobachteter zeitlich konstanter Heterogenität miteinander in Bezug gesetzt werden. Erhöht sich der Wortschatz um eine Standardabweichung zwischen t1 und t2, so resultiert dies in einer Erhöhung der Lesekompetenz um 0.227 Standardabweichungen, erhöht sich die phonologische Bewusstheit um eine Standardabweichung, führt dies zu einer zusätzlichen Erhöhung der Lesefähigkeiten um 0.142 Standardabweichungen zwischen t1 und t2.

Tabelle 5: Fixed-Effects-Modelle

	b	SE
Phonologische Bewusstheit	0.142**	0.049
RAN	-0.004	0.046
Wortschatz	0.227***	0.057

Hausman Test RE vs. FE: $\chi^2(3) = 51.06, p = 0.000$

$N = 626 (313/313); *p \leq 0.05, ** p \leq 0.01, *** p \leq 0.001$

7 Diskussion

7.1 Gesamtbild (Modell)

Die Resultate legen nahe, dass Fortschritte im Wortschatzbereich die Lesefähigkeit sowohl querschnittlich als auch längsschnittlich beeinflussen, dieser Effekt aber in den ersten zwei Schuljahren nur indirekt sichtbar wird. So zeigt das Cross-Lagged-Modell bis auf eine Ausnahme (das schnelle Benennen zu t2 wird durch die Lesefähigkeit zu t1 beeinflusst) ausschließlich autoregressive Effekte, das heißt, es zeigen sich nur longitudinale Effekte *innerhalb* einer Kompetenzgruppe. Betrachtet man hingegen die gegenseitigen Beziehungen der für das Lesen relevanten Faktoren jeweils zu einem einzelnen Zeitpunkt, so ergeben sich signifikante Zusammenhänge zwischen allen Faktoren. Überaus bedeutsam ist die Entwicklung dieser Abhängigkeiten: Insbesondere der Einfluss des Wortschatzes auf das (verstehende) Lesen nimmt prägnant zu, während sich der Einfluss des schnellen Benennens und der phonologischen Bewusstheit weit weniger verändert. Die Fixed-Effects-Modelle veranschaulichen sehr deutlich, wie sich der Effekt des Wortschatzes als Ganzes verstärkt. Diese Zunahme des Wortschatz-Effekts auf die Lesefähigkeit während der ersten Schuljahre findet sich auch in anderen Studien. Torgesen et al. (1997) wiesen nach Kontrolle des

autoregressiven Effekts einen Anstieg der Varianzerklärung des Wortschatzes aus der 2. respektive 4. Klasse auf das verstehende Lesen zwischen der 3. und der 5. Klasse von 24 auf 43 Prozent aus (ähnlich Bast & Reitsma, 1998, allerdings ohne Kontrolle des autoregressiven Effekts), und auch De Jong & Van der Leij (2002) interpretieren ihre Resultate dahingehend, obwohl sie keine substantielle Zunahme der Korrelationen zwischen dem verstehenden Lesen und dem linguistischen Verständnis (rezeptiver und produktiver Wortschatz, Hörverständnis) fanden, was in der vorliegenden Studie hingegen der Fall ist.

Eine mögliche Erklärung für den relativ geringen Einfluss des Wortschatzes auf das Lesen in der 1. Klasse ist, dass zu diesem Zeitpunkt das verstehende Lesen noch wenig entwickelt ist und erst mit zunehmender Dekodierfähigkeit an Bedeutung gewinnt (gemäß Biemiller, 2012 in der dritten Klasse). Dafür spricht die – im 1. Schuljahr durchaus übliche (z. B. Juska-Bacher et al., 2015) – geringe Anzahl Items, die besonders in den Satz- und Textlese-Tests überhaupt bearbeitet wurde. Dazu ist anzumerken, dass in der vorliegenden Studie die Leseleistung in der 1. Klasse deutlich tiefer liegt als jene der Studie von Fricke et al. (2016), während die Leistung in der 2. Klasse ähnlich ausfällt.

Im Cross-Lagged-Modell ebenfalls nicht signifikant war der längsschnittliche Einfluss der phonologischen Bewusstheit und des schnellen Benennens zu t1 auf das Leseverständnis zu t2. Diese Resultate entsprechen jenen von De Jong & Van der Leij (2002) sowie von Landerl & Wimmer (2008), die über das erste Schuljahr hinaus keinen signifikanten Einfluss der phonologischen Fähigkeiten auf die Rekodiergeschwindigkeit respektive Leseflüssigkeit fanden, die ihrerseits aber den größten Anteil an Varianzaufklärung am späteren Leseverständnis hatte. Der Einfluss der phonologischen Bewusstheit und des schnellen Benennens auf das Leseverständnis ist also indirekt und nimmt – im Falle der phonologischen Bewusstheit – im Verlauf der ersten Schuljahre stetig ab.

7.2 Das Konstrukt Lesen

In anderen Studien wurde die Schwierigkeit der Lesetests an die Schulstufe angepasst (z. B. De Jong & Van der Leij, 1999, 2002; Landerl & Wimmer, 2008) oder teilweise das rekodierende Lesen in das Konstrukt Lesen miteinbezogen (Reese et al., 2010). Die deutlich geringere Ladung des Pseudowortlesens im Vergleich zu den Variablen Wort-, Satz- und Textlesen auf den Faktor Lesen führte in der hier vorgestellten Studie zum Entscheid, dieses nicht in das Konstrukt Lesen zu integrieren. Dieses Vorgehen entspricht demjenigen der Studien von Ouellette (2006) und Swart et al. (2017), in der das Rekodieren anhand der Geschwindigkeit und Fehleranzahl beim Vorlesen von Pseudowörtern konzeptualisiert und nicht in ein gemeinsames Konstrukt mit dem verstehenden Lesen einbezogen wird. Ein Einbezug der Rekodiergeschwindigkeit in ein Gesamt-Lesekonstrukt ist eher heikel. Ennemoser et al. (2012) unterscheiden

klar zwischen Leseflüssigkeit, Lesegenauigkeit und Leseverständnis, wobei im ELFE II letzteres im Fokus steht. In diesem Sinne bildet das Konstrukt Lesen der vorliegenden Studie ausschließlich das Leseverständnis ab.

In den meisten Studien (z. B. De Jong & Van der Leij, 1999, 2002) wird das rekodierende Lesen mit dem Vorlesen von Listen mit Wörtern und/oder Pseudowörtern getestet und als eigenes Konstrukt positioniert, so auch hier. Die Entscheidung, in dieser Studie einzig das Pseudowortlesen einzubeziehen, beruht auf den Erkenntnissen aus der Studie von Verhoeven & Keuning (2018), die in einer vergleichenden Studie mit guten Lesern und Kindern, die von Dyslexie betroffen sind, nur für das Pseudowortlesen eine gute Prädiktionskraft und Sensitivität für eine mögliche Leseschwäche fanden, nicht aber für das isolierte Wortlesen.

7.3 Einzelne Wortschatzvariablen

Wenn man die Korrelationen der einzelnen Variablen des Lesens und des Wortschatzes zu t1 betrachtet, fällt auf, dass sie für das Satz- und das Textlesen höher sind als für das Wortlesen, was wiederum die oben beschriebene These untermauert, dass der Wortschatz zwar für das verstehende Lesen größerer Texteinheiten, weniger aber für das rekodierende Lesen entscheidend ist. Bei Kindern, die zu t1 bereits in der Lage waren, Sätze und kurze Texte zu lesen, zeigt sich ein Zusammenhang mit dem Wortschatz, nicht aber bei jenen, deren Lesefähigkeit nur für das Lesen von Wörtern ausreichte. Zu t2 sind die Korrelationen aller Wortschatzvariablen mit dem Lesen deutlich stärker, zu diesem Zeitpunkt sind alle Kinder imstande, verstehend auch Sätze und Texte zu lesen. Zu beiden Zeitpunkten korrelieren relationales und semantisches Wortwissen, analog zu Tannenbaum et al. (2006), stärker mit dem Lesen. Auffällig ist, dass die Zunahme der Korrelationsstärke vor allem beim relationalen und semantischen Wortwissen erfolgte, während der Wortschatzumfang sowohl quer- als auch längsschnittlich unterschiedliche Resultate zeitigte. Zu t1 waren die Korrelationen mit dem Lesen schwach bis nicht vorhanden (Wortlesen), und auch zu t2 waren die Korrelationen klar schwächer als die zwischen relationalem und semantischem Wortwissen und Lesen, ein Phänomen, das sich auch in einer Teilstudie mit bilingualen Kindern zeigte, wo vor allem das relationale Wortwissen eine hohe Kovarianz mit dem verstehenden Lesen zeigte (Röthlisberger et al., in Druck). Förderlich für das verstehende Lesen ist es also nicht nur, möglichst viele Wörter zu kennen, sondern besonders auch ein umfangreiches semantisches Wissen zu haben und über gute Verknüpfungen zwischen den Einträgen im mentalen Lexikon zu verfügen.

Ein indirekter Einfluss des Wortschatzumfanges über die phonologische Bewusstheit auf das Lesen (Juska-Bacher, 2015; Sénéchal et al., 2006) konnte in der vorliegenden Studie nicht bestätigt werden, obwohl zahlreiche Evidenzen

einen positiven Einfluss des Wortschatzes auf die phonologische Bewusstheit existieren (z. B. Goswami, 2001; Metsala, 1999).

8 Konklusion

Ganz offensichtlich ist der Wortschatz ein komplexes Konstrukt und seine Rolle bei der Entwicklung des frühen Lesens alles andere als trivial. Deshalb ist es bei der Auswertung sinnvoll, die üblicherweise in Entwicklungsstudien verwendeten Pfadanalysen um zusätzliche Modelle zu ergänzen. Allerdings bleibt die statistische Modellierung solch vielschichtiger Daten über die Zeit hinweg eine Herausforderung mit Potential für neue Methoden.

Potential ist auch vorhanden, was die Messmethoden im Wortschatzbe- reich, insbesondere des semantischen Wortwissens, angeht. Bisher besteht we- der Einigkeit, wie dieses am besten zu operationalisieren ist, noch nach welchen Kriterien – sollte tatsächlich mit Definitionen gearbeitet werden – die Itemaus- wahl erfolgen sollte, noch wie die Antworten zuverlässig bewertet werden. Die Entwicklung eines Bewertungsrasters mit Antwortmöglichkeiten und deren Bepunktung sowie die darauf abgestimmte Bewertung der Definitionen durch mehrere Rater und ggf. in mehreren Überarbeitungsschritten sind zeitintensiv und sollten nicht als „Nebenprodukte“ in individuellen Projekten entwickelt werden. Es erstaunt nicht, dass es bisher (nicht nur) im deutschsprachigen Raum kein standardisiertes Testinstrument gibt, es wäre aber für eine differen- zierte Untersuchung von Wortschatzkompetenzen umso wünschenswerter.

Danksagung

Das Projekt EnWoLe wurde vom Schweizerischen Nationalfonds ([http://p3.snf. ch/Project-173245](http://p3.snf.ch/Project-173245)) finanziert. Wir danken unserer Kollegin Ladina Brugger für wertvolle Anregungen und Unterstützung in diesem Projekt.

Literatur

Testinstrumente

- Glück, C. W. (2011): *Wortschatz- und Wortfindungstest für 6- bis 10-Jährige (WWT 6–10)*. München: Urban & Fischer.
- Lenhard, A., Lenhard, W., Segerer, R. & Suggate, S. (2015): *Peabody Picture Vocabulary Test – Revision IV*. (Deutsche Adaption). Frankfurt am Main: Pearson Assessment.
- Lenhard, W., Lenhard, A. & Schneider, W. (2018): *ELFE II. Ein Leseverständnistest für Erst- bis Siebtklässler – Version II*. (Manual). Göttingen: Hogrefe.

- Mayer, A. (2016): *TEPHOBE. Test zur Erfassung der phonologischen Bewusstheit und der Benennungsgeschwindigkeit*. München, Basel: Reinhardt.
- Moll, K. & Landerl, K. (2014): *SLRT-II Lese- und Rechtschreibtest. Weiterentwicklung des Salzburger Lese- und Rechtschreibtests (SLRT)*. Bern: Huber.
- Petermann, F. & Petermann, U. (2007): *Hamburg-Wechsler-Intelligenztest für Kinder – IV*. (Übersetzung und Adaptation des WISC-IV® von David Wechsler). Bern: Huber.
- Stock, C., Marx, P. & Schneider, W. (2017): *BAKO 1–4. Basiskompetenzen für Lese-Rechtschreibleistungen. Ein Test zur Erfassung der phonologischen Bewusstheit vom ersten bis vierten Grundschuljahr*. Göttingen: Beltz Test.

Forschungsliteratur

- Allison, P. D. (2009): *Fixed Effects Regression Models*. Thousand Oaks: Sage.
- Anderson, R. & Freebody, P. (1981): Vocabulary Knowledge. In J. Guthrie (Ed.), *Comprehension and Teaching: Research Reviews* (pp. 77–117). Newark: International Reading Association.
- Bast, J. & Reitsma, P. (1998): Analyzing the Development of Individual Differences in Terms of Matthew Effects in Reading: Results from a Dutch Longitudinal Study. *Developmental Psychology*, 34(6), 1373–1399. <https://doi.org/10.1037/0012-1649.34.6.1373>
- Biemiller, A. (2012): Teaching Vocabulary in the Primary Grades. In E. J. Kame'enui & J. F. Baumann (Eds.), *Vocabulary Instruction: Research to Practice* (pp. 34–50). New York: Guilford Press.
- Cain, K. & Oakhill, J. (2014): Reading Comprehension and Vocabulary: Is Vocabulary More Important for Some Aspects of Comprehension? *L'Année Psychologique/Topics in Cognitive Psychology*, 114(4), 647–662. <https://doi.org/10.4074/S0003503314004035>
- Castles, A. & Coltheart, M. (2004): Is There a Causal Link from Phonological Awareness to Success in Learning to Read? *Cognition*, 91(1), 77–111. [https://doi.org/10.1016/S0010-0277\(03\)00164-1](https://doi.org/10.1016/S0010-0277(03)00164-1)
- Clahsen, H. & Fleischhauer, E. (2014): Morphological Priming in Child German. *Journal of Child Language*, 41(6), 1305–1333. <https://doi.org/10.1017/S0305000913000494>
- De Jong, P. F. & Van der Leij, A. (1999): Specific Contributions of Phonological Abilities to Early Reading Acquisition. Results from a Dutch Latent Variable Longitudinal Study. *Journal of Educational Psychology*, 91(3), 450–476. <http://doi.org/10.1037/0022-0663.91.3.450>
- De Jong, P. F. & Van der Leij, A. (2002): Effects of Phonological Abilities and Linguistic Comprehension on the Development of Reading. *Scientific Studies of Reading*, 6(1), 51–77. https://doi.org/10.1207/S1532799XSSR0601_03
- Deacon, S. H. & Kirby, J. R. (2004): Morphological Awareness: Just “More Phonological”? The Roles of Morphological and Phonological Awareness in Reading Development. *Applied Psycholinguistics*, 25(2), 223–238. <https://doi.org/10.1017/S0124716404001117>
- Ennemoser, M., Marx, P., Weber, J. & Schneider, W. (2012): Spezifische Vorläuferfertigkeiten der Lesegeschwindigkeit, des Leseverständnisses und des Rechtschreibens: Evidenz aus zwei Längsschnittstudien vom Kindergarten bis zur 4. Klasse. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 44(2), 53–67. <https://doi.org/10.1026/0049-8637/a000057>

- Fricke, S., Szczerbinski, M., Fox-Boyer, A. & Stackhouse, J. (2016): Preschool Predictors of Early Literacy Acquisition in German-Speaking Children. *Reading Research Quarterly*, 51(1), 29–53. <https://doi-org/10.1002/rrq.116>
- Gana, K. & Broc, G. (2019): *Structural Equation Modeling with Lavaan*. Hoboken: Wiley.
- Goswami, U. (2001): Early Phonological Development and the Acquisition of Literacy. In S. Neuman & D. Dickinson (Eds.), *Handbook of Early Literacy Research* (pp. 111–125). New York: Guilford Press.
- Juska-Bacher, B., Beckert, C., Stalder, U. & Schneider, H. (2015): Die Bedeutung des Wortschatzes für basale Lesekompetenzen. *Didaktik Deutsch* 21(40), 20–39.
- Juska-Bacher, B., Brugger, L., Korthus, R. & Zangger, C. (2021): Definitionskompetenzen von Erst- und Zweitklässlern. Mit einem Ausblick auf die weitere Entwicklung. *Muttersprache*, 131(2), 117–137. <https://doi.org/10.53371/60203>
- Juska-Bacher, B. & Jakob, S. (2014): Wortschatzumfang und Wortschatzqualität und ihre Bedeutung im fortgesetzten Spracherwerb. *Zeitschrift für Angewandte Linguistik*, 61(1), 49–75. <https://doi.org/10.1515/zfal-2014-0016>
- Juska-Bacher, B., Röthlisberger, M., Brugger, L. & Zangger, C. (in Druck): Lesen im 1. Schuljahr: Die Bedeutung von Vorläuferkompetenzen und Wortschatz. In S. Gailberger & C. Sappok (Hrsg.), *Weiterführende Grundlagenforschung in der empirischen Leseforschung und Lesedidaktik. Theorie – Empirie – Didaktik*. Bochum: SLLD-B.
- Kim, Y.-S., Wagner, R. K. & Foster, E. (2011): Relations Among Oral Reading Fluency, Silent Reading Fluency, and Reading Comprehension: A Latent Variable Study of First-Grade Readers. *Scientific Studies of Reading*, 15(4), 338–362. <https://doi.org/10.1080/10888438.2010.493964>
- Kurland, B. F. & Snow, C. (1997): Longitudinal Measurement of Growth in Definitional Skill. *Journal of Child Language*, 24(3), 603–625. <https://doi.org/10.1017/S0305000997003243>
- Landerl, K., Freudenthaler, H. H., Heene, M., De Jong, P. F., Desrochers, A., Manolitsis, G., Parrila, R. & Georgiou, G. K. (2019): Phonological Awareness and Rapid Automatized Naming as Longitudinal Predictors of Reading in Five Alphabetic Orthographies with Varying Degrees of Consistency. *Scientific Studies of Reading*, 23(3), 220–234. <https://doi.org/10.1080/10888438.2018.1510936>
- Landerl, K. & Wimmer, L. (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>
- Little, T. D. (2013): *Longitudinal Structural Equation Modeling*. New York: The Guilford Press.
- Mayer, A. (2018): Benennungsgeschwindigkeit und Lesen. Rapid Automatized Naming (RAN) and Reading. *Forschung Sprache* (1), 20–42.
- Mayerl, J. & Andersen, H. (2019): Recent Developments in Structural Equation Modeling with Panel Data: Causal Analysis and Change over Time in Attitude Research. In J. Mayerl, T. Krause, A. Wahl & M. Wuketich (Hrsg.), *Einstellungen und Verhalten in der empirischen Sozialforschung* (S. 415–449). Wiesbaden: Springer Fachmedien. https://doi.org/10.1007/978-3-658-16348-8_17
- Metsala, J. L. (1999): Young Children's Phonological Awareness and Nonword Repetition as a Function of Vocabulary Development. *Journal of Educational Psychology*, 91(1), 3–19. <https://doi.org/10.1037/0022-0663.91.1.3>

- Ouellette, G. P. (2006): What's Meaning Got to Do with It: The Role of Vocabulary in Word Reading and Reading Comprehension. *Journal of Educational Psychology*, 98(3), 554–566. <https://doi.org/10.1037/0022-0663.98.3.554>
- Peña, E. D. (2007): Lost in Translation: Methodological Considerations in Cross-Cultural Research. *Child Development*, 78(4), 1255–1264. <https://doi.org/10.1111/j.1467-8624.2007.01064.x>
- Read, J. (2004): Plumbing the Depths: How Should the Construct of Vocabulary Knowledge be Defined? In P. Bogaards & B. Laufer (Eds.), *Vocabulary in a Second Language* (pp. 209–227). Amsterdam: John Benjamins.
- Reese, E., Suggate, S., Long, J. & Schaughency, E. (2010): Children's Oral Narrative and Reading Skills in the First 3 Years of Reading Instruction. *Reading and Writing*, 23(6), 627–644. <https://doi.org/10.1007/s11145-009-9175-9>
- Reinders, H. (2006): Kausalanalysen in der Längsschnittforschung. Das Crossed-Lagged-Panel Design. *Diskurs Kindheits- und Jugendforschung*, 1(4), 569–587.
- Röthlisberger, M., Schneider, H. & Juska-Bacher, B. (in Druck): Lesen von Kindern mit Deutsch als Erst- und Zweitsprache: Wortschatz als limitierender Faktor. *Zeitschrift für Grundschulforschung*. <https://doi.org/10.1007/s42278-021-00115-w>
- Rydland, V., Grøver Aukrust, V. & Fulland, H. (2012): How Word Decoding, Vocabulary and Prior Topic Knowledge Predict Reading Comprehension. A Study of Language-Minority Students in Norwegian Fifth Grade Classrooms. *Reading and Writing*, 25(2), 465–482. <https://doi.org/10.1007/s11145-010-9279-2>
- Schmitt, N. (2010): *Researching Vocabulary: A Vocabulary Research Manual*. Houndmills, Basingstoke, Hampshire, New York, NY: Palgrave Macmillan.
- Sénéchal, M., Ouellette, G. & Rodney, D. (2006): The Misunderstood Giant: On the Predictive Role of Early Vocabulary to Future Reading. In D. K. Dickinson & S. B. Neuman (Eds.), *Handbook of Early Reading Research* (pp. 173–182). New York: Guilford.
- Swart, N. M., Muijselaar, M. M. L., Steenbeek-Planting, E. G., Droop, M., De Jong, P. F. & Verhoeven, L. (2017): Differential Lexical Predictors of Reading Comprehension in Fourth Graders. *Reading and Writing*, 30(3), 489–507. <https://doi.org/10.1007/s11145-016-9686-0>
- Tannenbaum, K. R., Torgesen, J. K. & Wagner, R. K. (2006): Relationships between Word Knowledge and Reading Comprehension in Third-Grade Children. *Scientific Studies of Reading*, 10(4), 381–398. https://doi.org/10.1207/s1532799xssr1004_3
- Torgesen, J. K., Wagner, R. K., Rashotte, C. A., Burgess, S. & Hecht, S. (1997): Contributions of Phonological Awareness and Rapid Automatic Naming Ability to the Growth of Word-Reading Skills in Second- to Fifth-Grade Children. *Scientific Studies of Reading*, 1(2), 161–185. https://doi.org/10.1207/s1532799xssr0102_4
- Verhoeven, L. & Keuning, J. (2018): The Nature of Developmental Dyslexia in a Transparent Orthography. *Scientific Studies of Reading*, 22(1), 7–23. <http://doi.org/10.1080/10888438.2017.1317780>
- Vermeer, A. (2001): Breadth and Depth of Vocabulary in Relation to L1/L2 Acquisition and Frequency of Input. *Applied Psycholinguistics*, 22(2), 217–234. <http://doi.org/10.1017/S0142716401002041>

Die Wirksamkeit von Fördermaßnahmen zur Leseflüssigkeit in Abhängigkeit von Passung, Dosierung und Wiedergabetreue

Herausforderungen für Forschung und Praxis

Mareike Ehlert & Elmar Souvignier

Zusammenfassung: Flüssiges Lesen ist eine wesentliche Voraussetzung für Verständnisseleistungen. Sowohl in Modellen der Lesekompetenz als auch in Förderkonzepten wird die Leseflüssigkeit daher als eigenständige Komponente betrachtet. Im Beitrag ordnen wir das Konstrukt der Leseflüssigkeit theoretisch ein, skizzieren zentrale Befunde zu wirksamen Fördermaßnahmen und stellen mit dem Programm „Der Lesesportler“ eine Möglichkeit zur Umsetzung dieser Prinzipien in der Grundschule vor. Insgesamt legen empirische Befunde eine hohe Wirksamkeit gezielter Maßnahmen zur Förderung der Leseflüssigkeit nahe. Allerdings fallen dabei Effekte in der schulischen Praxis niedriger aus als in stärker kontrollierten Settings und sie weisen eine große Streuung auf. Vor diesem Hintergrund diskutieren wir anhand dreier Aspekte – Passung, Dosierung und Wiedergabetreue –, unter welchen Bedingungen Maßnahmen zur Förderung der Leseflüssigkeit ihre Wirksamkeit optimal entfalten.

Schlüsselwörter: Leseflüssigkeit, differenzierte Leseförderung, Implementation, Evidenzbasierung, Evaluation

Abstract: Reading fluency is regarded as an essential prerequisite for reading comprehension. Therefore, reading fluency has become a key component in effective reading interventions. In this article, we outline pivotal findings of reading fluency research on effective measures and present the program “The Reading Sportsperson” as one way to implement these principles in elementary schools. Overall, empirical findings suggest a high effectiveness of interventions promoting reading fluency. However, it is also clear that effect sizes of studies in school practice are generally lower and can vary significantly. In light of these ambiguous effects, we review potential factors that may moderate the effectiveness of fluency interventions. Specifically, we discuss three aspects – fit, dosage, and implementation fidelity – to address the conditions under which fluency interventions are effective.

Keywords: reading fluency; differentiated reading instruction; implementation; evidence-based teaching; evaluation

1 Einleitung

Wer Texte verstehen möchte, muss flüssig lesen können. Während sich die Forschung zunächst primär auf das Leseverständnis als zentralen Aspekt der Lesekompetenz konzentrierte, findet seit Beginn der 2000er Jahre auch die Leseflüssigkeit als Schlüsselkompetenz in der Leseforschung und -förderung verstärkte Beachtung. Dies illustrieren beispielsweise der Bericht des National Reading Panel (NICHD, 2000) und umfassende wissenschaftliche Untersuchungen zur Leseflüssigkeit (z. B. Rosebrock et al., 2017; Fuchs et al., 2001). Die Leseflüssigkeit selbst lässt sich, wie schon im Eingang dieses Bandes ausführlicher beschrieben wurde, als Zusammenspiel von vier interdependenten Komponenten verstehen (Rosebrock & Nix, 2006, 2011). Diese umfassen (1) die Lesegenauigkeit, definiert als Prozentsatz korrekt gelesener Wörter, (2) die Automatisierung von Leseprozessen, um kognitive Ressourcen für höhere Verstehensprozesse nutzen zu können, (3) die Lesegeschwindigkeit, unter der die Anzahl der korrekt gelesenen Wörter pro Minute verstanden wird, und schließlich (4) die Prosodie zur Hervorhebung zusammengehörender Sinneinheiten auf der Satz- bzw. Textebene.

Ein empirisches Argument für die Konzeptualisierung der Leseflüssigkeit als eigenständige Komponente in der Entwicklung der Lesekompetenz ist ihre hohe Aussagekraft als Prädiktor für das Leseverständnis (Gold, 2009; Klicpera et al., 1993; Silverman et al., 2013). Analog ist die Leseflüssigkeit als zentraler Teilprozess in kognitionspsychologischen Modellen zur Entwicklung der Lesekompetenz zu finden (z. B. Kintsch, 1998; Richter & Christmann, 2002). In diesen Modellen wird die Leseflüssigkeit zwischen der Worterkennung und dem Leseverständnis platziert, wodurch ihr eine Brückenfunktion zwischen diesen beiden Teilkompetenzen zugewiesen wird. Die vermittelnde Rolle der Leseflüssigkeit zwischen der Worterkennung und dem Leseverständnis zeigt sich beispielsweise in einer Studie von Silverman et al. (2013). Sie stellten fest, dass Viertklässlerinnen und Viertklässler, die zwar gut dekodieren, aber noch nicht flüssig lesen konnten, schlechter in Leseverständnistests abschnitten als diejenigen mit bereits ausgeprägten Fähigkeiten der Leseflüssigkeit. Für die Leseförderung bedeuten diese Befunde, dass zunächst hierarchieniedrigere Fähigkeiten beherrscht werden sollten, bevor darauf aufbauende Kompetenzen wie Verständnisleistungen eingeübt werden können. Entsprechend liegen mittlerweile einige effektive Förderprogramme vor, die Lesekompetenz als Zusammenspiel der Teilfertigkeiten Worterkennung, Leseflüssigkeit und Leseverständnis konzeptualisieren und passend zum jeweiligen Stand der Leseent-

wicklung differenzierte Förderangebote machen (z. B. Bertschi-Kaufmann & Hagendorf, 2008; Munser-Kiefer, 2014; Ritter & Scheerer-Neumann, 2009).

2 Förderung der Leseflüssigkeit

Zur Beantwortung der Frage, welche Methoden im Spezifischen für die Leseflüssigkeit förderlich sind, liegt eine Vielzahl an empirischen Befunden zur Wirksamkeit von Fördermaßnahmen vor, die die Lesegenauigkeit und die Lesegeschwindigkeit als zentrale Teilkomponenten fokussieren. Im folgenden Abschnitt skizzieren wir empirisch bewährte Methoden zur Förderung der Leseflüssigkeit und illustrieren am Beispiel des Förderprogramms „Der Lese-Sportler“, wie diese Prinzipien im Leseunterricht der Grundschule praktisch umgesetzt werden können.

2.1 Evidenzbasierte Methoden zur Förderung der Leseflüssigkeit

Zur Förderung der Lesegenauigkeit und -geschwindigkeit haben sich insbesondere das silbenbasierte und das laute Lesen zu vielzitierten Klassikern entwickelt (NICHD, 2000). Um gezielt die Lesegenauigkeit zu verbessern, werden beim *silbenbasierten Lesen* ganze Wörter in einzelne Silben zerlegt, z. B. durch das Einzeichnen von Silbenbögen. Zugrundeliegende Idee ist die automatisierte Erkennung orthografischer Repräsentationen (Ehri, 2005). Kindern, die noch langsam lesen, fällt es oft schwer, das Schriftbild von Wörtern als Ganzes zu erkennen, wodurch Lesefehler auftreten (Müller et al., 2013; Müller & Richter, 2017). Durch die Erfassung und Einprägung einzelner Silben werden Silben mit der Zeit nicht mehr Buchstabe für Buchstabe erlesen, sondern als ganze Einheiten im Sichtwortschatz abgespeichert. Diese Silbeneinheiten können dann leichter zu ganzen Wörtern zusammengezogen werden, wodurch Lesefehler sukzessive vermieden werden. Eine systematische Förderung des Erkennens von Silben zeigt auch in (quasi-)experimentellen Studien erfolgsversprechende Ergebnisse (Müller & Richter, 2017; Müller et al., 2020; Ritter & Scheerer-Neumann, 2009). So stellten beispielsweise Müller et al. (2020) bei leseschwachen Kindern positive Effekte eines silbenbasierten Trainings auf die Fähigkeiten der phonologischen Rekodierung, der Worterkennung und sogar auf das textbasierte Leseverständnis fest.

Ergänzend zum silbenbasierten Lesen haben sich *Lautleseverfahren* zur Förderung der Leseflüssigkeit bewährt. Es wird angenommen, dass sich Leserinnen und Leser durch das laute Aussprechen vom buchstabenweisen Lesen lösen und das Erkennen ganzer Silben und Wörter einüben. Dabei zeigt sich, dass Schülerinnen und Schüler insbesondere von solchen Fördermaßnahmen profitieren, in denen eine unmittelbare Rückmeldung von Lesefehlern enthalten ist (Top-

ping, 1999, 2006). Dies ist z. B. durch Vorlesen in Partnerarbeit möglich, bei dem Lesefehler direkt von einer zweiten Person korrigiert werden können. Umgesetzt wird dieses Prinzip in der Methode des „Paired Reading“ (Topping, 1999). Durch das Lernen am Modell eines lesestärkeren Kindes werden Lesetempo und Betonung verbessert und Leserinnen und Leser gewinnen zunehmend an Sicherheit beim Lesen. Lautleseverfahren mit einer solchen Korrekturkomponente können auch in empirischen Studien positive Effekte auf die Leseflüssigkeit aufweisen (Gold et al., 2013; Rosebrock et al., 2010).

Eine weitere evidenzbasierte Form des Lautleseverfahrens, die bereits im einführenden Kapitel dieses Bands näher beschrieben wurde, ist das „Repeated Reading“ (Samuels, 1979). Beim Repeated Reading wird derselbe Text wiederholt laut vorgelesen, wodurch Leserinnen und Leser ihren Sichtwortschatz erweitern und unbekannte Wörter beim erneuten Lesen schneller wiedererkennen. Die Wirksamkeit des Repeated Reading ist durch eine Vielzahl an Metaanalysen und Überblicksarbeiten bestätigt worden (Hattie, 2009; Rosebrock et al., 2010; Therrien, 2004), weswegen sie oft als Methode der Wahl zur Förderung der Leseflüssigkeit herangezogen wird. Im Vergleich zu Kontrollgruppen finden sich dabei für Ansätze des wiederholenden lauten Lesens im Mittel moderate Effekte um $d = 0.41$ (NICHD, 2000).

Insgesamt besteht vor dem Hintergrund dieser Befundlage zunächst hinreichende Evidenz für die Wirksamkeit des silbenbasierten Lesens und von Lautleseverfahren zur Förderung von Leseflüchtigkeitsprozessen. Gleichzeitig weisen die Befunde eine hohe Heterogenität auf. So berichtet Therrien (2004) in seiner Metaanalyse zwar von der generellen Wirksamkeit von Repeated-Reading-Interventionen, macht aber auch auf variierende Effekte aufmerksam. Bei Schülerinnen und Schülern mit Lernschwierigkeiten stellt er eine gemittelte Effektstärke von $d = 0.79$ fest, während diese bei Schülerinnen und Schülern ohne Lernschwierigkeiten deutlich geringer ausfällt ($d = 0.59$). Abweichend davon stellen Chard et al. (2009) in ihrem Review eine ausbleibende Wirksamkeit von Repeated-Reading-Interventionen bei leistungsschwachen Schülerinnen und Schülern fest. Sie kommen zu dem Schluss, dass sich wiederholtes Lesen möglicherweise nicht für alle Schülerinnen und Schüler gleichermaßen anbietet. Während also die Ansätze des Repeated Reading und des silbenbasierten Lesens generell als Schlüssel zur Entwicklung der Leseflüssigkeit betrachtet werden, deuten heterogene Befunde aus Studien in der schulischen Praxis darauf hin, dass deren Wirksamkeit möglicherweise von spezifischen Bedingungsfaktoren moderiert wird.

2.2 Das Lese-Sportler-Programm als Ansatz zur Förderung der Leseflüssigkeit

Auf Basis der vorgestellten Bedingungsmodelle, die die Entwicklung der Lesekompetenz als mehrstufigen Prozess beschreiben, wurde mit dem „Lese-Sportler“ ein differenziertes Programm zur Förderung der drei zentralen Teilprozesse der Lesegenauigkeit, der Lesegeschwindigkeit und des Leseverständnisses entwickelt (Hebbecke et al., 2020; Kawohl, 2015). Für jeden Teilprozess werden im Lese-Sportler-Programm Materialien¹ bereitgestellt, die im Leseunterricht der zweiten bis vierten Klasse eingesetzt werden können. Die drei Fördermethoden des Lese-Sportler-Programms sind jeweils nach Sportarten (Slalom, Sprint und Kanu) benannt. Die Analogie zum Sport verdeutlicht, dass auch beim Lesen „Trainingsziele“ nur durch intensives Üben erreicht werden können. Die Durchführung der Methoden erfolgt für die Förderung der Leseflüssigkeit (Sprint) und des Leseverständnisses (Kanu) in kooperativer Partnerarbeit, indem zwei vergleichbar lesestarke Kinder wechselseitig entweder die Rolle des übenden „Sportler-Kindes“ oder die Rolle des Feedback gebenden „Trainer-Kindes“ einnehmen. Durch die Rollenwechsel wird ein hohes Maß an aktiver Übungszeit erreicht und die Kinder setzen sich intensiver mit den Methoden auseinander, wenn sie die anleitende Rolle einnehmen. Beim Lese-Slalom zur Förderung der Lesegenauigkeit arbeitet hingegen ein leseschwächeres Kind mit einem lesestarken Kind zusammen. Hier zeigte sich, dass das Trainerkind bei einer homogenen Zusammensetzung der Paare mit der Erkennung und Korrektur von Lesefehlern überfordert war. Entsprechend trainiert in dieser Methode ein lesestärkeres Trainer-Kind mit einem leseschwächeren Sportler-Kind zusammen. Um Lehrkräften eine noch weitergehende Form der Differenzierung im Leseunterricht zu ermöglichen, sind die drei Methoden jeweils in sechs Schwierigkeitsstufen untergliedert. Die Schwierigkeitsstufen unterscheiden sich z. B. in der sprachlichen Ebene (Wort- oder Textebene) und dem Schwierigkeitsgrad der Wörter (Frequenz, Anzahl der Silben). Im Folgenden stellen wir Ziele und Ablauf zweier Methoden, Lese-Slalom und Lese-Sprint, genauer vor, da diese gezielt Leseflüchtigkeitsprozesse fördern. Der Aufbau und die theoretischen Prinzipien des Lese-Sportler-Programms sind in Tabelle 1 zusammenfassend dargestellt.

1 Das Lese-Sportler-Material ist unter folgendem Link frei verfügbar: <https://go.wwu.de/di2download>.

Tabelle 1: Aufbau des Lese-Sportler-Programms

	Lese-Slalom	Lese-Sprint	Lese-Kanu
			
<i>Förderziel</i>	Lesegenauigkeit	Lesegeschwindigkeit	Leseverständnis
<i>Inhaltlicher Fokus</i>	Silbenbasiertes Lesen: Einzeichnen von Silbenbögen und fehlerfreies lautes Vorlesen (Paired Reading)	Wiederholtes Lautes Lesen: Wortlisten und Texte zweimal so schnell und so genau wie möglich lesen (Repeated Reading)	Strategieorientiertes Lesen: Organisations- und Elaborationsstrategien anwenden und Fragen zum Text beantworten (Reciprocal Teaching)
<i>Zusammensetzung der Lesepaare</i>	heterogen: Ein lesestärkeres Trainer-Kind unterstützt das Sportler-Kind.	homogen: Die beiden Kinder wechseln sich in der Rolle des übenden Sportler-Kindes und des anleitenden Trainer-Kindes ab.	

(cc-by-nc 4.0 Arbeitseinheit Diagnostik und Evaluation im schulischen Kontext / Universität Münster)

Zur Förderung der Lesegenauigkeit werden die Prinzipien des silbenbasierten Lesens und des Paired Reading mit dem „Lese-Slalom“ umgesetzt. Ziel bei der Methode des Lese-Slaloms ist es, eine Liste von Wörtern fehlerfrei vorzulesen. Ein lesestarkes Trainer-Kind und ein leseschwächeres Sportler-Kind arbeiten dabei in Partnerarbeit. Während das Trainer-Kind eine Liste von Wörtern langsam und laut vorliest, zeichnet das Sportler-Kind Silbenbögen unter den Wörtern ein. In Anlehnung an das Prinzip des Paired Reading lesen anschließend beide Kinder die Wortliste gemeinsam laut vor. Fühlt sich das Sportler-Kind beim Lesen sicher, kann es ein Allein-Lese-Signal geben (z. B. durch ein Klopfen auf den Tisch) und alleine weiterlesen. Lesefehler werden, sofern sie nicht selbst vom Sportler-Kind verbessert werden, durch das Trainer-Kind korrigiert. Dadurch wird im Sinne des Paired-Reading-Ansatzes sichergestellt, dass nicht flüchtig über Fehler hinweggelesen werden kann.

Zur Förderung der Lesegeschwindigkeit wird beim „Lese-Sprint“ die Methode des Repeated Reading angewandt. Ziel beim „Lese-Sprint“ ist es, möglichst viele Wörter einer Liste bzw. eines Textes innerhalb einer Minute korrekt vorzulesen. Dabei arbeiten zwei ähnlich lesestarke Schülerinnen und Schüler in Partnerarbeit. Während das Sportler-Kind vorliest, markiert das Trainer-Kind Lesefehler und die Anzahl der gelesenen Wörter pro Minute. Nach dem zweiten Vorlesen desselben Textes werden die Rollen getauscht. In aller Regel werden bereits beim zweiten Durchlauf mehr Wörter gelesen, weniger Lesefehler gemacht und es entstehen unmittelbare Erfolgserlebnisse. Neben den beiden

Methoden des „Lese-Slalom“ und des „Lese-Sprint“, mit denen die Leseflüssigkeit gefördert wird, wird beim Lese-Sportler zudem die Methode des „Lese-Kanu“ zur Förderung des Leseverständnisses angeboten. Das „Lese-Kanu“ orientiert sich an Prinzipien strategieorientierter Förderung, wie sie beispielsweise bei der Methode des „Reciprocal Teaching“ (Palincsar & Brown, 1984) umgesetzt werden.

Zur Wirksamkeit des Lese-Sportler-Programms gibt es mittlerweile mehrere längsschnittliche Studien aus der schulischen Praxis (für eine Studienübersicht s. Hebecker et al., 2020). Kawohl (2015) konstatiert nach einem Schuljahr positive Effekte des Lese-Sportlers auf die Leseflüssigkeit für Kinder der dritten und vierten Klasse ($d = 0.20$; $N = 110$). In einigen Studien zeigen sich jedoch keine direkten Effekte der Fördermaßnahmen auf die Leseleistungen der Schülerinnen und Schüler (Hebecker & Souvignier, 2018; Peters et al., 2021). So berichten Peters et al. (2021) bei der Auswertung von sechs Interventionsstudien keine signifikant positiven Effekte einer kombinierten Förderung bestehend aus einer Lernverlaufsdagnostik und dem Lese-Sportler auf die Leseleistungen und -motivation leseschwacher Schülerinnen und Schüler im Vergleich zur Kontrollgruppe ($N = 1346$ Schülerinnen und Schüler aus 264 Klassen). Ähnlich wie dies bei dem Forschungsüberblick in Abschnitt 2.1 deutlich wurde, fallen also auch die Befunde zum Lese-Sportler-Programm heterogen aus.

3 Wirksamkeitsbedingungen von Maßnahmen zur Förderung der Leseflüssigkeit

Bei der praktischen Umsetzung prinzipiell wirksamer Fördermethoden im schulischen Alltag liegen mitunter deutlich geringere Effekte als in kontrollierten Interventionsstudien vor (vgl. auch Slavin et al., 2008). Dabei legt auch die vorgefundene Heterogenität der Effekte nahe, die Frage nach der Wirksamkeit von Maßnahmen vor dem Hintergrund zu betrachten, wie die Programme zur Förderung von Leseflüssigkeit in der schulischen Praxis konkret umgesetzt werden (Schneider & Randel, 2009). In den folgenden Abschnitten fokussieren wir daher drei Faktoren, die die Wirksamkeit von Maßnahmen zur Förderung der Leseflüssigkeit potenziell moderieren können. Zunächst diskutieren wir mit Rückgriff auf das Prinzip der Passung, welche Eigenschaften der Lernenden eine Voraussetzung dafür darstellen, dass sie von Maßnahmen zur Förderung der Leseflüssigkeit profitieren und welche Chancen sich durch diagnosebasierte Förderentscheidungen ergeben. Zweitens thematisieren wir mit dem Aspekt der Dosierung ein Merkmal, das bislang eher wenig Beachtung in der Literatur findet. Wir diskutieren, wie Dosierung konzeptualisiert und erfasst werden kann und beschreiben, welchen Einfluss die Dosierung von Fördermaßnahmen auf die Wirksamkeit der Intervention haben kann. Abschließend erörtern wir

anhand des Konzepts der Wiedergabetreue den Einfluss der Programmanwendung durch die Lehrkraft auf die Wirksamkeit der Fördermaßnahme und skizzieren die eher uneinheitlichen Befunde zum Zusammenhang von Wiedergabetreue und Schülerleistungen.

3.1 Passung

Die Entwicklung der Lesekompetenz wird durch eine Vielzahl an kognitiven, motivationalen und kontextuellen Faktoren beeinflusst. Entsprechend fallen die Vorläuferfähigkeiten und die frühen Lesefähigkeiten von Grundschülerinnen und Grundschulern bereits bei Schuleintritt sehr unterschiedlich aus (Frith, 1986; NICHD, 2000). Dass diese Heterogenität auch am Ende der Grundschulzeit persistiert, wird beispielsweise in den Ergebnissen der IGLU-Studie deutlich: So fiel die Leistungsstreuung bei deutschen Schülerinnen und Schülern der vierten Klasse im internationalen Vergleich besonders hoch aus (Hussmann et al., 2017). Pierce et al. (2007) identifizierten unter leseschwachen Kindern der zweiten und dritten Klassenstufe vier verschiedene Gruppen, die beim Lesen unterschiedliche Leistungsmuster aufwiesen. Cluster 1 erzielte unterdurchschnittliche Werte in der Worterkennung und der Lesegenauigkeit, aber durchschnittliche Werte im Wortschatz und im Dekodieren. In Cluster 2 lagen die Leistungen bei der Worterkennung und im Wortschatz noch niedriger, während die Fähigkeiten der Lesegenauigkeit und des Dekodierens unauffällig ausfielen. Cluster 3 zeigte, ähnlich wie Cluster 1, unterdurchschnittliche Werte in der Worterkennung und der Lesegenauigkeit, aber deutlich niedrigere Fähigkeiten im Wortschatz. Cluster 4 zeigte schließlich in allen vier Bereichen Schwächen. Vor dem Hintergrund dieser unterschiedlichen Leistungsmuster argumentieren Pierce et al. (2007), dass diese vier Gruppen jeweils eigene Förderbedürfnisse haben, sodass Methoden zur Leseförderung vermutlich nicht bei allen Kindern in gleicher Weise funktionieren. Um die individuellen Lernstände und -entwicklungen der Schülerinnen und Schüler zu berücksichtigen, wird in einigen Interventionen explizit das Prinzip der Passung zum Untersuchungsgegenstand gemacht (Bernard et al., 2019; Connor, 2019). Dabei zeigte sich, dass sich die Leistungen der Schülerinnen und Schüler insbesondere dann verbessern, wenn eine enge Passung zwischen dem Lernstand der Kinder und der eingesetzten Fördermethode bzw. der Schwierigkeit eingesetzter Texte gegeben war. Wie hoch die Effekte prinzipiell wirksamer Methoden wie silbenbasiertes oder wiederholtes Lesen ausfallen, scheint also davon abzuhängen, wie passend sie für die individuellen Lernvoraussetzungen eines Kindes sind. Im Folgenden sollen einige Befunde zu der Frage vorgestellt werden, welche konkreten Merkmale der Lernenden eine Voraussetzung für die Wirksamkeit dieser Methoden darstellen.

Karageorgos et al. (2019) untersuchten differentielle Effekte individueller Lernvoraussetzungen von Leserinnen und Leser auf die Wirksamkeit einer Intervention zum Wortlesen. Die Befunde legen nahe, dass Effekte auf die Lesegeschwindigkeit davon abhängen, ob die Viertklässlerinnen und Viertklässler bereits über grundlegende Fähigkeiten der Lesegenauigkeit verfügen. Damit konnten Karageorgos et al. (2019) Befunde einer ähnlich angelegten Studie aus Dänemark replizieren (Poulsen et al., 2015). Auch Fuchs et al. (2021) wiesen Moderatoreffekte von Fähigkeiten der phonologischen Bewusstheit auf die Wirksamkeit des Repeated Reading nach. Sie verglichen den Lernzuwachs von Schülerinnen und Schülern, die anhand kooperativer, silbenbasierter Lautleseverfahren – mit und ohne Repeated-Reading-Komponente – die Worterkennung und die Leseflüssigkeit trainierten, mit einer Kontrollgruppe. Dabei zeigten sich in beiden Experimentalgruppen stärkere Effekte der Trainings auf die Entwicklung der phonologischen Bewusstheit und der Worterkennung für diejenigen Schülerinnen und Schüler, die vor der Intervention geringe Fähigkeiten der phonologischen Bewusstheit aufwiesen. Gleichzeitig wird beim Vergleich der beiden Interventionen deutlich, dass die leistungsschwächeren Schülerinnen und Schüler *ohne* Repeated Reading-Komponente einen stärkeren Lernzuwachs erreichten. Diese differentiellen Effekte stehen in Einklang mit den Befunden von Chard et al. (2009), dass die Methode des Repeated Reading nicht die Methode der Wahl für die leseschwächsten Schülerinnen und Schüler sein sollte. So resümieren Fuchs et al. (2021): „If teachers are most concerned about their weaker readers, then adoption of [a] program that emphasizes PA [phonological awareness] and word recognition makes sense.“ (S. 18). Eine zentrale Schlussfolgerung dieser Studienergebnisse für die Leseförderung in der unterrichtlichen Praxis ist die Notwendigkeit einer stärkeren Differenzierung: Leserinnen und Leser, die noch viele Fehler machen, sollten zunächst mit Methoden zur Förderung der Lesegenauigkeit trainieren, also z. B. mit dem Ansatz des silbenbasierten Lesens. Umgekehrt sollten für Leserinnen und Leser, die schon genau, aber noch nicht angemessen schnell lesen, Maßnahmen zur Förderung der Lesegeschwindigkeit wie das Repeated Reading angeboten werden.

Um Entscheidungen über solchermaßen passende Förderangebote machen zu können, ist eine Diagnose des Lernstands bei Schülerinnen und Schülern eine Voraussetzung. Als beispielgebend dafür können die Arbeiten von Connor (2019) gelten, in deren Arbeitsgruppe das Online-Tool „Assessment-to-Instruction“ (A2i) entwickelt wurde. A2i stellt Lehrkräften diagnostische Informationen über die Lesefertigkeiten ihrer Schülerinnen und Schüler zur Verfügung und verknüpft diese mit Empfehlungen zu passenden – differenzierten – Förderangeboten. Die Ergebnisse aus sieben Kontrollgruppenstudien weisen darauf hin, dass sich die Leistungen derjenigen Schülerinnen und Schüler innerhalb eines Schuljahres signifikant verbesserten ($d = 0.2-0.4$), deren Lehrkräften diagnostische Informationen durch das Online-Tool A2i zur Verfügung

standen. Diese Effekte kumulieren sich im Verlauf von der ersten bis zur dritten Klassenstufe zu einer Effektstärke von $d = 0.7$ (Connor, 2019). Im sonderpädagogischen Kontext wurden bereits seit den 1980er Jahren Untersuchungen zur Wirksamkeit einer Kombination aus Lernverlaufsdagnostik und an Lernstand und Leistungsentwicklung von Kindern angepasster Förderung durchgeführt (für einen Überblick s. Stecker et al., 2005). Positive Effekte einer passgenauen Förderung unterstreichen auch hier, wie sinnvoll ein gezieltes Zusammenspiel aus Diagnose und Förderung ist.

Auch bei dem Lese-Sportler-Programm wird eine solche Kombination aus differenziertem Fördermaterial und einer begleitenden Diagnostik realisiert. Dazu wird die internetbasierte Lernverlaufsdagnostik „quop“ genutzt (Souvignier et al., 2021). Bei diesem diagnostischen Ansatz bearbeiten die Schülerinnen und Schüler alle drei Wochen einen zehninütigen Lesetest am Computer, bei dem die drei Teilkompetenzen Lesegenauigkeit, Lesegeschwindigkeit, Leseverständnis erfasst werden. Anhand der Testergebnisse wird unmittelbar eine Rückmeldung gegeben, welcher Aspekt der Lesekompetenz im Fokus der Förderung eines Kindes stehen sollte. Aufgrund des kontinuierlichen Wechselspiels zwischen Förderung und Diagnose wird zudem kurzfristig festgestellt, ob Anpassungen im Hinblick auf die gewählte Fördermethode angezeigt sind. Aktuelle Studien zur Implementation dieses Ansatzes zeigen, dass die Kombination aus dem „Lese-Sportler“ mit der Lernverlaufsdagnostik „quop“ sich prinzipiell in der schulischen Praxis umsetzen lässt (Förster et al., 2018; Hebecker & Souvignier, 2018; Souvignier et al., 2016). Befragungen der Lehrkräfte weisen allerdings auch darauf hin, dass die Implementation eine Veränderung unterrichtlicher Routinen erfordert.

Dass eine hohe Passung zwischen individuellem Lernstand des Kindes und angebotener Fördermaßnahme entscheidend für die Wirksamkeit von Fördermaßnahmen ist, ist zunächst wenig überraschend. Für Maßnahmen zur Förderung der Leseflüssigkeit bedeutet das, dass sie idealerweise dann zum Einsatz kommen sollten, wenn Kinder hinreichend genau lesen können, Verbesserungen hinsichtlich der Leseflüssigkeit aber noch nötig sind, um Prozesse sinnentnehmenden Lesens erleichtern zu können. Diagnostische Informationen stellen dabei für Lehrkräfte eine objektive Entscheidungsgrundlage dar, um individuell passende Methoden auszuwählen. Eine für den schulischen Alltag praktikable Methode ist dabei die Implementation computerbasierter Assessments. Aus einer Forschungsperspektive ist festzuhalten, dass weiterer Bedarf an Moderatoranalysen zu Merkmalen von Lernenden besteht, um eine optimale Passung zwischen Förderangeboten zur Leseflüssigkeit und individuellen Lernvoraussetzungen ausmachen zu können (Fuchs & Fuchs, 2019).

3.2 Dosierung

Die Dosierung – im Sinne des aufgewendeten Umfangs, in dem Leserinnen und Leser trainieren – ist ein zweiter potenzieller Einflussfaktor auf die Wirksamkeit. Bereits die Gedächtnisexperimente von Ebbinghaus (1885) wiesen auf den Wert höherer Lernzeit hin und auch Snoddy (1926) identifizierte mit seinem „Power Law of Practice“ einen Lernkurveneffekt: Die Reaktionszeit für die Bearbeitung von Aufgaben nahm linear mit der Anzahl durchgeführter Versuche ab. In ähnlicher Weise stellten Ericsson et al. (1993) beim Vergleich von professionellen Musikerinnen und Musikern mit Hobby-Musizierenden resümierend fest, dass die Hochleistungen der Expertinnen und Experten vielmehr das Ergebnis intensiven Übens als angeborenen Talents sind.

Als Operationalisierung für die Dosierung von Interventionen zur Leseförderung liegt es nahe, Angaben zum zeitlichen Umfang entsprechender Programme zu erheben (Hammerschmidt-Snidarich et al., 2019; Torgesen, 1998). Bei einem Großteil der Studien zur Wirksamkeit von Fördermaßnahmen der Leseflüssigkeit wird die Dosierung in diesem Sinne erfasst (Maki et al., in Druck). Alternativ bzw. ergänzend dazu schlagen Hammerschmidt-Snidarich et al. (2019) vor, die Dosierung anhand der Anzahl gelesener Wörter zu erfassen. Die Unterschiede zwischen der Konzeptualisierung der Dosierung über die Lesezeit oder die Lesemenge liegen auf der Hand: Schnellere Leserinnen und Leser lesen in der gleichen Zeit mehr Wörter als langsamere Leserinnen und Leser. Tatsächlich beobachteten Hammerschmidt-Snidarich et al. (2018), dass lesestarke Kinder bei der Durchführung eines Repeated-Reading-Trainings mehr als doppelt so viele Wörter lasen wie leseschwächere Kinder, wenn die Kinder an der gleichen Anzahl an Sitzungen teilnahmen. Dass die Konzeptualisierung von Dosierung Konsequenzen für die Interpretation der Effekte in Interventionsstudien haben kann, illustrieren Hammerschmidt-Snidarich et al. (2019) in ihrer Studie zum Wirksamkeitsvergleich von Repeated Reading und Continuous Reading, dem Lesen vieler, aber verschiedener Texte. Dabei erfassen sie als Dosierungsmaße sowohl die Trainingszeit als auch die Anzahl der gelesenen Wörter. Unter der Bedingung, dass langsamere und schnellere Leserinnen gleich viele Wörter lasen, fielen die Trainingseffekte vergleichbar aus. Bei identischer Förderzeit erzielten leistungsstärkere Kinder hingegen höhere Leistungszuwächse im Lesen. Diese Befunde verdeutlichen, dass eine Operationalisierung der Dosierung über die reine Übungszeit potenziell damit einhergeht, dass sich interindividuelle Leistungsunterschiede vergrößern (Matthäus-effekt, Stanovich, 2009). Im Hinblick auf das in Abschnitt 3.1. vorgeschlagene Ziel einer hohen Passung läge es daher nahe, leseschwächeren Kindern bezüglich des Übungsumfangs intensivere Angebote zur Verfügung zu stellen, was z. B. im Rahmen von zusätzlichen Förderstunden realisiert werden könnte (Snidarich, 2015).

Gleichzeitig wird bei der Suche nach verfügbaren Studien zur Dosierung auch ein Mangel an empirischer Evidenz deutlich. In einem aktuellen meta-analytischen Review werteten Maki et al. (in Druck) mehrere Interventionsstudien zur Förderung der Leseflüssigkeit aus. Generell stellten sie hier eine hohe Wirksamkeit fest ($d = 0.44$). Angaben zur Dosierung der Fördermaßnahme wurden allerdings nur in fünf Studien erfasst – jeweils operationalisiert über die Lesezeit. Keine Studie untersuchte direkte Effekte der Menge gelesener Texte auf die Lernleistungen der Schülerinnen und Schüler. Jenseits der differenzierten Befunde von Hammerschmidt-Snidarich et al. (2019) zur Bedeutung des konkreten Umfangs gelesener Texte spricht mit Blick auf die Studienlage dennoch vieles für den Einfluss der Übungszeit auf die Wirksamkeit von Fördermaßnahmen. Demzufolge ist ein möglichst hohes Maß an Übungszeit anzustreben. Entsprechend sind vorliegende Interventionen zur Förderung der Leseflüssigkeit in aller Regel als zeitintensive Maßnahmen konzipiert. So wird beim Lese-Sportler-Programm ein Übungsumfang von 20 Minuten an drei Tagen pro Woche empfohlen. In ähnlicher Weise sollen Schülerinnen und Schüler bei anderen Förderprogrammen im Laufe eines Schuljahres zwischen 18 Sitzungen (Ritter & Scheerer-Neumann, 2009) und ca. 40 Sitzungen (Rosebrock et al., 2010) mit den Materialien arbeiten. Bislang mangelt es allerdings an Studien zu der Frage, welchen Effekt solche konkreten Empfehlungen zur Dosierung auf die Lernleistungen haben. Die Annahme liegt nahe, dass es hinsichtlich des Übungsumfangs einen abnehmenden Grenznutzen gibt, bei dem die Grundregel „je mehr, desto besser“ ab einem bestimmten Zeitpunkt nicht länger gilt. Dennoch: Betrachtet man die geringe Anzahl an Minuten, die Lehrkräfte laut Selbstauskünften in einer Woche für den Leseunterricht aufwenden (136 Minuten, internationaler Mittelwert: 160 Minuten, vgl. Hussmann et al., 2017, S. 280), so erscheint eine Intensivierung in jedem Falle erforderlich, zumal davon ausgegangen werden kann, dass evidenzbasierte Methoden im Regelunterricht eher selten zum Einsatz kommen (Philipp, 2014).

Insgesamt wird deutlich, dass nicht nur eine hohe Passung zwischen dem individuellen Lernstand und der Fördermethode, sondern auch eine individuell angemessene Dosierung im Hinblick auf die Intensität von Förderprogrammen wichtig für die optimale Entwicklung der Lesekompetenz ist. Forschungsseitig besteht ein erheblicher Bedarf an Studien, die sich dem Thema der Dosierung explizit widmen. Dabei müssten verschiedene Interventionskomponenten unter Bedingungen verglichen werden, bei denen die Dosierung sowohl als Lesezeit als auch als Anzahl der gelesenen Wörter gemessen wird (Hammerschmidt-Snidarich et al., 2019).

3.3 Wiedergabetreue

Mit den in Abschnitt 2 vorgestellten evidenzbasierten Methoden wirksamer Leseförderung und dem Lese-Sportler-Programm werden konkrete Vorgehensweisen für eine wirksame Förderung der Leseflüssigkeit beschrieben. Es scheint naheliegend, dass von einer Wirksamkeit solcher konkreten Methoden und Programme nur dann ausgegangen werden kann, wenn sie mit einem hohen Maß an Wiedergabetreue (fidelity) umgesetzt werden (Noell, 2014). Im Umkehrschluss bedeutet das, dass sich heterogene Befunde zur Wirksamkeit solcher Fördermaßnahmen möglicherweise durch Unterschiede hinsichtlich der Wiedergabetreue erklären lassen.

Das Konstrukt der Wiedergabetreue selbst, entsprechende Messansätze und auch die Variablenkonstruktion weisen eine große Vielfalt auf (Dane & Schneider, 1998). So werden mit den Begriffen *treatment integrity* (Yeaton & Sechrest, 1981), *treatment fidelity* (Moncher & Prinz, 1991) und *implementation fidelity* (O'Donnell, 2008) Nuancierungen vorgenommen, die allerdings nicht einheitlich verwendet werden (für einen Überblick s. Swanson et al., 2013; Dusenbury et al., 2003). Weitgehender Konsens besteht mit Blick auf die von Dane und Schneider (1998) vorgenommene Unterscheidung von fünf Facetten der Wiedergabetreue: 1. „adherence“ (Genauigkeit, mit der die Kernelemente eines Programms umgesetzt werden), 2. „dosage“ (Häufigkeit und Dauer der Teilnahme), 3. „quality of delivery“ (Qualität der Umsetzung), 4. „participant responsiveness“ (Beteiligung und Engagement der Teilnehmenden) und 5. „program differentiation“ (Differenziertheit des Programms und Unterschiedlichkeit des Programms zu ‚Business-as-usual‘-Settings). Während beispielsweise der Aspekt der „adherence“ die Perspektive einer präzisen Umsetzung vorgegebener Programminhalte in den Fokus nimmt, betont „participant responsiveness“, dass eine hohe Wiedergabetreue auch dann vorliegen kann, wenn situationsangemessene Anpassungen vorgenommen werden. Gerade mit Blick auf die schulische Praxis wird die Uneinheitlichkeit des Konstrukts der Wiedergabetreue deutlich: Wenn eine Innovation gemäß ihrer Konzeption in der unterrichtlichen Praxis umgesetzt wird (Mihalic, 2004; O'Donnell, 2008), dann kann das bedeuten, dass zentrale Wirkprinzipien in einer an den individuellen Bedürfnissen der Schülerinnen und Schüler angepassten Weise vermittelt werden. Es ist aber auch denkbar, nur dann von einer hohen Wiedergabetreue zu sprechen, wenn Maßnahmen im Hinblick auf Ablauf und Umfang exakt konkreten Vorgaben entsprechen. Forschungsseitig findet sich vor allem eine Orientierung an der Frage, ob konkrete Vorgaben (Inhalte, Dosierung) umgesetzt wurden. Während Fragen zu einer angemessenen Adaptivität oder Qualität bei der Durchführung von Maßnahmen Interpretationsspielräume (und damit potenziell Einbußen hinsichtlich der Beurteilerübereinstimmung) mit sich bringen,

führt ein Vergleich zwischen konkreten Soll- und Ist-Werten typischerweise zu reliableren Ergebnissen (O'Donnell, 2008).

Konkret liegen einzelne Befunde zum Zusammenhang zwischen der Wiedergabetreue und dem Lernzuwachs der Schülerinnen und Schüler vor. So stellt Kuhn (2018) für die Sekundarstufe fest, dass sich ein Teil der Varianz zwischen Klassen, die das gleiche Treatment erhielten, durch Unterschiede in der Wiedergabetreue, also z. B. durch einen unterschiedlichen Umfang in der Durchführung der Interventionsmaßnahmen erklären ließ. Ähnliche Befunde berichten auch Vadasy und Sanders (2009) für Grundschülerinnen und Grundschüler. Je mehr sich die Lehrkräfte am empfohlenen Durchführungsumfang und an den Vorgaben des Handbuchs zur Durchführung einer Repeated-Reading-Intervention orientierten, desto höher fielen die Lernzuwächse ihrer Schülerinnen und Schüler aus. Nunnery et al. (2006) erfassen Wiedergabetreue anhand mehrerer Komponenten wie z. B. Dosierung und Qualität. Im Gegensatz zu den bereits genannten Studien berichten sie allerdings, dass die Wiedergabetreue die Varianz zwischen den Leistungsunterschieden nur für Schülerinnen und Schüler mit Behinderung, nicht aber für die Gesamtstichprobe erklären kann. So ist der konkrete Zusammenhang zwischen der Wiedergabetreue und den Leistungen der Schülerinnen und Schüler noch unklar. Capin et al. (2018) stoßen in ihrem Review zur Wiedergabetreue in Leseinterventionen auch nur auf vier Studien, die überhaupt Zusammenhänge zwischen der Wiedergabetreue und der Wirksamkeit der Intervention analysieren. Diese knappen Befunde weisen darauf hin, dass sich bisher nur eine geringe Anzahl von Studien diesem Thema explizit widmet und an dieser Stelle weiterer Forschungsbedarf besteht.

Befunde aus Befragungen und Beobachtungen zum Einsatz der Kombination aus Lernverlaufdiagnostik und Lese-Sportler-Programm legen nahe, dass bei diesem umfassenden Ansatz die Variation zwischen Lehrkräften, die die Maßnahme wie geplant umsetzten und Lehrkräften, die größere Veränderungen vornahmen, hoch ausfiel (Hebbecker & Souvignier, 2018). So gaben einige Lehrkräfte an, durch Auslassen von Inhalten und Anpassungen der Methoden Änderungen bei der Durchführung vorgenommen zu haben. Vor dem Hintergrund einer solchen Diskrepanz zwischen den Anforderungen des Konzepts und der tatsächlich gelebten Schulpraxis ist ein Rückschluss auf die Wirksamkeit der Methoden nicht sinnvoll.

Um ein hohes Maß an Wiedergabetreue in der unterrichtlichen Praxis zu erzielen, empfehlen Hondrich et al. (2016) konkrete Materialien zu den Förderkonzepten zur Verfügung zu stellen. Dies entspricht dem von Lipowsky und Rzejak (2019) in ihrer Analyse wirksamer Fortbildungsmaßnahmen genannten Prinzip einer starken Orientierung an den Bedürfnissen der Zielpersonen im Implementationsprozess. Entsprechend werden beim Lese-Sportler-Programm zusätzlich zu den Aufgabenheften mittlerweile auch kindgerechte Ablaufpläne

der Methoden für den Aushang im Klassenzimmer sowie einseitige Kurzbeschreibungen der Methoden für die Lehrkräfte bereitgestellt. Zudem wurden Erklärvideos² für Lehrkräfte entwickelt, in denen neben theoretischen Grundlagen auch unterrichtspraktische Fragen thematisiert werden.

Zusammenfassend lassen sich für den Aspekt der Wiedergabetreue zwei Perspektiven ableiten: Zunächst erscheint es wichtig, dass die Wiedergabetreue in Evaluationsstudien standardmäßig – und unter Beachtung mehrerer Facetten des Konstrukts – erfasst wird, um Aussagen über die Wirksamkeit einer Intervention machen zu können (O’Donnell, 2008). Zweitens deutet sich an, dass Lehrkräfte gezielter Unterstützung bedürfen, um Maßnahmen zur Förderung der Leseflüssigkeit im Unterricht wiedergabetreu einsetzen zu können (Hebbecke & Souvignier, 2018). Hier empfiehlt es sich, bereits bei der Entwicklung der Intervention eine enge Orientierung an den Adressatinnen und Adressaten und den schulischen Rahmenbedingungen vorzunehmen und Lehrkräften praktikable Materialien und Informationsangebote zur Verfügung zu stellen (Lipowsky & Rzejak, 2019).

4 Fazit

Eine Vielzahl an wissenschaftlichen Studien zur Förderung der Leseflüssigkeit gibt Hinweise auf wirksame Förderansätze für den Leseunterricht. Insbesondere liefern (quasi-)experimentelle Studien hinreichend Evidenz dafür, dass Methoden wie das silbenbasierte Lesen und (wiederholtes) lautes Lesen sowie die Organisationsform des Partnerlesens mit unmittelbarem Feedback als zentrale Wirkkomponenten gelten können (NICHD, 2000). Gleichzeitig weisen die Befunde von Studien in der unterrichtlichen Praxis inkonsistente Ergebnismuster und mitunter auch geringe Effekte aus (Souvignier, 2020). Vor diesem Hintergrund stellt sich die Frage nach moderierenden Faktoren, beispielsweise mit Blick auf die Implementationsbedingungen, die die Wirksamkeit der Maßnahmen beeinflussen können (Fuchs & Fuchs, 2019; Souvignier, 2020).

Mit der Passung, der Dosierung und der Wiedergabetreue sind wir auf drei dieser potenziell moderierenden Faktoren eingegangen. Auf den ersten Blick mag die Auseinandersetzung mit diesen Aspekten überraschen. Denn es erscheint unmittelbar plausibel, dass auf die individuellen Bedürfnisse zugeschnittene Maßnahmen besonders wirksam sind, dass es grundsätzlich förderlich ist, eine möglichst große Zahl an Übungsmöglichkeiten anzubieten und dass Maßnahmen umso wirksamer sind, je enger man sich an Empfehlungen zu deren Durchführung hält. Trotz dieser theoretischen Nachvollziehbarkeit feh-

2 Unser YouTube-Kanal: <https://www.wwu.de/PsyIPBE.Download/YT>.

len empirische Studien zum Einfluss moderierender Faktoren. Diese erscheinen insbesondere vor dem Hintergrund erforderlich, dass Untersuchungen zur Bestandsaufnahme schulischer Praxis zeigen, dass differenzierte Förderung und die Nutzung evidenzbasierter Methoden eher selten sind (Philipp, 2014).

Ziel bei der Entwicklung des Lese-Sportler-Programms war daher zunächst, drei evidenzbasierte Methoden für eine differenzierte Leseförderung so aufzubereiten, dass sie sich leicht im unterrichtlichen Alltag umsetzen lassen. Damit eine optimale Passung zwischen dem individuellen Lernstand und der angebotenen Fördermethode erreicht wird, wurde das Material mit der Lernverlaufsdiagnostik quop verknüpft, anhand derer über das Schuljahr hinweg fortlaufend geprüft wird, welche Fördermethode passend für ein Kind ist (Souvignier et al., 2021). Das Lese-Sportler-Programm wurde frei verfügbar über unsere Website zugänglich gemacht, sodass Lehrkräfte unmittelbar auf die jeweiligen Materialien zugreifen können. Frei editierbare Vorlagen ermöglichen es Lehrkräften, eigene Texte einzubinden. Informationsmaterialien wie ein Handbuch für Lehrkräfte vermitteln Anwendungswissen zur Nutzung des Förderkonzepts. Schließlich wurden Erklärvideos erstellt, anhand derer Informationen zu theoretischen Hintergründen und zur praktischen Anwendung gegeben werden. Ein solches rahmendes Konzept aus Diagnostik, Material und Information soll dazu beitragen, die Implementation des Lese-Sportler-Programms – nicht zuletzt im Hinblick auf Passung, Dosierung und Wiedergabetreue – in die schulische Praxis zu unterstützen.

Literatur

- Bernard, R. M., Borokhovski, E., Schmid, R. F., Waddington, D. I. & Pickup, D. I. (2019): Twenty-First Century Adaptive Teaching and Individualized Learning Operationalized as Specific Blends of Student-Centered Instructional Events: A Systematic Review and Meta-Analysis. *Campbell Systematic Reviews*, 15(1–2), 1–35. <https://doi.org/10.1002/cl2.1017>
- Bertschi-Kaufmann, A. & Hagendorf, P. (2008): *Lesefertigkeiten. Lesegeläufigkeit. Lesestrategien* (2. Aufl.). *Lesen – das Training ab Klasse 5*. Seelze: Friedrich.
- Capin, P., Walker, M. A., Vaughn, S. & Wanzek, J. (2018): Examining How Treatment Fidelity Is Supported, Measured, and Reported in K-3 Reading Intervention Research. *Educational Psychology Review*, 30(3), 885–919. <https://doi.org/10.1007/s10648-017-9429-z>
- Chard, D. J., Ketterlin-Geller, L. R., Baker, S. K., Doabler, C. & Apichatabutra, C. (2009): Repeated Reading Interventions for Students with Learning Disabilities: Status of the Evidence. *Exceptional Children*, 75(3), 263–281. <https://doi.org/10.1177/001440290907500301>
- Connor, C. M. (2019): Using Technology and Assessment to Personalize Instruction: Preventing Reading Problems. *Prevention Science*, 20(1), 89–99. <https://doi.org/10.1007/s11121-017-0842-9>
- Dane, A. V. & Schneider, B. H. (1998): Program Integrity in Primary and Early Secondary Prevention: Are Implementation Effects out of Control? *Clinical Psychology Review*, 18(1), 23–45. [https://doi.org/10.1016/s0272-7358\(97\)00043-3](https://doi.org/10.1016/s0272-7358(97)00043-3)

- Dusenbury, L., Brannigan, R., Falco, M. & Hansen, W. B. (2003): A Review of Research on Fidelity of Implementation: Implications for Drug Abuse Prevention in School Settings. *Health Education Research*, 18(2), 237–256. <https://doi.org/10.1093/her/18.2.237>
- Ebbinghaus, H. (1885): *Über das Gedächtnis: Untersuchungen zur experimentellen Psychologie*. Leipzig: Duncker & Humblot.
- Ehri, L. C. (2005): Learning to Read Words: Theory, Findings, and Issues. *Scientific Studies of Reading*, 9(2), 167–188. https://doi.org/10.1207/s1532799xssr0902_4
- Ericsson, K. A., Krampe, R. T. & Tesch-Römer, C. (1993): The Role of Deliberate Practice in the Acquisition of Expert Performance. *Psychological Review*, 100(3), 363–406. <https://doi.org/10.1037//0033-295X.100.3.363>
- Förster, N., Kawohl, E. & Souvignier, E. (2018): Short- and Long-Term Effects of Assessment-Based Differentiated Reading Instruction in General Education on Reading Fluency and Reading Comprehension. *Learning and Instruction*, 56, 98–109. <https://doi.org/10.1016/j.learninstruc.2018.04.009>
- Frith, U. (1986): A Developmental Framework for Developmental Dyslexia. *Annals of Dyslexia*, 36(1), 67–81. <https://doi.org/10.1007/BF02648022>
- Fuchs, D. & Fuchs, L. S. (2019): On the Importance of Moderator Analysis in Intervention Research: An Introduction to the Special Issue. *Exceptional Children*, 85(2), 126–128. <https://doi.org/10.1177/0014402918811924>
- Fuchs, D., Cho, E., Toste, J. R., Fuchs, L. S., Gilbert, J. K., McMaster, K. L., ... Thompson, A. (2021): A Quasiexperimental Evaluation of Two Versions of First-Grade PALS: One with and One without Repeated Reading. *Exceptional Children*, 87(2), 141–162. <https://doi.org/10.1177/0014402920921828>
- Fuchs, L. S., Fuchs, D., Hosp, M. K. & Jenkins, J. R. (2001): Oral Reading Fluency as an Indicator of Reading Competence: A Theoretical, Empirical, and Historical Analysis. *Scientific Studies of Reading*, 5(3), 239–256. https://doi.org/10.1207/S1532799XSSR0503_3
- Gold, A. (2009): Leseflüssigkeit. In A. Bertschi-Kaufmann & C. Rosebrock (Hrsg.), *Lesesozialisation und Medien. Literalität: Bildungsaufgabe und Forschungsfeld* (S. 151–164). Weinheim: Juventa.
- Gold, A., Behrendt, S., Lauer-Schmaltz, M. & Rosebrock, C. (2013): Förderung der Leseflüssigkeit in dritten Grundschulklassen. In C. Rosebrock & A. Bertschi-Kaufmann (Hrsg.), *Lesesozialisation und Medien. Literalität erfassen: bildungspolitisch, kulturell, individuell* (S. 203–218). Weinheim, Basel: Beltz-Juventa.
- Hammerschmidt-Snidarich, S. M., Maki, K. E. & McComas, J. J. (2018): Repeated Reading to Improve Oral Reading Fluency: The Predictive Effects of Student Inputs. Manuscript in preparation.
- Hammerschmidt-Snidarich, S. M., Maki, K. E. & Adams, S. R. (2019): Evaluating the Effects of Repeated Reading and Continuous Reading Using a Standardized Dosage of Words Read. *Psychology in the Schools*, 56(5), 635–651. <https://doi.org/10.1002/pits.22241>
- Hattie, J. A. (2009): *Visible Learning: A Synthesis of over 800 Meta-Analyses Relating to Achievement*. London: Routledge.
- Hebbecke, K. & Souvignier, E. (2018): Formatives Assessment im Leseunterricht der Grundschule – Implementation und Wirksamkeit eines modularen, materialgestützten Konzepts. *Zeitschrift für Erziehungswissenschaft*, 21(4), 735–765. <https://doi.org/10.1007/s11618-018-0834-y>

- Hebbecke, K., Förster, N., Forthmann, B., Heyne, L., Peters, M., Salaschek, M. & Souvignier, E. (2020): Diagnostik, Feedback und differenzierte Leseförderung. Umsetzung evidenzbasierter Konzepte im schulischen Alltag. Retrieved from www.leseforum.ch
- Hondrich, A. L., Hertel, S., Adl-Amini, K. & Klieme, E. (2016): Implementing Curriculum-Embedded Formative Assessment in Primary School Science Classrooms. *Assessment in Education: Principles, Policy & Practice*, 23(3), 353–376. <https://doi.org/10.1080/0969594X.2015.1049113>
- Hussmann, A., Wendt, H., Bos, W., Bremerich-Vos, A., Kasper, D., Lankes, E.-M., ... Valtin, R. (Hrsg.) (2017): *IGLU 2016: Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster, New York: Waxmann.
- Karageorgos, P., Müller, B. & Richter, T. (2019): Modelling the Relationship of Accurate and Fluent Word Recognition in Primary School. *Learning and Individual Differences*, 76, Article 101779. <https://doi.org/10.1016/j.lindif.2019.101779>
- Kawohl, E. (2015): *Diagnosebasierte individuelle Leseförderung in der Grundschule*. Dissertation. Hamburg: Kovac.
- Kintsch, W. (1998): *Comprehension. A Paradigm for Cognition*. Cambridge: University Press.
- Klicpera, C., Gasteiger-Klicpera, B. & Schabmann, A. (1993): *Lesen und Schreiben – Entwicklung und Schwierigkeiten: Die Wiener Längsschnittuntersuchungen über die Entwicklung, den Verlauf und die Ursachen von Lese- und Schreibschwierigkeiten in der Pflichtschulzeit*. Bern: Huber.
- Kuhn, D. (2018): *Implementationsgrad und Wirksamkeit von Unterrichtsinterventionen zur Förderung der selbstbestimmten Lernmotivation*. Dissertation. Pädagogische Hochschule, Ludwigsburg.
- Lipowsky, F. & Rzejak, D. (2019): Was macht Fortbildungen für Lehrkräfte erfolgreich? – Ein Update. In B. Groot-Wilken & R. Körber (Hrsg.), *Nachhaltige Fortbildungen für Lehrerinnen und Lehrer: Ideen, Entwicklungen, Konzepte* (S. 15–56). Bielefeld: WBV.
- Maki, K. E., Hamerschmidt-Snidarich, S. M. & Adams, S. R. (in Druck): Novel Research Synthesis of Reading Fluency Interventions: Review of Dosage & Effectiveness of Reading Fluency Intervention. Manuscript submitted for publication.
- Mihalic, S. (2004): The Importance of Implementation Fidelity. *Emotional & Behavioral Disorders in Youth*, 4(4), 83–86.
- Moncher, F. J. & Prinz, R. J. (1991): Treatment Fidelity in Outcome Studies. *Clinical Psychology Review*, 11(3), 247–266. [https://doi.org/10.1016/0272-7358\(91\)90103-2](https://doi.org/10.1016/0272-7358(91)90103-2)
- Müller, B. & Richter, T. (2017): Förderung hierarchieniedriger Leseprozesse. In M. Philipp (Hrsg.), *Handbuch: Schriftspracherwerb und weiterführendes Lesen und Schreiben* (S. 236–251). Weinheim, Basel: Beltz Juventa.
- Müller, B., Križan, A., Hecht, T., Richter, T. & Ennemoser, M. (2013): Leseflüssigkeit im Grundschulalter: Entwicklungsverlauf und Effekte systematischer Leseförderung. *Lernen und Lernstörungen*, 2(3), 131–146. <https://doi.org/10.1024/2235-0977/a000039>
- Müller, B., Richter, T. & Karageorgos, P. (2020): Syllable-Based Reading Improvement: Effects on Word Reading and Reading Comprehension in Grade 2. *Learning and Instruction*, 66, Article 101304. <https://doi.org/10.1016/j.learninstruc.2020.101304>
- Munser-Kiefer, M. (2014): *Leseförderung im Leseteam in der Grundschule: Eine Interventionsstudie zur Förderung von basaler Lesefertigkeit und (meta-)kognitiven Lesestrategien*. Münster: Waxmann.

- NICHD = National Institute of Child Health and Human Development (2000): *Teaching Children to Read: An Evidence-Based Assessment of the Scientific Research Literature on Reading and its Implications for Reading Instruction*. Washington: U.S. Government Printing Office.
- Noell, G. H. (2014): Research Examining the Relationships among Consultation Process, Treatment Integrity, and Outcome. In W. P. Erchul & S. M. Sheridan (Eds.), *Handbook of Research in School Consultation* (pp. 323–341). New York: Routledge.
- Nunnery, J. A., Ross, S. M. & McDonald, A. (2006): A Randomized Experimental Evaluation of the Impact of Accelerated Reader/ Reading Renaissance: Implementation on Reading Achievement in Grades 3 to 6. *Journal of Education for Students Placed at Risk (JESPAR)*, 11(1), 1–18. https://doi.org/10.1207/s15327671espr1101_1
- O'Donnell, C. L. (2008): Defining, Conceptualizing, and Measuring Fidelity of Implementation and Its Relationship to Outcomes in K–12 Curriculum Intervention Research. *Review of Educational Research*, 78(1), 33–84. <https://doi.org/10.3102/0034654307313793>
- Palincsar, A. S. & Brown, A. L. (1984): Reciprocal Teaching of Comprehension-Fostering and Comprehension-Monitoring Activities. *Cognition and Instruction*, 1(2), 117–175. https://doi.org/10.1207/s1532690xci0102_1
- Peters, M. T., Förster, N., Hebecker, K., Forthmann, B., Souvignier, E. (2021): Effects of Data-Based Decision Making on Low-Performing Readers in General Education Classrooms: Cumulative Evidence from Six Intervention Studies. *Journal of Learning Disabilities*, 1–15. <https://doi.org/10.1177/002221942111011580>
- Philipp, M. (2014): Leseunterricht in der Grundschule – vom Ist-Zustand und vom Soll-Zustand. Was Beobachtungsstudien lehren. In R. Valtin & I. Tarelli (Hrsg.), *DGLS-Beiträge: Vol. 16. Lesekompetenz nachhaltig stärken: Evidenzbasierte Maßnahmen und Programme* (S. 122–166). Berlin: Deutsche Gesellschaft für Lesen und Schreiben.
- Pierce, M., Katzir, T., Wolf, M. & Noam, G. G. (2007): Clusters of Second and Third Grade Dysfluent Urban Readers. *Reading and Writing*, 20(9), 885–907. <https://doi.org/10.1007/s11145-007-9058-x>
- Poulsen, M., Juul, H. & Elbro, C. (2015): Multiple Mediation Analysis of the Relationship between Rapid Naming and Reading. *Journal of Research in Reading*, 38(2), 124–140. <https://doi.org/10.1111/j.1467-9817.2012.01547.x>
- Richter, T. & Christmann, U. (2002): Lesekompetenz: Prozessebenen und interindividuelle Unterschiede. In N. Groeben & B. Hurrelmann (Hrsg.), *Lesesozialisation und Medien. Lesekompetenz: Bedingungen, Dimensionen, Funktionen* (S. 25–58). Weinheim: Juventa.
- Ritter, C. & Scheerer-Neumann, G. (2009): *PotsBlitz: Das Potsdamer Lesetraining – Förderung der basalen Lesefähigkeiten*. Köln: ProLog Therapie- und Lernmittel.
- Rosebrock, C. & Nix, D. (2006): Forschungsüberblick: Leseflüssigkeit (Fluency) in der amerikanischen Leseforschung und -didaktik. *Didaktik Deutsch*, 20, 90–112.
- Rosebrock, C. & Nix, D. (2011): *Grundlagen der Lesedidaktik und der systematischen schulischen Leseförderung* (4., korr. und erg. Aufl.). Baltmannsweiler: Schneider-Verlag Hohengehren.
- Rosebrock, C., Nix, D., Rieckmann, C. & Gold, A. (2017): *Leseflüssigkeit fördern: Lautleseverfahren für die Primar- und Sekundarstufe* (5. Aufl.). Praxis Deutsch. Seelze: Klett Kallmeyer.
- Rosebrock, C., Rieckmann, C., Nix, D. & Gold, A. (2010): Förderung der Leseflüssigkeit bei lese schwachen Zwölfjährigen. *Didaktik Deutsch* (28), 33–58.
- Samuels, S. J. (1979): The Method of Repeated Readings. *The Reading Teacher* (32), 403–408.

- Schneider, C. M. & Randel, B. (2009): Research on Characteristics of Effective Professional Development Programs for Enhancing Educators' Skills in Formative Assessment. In G. J. Cizek & H. L. Andrade (Eds.), *Handbook of Formative Assessment* (pp. 251–276). New York, Abingdon, Oxon, England: Routledge.
- Silverman, R. D., Speece, D. L., Harring, J. R. & Ritchey, K. D. (2013): Fluency Has a Role in the Simple View of Reading. *Scientific Studies of Reading*, 17(2), 108–133. <https://doi.org/10.1080/10888438.2011.618153>
- Slavin, R. E., Cheung, A., Groff, C. & Lake, C. (2008): Effective Reading Programs for Middle and High Schools: A Best-Evidence Synthesis. *Reading Research Quarterly*, 43(3), 290–322. <https://doi.org/10.1598/RRQ.43.3.4>
- Snidarich, S. (2015): *Examining Effects of a Repeated Reading Intervention and Predictive Effects of Student Inputs*. Dissertation. Retrieved from the University of Minnesota Digital Conservancy, <https://hdl.handle.net/11299/177096>.
- Snoddy, G. S. (1926): Learning and Stability: A Psychophysiological Analysis of a Case of Motor Learning with Clinical Applications. *Journal of Applied Psychology*, 10(1), 1–36. <https://doi.org/10.1037/h0075814>
- Souvignier, E. (2020): Interventionsforschung im Kontext Schule. In T. Hascher, T.-S. Idel & W. Helsper (Hrsg.), *Handbuch Schulforschung* (S. 1–17). Wiesbaden: Springer Fachmedien. https://doi.org/10.1007/978-3-658-24734-8_9-1
- Souvignier, E., Förster, N. & Kawohl, E. (2016): Implementation eines Förderkonzeptes zur diagnosebasierten individuellen Leseförderung in der Grundschule. In M. Philipp & E. Souvignier (Hrsg.), *Implementation von Lesefördermaßnahmen: Perspektiven auf Gelingenbedingungen und Hindernisse* (S. 77–98). Münster: Waxmann.
- Souvignier, E., Förster, N., Hebbeker, K. & Schütze, B. (2021): quop: An Effective Web-Based Approach to Monitor Student Learning Progress in Reading and Mathematics in Entire Classrooms. In S. Jornitz & A. Wilmers (Eds.), *International Perspectives on School Settings, Education Policy and Digital Strategies. A Transatlantic Discourse in Education Research* (pp. 283–298). Leverkusen: Budrich.
- Stanovich, K. E. (2009): Matthew Effects in Reading: Some Consequences of Individual Differences in the Acquisition of Literacy. *Journal of Education*, 189(1–2), 23–55. <https://doi.org/10.1177/0022057409189001-204>
- Stecker, P. M., Fuchs, L. S. & Fuchs, D. (2005): Using Curriculum-Based Measurement to Improve Student Achievement: Review of Research. *Psychology in the Schools*, 42(8), 795–819. <https://doi.org/10.1002/pits.20113>
- Swanson, E., Wanzek, J., Haring, C., Ciullo, S. & McCulley, L. (2013): Intervention Fidelity in Special and General Education Research Journals. *The Journal of Special Education*, 47(1), 3–13. <https://doi.org/10.1177/0022466911419516>
- Therrien, W. J. (2004): Fluency and Comprehension Gains as a Result of Repeated Reading. *Remedial and Special Education*, 25(4), 252–261. <https://doi.org/10.1177/07419325040250040801>
- Topping, K. J. (1999): *Paired Reading, Spelling, and Writing: The Handbook for Teachers and Parents*. London: Cassell.
- Topping, K. J. (2006): Paired Reading: Impact of a Tutoring Method on Reading Accuracy, Comprehension and Fluency. In T. Rasinski (Ed.), *Fluency Instruction: Research-Based Best Practices* (pp. 173–191). New York, NY: Guilford Press.
- Torgesen, J. K. (1998): Catch Them Before They Fall: Identification and Assessment to Prevent Reading Failure in Young Children. *American Educator*, 22, 32–41.

- Vadasy, P. F. & Sanders, E. A. (2009): Supplemental Fluency Intervention and Determinants of Reading Outcomes. *Scientific Studies of Reading*, 13(5), 383–425. <https://doi.org/10.1080/10888430903162894>
- Yeaton, W. H. & Sechrest, L. (1981): Critical Dimensions in the Choice and Maintenance of Successful Treatments: Strength, Integrity, and Effectiveness. *Journal of Consulting and Clinical Psychology*, 49(2), 156–167. <https://doi.org/10.1037//0022-006X.49.2.156>

Ein Lesestrategietraining im Distanzunterricht?

Wie sich das veränderte Lernumfeld während der Corona-Pandemie auf die Entwicklung des Leseverstehens von Schülerinnen und Schülern auswirkt

Johannes Wild, Elisabeth Kraus, Maria Steinert, Sven Hilbert
& Anita Schilcher

Zusammenfassung: Lesestrategietrainings gelten als effektives Mittel, um das Leseverstehen von Schülerinnen und Schülern zu fördern. Während zur Implementation in klassischen Unterrichtssettings inzwischen zahlreiche Studien vorliegen, sind nur wenige belastbare Forschungsergebnisse zu digitalen Trainingsumgebungen vorhanden. Zur Durchführung eines klassischen Lesetrainings in der Form von Distanzunterricht liegen bislang keine Ergebnisse im deutschsprachigen Raum vor. Die vorliegende Studie untersucht an einer Stichprobe von 1.531 Grundschulkindern, wie sich die mit dem Distanzunterricht während der Corona-Pandemie einhergehenden veränderten Durchführungsbedingungen des FiLBY-Lesetrainings auf den Lernfortschritt der Schülerinnen und Schüler auswirken. Es wird deutlich, dass die Vermittlung der Lesestrategien durch die Lehrkraft selbst ein wichtiges Element für erfolgreichen Distanzleseunterricht ist. Erklärvideos und Erklärtexte können die Strategiemodellierung durch die Unterrichtenden nicht vollständig ersetzen.

Schlüsselwörter: Lesestrategien, Modellieren, Grundschule, Längsschnitt, Distanzunterricht

Abstract: Reading strategy trainings are considered an effective approach to promote reading comprehension. While there are now numerous studies on the implementation of reading strategies in classical classroom settings, there are only a few reliable research results on digital trainings. So far, there are no results on the implementation of classical reading strategy programs in the form of distance learning in German-speaking countries. The present study investigates a sample of 1.531 students and examines how the changed conditions of the implementation of the FiLBY (Fach-

integrierte Leseförderung Bayern) reading training that have accompanied distance learning affect the learning progress of the students during the Corona pandemic. It is evident that modeling the reading strategies by the teacher is an important element for successful distance reading instruction. It cannot be substituted by explanatory videos or texts.

Keywords: reading strategies, modeling, elementary school, longitudinal, distance learning

1 Einleitung

Bereits in den 1970er Jahren zeigte die wegweisende Studie von Durkin (1978), dass die systematische Förderung des Leseverstehens in amerikanischen Grundschulklassen bis dahin praktisch keine Rolle gespielt hatte. Statt die Schülerinnen und Schüler etwa durch die Vermittlung von Lesestrategien effektiv zu unterstützen, überprüften die Lehrkräfte in den von Durkin untersuchten Unterrichtsstunden lediglich das Leseverstehen oder gaben allgemeine Hinweise zur Aufgabenanforderung. In Folge der Studie erlebte die unterrichtsbezogene Leseforschung international einen regelrechten Aufschwung, sodass heute zahlreiche Befunde sowohl im Bereich der Grundlagen- wie auch der Interventionsforschung vorliegen (Pearson & Cervetti, 2017; Okkinga, van Steensel, van Gelderen, van Schooten, Slegers & Arends, 2019). In Deutschland rückten vor allem die schlechten Ergebnisse der deutschen Schülerinnen und Schüler in der ersten PISA-Studie die Förderung des Leseverstehens durch Lesestrategien in das Bewusstsein der Forschung (exemplarisch z. B. Willenberg, 2004).

Obgleich die Forschungsbefunde inzwischen auch in der Schulpraxis Früchte tragen – international verfügen die deutschen Schülerinnen und Schüler beispielsweise in der Sekundarstufe über das höchste Lesestrategiewissen (Diedrich, Schiepe-Tiska, Ziernwald, Tupac-Yupanqui, Weis, McElvany & Reiss, 2019) –, ist der Anteil schwach Lesender sowohl in der Primar- als auch in der Sekundarstufe noch immer hoch (Bremerich-Vos, Wendt & Bos, 2017; McElvany, Kessels, Schwabe & Kasper, 2017; Weis, Doroganova, Hahnel, Becker-Mrotzek, Lindauer, Artelt & Reiss, 2019). Diese Befunde deuten darauf hin, dass ein Transfer der gelernten Strategien in das prozedurale Wissen nicht ausreichend gelingt, sodass die Strategien im Alltag nicht gewinnbringend eingesetzt werden können. Die langfristige Integration von Lesestrategien in den Schulalltag anstelle lediglich kurzfristiger und isolierter Maßnahmen wird von Pearson und Cervetti (2017) deshalb zurecht als Achillesferse von Lesestrategieprogrammen bezeichnet. Das hier vorgestellte Leseförderprogramm FiLBY (*Fachintegrierte Leseförderung Bayern*) will Lesestrategien langfristig in das

Repertoire von Grundschülerinnen und Grundschulern integrieren. Es wurde in Kooperation mit dem Bayerischen Staatsministerium für Unterricht und Kultus als langfristiges Training über mehrere Schuljahre hinweg angelegt und ist mit einer intensiven Schulung und Begleitung der beteiligten Lehrkräfte verbunden. In der dritten Jahrgangsstufe liegt der Fokus auf der Vermittlung von Lesestrategien. Da die reguläre Durchführung des FiLBY-Trainings in dieser Jahrgangsstufe aufgrund der Corona-Pandemie nur eingeschränkt möglich war, nahm ein Teil der Lehrkräfte Anpassungen vor (etwa Einsatz von Erklärvideos oder Erklärtexten zur Strategievermittlung, Ersatz des Strategietrainings durch ein Lautleseverfahren), um unter den veränderten Unterrichtsbedingungen dennoch mit FiLBY weiterarbeiten zu können. Im Folgenden werden zunächst theoretische Grundlagen des Trainings beschrieben, bevor Anforderungen an einen erfolgreichen Distanz-Leseunterricht abgeleitet werden. Methodik und Ergebnisse der Untersuchung des FiLBY-Trainings im Distanzunterricht werden im Anschluss präsentiert.

2 Leseverstehen durch Lesestrategien fördern

Neuere Definitionsversuche von Lesestrategien, etwa von Anmarkrud und Bräten (2012) oder Okkinga et al. (2019), betonen deren instrumentellen Charakter für das Leseverstehen. Im Gegensatz zu den automatisierten Lesefertigkeiten, die durch die Optimierung des Dekodierens zwar ebenfalls zum Leseverstehen beitragen, in der Regel aber ohne eine bewusste Kontrolle ablaufen, werden Strategien planvoll, d. h. bewusst und zielgerichtet, eingesetzt und kontrolliert (Afflerbach, Pearson & Paris, 2008; Philipp, 2010; Almasi & Hart, 2019; Lenhard, 2019; Rosebrock & Nix, 2020). Sie können „darauf abzielen, dass man seine Motivation beeinflusst oder aber gezielt neues Wissen auswählt, sich aneignet, es organisiert oder in sein eigenes Wissen integriert“ (Philipp, 2017, S. 117).

Dabei kann eine Strategie sowohl einzelne Handlungen umfassen als auch komplexere Handlungssequenzen (Duffy, 1993; Philipp, 2017). Lesestrategien können nach unterschiedlichen Gesichtspunkten klassifiziert werden. Gängige Kategorien orientieren sich etwa an dem Zeitpunkt ihres Einsatzes (*vor, während, nach dem Lesen*), dem Bezugspunkt der Aktivität (*kognitiv, metakognitiv, ressourcenbezogen*), der zugeordneten Funktion (z. B. *ordnend, wiederholend, elaborierend*), der Verarbeitungstiefe (*Oberflächen-, Tiefenstrategien*) oder an der Spezifität (*allgemein, domänenspezifisch, aufgabenspezifisch*). Hinzu kommen spezielle Anforderungen unterschiedlicher Textarten, etwa diskontinuierlicher und kontinuierlicher Texte oder von Hypertexten (Philipp, 2017). Leserinnen und Leser müssen deshalb nicht nur in der Lage sein, erworbene Strategien situationsspezifisch auszuwählen, sondern diese auch an die jeweiligen

Anforderungen anzupassen (Duffy, 1993; Artelt & Dörfler, 2011; Wild & Schilcher, 2019).

Die im Folgenden zitierten Metaanalysen unterstreichen die generelle Wirksamkeit von Lesestrategietrainings für das Leseverstehen, zeigen aber auch deren Sensitivität für die Art und Weise der Durchführung und Evaluation. Hinsichtlich der Durchführung erweist es sich als sinnvoll, über einen längeren Förderzeitraum mindestens zwei Strategien in einer Kombination aus direkter Instruktion sowie reziproken Lehren und Lernen einzuführen und systematisch zu üben. Als besonders gewinnbringend erweisen sich dabei Fragestrategien, das Zusammenfassen und das Überwachen des eigenen Leseprozesses (Souvignier & Antoniou, 2007; Edmonds, Vaughn, Wexler, Reutebuch, Cable, Klingler Tackett & Wick Schnakenberg, 2009; Slavin, Lake, Chambers, Cheung & Davis, 2009; Mayer & Marks, 2019). Duffy (1993) weist jedoch darauf hin, dass es weniger darauf ankomme, welche konkreten Strategien Schülerinnen und Schüler lernen, sondern dass es das strategische Denken an sich ist, das sie lernen müssten: „what is basic in strategy instruction, particularly for low achievers, is not whether teachers are teaching the right strategies but, rather, whether students are developing an integrated concept of what it means to be strategic“ (Duffy, 1993, S. 231). Dazu reicht es nicht, wenn die unterrichtenden Lehrkräfte die Schülerinnen und Schüler von der Notwendigkeit von Lesestrategien überzeugen, es kommt vielmehr darauf an, ihnen die einzelnen Strategien sowie die Bedingungen und Ziele ihres Einsatzes umfassend zu erklären (Wild & Schilcher, 2019).

Im angloamerikanischen Raum hat sich hinsichtlich der Strategievermittlung das sechsstufige Vorgehen des Self-Regulated Strategy Developments (SRSD) etabliert, bei dem das Modellieren der Strategien durch die Lehrkraft im lauten Denken ein zentrales Element ist und die Verantwortung für die Strategie Schritt für Schritt in die Hand der Schülerinnen und Schüler übergeht (Lienemann & Ried, 2006; vgl. Pissarek, 2018). Am Modell der Lehrkraft erhalten die Lernenden zunächst Einsichten in die Denkvorgänge eines Experten und vollziehen die Strategieanwendung sowie deren Begründungskontexte gedanklich nach, um sie später selbstständig und flexibel einsetzen zu können (vgl. auch Bandura, 1979). Sie erwerben außerdem Hintergrundwissen zu den Lesestrategien. Nach einer Memorierphase, die dem Verankern der Strategie im deklarativen Wissen dient, üben die Kinder unter Anleitung der Lehrperson, die Strategien auf neue Texte zu übertragen. Um einer ineffizienten Nutzung und motivationalen Defiziten vorzubeugen, strukturiert diese dabei das Lernsetting stark vor, korrigiert oder lobt und gibt Tipps für die Weiterarbeit (vgl. Hasselhorn & Gold, 2013; Philipp, 2015). Je besser die Kinder die Strategiearbeit beherrschen, desto stärker nimmt sich die Lehrkraft zurück. Deshalb erweist sich in Übungsphasen das Arbeiten in Kleingruppen, in denen sich die Schülerinnen und Schüler gegenseitig unterstützen und beraten können, als

sehr günstig (Brown & Palinscar, 1989; Mayer & Marks, 2019; zum computerbasierten kooperativen Lernen vgl. Chen, Wang, Kirschner & Tsai, 2018). Dennoch sind die Kinder auch hier auf die Unterstützung der Lehrperson angewiesen, da die Mitschülerinnen und Mitschüler oft vor vergleichbaren Problemen stehen und deshalb die Probleme ihrer Mitschülerinnen und Mitschüler nicht lösen können.

Der Lehrkraft kommt also bei der Strategievermittlung eine zentrale Rolle zu, sowohl bei der Einführung als auch beim Üben. Qualitative Studien zeigen jedoch, dass Lehrerinnen und Lehrer die größten Probleme bei der Umsetzung von Interventionen haben, die die Prinzipien der Strategieinstruktion, des Modellierens und der Gruppenarbeit nutzen (Okkinga et al., 2019). Nicht zuletzt deshalb betonen Souvignier, Förster und Kawohl (2016), dass Lehrkräfte mit Hilfe gut strukturierter Materialien bei der Umsetzung komplexer Trainingsprogramme unterstützt werden müssten.

3 Das Lesetraining FiLBY (Fachintegrierte Leseförderung Bayern)

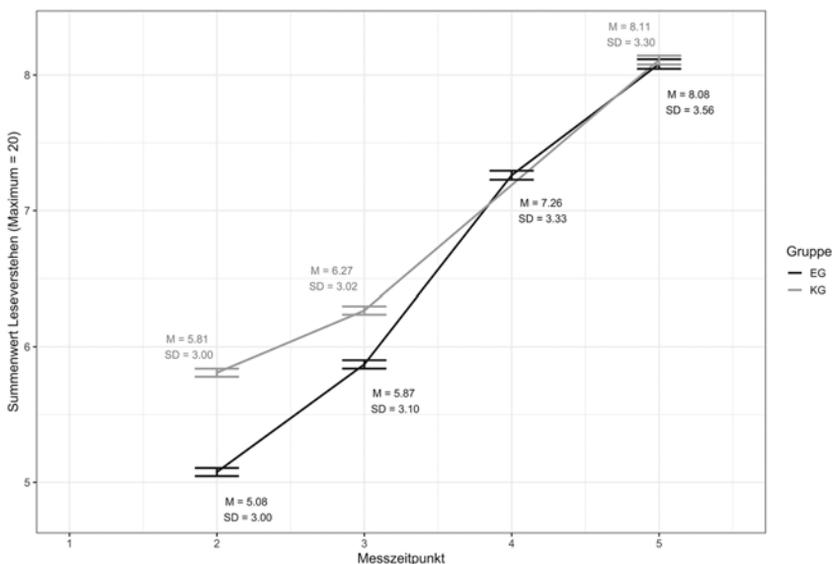
In FiLBY werden den Lehrkräften daher nicht nur flächendeckend umfangreiche Trainingsmaterialien kostenfrei zur Verfügung gestellt, die Implementation wird darüber hinaus durch ein wissenschaftliches Team der Universität Regensburg betreut sowie durch eine interaktive Fortbildungsplattform unterstützt. Das FiLBY-Training erstreckt sich über insgesamt drei Schuljahre, in denen jeweils sowohl mit narrativen Texten (Klassenlektüre, erstes Schulhalbjahr) als auch mit eigens verfassten, altersangemessenen Sachtexten (zweite Schuljahreshälfte) trainiert wird.

Während in der zweiten Jahrgangsstufe (FiLBY-2; Schilcher, Wild & Steinert, 2019) die Leseflüssigkeit als Basis des Leseverstehens fokussiert wird, steht in der dritten Jahrgangsstufe (FiLBY-3; Wild, Schilcher & Steinert, 2020) die Vermittlung effektiver kognitiver Lesestrategien für literarische Texte (*Vorwissen aktivieren, Figur, Ereignis und Situation untersuchen*) und Sachtexte (*Vorwissen aktivieren, Überfliegen, Visualisieren*) im Vordergrund. Es handelt sich dabei um kognitive (ordnende, elaborierende und wiederholende) Strategien, die durch Modellieren der Lehrkraft im lauten Denken für unterschiedliche Phasen des Leseprozesses eingeführt und anschließend über mindestens vier Wochen hinweg täglich geübt werden. Im Sinne eines Scaffoldings wird der Strategieeinsatz zunächst von der Lehrkraft gezeigt und mit ihr gemeinsam geübt, wobei auch kooperative Lernphasen eingeplant sind. Im ersten Schulhalbjahr werden also literarische Strategien an einer Ganzschrift trainiert, im zweiten Schulhalbjahr Strategien für Sachtexte (zum Material vgl. <https://www.lesen.bayern.de/filby3>). Eine Besonderheit der 60 zur Verfügung gestellten Sachtexte ist, dass sie alle gleich lang und gleich schwer sind. In der vierten

Jahrgangsstufe (FiLBY-4; Steinert, Schilcher & Wild, 2021) kommen zusätzlich metakognitive Strategien hinzu. Die Lehrkräfte nehmen jeweils am Anfang eines Schuljahres an einer zweitägigen Fortbildung teil und erhalten eine Einführung in das Training, unterstützt durch eine Online-Fortbildungsplattform. Darüber hinaus bekommen sie Lehrerhandreichungen, die die didaktischen Grundlagen des Trainings nochmals zusammenfassen sowie Stundenverläufe und weiteres Unterrichtsmaterial enthalten (z. B. Moderationskarten zum Modellieren der Strategien). Die Begleitevaluation begann im Schuljahr 2018/19 in der zweiten Jahrgangsstufe (Messzeitpunkte 1–3) und wird im Schuljahr 2020/21 abgeschlossen sein.

Erste Analysen zeigen, dass das FiLBY-Training von der zweiten Klasse an positive Effekte für das Leseverstehen hatte, sodass die anfänglich schwächere Experimentalgruppe (EG) im Vergleich zur Kontrollgruppe (KG) signifikant größere Fortschritte machte und bis zum Ende der dritten Jahrgangsstufe ihren Rückstand bereits aufgeholt hatte (Messzeitpunkt 5). Ähnliche Ergebnisse zeigen sich hinsichtlich der Leseflüssigkeit (Schilcher, Wild, Kraus & Hilbert, in Druck; zum Leseverstehen vgl. Abb. 1).

Abbildung 1: Entwicklung des Leseverstehens der FiLBY-Experimentalgruppe (EG) im Vergleich zur Kontrollgruppe (KG) mit regulärem Leseunterricht (Messzeitpunkte 2/3: Mitte bzw. Ende der zweiten Jahrgangsstufe, Messzeitpunkte 4/5: Mitte bzw. Ende der dritten Jahrgangsstufe). Zu Messzeitpunkt 4 lagen keine Daten der KG vor. Das Leseverstehen wurde mittels des Bayerisches Lesetest BYLET erfasst (vgl. Teilkap. 6.2)



M = Mittelwert, SD = Standardabweichung. Darstellung in Rohpunkten, da die Normierung pro Messzeitpunkt erfolgte. Die Balken zeigen die Standardfehler der Gruppenmittelwerte an.

Im Februar 2020 wurde das Training der dritten Klassen jedoch durch die weltweite Corona-Pandemie unterbrochen. Die Einschränkung des regulären Unterrichtsbetriebs durch das Pandemiegeschehen wirkte sich auch auf die Entwicklung der Lesekompetenz der an FiLBY teilnehmenden Schülerinnen und Schüler aus (vgl. Abb. 1: Messzeitpunkt 4 und 5): Die Kinder der Experimentalgruppe entwickelten sich analog zur Kontrollgruppe. Viele Lehrkräfte unterbrachen aufgrund der Pandemie das FiLBY-Training, ein Teil führte dieses jedoch als Bestandteil ihres Distanzunterrichts weiter und nahm Anpassungen vor, um der veränderten Lehrsituation gerecht zu werden.

4 Lehren und Lernen in Distanz

Der Distanzunterricht kann als eine Form der Fernlehre bzw. ‚distance education‘ verstanden werden. Diese „umfasst alle Formen des Lehrens und Lernens, bei denen eine Betreuung durch Lehrende regelmäßig über Distanzen erfolgt“ (Kerres, 2018, S. 520). Im Unterschied zum vorliegenden Training handelt es sich dabei aber häufig um kommerzielle Angebote, in denen Lerninhalte aufbereitet werden (Fickermann & Edelstein, 2020). Die Lernumgebung zeichnet sich vor allem durch eine zerdehnte Kommunikationssituation aus, das heißt zeitliche, räumliche und/oder emotionale Trennung der Beteiligten (Ehlich, 1984), wengleich diese zum Teil mit Hilfe analoger (z. B. Wochenplan) oder digitaler Hilfsmittel (z. B. Videokonferenz) überbrückt werden kann. Eine Untersuchung von Wacker, Unger und Rey (2020) während der ersten Schulschließungen zeigt, dass die Vermittlung von Lerninhalten, Arbeitsaufträgen und Feedback in den meisten Fällen per E-Mail geschah, gefolgt von Lernplattformen, Cloudspeicher oder über Homepages. Der Distanzunterricht während der Schulschließungen kann also überwiegend als ein wenig betreutes Lernsetting charakterisiert werden. Dies ist deshalb problematisch, da es vielen Schülerinnen und Schülern nicht gelingt, sich ohne Unterstützung der Lehrkraft selbst zu organisieren und zu lernen (Porsch & Porsch, 2020).

Für das erfolgreiche Lernen zuhause benötigen die Schülerinnen und Schüler eine gut strukturierte Lernumgebung, ansprechbare Hilfspersonen, die beispielsweise bei Problemen weiterhelfen können, sowie konstruktives Feedback auf ihre Arbeit (Stahns, Rieser & Lankes, 2017; Dumont, 2019; Köller, Fleckenstein, Guill & Meyer, 2020; Huber & Helm, 2020). Das setzt konkrete, erreichbare Ziele, beispielsweise über Tages- oder Wochenpläne, voraus (Klieme, 2020). Lehrerinnen und Lehrer müssen also wie im Präsenzunterricht die „Vorbereitung, Begleitung und Nachbearbeitung der häuslichen Arbeiten“ (Köller et al., 2020, S. 164) gewährleisten. Gerade in komplexen (digitalen) Lernumgebungen, wie sie beim Distanzlernen häufig auftreten, benötigen die Kinder demnach Kontaktmöglichkeiten (z. B. Videotreffen, Chat, Telefon-

sprechstunde) sowie eine ausreichende Unterstützung beim Lernen (z. B. vorstrukturiertes Webquest), um motiviert, konzentriert und zielgerichtet zu arbeiten (Segers & Verhoeven, 2009; Kunter & Ewald, 2016; Klieme, 2020; Köller et al., 2020). Überlassen Lehrpersonen dies den Erziehungsberechtigten, können sozioökonomisch bedingte Leistungsdisparitäten verschärft werden (McElvany, Becker & Lüttke, 2009; Niederbacher & Neuenschwander, 2020; Holzberger, Reinhold, Lüttke & Seidel, 2020). Weder sind die Eltern darauf vorbereitet, die Rolle der Lehrkraft zu übernehmen, noch fühlen sie sich dazu in der Lage (Besa, Gensler, Gesang & Röhrig, 2020).

Soll ein Lesetraining wie FiLBY-3 unter den Bedingungen des Distanzlernens gelingen und ein negativer Effekt auf das Leseverstehen, wie er etwa häufig in unterrichtsfreien Zeiten bei Kindern aus bildungsfernen Familien auftritt (Stanovich, 2000), und die Lesemotivation vermieden werden, reicht es nicht aus, die Kinder lediglich zum Lesen aufzufordern bzw. ihnen im Rahmen von Wochenplänen etc. „Lesezeit“ abzuverlangen oder das Lesen von Texten als Aufgabe zu geben, da gerade schwach Lesende Leseaufgaben zu vermeiden suchen (Stalder, 2013; Nonte, Hartwich & Willems, 2018). Lehrerinnen und Lehrer sollten systematisch und strukturiert mit dem vorliegenden Trainingsmaterial arbeiten.

Das Modellieren der Strategien im lauten Denken sollte durch die Lehrkraft selbst erfolgen, da die emotionale Beziehung zur Lehrkraft für das Lernen eine wichtige Rolle spielt (Casale, Börnert-Ringleb & Hillenbrand, 2020; Rubach & Lazarides, 2021). Die Bedingungen ihrer Anwendung sollten dabei ebenfalls geklärt werden. Zwar können unterstützend entsprechende Erklärvideos o. Ä. eingesetzt werden, etwa zur Wiederholung, jedoch ist davon auszugehen, dass Kinder dieses Alters sich komplexe Lesestrategien wie die vorliegenden nicht selbstständig aneignen können. Auch die inhaltliche Kohärenz zwischen einzelnen Trainingseinheiten geht ohne Unterstützung der Lehrperson verloren, da für den Erfolg zentrale Aspekte fehlen. Auch im Distanzunterricht sollten sich nach der Einführung von Strategien gemeinsame Übungsphasen mit der Lehrkraft anschließen, bei denen die Ziele des Trainings nochmals betont, Probleme der Kinder geklärt oder Strategien wiederholt bzw. angepasst werden können (Okkinga et al., 2019). Dies unterstützt die Schülerinnen und Schüler einerseits dabei, das Lernen zu organisieren, zum anderen können Strategienutzungsdefizite erkannt und vermieden werden. Deshalb ist es vor allem am Anfang des Trainings wichtig, dass die Lernenden durch die Lehrkraft selbst zeitnah und unmittelbar Feedback erhalten (Dumont, 2019; Köller et al., 2020) und die Beziehung aufrecht erhalten wird, indem sie als Ansprechpartner zur Verfügung steht, lobt oder korrigiert (Casale et al., 2020).

Es bietet sich im Rahmen des Distanzunterrichts an, mit den einzelnen Kindern Kontakt zu halten und auch auf kooperative Lernformen in Kleingruppen (z. B. in Break-out-Rooms) zu setzen, damit die Kinder nach und nach selbst

Aufgaben wie das Feedback oder die nochmalige Erklärung übernehmen. Ein solides Wissen über den Einsatz von digitalen Medien im Unterricht ist also eine Voraussetzung, um die zuvor dargestellten Ansprüche zu realisieren (Reiss, 2020). Welche Anpassungen die Lehrkräfte in Hinblick auf die Durchführung von FiLBY im Distanzunterricht konkret vornahmen und wie erfolgreich diese waren, wird im Ergebnisteil berichtet.

5 Zusammenfassung und Forschungsfragen

Als zentrale Gelingensfaktoren für den Unterricht zuhause lassen sich aufgrund der zuvor dargestellten Ergebnisse zusammenfassend drei Aspekte identifizieren: eine strukturierte Lernumgebung, die Verfügbarkeit von Hilfspersonen sowie konstruktives Feedback. Die Lehrkräfte sollten Lesestrategien deshalb nach Möglichkeit persönlich einführen und die Kinder während der anschließenden Übungsphasen konstruktiv begleiten, auch während kooperativer Arbeitsphasen. Die folgende Untersuchung beschäftigt sich daher damit, welche Anpassungen die an FiLBY beteiligten Lehrkräfte (1) generell und (2) in Bezug auf das Training vorgenommen haben:

Um alle weiteren Analysen in den chronologischen Verlauf einzubetten, wird zunächst auf Basis von Fragebogendaten dargestellt, welche Unterrichtselemente die Lehrpersonen in ihrem Distanzunterricht integrierten, wie den Schülerinnen und Schülern das benötigte Arbeitsmaterial zur Verfügung gestellt wurde, wie der Unterricht an sich organisiert wurde und wie die Kinder beim Lernen unterstützt wurden (etwa Kontrolle von Lernergebnissen, Feedback). Daneben wird ausgewertet, welches inhaltliche bzw. organisatorische Unterstützungspotenzial die Lehrpersonen den Erziehungsberechtigten zumaßen und ob diese Anpassungen bzw. Einschätzungen der Lehrkräfte in Zusammenhang mit einer Fortführung des FiLBY-Trainings in der Pandemiesituation standen.

Anschließend wird untersucht, welche Anpassungen in Hinblick auf das FiLBY-3-Training im Distanzunterricht vorgenommen wurden und wie förderlich diese hinsichtlich des Leseverstehens waren. Damit ergeben sich zwei maßgebliche Forschungsfragen:

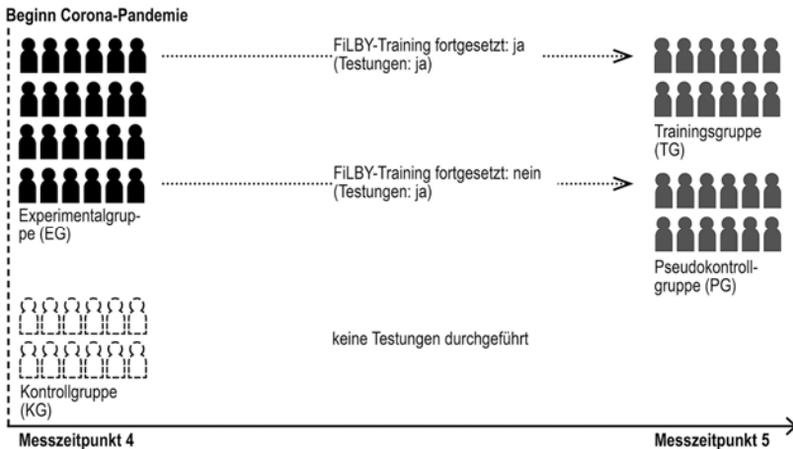
1. Wie gestalteten Grundschullehrkräfte ihren Leseunterricht (un)abhängig von FiLBY-3 während der pandemiebedingten Schulschließungen im Frühjahr 2020?
2. Wie hängen die von den Lehrkräften im Rahmen der Durchführung von FiLBY-3 ergriffenen Anpassungen während des Distanzunterrichts mit der Entwicklung des Leseverstehens ihrer Schülerinnen und Schüler zusammen?

6 Methode

6.1 Stichprobe und Beschreibung des Studienablaufs

An der longitudinalen Evaluation der FiLBY-Studie nehmen insgesamt ca. 8.500 bayerische Schülerinnen und Schüler teil. Davon entfallen etwa 7.000 Kinder auf die Experimentalgruppe und 1.500 Kinder auf die Kontrollgruppe (KG), die kein systematisches Lesetraining erhält, sondern im Rahmen des regulären Unterrichts Lesen trainiert. Aufgrund der unvermittelten Schulschließungen war es vielen Lehrkräften nicht mehr möglich, die Testungen durchzuführen, weshalb nur von einem Teil der Stichprobe Daten vorliegen. Ein großer Teil der Lehrkräfte der Experimentalgruppe konnte während der Schulschließungen die FiLBY-3-Materialien nicht wie geplant für das Lesetraining einsetzen, führte die Testungen aber dennoch durch. Die Experimentalgruppe wurde daher für die folgenden Analysen geteilt: Die Schülerinnen und Schüler, die auch im Distanzunterricht weiterhin mit FiLBY arbeiteten, werden als Trainingsgruppe (TG) bezeichnet. Diejenigen Klassen, die seit Beginn der Schulschließungen nicht mehr mit den FiLBY-Materialien arbeiteten, wurden zu einer Pseudokontrollgruppe (PG) zusammengefasst (vgl. Abb. 2).

Abbildung 2: Stichprobenaufteilung unter den veränderten Durchführungsbedingungen der Corona-Pandemie



Die Analysen umfassen die beiden Messzeitpunkte der dritten Jahrgangsstufe (Messzeitpunkte 4 bzw. 5: März bzw. Juni/Juli 2020). Für diesen Zeitraum liegen von einer Teilstichprobe von 85 Lehrkräften der TG/PG und ihren dritten Klassen Informationen zur Leseleistung der Schülerinnen und Schüler ($N = 1.531$) sowie zur Fortführung des FiLBY-Trainings während der Schulschließungen vor. Das Leseverstehen wurde zu beiden Messzeitpunkt erfasst, also vor

bzw. nach den Schulschließungen, die Fortführung des FiLBY-Trainings einmalig durch Befragung der Lehrkräfte. Tabelle 1 gibt eine Übersicht über die Zusammensetzung der Stichprobe zu den beiden Messzeitpunkten.

Tabelle 1: Übersicht der Klassen und Schülerzahlen zu den Messzeitpunkten 4 und 5. Schülerinnen und Schüler mit fehlenden Werten zu einem der beiden Messzeitpunkte (MZP) wurden nicht ausgeschlossen

	Pseudokontrollgruppe (PG)		Trainingsgruppe (TG)		Gesamt	
	Klassen	Schüler/-innen	Klassen	Schüler/-innen	Klassen	Schüler/-innen
MZP 4	22	371	27	433	49	804
MZP 5	33	578	41	697	74	1.275
Gesamt	38	695	47	836	85	1.531

Die demografischen Daten liegen von 1.227 Schülerinnen und Schülern der Stichprobe vor. Die Mehrheit der Kinder (84.4%) gab an, in Deutschland geboren zu sein. Lediglich 12.7% waren nicht in Deutschland geboren, rund 3.0% machten keine Angabe oder wussten es nicht. Das Geschlechterverhältnis ist ausgewogen (weiblich = 49.9%). Es handelt sich bei den beteiligten Lehrkräften überwiegend um erfahrene Lehrpersonen, die im Median 11–15 Jahre in ihrem Beruf tätig waren.

6.2 Messinstrumente

Das Leseverstehen wurde mit dem Bayerischen Lesetest (BYLET) erfasst. Es handelt sich dabei um einen neu entwickelten Leseverstehetest mit drei Parallelversionen, bei dem die Schülerinnen und Schüler 20 Multiple-Choice-Fragen zu vier in ihrer Komplexität ansteigenden Textabschnitten beantworten. Gegenstand des Texts ist eine trainingsferne Science-Fiction-Geschichte: Eine Weltraumcrew erkundet einen neuen Planeten und erlebt ein Abenteuer. Da der Test zusätzlich eine steigende Aufgabenschwierigkeit innerhalb der Textabschnitte aufweist, kann er das Leseverstehen innerhalb eines breiten Leistungsspektrums erfassen. Ein Beispielitem kann unter https://is.gd/bylet_sample (Username: *filby*, Passwort: *lesetraining*) eingesehen werden.

Eine Pilotstudie (Kraus, Wild, Schilcher & Hilbert, 2021) zeigt, dass die Art der Konzeption vorwissensbedingte und geschlechtsspezifische Verzerrungen vermeidet und deshalb eine gute Voraussetzung für eine faire Messung des Leseverstehens bietet. Zur Auswertung wurden die 20 Multiple-Choice-Fragen als richtig oder falsch kodiert und mittels Raschmodell psychometrisch modelliert (Rasch, 1960; Andersen, 1973). Zur externen Validierung wurde die Korrelation mit der Leseflüssigkeit geprüft, erfasst durch das Salzburger Lesescreening 2-9

(SLS 2-9; Wimmer & Mayringer, 2016). Der Zusammenhang von BYLET und SLS betrug $r = .43$, sodass von einer validen Messung der Lesekompetenz ausgegangen werden kann. Zur besseren Interpretierbarkeit wurden die Testwerte im Folgenden auf die IQ-Skala ($M = 100$, $SD = 15$) normiert. Die Ergebnisse der psychometrischen Analysen können im OSF-Repositorium (<https://osf.io/9cspz/>) eingesehen werden.

Die unterrichtenden Lehrkräfte wurden im Juli 2020 mittels Online-Fragebogen befragt. Dieser erfasst neben allgemeineren Fragen zur Gestaltung und Bewertung des Distanzunterrichts auch Aspekte der Umsetzung des Lesestrategietrainings FiLBY sowie Fragen zur Ausstattung und Erfahrung der Lehrkraft mit digitalen Geräten und Medien.

6.3 Statistische Analysen

Alle statistischen Analysen wurden mit der Statistiksoftware R (R Core Team, 2020) durchgeführt. Zur Beantwortung der ersten Forschungsfrage werden deskriptive Statistiken erstellt. Konkret werden die Häufigkeiten der Variablen zur Strukturiertheit des Unterrichts, zur Unterstützung sowie zum Stellenwert des Lesens bei den Lehrpersonen berechnet und mit den Schülerleistungen in Beziehung gesetzt.

Zur Untersuchung der zweiten Forschungsfrage werden zwei hierarchisch lineare Regressionsmodelle geschätzt. Die Datenstruktur ist in zweifachem Sinne hierarchisch, also geschachtelt. Zum einen wurden dieselben Kinder zweimal getestet, sodass von jedem Kind jeweils zwei Lesetestergebnisse vorliegen (Messungen geschachtelt in Kindern). Gleichzeitig ist anzunehmen, dass Kinder, die dieselbe Klasse besuchten, sich in ihren Entwicklungen gleichen, sodass auch die Schachtelung von Kindern in Schulklassen berücksichtigt werden muss. Durch die Schachtelung entstehen Abhängigkeiten in den Messwerten, die nicht auf das Training zurückzuführen sind und deshalb für die Interpretation der Ergebnisse nicht relevant sind. Um diese Abhängigkeiten statistisch zu kontrollieren, werden anstelle von multiplen linearen Regressionen, hierarchische lineare Regressionen berechnet. Dazu wurden die Daten aus einer multivariaten Darstellung in eine univariate überführt, wobei die Testwerte der verschiedenen Messwerte jedem Kind über eine ID-Variable und jedem Messzeitpunkt über eine Messzeitpunktvariable zugeordnet werden. Die PG dient als Konstante, sodass positive oder negative Veränderungen zwischen den Messzeitpunkten als Regressionsgewicht beschrieben werden können. Die Veränderungen werden dabei als Interaktionseffekte zwischen der Messzeitpunktvariable und weiteren Prädiktoren modelliert, sodass die Regressionsgewichte die unterschiedliche Entwicklung der Gruppen in Abhängigkeit weiterer Prädiktoren beschreiben (Hilbert, Stadler, Lindl, Naumann & Bühner, 2019). Ein erstes hierarchisches lineares Modell untersucht, ob sich eine Verwendung der

FiLBY-Materialien im Distanzunterricht gegenüber anderen Materialien als wirksamer für die Entwicklung des Leseverstehens erweist. Mit Hilfe eines zweiten Modells wird analysiert, welche Umsetzungsvarianten des FiLBY-3-Trainings sich wie auf die Entwicklung des Leseverstehens auswirken. Zur Korrektur für die sich so ergebenden multiplen Signifikanztests der Regressionskoeffizienten wird eine Bonferroni-Holm-Korrektur vorgenommen (Holm, 1979).

7 Ergebnisse

7.1 Forschungsfrage 1

In Hinblick auf die erste Forschungsfrage wird analysiert, wie Grundschullehrkräfte während der Schulschließungen Arbeitsaufträge vermittelten, welche organisatorischen Informationen sie beifügten und welche Unterstützungsangebote den Kindern zur Verfügung standen. Es wird darüber hinaus untersucht, wie die Lehrkräfte die Unterstützungsfähigkeiten der Eltern einschätzten und wie viel Unterrichtszeit für das Lesen veranschlagt wurde. Dabei wird verglichen, wie sich die Lehrkräfte der TG von denen der PG unterscheiden (für eine detaillierte Übersicht vgl. OSF: <https://osf.io/9cspz/>).

Die Lehrkräfte beider Gruppen nutzten jeweils mehrere Kommunikationskanäle, um mit den Kindern bzw. den Erziehungsberechtigten in Kontakt zu bleiben und das Lernen zuhause zu organisieren, z. B. Clouddienste, E-Mails oder Videokonferenzen. Im Vergleich zur PG ließen die Lehrkräfte der TG die Arbeitsmaterialien (z. B. die FiLBY-Trainingshefte) häufiger abholen oder verschickten diese per E-Mail. Die Lehrenden der PG verschickten eher Kopien bzw. brachten diese den Kindern selbst vorbei.

Die Arbeitsmaterialien versahen die Lehrkräfte überwiegend mit zusätzlichen Informationen, z. B. Hinweise zur inhaltlichen Strukturierung. Die Unterschiede zwischen TG und PG fallen diesbezüglich gering aus. In beiden Gruppen bildeten allgemeinere Unterstützungsangebote wie Hinweise zu Kontaktmöglichkeiten oder Differenzierungsmöglichkeiten einen Schwerpunkt. Nur wenige der Lehrkräfte boten den Kindern allerdings konkrete Strategien zur Erarbeitung, Anregungen zur Lernreflexion oder Lernziele an. Auffällig ist, dass die Lehrperson der TG häufiger angab, auf Strategien und schriftliches Feedback zu setzen.

Fast alle Lehrkräfte schätzten das Unterstützungspotenzial der Eltern bei der Lernorganisation befriedigend bis gut ein und trauten ihnen auch die inhaltliche Unterstützung zu. Die Lehrerinnen und Lehrer der Pseudokontrollgruppe trauten den Erziehungsberechtigten in der Tendenz etwas mehr zu als ihre Kolleginnen und Kollegen der FiLBY-Gruppe. Ein systematischer Zusammenhang der Bearbeitungshinweisen und dem eingeschätzten Unterstützungs-

potenzial der Eltern kann jedoch nur hinsichtlich der Formulierung von Lernzielen festgestellt werden: Die Lehrpersonen gaben seltener Lernziele an, wenn sie die Eltern als kompetent empfanden ($r = -.2$). Zwischen der Wahl der Kommunikationswege und zusätzlich angebotenen Informationen bestehen keine Zusammenhänge. Die teilnehmenden Lehrkräfte fühlten sich überwiegend sicher bis sehr sicher im Umgang mit Medien.

Zusammenfassend lässt sich sagen, dass Lehrkräfte, die mit den FiLBY-Materialien weiterarbeiteten, dem Lesen einen höheren Stellenwert zumaßen, mehr Zeit für das Lesen reservierten und die Schülerinnen und Schüler besser informierten. Dafür war der persönliche Kontakt weniger umfangreich. Diese Lehrerinnen und Lehrer setzten stattdessen stärker auf das selbstständige Arbeiten der Kinder mit den vorstrukturierten FiLBY-Materialien und schriftliches Feedback. Systematische Zusammenhänge mit der positiveren Einschätzung der eigenen Medienkompetenz oder dem geringer eingeschätztem Unterstützungspotenzial der Eltern lassen sich jedoch nicht nachweisen (s. Tabelle 2).

Tabelle 2: Zusammenhänge zwischen Einschätzung von Medienkompetenz und organisatorischem Unterstützungspotenzial der Eltern und Art des Feedbacks in der FiLBY-Gruppe (Kendall- τ -Koeffizient)

Korrelation nach Kendall	Musterlösung	Korrektur schriftlich	Korrektur mündlich
Medienkompetenz der Lehrkräfte	-.01	-.11	.23
Unterstützungspotenzial Eltern	-.11	.17	.20

7.2 Forschungsfrage 2

Zur Untersuchung der zweiten Forschungsfrage wird analysiert, ob eine Verwendung der FiLBY-Materialien gegenüber der Verwendung anderer Lesefördermaterialien während der Schulschließungen generell förderlicher war. In der hierarchisch linearen Regression gibt die Konstante an, mit welcher Leistung die PG gestartet ist. Es wurden vier Effekte geschätzt: Der Effekt „Messzeitpunkt“ ist ein Schätzer für den Fortschritt der Kinder der PG während des Erhebungszeitraums. Der Effekt „Gruppe“ zeigt den mittleren Leistungsunterschied zwischen PG und TG vor dem Training an, der Interaktionseffekt zwischen Messzeitpunkt und Gruppe „Messzeitpunkt \times Gruppe“ gibt an, wie sich die TG im Vergleich zur PG entwickelt hat. Die Gesamtveränderung der TG ergibt sich also aus der Summe „Messzeitpunkt“ + „Messzeitpunkt \times Gruppe“. Tabelle 3 zeigt, dass sich kein signifikanter Vorteil der TG gegenüber der PG während der Schulschließungen ergibt. Die Leseförderung der TG war demnach der der PG während des Distanzunterrichts scheinbar nicht überlegen. Die Intraklassenkorrelation (ICC) gibt im Modell an, wie viel Prozent der Fehlervarianz auf die geschachtelte Datenstruktur zurückzuführen war. Der (hohe)

Wert von 0.55 zeigt, dass die meisten Schülerinnen und Schüler sich während des Erhebungszeitraums ähnlich entwickelten (leistungsstarke Schülerinnen und Schüler gehören weiterhin zu den leistungsstarken) und die Leistungsentwicklung zudem mit der Klassenstruktur assoziiert war (z. B. Klassen mit großem bzw. kleinem Leistungszuwachs).

Tabelle 3: Fortschritte im Leseverstehen in Abhängigkeit der Gruppenzugehörigkeit

	Modelle			
	Nullmodell		Volles Modell	
<i>Feste Effekte</i>	<i>b</i>	<i>p</i> -Wert	<i>b</i>	<i>p</i> -Wert
Konstante	100.01	< .001	99.99	< .001
Gruppe	-	-	-0.93	.472
Messzeitpunkt	-	-	0.50	.516
Messzeitpunkt x Gruppe	-	-	0.58	.576
<i>Zufällige Effekte</i>	σ^2		σ^2	
ICC _{Subjekt:Klasse}	.54		.55	
R ² _{marginal}	-		.001	
R ² _{konditional}	-		.55	

Anmerkung. *b* = Regressionskoeffizient, TG = Trainingsgruppe, MZP = Messzeitpunkt, ICC = Intraklassenkorrelation, σ = Varianz, *x* bezeichnet Interaktionseffekte, N = 1495 Schülerinnen und Schüler, 85 Klassen.

Die Fragebogenerhebung demonstrierte, dass die Lehrkräfte spezifische Anpassungen des FiLBY-3-Trainings vorgenommen hatten. In einer zweiten Analyse wird daher untersucht, welche Elemente der Trainingsdurchführung die Wirksamkeit des FiLBY-Trainings im Distanzunterricht beeinflussten. Dazu wird die TG nochmals nach der Art der Trainingsanpassung unterteilt: Eine Gruppe bilden dabei die Schülerinnen und Schüler, denen die Strategien wie im Training vorgesehen von der Lehrkraft noch im Präsenzunterricht persönlich vermittelt worden waren (11.0%). Rund ein Viertel der Schülerinnen und Schüler arbeitete mit den Trainingsheften, ohne dass zuvor Lesestrategien vermittelt worden waren (26.7%). Etwa ein Fünftel der Kinder sollte sich die Strategien durch Erklärvideos aneignen (22.4%), ein kleinerer Teil der Lernenden erhielt dazu Erklärtexte (12.8%; vgl. Tabelle 4). Andere Einzelfallnennungen sind in dieser Analyse nicht berücksichtigt.

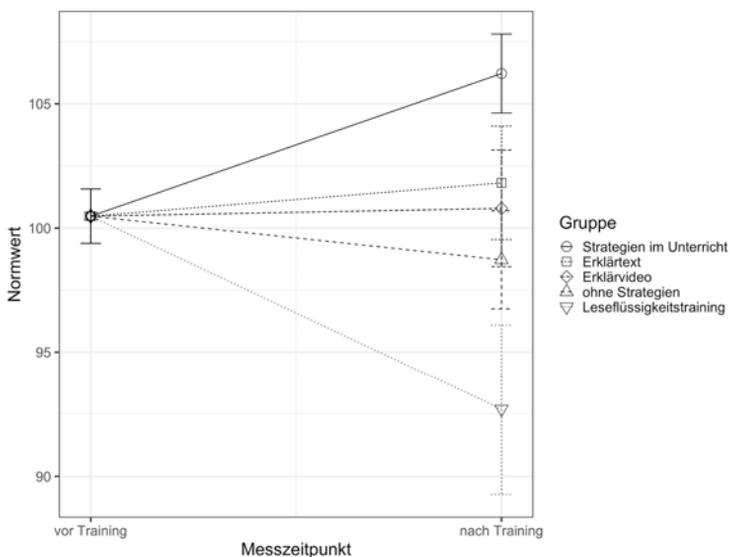
Tabelle 4: Verteilung der FILBY-Gruppe über die Kategorien der Arbeit mit den FILBY-Materialien (absolut und in Prozent)

	Strategie-einführung im Unterricht	Erklär-text f. Strategien	Erklär-video f. Strategien	Lese-flüssig-keits-training	ohne Stra-te-gien	Strategien per Video-konferenz	gesamt
Anzahl Schüler/-innen	92 (11.0%)	107 (12.8%)	187 (22.4%)	63 (7.5%)	223 (26.7%)	0 (0%)	672
Anzahl Klassen	5 (13.9%)	5 (13.9%)	10 (27.8%)	3 (8.3%)	13 (36.1%)	0 (0%)	36

Anzahl Missings (Kategorie: Sonstige): 164 Schüler/-innen, 11 Klassen

Als Referenzkategorie (Konstante) dient in diesem Modell die Strategieeinführung durch die Lehrkraft selbst im Präsenzunterricht. Die Gruppe machte während des Erhebungszeitraums 5.74 BYLET-Normpunkte Fortschritt („Messzeitpunkt“). Die Interaktionseffekte der Prädiktoren (Anpassungen des Trainings) und des Messzeitpunkts können also als Entwicklungsunterschied zu dieser Gruppe interpretiert werden. Der Interaktionseffekt aus Messzeitpunkt und Erklärtext („Messzeitpunkt \times Erklärtext“; $b = -5.43$) bedeutet beispielsweise, dass Kinder, die mit Erklärvideos arbeiteten, weniger profitieren konnten als Kinder, die Strategien im Unterricht vermittelt bekamen. Sie steigerten sich im Vergleich dazu lediglich um 0.31 Normpunkte ($0.31 = 5.74 - 5.43$; Werte „Messzeitpunkt“ und „Messzeitpunkt \times Erklärtext“).

Abbildung 3: Entwicklung des Leseverstehens in Abhängigkeit der Anpassungen des FILBY-3 Trainings in der EG während der Corona-Pandemie



Die Balken kennzeichnen die Standardfehler der Regressionsgewichtsschätzungen.

Abbildung 3 zeigt, dass die von den Lehrkräften genutzten Alternativen im Vergleich zur Strategieeinführung in Präsenz signifikant kleinere Zuwächse im Leseverstehen aufweisen – lediglich der Einsatz von Erklärtexten hat keine signifikanten Nachteile. Weder der Einsatz von Erklärvideos kann die Vermittlung durch die Lehrkraft ersetzen, noch erweist sich der Verzicht auf Strategien als sinnvoll. Als besonders nachteilig entpuppt sich ein erneutes Training der Leseflüssigkeit, zumal die teilnehmenden Kinder diese bereits in der zweiten Jahrgangsstufe trainiert hatten. Hier beträgt der Effekt fast eine Standardabweichung (ca. 15 Normpunkte). Die genauen Modellkennwerte und Regressionsgewichte der Analyse sind aus Tabelle 5 ersichtlich.

Tabelle 5: Fortschritte im Leseverstehen nach Durchführungsmodalität des FiLBY-3-Trainings

	Modelle				
	Nullmodell		Volles Modell		
Feste Effekte	<i>b</i>	<i>p</i> -Wert	<i>b</i>	<i>p</i> -Wert	<i>p</i> -Wert korrigiert*
Konstante	99.77	< .001	100.48	< .001	< .001
Messzeitpunkt	-	-	5.74	< .001	.001
Messzeitpunkt x Erklärtext	-	-	-4.40	.055	.055
Messzeitpunkt x Erklärvideo	-	-	-5.43	.022	.044
Messzeitpunkt x ohne Strategien	-	-	-7.50	< .001	< .001
Messzeitpunkt x Leseflüssigkeitstraining	-	-	-13.54	< .001	< .001
Zufällige Effekte	σ^2		σ^2		
ICC _{Subjekt:Klasse}	.55		.50		
R ² _{marginal}	-		.04		
R ² _{konditional}	-		.52		

Anmerkung: * Bonferroni-Holm-Korrektur, *b* = Regressionskoeffizient, MZP = Messzeitpunkt, ICC = Intraklassenkorrelation, σ = Varianz, *x* bezeichnet Interaktionseffekte, N = 669 Schülerinnen und Schüler, 36 Klassen.

8 Diskussion

Ziel der Untersuchung war es, die Einflüsse des Lernens zuhause auf die Entwicklung des Leseverstehens von Grundschülerinnen und Grundschulern zu untersuchen.

Die Lehrkräfte reagieren in Hinblick auf die Leseförderung durchwegs flexibel auf die Situation des Distanzlernens. Die Lehrkräfte, die während der Schulschließungen mit FiLBY weiterarbeiteten, maßen der Leseförderung dabei einen höheren Stellenwert bei und investierten mehr Zeit in das Lesen als diejeni-

gen, die das FiLBY-Material nicht einsetzten. Beide Gruppen nutzten verschiedene Kommunikationswege und versuchten größtenteils, den persönlichen Kontakt aufrechtzuerhalten. Da das FiLBY-Material für Schülerinnen und Schüler gut vorstrukturiert ist, setzten die Lehrpersonen der Trainingsgruppe stärker auf die Selbstständigkeit der Kinder und die Unterstützung durch die Erziehungsberechtigten. Wie bereits in anderen Untersuchungen (z. B. Besa, Gensler, Gesang & Röhrig, 2020) spielt die Unterstützung durch die Erziehungsberechtigten keine Rolle für die Entwicklung der Schülerinnen und Schüler.

Die weiteren Analysen zeigen, dass sich die veränderten Lernbedingungen in der Phase der Schulschließungen ungünstig auf den Fortschritt der Kinder im Leseverstehen auswirkten. Es gelingt den Kindern im Distanzunterricht offensichtlich nicht, sich Lesestrategien selbstständig anzueignen bzw. diese ohne Hilfe zu üben, sodass die Trainingsgruppe zunächst keine Vorteile aufzuweisen scheint. Die teilnehmenden Lehrerinnen und Lehrer passten jedoch das FiLBY-3-Training ihrer Situation an. Dass während der Schulschließungen die technischen Mittel noch begrenzt waren, viele Lehrkräfte Videokonferenzsoftware kaum nutzen konnten, um die Strategien einzuführen oder gemeinsam mit Schülerinnen und Schüler zu üben, erschwerte den Lehrkräften das systematische Training. Es ist deshalb zu vermuten, dass vor allem die Lehrkräfte das Training nicht durchführten, die noch nicht damit begonnen hatten.

Die Lehrkräfte, die mit FiLBY-3 weiterarbeiteten, ersetzten die Strategieeinführung in Präsenz durch Alternativen wie beispielsweise das Lesen ohne Strategien, Erklärtexte und Erklärvideos sowie weitere Leseflüssigkeitsübungen analog zu FiLBY-2, die im Vergleich jedoch ungünstige Effekte für das Leseverstehen aufweisen. Möglicherweise auch deshalb, weil aufgrund der fehlenden Kontrolle der Lehrkraft nicht alle Lernenden in der Lage waren, die mit den Anpassungen verbundenen Lernchancen selbstständig wahrzunehmen. Dies betont die Rolle der Lehrkraft für die Strategievermittlung: Das effektive Lernen zuhause erfordert ein hohes Maß an Selbstregulation (Köller et al., 2020) und stellt hohe Ansprüche an strukturelle (wie z. B. Verfügbarkeit schnellen Internets) sowie soziale (z. B. Ansprechbarkeit der Lehrkraft) Ressourcen (Casale et al., 2020; Reiss, 2020), was viele Kinder zu überfordern scheint. Insbesondere scheint außerdem die Kontinuität des Trainings für den Erfolg wichtig zu sein. Das zentrale Element einer gelingenden Förderung des Leseverstehens während des Distanzunterrichts ist demnach die Einführung der Lesestrategien durch die Lehrkraft. Dies deutet darauf hin, dass schwächere Leserinnen und Leser zwar in besonderem Maße auf die Hilfe der Lehrperson angewiesen sind (Schilcher, Wild, Kraus & Hilbert, in Druck), für den Erwerb von Lesestrategien jedoch alle Kinder das Expertenmodell der Lehrkraft benötigen. Sind die Strategien eingeführt, können weitere Übungsphasen mit einem geringeren Grad an Unterstützung gestaltet werden.

Bemerkenswert sind die geringen Fortschritte im Leseverstehen, die diejenigen Klassen erreichten, die anstelle des Lesestrategietrainings weiter die Leseflüssigkeit trainierten. Zwar besteht ein theoretischer Zusammenhang zwischen Leseflüssigkeit und Leseverstehen, wenn aber die basalen Prozesse hinreichend flüssig bewältigt werden, trägt ein weiteres Training der Leseflüssigkeit nicht mehr maßgeblich zum Leseverstehen bei (Afflerbach, Pearson & Paris, 2008). Weitere, hier nicht berichtete, Analysen zeigen allerdings, dass ein solches Training durchaus noch signifikant positive Effekte bezüglich der Leseflüssigkeit aufweisen kann (vgl. OSF: <https://osf.io/9cspz/>). Dies dürfte u. a. auf einen Aufholeffekt der schwächeren Schülerinnen und Schüler zurückzuführen sein.

Bei der Interpretation der Ergebnisse müssen folgende Einschränkungen bedacht werden: Aufgrund der akuten Pandemiesituation konnten nur die Daten zweier Messzeitpunkte herangezogen werden, was die untersuchte Gruppe auf Kinder der dritten Jahrgangsstufe beschränkt. Es konnten außerdem nur Lehrpersonen der Experimentalgruppe befragt werden. Die berichteten Ergebnisse beruhen überwiegend auf der Perspektive dieser Lehrkräfte. Darüber hinaus ist denkbar, dass die für den Distanzunterricht entwickelten Lernangebote von den Schülerinnen und Schülern nicht wie vorgesehen genutzt wurden, etwa wenn hilfreiche Erklärvideos zur Strategievermittlung zwar zur Verfügung gestellt, aber von den Kindern nicht abgerufen wurden. Weiterer Forschungsbedarf ergibt sich außerdem in Hinblick auf verschiedene Schülergruppen, etwa mit unterschiedlichem Leistungsniveau, sozioökonomischen Hintergrund oder Migrationshintergrund.

Literatur

- Afflerbach, P., Pearson, P. D. & Paris, S. (2008): Clarifying Differences Between Reading Skills and Reading Strategies. *The Reading Teacher*, 61(5), 364–373. doi: 10.1598/RT.61.5.1
- Almasi, J. & Hart, S. (2019): Best Practices in Narrative Text Comprehension Instruction. In L. Morrow & L. Gambrell (Eds.), *Best Practices in Literacy Instruction* (pp. 221–249). New York, London: Guilford Press.
- Andersen, E. B. (1973): A Goodness of Fit Test for The Rasch Model. *Psychometrika*, 38(1), 123–140. doi: 10.1007/BF02291180
- Anmarkrud, Ø. & Bråten, I. (2012): Naturally-Occurring Comprehension Strategies Instruction in 9th-Grade Language Arts Classrooms. *Scandinavian Journal of Educational Research*, 56(6), 591–623. doi: 10.1080/00313831.2011.621134
- Artelt, C. & Dörfler, T. (2011): Förderung von Lesekompetenz als Aufgabe aller Fächer. In Bayerisches Staatsministerium für Unterricht und Kultus (Hrsg.), *ProLesen. Auf dem Weg zur Leseschule – Leseförderung in den gesellschaftswissenschaftlichen Fächern* (S. 13–36). Donauwörth: Auer.
- Bandura, A. (1979): *Sozial-kognitive Lerntheorie*. Stuttgart: Klett-Cotta.

- Besa, K., Gensler, A., Gesang, J. & Röhrig, E. (2020): Erforschung der Zusammenhänge von elterlichem Belastungsempfinden und binnendifferenzierenden digitalen Lernmaterialien in Zeiten der Corona-Pandemie. *PFLB PraxisForschungLehrer*innenBildung*, 2(6), 45–58. doi: 10.4119/pflb-3928
- Bremerich-Vos, A., Wendt, H. & Bos, W. (2017): Lesekompetenzen im internationalen Vergleich: Testkonzeption und Ergebnisse. In A. Hußmann, H. Wendt, W. Bos, A. Bremerich-Vos, D. Kasper, E. Lankes, N. McElvany, T. C. Stubbe & R. Valtin (Hrsg.), *IGLU 2016. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 79–142). Münster: Waxmann.
- Brown, A. L. & Palinscar, A. S. (1989): Guided Cooperative Learning and Individual Knowledge Acquisition. In L. B. Resnick (Ed.), *Knowing, Learning, And Instruction: Essays in Honor of Robert Glaser* (pp. 393–451). Hillsdale, NJ: Erlbaum.
- Casale, G., Börnert-Ringleb, M. & Hillenbrand, C. (2020): Fördern auf Distanz? Sonderpädagogische Unterstützung im Lernen und in der emotionalen-sozialen Entwicklung während der Schulschließungen 2020. *Zeitschrift für Heilpädagogik*, 71(5), 254–267.
- Chen, J., Wang, M., Kirschner, P. & Tsai, C. (2018): The Role of Collaboration, Computer Use, Learning Environments, and Supporting Strategies In CSCL: A Meta-Analysis. *Review of Educational Research*, 88(6), 799–843. doi: 10.3102/0034654318791584
- Diedrich, J., Schiepe-Tiska, A., Ziernwald, L., Tupac-Yupanqui, A., Weis, M., McElvany, N. & Reiss, K (2019): Lesebezogene Schülermerkmale in PISA 2018: Motivation, Leseverhalten, Selbstkonzept und Lesestrategiewissen. In K. Reiss, M. Weis, E. Klieme, O. Köller (Hrsg.), *PISA 2018. Grundbildung im internationalen Vergleich* (81–110). Münster: Waxmann.
- Duffy, G. (1993): Rethinking Strategy Instruction: Four Teachers' Development and Their Low Achievers' Understandings. *The Elementary School Journal*, 93(3), 231–247. doi: 10.1086/461724
- Dumont, H. (2019): Neuer Schlauch für alten Wein? Eine konzeptuelle Betrachtung von individueller Förderung im Unterricht. *Zeitschrift für Erziehungswissenschaft*, 22(2), 249–277. doi: 10.1007/s11618-018-0840-0
- Durkin, D. (1978): What Classroom Observations Reveal about Reading Comprehension Instruction. *Reading Research Quarterly*, 14(4), 481–533.
- Edmonds, M. S., Vaughn, S., Wexler, J., Reutebuch, C., Cable, A., Klingler Tackett, K. & Wick Schnakenberg, J. (2009): A Synthesis of Reading Interventions and Effects on Reading Comprehension Outcomes for Older Struggling Readers. *Review of Educational Research*, 79(1), 262–300. doi: 10.3102/0034654308325998
- Ehlich, Konrad (1984): Zum Textbegriff. In A. Rothkegel & B. Sandig (Hrsg.), *Text – Textsorten – Semantik. Linguistische Modelle und maschinelle Verfahren* (S. 9–25). Hamburg: Buske.
- Fickermann, D. & Edelstein, B. (2020): „Langsam vermisste ich die Schule ...“ Schule während und nach der Corona-Pandemie. *Die Deutsche Schule*, 16. Beiheft, 9–36. doi: 10.31244/9783830992318.01
- Hasselhorn, M. & Gold, A. (2013): *Pädagogische Psychologie. Erfolgreiches Lernen und Lehren*. Stuttgart: Kohlhammer.
- Hilbert, S., Stadler, M., Lindl, A., Naumann, F. & Bühner, M. (2019): Analyzing Longitudinal Intervention Studies with Linear Mixed Models. *Testing, Psychometry, Methodology in Applied Psychology*, 26, 101–119. doi: 10.4473/TPM26.1.6
- Holm, S. (1979): A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics*, 6(2), 65–70.

- Holzberger, D., Reinhold, S., Lüdtke, O. & Seidel, T. (2020): A Meta-Analysis on the Relationship Between School Characteristics and Student Outcomes in Science and Math – Evidence from Large-Scale Studies. *Studies in Science Education*, 56(1), 1–34. doi: 10.1080/03057267.2020.1735758
- Huber, S. G. & Helm, C. (2020): Lernen in Zeiten der Corona-Pandemie. Die Rolle familiärer Merkmale für das Lernen von Schüler*innen: Befunde vom Schul-Barometer in Deutschland, Österreich und der Schweiz. *Die Deutsche Schule*, 16. Beiheft, 37–60. doi: 10.31244/9783830992318.02
- Kerres, M. (2018): *Mediendidaktik. Konzeption und Entwicklung digitaler Lernangebote*. Oldenburg: de Gruyter.
- Klieme, E. (2020): Teaching Quality – Also and Especially under the Constraints of the Pandemic? *Die Deutsche Schule*, 16. Beiheft, 117–135. doi: 10.31244/9783830992318.07
- Köllner, O., Fleckenstein, J., Guill, K. & Meyer, J. (2020): Educational Challenges of Home Learning. *Die Deutsche Schule*, 16. Beiheft, 163–174. doi: 10.31244/9783830992318.10
- Kraus, E., Wild, J., Schilcher, A., Hilbert, S. (2021): *Research Report zur Pilotierung des Bayerischen Lesetests (BYLET)*. https://osf.io/bmp54/?view_only=3db22c0b3cec4383820514243e431b1b.
- Kunter, M. & Ewald, S. (2016): Bedingungen und Effekte von Unterricht: Aktuelle Forschungsperspektiven aus der pädagogischen Psychologie. In N. McElvany, W. Bos, H. G. Holtappels, M. Gebauer & F. Schwabe (Hrsg.), *Bedingungen und Effekte guten Unterrichts* (S. 9–31). Münster: Waxmann.
- Lenhard, W. (2019): Gegenwärtiger Stand der empirischen Unterrichtsforschung zur Vermittlung von Lesekompetenz. In M. Kämper-van den Boogaart & K. H. Spinner (Hrsg.), *Lese- und Literaturunterricht. Teil 2: Kompetenzen und Unterrichtsziele, Methoden und Unterrichtsmaterialien* (S. 305–329). Baltmannsweiler: Scheider Hohengehren.
- Lienemann, T. O. & Reid, R. (2006): Self-Regulated Strategy Development for Students with Learning Disabilities. *Teacher Education and Special Education*, 29(1), 3–11. doi: 10.1177/088840640602900102
- Mayer, A. & Marks, D. (2019): Förderung des Textverständnisses durch die Vermittlung von Verstehensstrategien – Eine Metaanalyse zur Effektivität. *Forschung Sprache*, 7 (1), 4–36.
- McElvany, N., Becker, M. & Lüdtke, O. (2009): Die Bedeutung familiärer Merkmale für Lesekompetenz, Wortschatz, Lesemotivation und Leseverhalten. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 41(3), 121–131. doi: 10.1026/0049-8637.41.3.121
- McElvany, N., Kessels, U., Schwabe, F. & Kasper, D. (2017): Geschlecht und Lesekompetenz. In A. Hußmann, H. Wendt, W. Bos, A. Bremerich-Vos, D. Kasper, E. Lankes, N. McElvany, T. C. Stubbe & R. Valtin (Hrsg.), *IGLU 2016. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 177–194). Münster: Waxmann.
- Niederbacher, E. & Neuenschwander, M. (2020): Herkunftsbedingte Leistungsdisparitäten: Die Rolle von Selbstwirksamkeitsüberzeugungen und Unterstützungshandlungen von Eltern und Leistungserwartungen von Lehrpersonen. Generalisierbarkeit eines Mediationsmodells für einsprachige und fremd- bzw. mehrsprachige Schülerinnen und Schüler. *Zeitschrift für Erziehungswissenschaft*, 23(4), 739–767. doi: 10.1007/s11618-020-00955-9
- Nonte, S., Hartwich, L. & Willems, A. S. (2018): Promoting Reading Attitudes of Girls and Boys: A New Challenge for Educational Policy? Multi-Group Analyses Across Four European Countries. *Large-Scale Assessments in Education*, 6(5), 1–22. doi: 10.1186/s40536-018-0057-y

- OECD (2019): *PISA 2018 Results (Volume III): What School Life Means for Students' Lives.* Paris: OECD. doi: 10.1787/acd78851-en
- Okkinga, M., van Steensel, R., van Gelderen, A. J., van Schooten, E., Slegers, P. J. & Arends, L. R. (2019): Effectiveness of Reading-Strategy Interventions in Whole Classrooms: A Meta-Analysis. *Educational Psychology Review*, 30 (4), 1215–1239. doi: 10.1007/s10648-018-9445-7
- Pearson, P. D. & Cervetti, G. N. (2017): The Roots of Reading Comprehension Instruction. In S. E. Israel (Ed.), *Handbook of Research on Reading Comprehension* (pp. 12–56). New York, NY: The Guilford Press.
- Philipp, M. (2010): Peer Assisted Learning in der Lesedidaktik am Beispiel Lesestrategie-Trainings. Teil 1: Definitionen, empirische Evidenz und Programmvergleich. *www.leseforum.ch*, 3, 1–15.
- Philipp, M. (2015): *Lesestrategien: Bedeutung, Formen und Vermittlung*. Weinheim: Beltz Juventa.
- Philipp, M. (2017): Lesestrategieinsatz bei Sach- und literarischen Texten. Unterschiedliche Textsorten, unterschiedliche Anforderungen, unterschiedliche Strategien? In D. Scherf (Hrsg.), *Inszenierungen literalen Lernens: kulturelle Anforderungen und individueller Kompetenzerwerb* (S. 116–125). Baltmannsweiler: Schneider Hohengehren.
- Pissarek, M. (2018): Zum Begriff der Lesekompetenz – Förderung von Lesekompetenz bei jungen Erwachsenen. In P. Resinger (Hrsg.), *Förderung der Lesekompetenz bei Jugendlichen in Ausbildung. Grundlagen – Konzepte – Praxisbeispiele* (S. 5–14). Innsbruck: Teatum.
- Porsch, R. & Porsch, T. (2020): Homeschooling as an Exceptional Situation. Findings from a Nationwide Survey of Parents with Primary School Children. *Die Deutsche Schule*, 16. Beiheft, 61–78. doi: 10.31244/9783830992318.03
- R Core Team (2020): *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. <https://www.R-project.org/>.
- Rasch, G. (1960): *Studies In Mathematical Psychology: I. Probabilistic Models for Some Intelligence and Attainment Tests*. Kopenhagen: Nielsen & Lydiche.
- Reiss, K. (2020): Lernen mit digitalen Medien: das Beispiel des Fachs Mathematik. In K. Kaspar, M. Becker-Mrotzek, S. Hofhues, J. König & D. Schmeinck (Hrsg.), *Bildung, Schule, Digitalisierung* (S. 284–290). Münster: Waxmann.
- Rosebrock, C. & Nix, D. (2020): *Grundlagen der Lesedidaktik und der systematischen schulischen Leseförderung* (9., aktual. Aufl.). Baltmannsweiler: Schneider.
- Rubach, C. & Lazarides, R. (2021): Wechselseitige Effekte zwischen Lehrkräften, Eltern und Lernenden: Zusammenhänge von kognitiver und emotionaler Unterstützung und der Motivation von Lernenden. In G. Hagenauer & D. Raufelder (Hrsg.), *Soziale Eingebundenheit. Sozialbeziehungen im Fokus von Schule und Lehrer*innenbildung* (S. 287–302). Münster: Waxmann.
- Schilcher, A., Wild, J. & Steinert, M. (2019): *FiLBY-2 Lehrerhandreichung. Fachintegrierte Leseförderung Bayern*. Regensburg: o. V. <https://fibs.alp.dillingen.de>
- Schilcher, A., Wild, J., Kraus, E., Hilbert, S. (in Druck): Kann das Leseförderprogramms „FiLBY-2“ den Einfluss des sozio-ökonomischen Hintergrunds reduzieren? *SLLD-Z*.
- Segers, E. & Verhoeven, L. (2009): Learning In a Sheltered Internet Environment: The Use of Webquests. *Learning and Instruction*, 19, 423–432. doi: 10.1016/j.learninstruc.2009.02.017

- Slavin, R. E., Lake, C., Chambers, B., Cheung, A. & Davis, S. (2009): Effective Reading Programs for The Elementary Grades: A Best-Evidence Synthesis. *Review of Educational Research*, 79(4), 1391–1466. doi: 10.3102/0034654309341374
- Souvignier, E. & Antoniou, F. (2007): Förderung des Leseverständnisses bei Schülerinnen und Schülern mit Lernschwierigkeiten – eine Metaanalyse. *Vierteljahrszeitschrift für Heilpädagogik und ihre Nachbargebiete*, 76(1), 46–62.
- Souvignier, E., Förster, N. & Kawohl, E. (2016): Implementation eines Förderkonzeptes zur diagnosebasierten individuellen Leseförderung in der Grundschule. In M. Philipp & E. Souvignier (Hrsg.), *Implementation von Lesefördermaßnahmen. Perspektiven auf Gelingenbedingungen und Hindernisse* (S. 77–98). Münster: Waxmann.
- Stahns, R., Rieser, S. & Lankes, E. (2017): Unterrichtsführung, Sozialklima und kognitive Aktivierung im Deutschunterricht in vierten Klassen. In A. Hußmann, H. Wendt, W. Bos, A. Bremerich-Vos, D. Kasper, E. Lankes, N. McElvany, T. C. Stubbe & R. Valtin (Hrsg.), *IGLU 2016. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 251–278). Münster: Waxmann.
- Stalder, U. (2013): *Leselust in Risikogruppen. Gruppenspezifische Wirkungszusammenhänge*. Wiesbaden: Springer.
- Stanovich, K. E. (2000): *Progress in Understanding Reading: Scientific Foundations and New Frontiers*. New York: Guilford.
- Steinert, M., Schilcher, A. & Wild, J. (2021): *FiLBY-4 Lehrerhandreichung. Fachintegrierte Leseförderung Bayern*. Regensburg: o. V. <https://fibs.alp.dillingen.de>
- Wacker, A., Unger, V. & Rey, T. (2020): „Sind doch Corona-Ferien, oder nicht?“ Befunde einer Schüler*innenbefragung zum „Fernunterricht“. *Die Deutsche Schule*, 16. Beiheft, 79–94. doi: 10.31244/9783830992318.04
- Weis, M., Doroganova, A., Hahnel, C., Becker-Mrotzek, M., Lindauer, T., Artelt, C. & Reiss, K. (2019): Lesekompetenz in PISA 2018 – Ergebnisse in einer digitalen Welt. In K. Reiss, M. Weis, E. Klieme, O. Köller (Hrsg.), *PISA 2018. Grundbildung im internationalen Vergleich* (S. 47–80). Münster: Waxmann.
- Wild, J. & Schilcher, A. (2019): Grundlagen einer systematischen schulischen Leseförderung. In G. Stückl & M. Wilhelm (Hrsg.), *Lehren und Lernen in der bayerischen Grundschule* (S. 1–17). Kronach, Köln: Carl Link.
- Wild, J., Schilcher, A. & Steinert, M. (2020): *FiLBY-3 Lehrerhandreichung. Fachintegrierte Leseförderung Bayern*. Regensburg: o. V. <https://fibs.alp.dillingen.de>
- Willenberg, H. (2004): Lesestrategien. Vermittlung zwischen Eigenständigkeit und Wissen. *Praxis Deutsch*, 187, 6–15.
- Wimmer, H. & Mayringer, H. (2016): *Salzburger Lese-Screening für die Schulstufen 2–9: SLS 2–9*. Bern: Huber.

Synthesen schreiben – eine textbasierte Längsschnittuntersuchung studentischer Lese- und Schreibperformanz

Alex Rickert & Maik Philipp

Zusammenfassung: In der Forschung zum Schreiben von Diskurssynthesen ist noch wenig geklärt, wie sich kognitive Prozesse des Auswählens, Verknüpfens und Organisierens in Texten niederschlagen, wie diese miteinander zusammenhängen und wie sie mit der Textqualität korrelieren. In unserer Untersuchung haben wir diese Fragen adressiert, indem wir ausgehend von Textprodukten kognitive Prozesse rekonstruiert haben. Dafür haben wir Diskurssynthesen von 25 Lehramtsstudierenden der Kindergarten-, Primar und Sekundarstufe zu zwei Messzeitpunkten analysiert und daraus Rückschlüsse auf die genannten kognitiven Prozesse gezogen. Die Resultate zeigen, dass das Verknüpfen und Organisieren von Inhalten den Studierenden am meisten Schwierigkeiten bereiteten, Verknüpfungsprozesse aber gleichzeitig am stärksten mit der Textqualität zusammenhängen. Zudem zeigt die Empirie, dass Auswahl-, Verknüpfungs- und Organisationsprozesse entgegen theoretischen Annahmen wenig miteinander verbunden zu sein scheinen.

Schlüsselwörter: Diskurssynthese, kognitive Prozesse, Auswählen, Verknüpfen, Organisieren

Abstract: In research on writing discourse syntheses, there is still little consensus of how cognitive processes of selecting, connecting and organizing are reflected in texts, how they are interrelated, and how they correlate with text quality. In our study, we addressed these questions by means of reconstructing cognitive processes by analyzing written products. For this sake, we examined discourse syntheses of 25 student teachers at kindergarten, primary and secondary school level at two measurement time points and drew conclusions about the cognitive processes. The results show that the connection and organization of contents caused the most difficulties for the test persons, but the process of connecting was also most strongly related to text quality. Moreover, empirical evidence shows that, contrary to theoretical assumptions, processes of selection, connection and organization seem to be little interrelated.

Keywords: discourse synthesis, cognitive processes, selecting, connecting, organizing

1 Einleitung

Die Fähigkeit, aus verschiedenen Texten zu lernen und Informationen schriftlich weiterzuverarbeiten, gewinnt in Bildungs- und Berufskontexten an Bedeutung. Längst hat sich ein neues Forschungsfeld etabliert, das den versierten Umgang mit multiplen Texten zu kartieren beginnt: dasjenige der „multiple document literacy“ (List & Alexander, 2019). Diese Fähigkeit, mehrere einzelne Dokumente zu Zwecken der Bewältigung von Aufgaben sinnvoll zu nutzen, zieht aus Produkt- und Prozessperspektive immer mehr Aufmerksamkeit auf sich (Primor & Katzir, 2018), zuletzt sogar bei PISA 2018. Auch die Schreibforschung hat sich unter dem Stichwort „Diskurssynthese“ dem Hybrid aus Lesen und Schreiben intensiver gewidmet (van Ockenburg et al., 2019).

Einigkeit besteht darin, dass hybride Lese- und Schreibprozesse als interaktive strategische Prozesse zusammenwirken müssen, damit Personen erfolgreich einen eigenen Text herstellen können, der auf mehreren gelesenen Texten basiert und zur Aufgabenstellung passt (Philipp, 2020). Es sind also Lese- und Schreibforschung gleichermaßen gefragt, Prozesse und Produkte theoretisch zu fassen, empirisch zu erforschen und didaktisch zu fördern.

Dieser Beitrag greift dies auf. Er widmet sich im Sinne der Grundlagenforschung der Frage, ob sich theoretisch postulierte Prozesse in studentischen Diskurssynthesen rekonstruieren lassen, wie sie sich im kurzen Längsschnitt verändern und wie solche produktbezogenen Maße – darunter auch solche aus dem Kontext der Schreibforschung – zusammenhängen. Hierfür haben wir 25 Studierende zweimal im Abstand von drei Monaten dieselben Syntheseaufgaben bearbeiten lassen und die Produkte mehrdimensional analysiert. Aus produktbezogenen Analysen haben wir Rückschlüsse auf kognitive Prozesse des Auswählens, Verknüpfens und Organisierens gezogen (Primor & Katzir, 2018; Segev-Miller, 2007).

Der Beitrag hat folgenden Aufbau: Wir gehen zunächst auf die Trias kognitiver Prozesse bei den Diskurssynthesen ein (Kapitel 2) und verdeutlichen die kognitiv anspruchsvollen Erfordernisse für das Lesen und Schreiben. Dabei greifen wir auf die in der Studie zum Einsatz gelangte Aufgabe und auf einen studentischen Text aus dem Korpus zurück. Im Kapitel 3 legen wir unsere Methodik ausführlich dar und präsentieren im Kapitel 4 die Befunde zu unseren sechs Fragestellungen. Das kurze Fazit in Kapitel 5 bündelt und diskutiert die Essenz, wobei Forschungsdesiderate skizziert werden.

2 Kognitive Prozesse beim Schreiben von Diskurssynthesen

Schreibprozesse, die mehrere konvergierende oder konfligierende Texte oder Dokumente zur Grundlage haben, werden als „writing from sources“, „materialgestütztes Schreiben“ oder „schriftliches Synthetisieren“ bezeichnet. Ein prototypisches Resultat, das aus solchen Lese- und Schreibprozessen resultiert, ist die *Diskurssynthese*. Unter Diskurssynthese versteht Nelson (2001) einen Diskurs (verbale Abhandlung eines Themas), der durch Synthese (Kombinieren und Abgleichen) anderer Diskurse hervorgeht. Verfasserinnen und Verfasser von Diskurssynthese konstruieren einen Text mit eigener Struktur, der auf Inhalte aus anderen Texten zurückgreift und diese für eigene Textziele relevant setzt und transformiert.

2.1 Ein Einstiegsbeispiel

Dies lässt sich leicht mit einem Beispiel verdeutlichen, das wir zur Entfaltung der Thematik im Folgenden nutzen möchten. Es handelt sich dabei um die in unserer Studie zum Einsatz gelangte Aufgabe (s. Kasten 1) und eine aus dem Pool der entstandenen Synthesen stammende Lösung (Kasten 2). Wir nutzen dieses Beispiel, um im Folgenden auf die Erfordernisse des materialgestützten Schreibens einzugehen, wenn es um die Prozesse geht, welche für diesen Umgang mit multiplen Dokumenten typisch sind. Unterschiede zwischen den Bezugstexten hinsichtlich Textschwierigkeit werden in Teilkapitel 3.3.1 erläutert.

Kasten 1: Arbeitsauftrag und Bezugstexte für eine Aufgabe zum schriftlichen Synthetisieren

Arbeitsauftrag

Bitte schreiben Sie eine Synthese über die beiden unteren Texte, welche die geäußerten Standpunkte von F. Mole und M. Crahay berücksichtigt. Ihre Synthese soll diese Frage beantworten: Wofür gibt es die Schule und was sind ihre Ziele? Erklären Sie ihre Funktion. Bedenken Sie bitte, dass eine Synthese die Funktion hat, die Positionen aus beiden Texten angemessen wiederzugeben und Gemeinsamkeiten und Unterschiede in den Positionen zu verdeutlichen. Ihr Text sollte rund 2000 Zeichen (inkl. Leerzeichen) umfassen.

Text 1

Mole, F. (2013). Die einzelne Schule: kollektive oder individuelle Emanzipation? Von den Anfängen des 20. Jahrhunderts bis in die 1930er. *La Télémaque*, 103–104.

[1] „Emanzipation durch Schule.“ Zusätzlich zu seiner [2] intellektuellen Bedeutung hat dieser Ausdruck eine [3] soziale Bedeutung, die man entweder aus einer [4] kollektiven Perspektive oder von einem [6] persönlichen Standpunkt verstehen kann. Gemäß der [4] kollektiven Perspektive [5] sollte Schule junge Menschen ermutigen, alle charakterlichen Züge der Unterordnung zu verbannen, um politisch autonomer zu werden. [6] Gemäß dem individuellen Standpunkt [7] gibt Schule den Individuen eine Möglichkeit, um sich von ihrem sozialen Hintergrund zu befreien. [8] Es gibt zwei verschiedene französische republikanische Sichtweisen der Bildung:

[9] 1) Schule muss die Bedingungen schaffen, die man benötigt, um politische und soziale Demokratie zu fördern, [10] indem man die intellektuellen Ressourcen der Menschen entwickelt. Erstgenannte wird als kollektives, beherrschtes Gebilde betrachtet, das frei von jeder Art von Autorität ist und seine Souveränität ausübt. Ein Jahrhundert nach Concordet sagte Jaurès¹, dass Bildung Individuen erlaubt, zu agieren, um eine bessere Republik und eine gerechtere Gesellschaft zu schaffen.

[11] 2) Schulen müssen Individuen hervorbringen, die fähig sind, menschliche Belange in einer Gesellschaft zu behandeln, in der Funktionen und Status nicht das Ergebnis von Zufall sind. [12] Seit dem Beginn des 20. Jahrhunderts haben Reformatoren die Grenzen kritisiert, die zwischen verschiedenen Kontexten bestehen, und für eine stärker vereinte Art der Bildung plädiert, indem man die Gemeinschaft erweitert, aus der die Mitglieder der Elite rekrutiert werden und das Land sein Humankapital² entwickelt, und gleichzeitig alle Individuen gerecht behandelt.

¹ Concordet schlug vor, durch allgemeine Bildung den Menschen die Hilfsmittel zu geben, um „eine perfektere Verfassung zu erhalten“, und Jaurès behauptete, dass der „Sozialismus eine Konsequenz der sukzessiven Bürgerrechte zu sein scheint“ (Der sozialistische Staat und Funktionäre, *La Revue Socialiste*, 21, 124, April 1985, S. 387–388).

² F. Buisson, L. Gérard-Varet, J. Bouveri, Gesetzentwurf zu den gleichen Rechten in der Bildung für Kinder, (Frankreich: Repräsentantenhaus, 9. Legislatur, Sitzung von 1910, Nr. 3265), S. 8.

Text 2

Crahay, M. (2000). *Können Schulen fair und effizient sein? Von der Chancengleichheit zur Gleichheit der Ergebnisse*. Brüssel: De Boeck & Larcier.

Gemäß Dubet und Martuchelli (1996) haben [13] alle Schulsysteme drei wichtige *Funktionen* zu erfüllen und können danach definiert werden, wie sie Funktionen hierarchisieren und artikulieren.

[14] Die erste Funktion – *Bildung* – ist verbunden mit den [15] persönlichen Projekten der Schüler, die, [16] insofern sie zur Reflexion fähig sind, [15] diese Projekte selbst regulieren und bestimmen, [16] indem sie Situationen, in denen sie sich befinden, rational analysieren. Durkheim (1990) hat speziell diese Dimension hervorgehoben, indem er daran erinnerte, dass Schulen das Produkt eines christlichen Projekts sind, Individuen in ein anderes Leben zu überführen.

[17] Die zweite Funktion ist *Sozialisation*. [18] Schulen sollten das Aufkommen von Individuen fördern, die an die Gesellschaft angepasst sind. [19] Für dieses Ziel müssen Schulen die Normen, Gewohnheiten, Wissensbestände und Werte integrieren, die von der sozialen Gruppe geschätzt wird, in der diese Individuen Mitglieder werden sollen.

[20] Die dritte Funktion ist *Verteilung*. Dies [22] „betrifft die Rollen, die die Schule erfüllt bezüglich der Abschlüsse, die einen sozialen Nutzen haben aufgrund des Fakts, dass bestimmte Jobs, Positionen und Status das Vorrecht der Akademiker sind. [21] Die Schule verteilt *Vermögensbestände*, die einen bestimmten Wert in den Arbeitsmärkten und der sozialen Hierarchie haben“ (Dubet und Martucelli, 1996, S. 23).

(Quelle: Übersetzung und leichte Modifikation von Escorcía et al., 2017, S. 269 f.; Zahlen in eckigen Klammern geben an, welche Propositionen aus der Musterlösung unserer Studie an welcher Stelle im Text vorkommen; das Lesemodell ist dargestellt in Abbildung 1 in Teilkap. 3.3.1)

Kasten 2: Authentische studentische Synthese zur Aufgabe aus Kasten 1

Synthese (studentischer Text)

F. Mole beschreibt in seinem Text die Funktion der Schule aus zwei Perspektiven. Aus der [4] kollektiven Perspektive betrachtet, soll die Schule junge Menschen dazu [5] ermutigen, charakteristische Züge der [37] Unterordnung zu verbannen, um sich so politisch autonom in der Gesellschaft zu bewegen. Dies steht im Widerspruch zur Aussage von M. Crahay, der als eine [13] der drei Funktionen der Schule die [17] [34] [35] Sozialisation sieht, welche [18] Individuen hervorbringen soll, die an die Gesellschaft angepasst sind. Die Individuen sollen sich also nach M. Crahay unterordnen und [19] sich den Werten, Normen und Gewohnheiten und Wissensbeständen der entsprechenden Gesellschaft anpassen und nicht völlig autonom verfügen können. Jedoch beschreibt auch Crahay, dass es dennoch zur Aufgabe der Schule gehört, die Schüler und Schülerinnen in ihrer [15] persönlichen Entwicklung zu unterstützen, damit sie sich selbstreguliert und selbstbestimmt entwickeln können, jedoch unter der Voraussetzung, dass sie sich den bereits erwähnten gesellschaftlichen Kriterien fügen.

[31] [6] Bei der individuellen Perspektive liegt das [7] Augenmerk auf der Befreiung vom sozialen Hintergrund. [32] Es soll nach Mole unter anderem [11] kein Zufall sein, welche Funktion oder welchen Status ein Individuum erlangt. [29] [30] Diesen Punkt widerlegt Crahay in der Beschreibung der dritten Funktion der Schule, der [20] Verteilung. Hiernach sind die Rollen betroffen, die die Schule gemäß den [22] Qualifikationen erfüllt, die vor dem Hintergrund des Faktes einen sozialen Nutzen haben, dass bestimmte Jobs, Positionen und Status das Vorrecht der Akademiker sind. Zudem sei es Aufgabe der Schule, [21] Vermögensbestände zu verteilen, die einen bestimmten Wert in den Arbeitsmärkten und der sozialen Hierarchie haben. Dieser Ansicht nach wird eine Loslösung vom sozialen Hintergrund nicht angestrebt, sondern orientiert sich stark nach dem sozialen Hintergrund des einzelnen Individuums.

Abschließend soll aber erwähnt werden, dass Mole wie auch Crahay sich aber in einem Punkt einig sind und zwar, dass [23] die Schule die ganz allgemeine Aufgabe hat, Menschen in ihrer Entwicklung zu begleiten und zu fördern.

(Eckige Zahlen geben die Propositionen aus der Benchmarklösung in Abbildung 1 an; Zahlen von 1 bis 12 beziehen sich auf den Bezugstext 1, Zahlen von 13 bis 22 auf den Bezugstext 2 und Zahlen von 23 bis 38 auf das integrierte mentale Modell. Die Unterstreichungen markieren Textstellen mit Bezügen zur Benchmarklösung.)

2.2 Drei Prozesse im Verbund

Das Verfassen von Diskurssynthesen ist hochgradig strategisch, was aus der Zahl und Vernetztheit kognitiver und metakognitiver Prozesse beim Lesen und beim Schreiben resultiert. Spivey (1990) benennt drei Prozesse: das Auswählen, das Organisieren und das Verknüpfen. Wir werden diese drei Prozesse im Folgenden beschreiben und dabei zunächst das Lesen und danach das Schreiben thematisieren.

2.2.1 Auswählen

Das Auswählen bezeichnet einen kognitiven Gewichtungsvorgang, der die Selektion der textuellen Inhaltseinheiten (Propositionen) steuert. Diesem Gewichtungsvorgang liegt also die Analyse der in den Texten enthaltenen Propositionen zugrunde (Philipp, 2018). Im Falle des *Lesens* gilt es, aus Texten jene

Bedeutungseinheiten zu extrahieren, die in einer Hierarchie eine über- oder untergeordnete Position aufweisen. Deshalb unterscheidet man in der Forschung zwischen sogenannten intratextuellen Mikro- und Makropropositionen (Kintsch & van Dijk, 1978) – im Falle des Texts 2 aus dem Kasten 1 wären solche Makropropositionen die drei explizit genannten Funktionen der Schule, die sogar noch typografisch durch Kursivierungen hervorgehoben wurden.

Die Gewichtung von Propositionen beim Lesen vollzieht sich häufig auf der Basis von textseitig rekonstruierter *Wichtigkeit*. Verknüpft ausgedrückt geht es darum, dass Leserinnen und Leser die Makrostruktur des Textes analysebasiert als propositionale Struktur unter Vorwissensbezug rekonstruieren, wobei sie auch implizite Aussagen durch Inferenzen anreichern (Spivey, 1990; Kintsch & van Dijk, 1978). Eine andere Möglichkeit, Informationen in die eigene Verarbeitung einzubeziehen, besteht darin, die *Relevanz* als Gradmesser für die Selektion zu nutzen (McCrudden, 2018). Es geht hier um eine Verwertbarkeit von Informationen auf der Basis textexterner, zielbezogener Einschätzungen.

Die Auswahlprozesse des Lesens bilden die Basis für das *Schreiben*, wie es theoretische Modelle des aufgabenbasierten Lesens explizit postulieren. Das MD-TRACE-Modell (Rouet & Britt, 2011) etwa nimmt an, dass Zielvorstellungen (Aufgabenmodelle) für einen Text wie eine Diskurssynthese als metakognitive Filter fungieren und Prüfungs- und Entscheidungsprozesse steuern, um adäquate Informationen für den eigenen Text zu nutzen.

Ein in der Forschung relativ wenig bearbeitetes Feld bezüglich der Auswahl an Informationen ist, welche Informationen aus welchen Texten ausgewählt werden. Zwar ist der Abdeckungsgrad von Makropropositionen ein recht übliches Maß (Philipp, 2020; Primor & Katzir, 2018). Gleichwohl ist es unbefriedigend, dass diese Summenwerte notwendigerweise wenig darüber aussagen, aus welchen Texten Personen welche Inhalte auswählen. Dadurch werden nämlich spezifische textseitige Merkmale bezogen auf die Einzeltextebene nicht mehr sichtbar. Daher widmen wir uns in dieser Studie folgender *erster Forschungsfrage*: Welchen Anteil an Propositionen übernehmen Studierende aus welcher hierarchischen Ebene der beiden Bezugstexte? Wir untersuchen hierfür genauer, ob die in Bezugstexten enthaltenen Propositionen in studentischen Synthesen explizit oder implizit vorhanden sind und ob sie inhaltlich korrekt übernommen wurden.

2.2.2 Verknüpfen

Der zweite Prozessverbund ist das inferenzbasierte Verknüpfen, das sachlogisch auf dem Auswählen aufbaut und das Anwenden von abstrahierenden Makroregeln des Textverstehens erfordert (Kintsch & van Dijk, 1978). Inferenzen sind auf Schlussfolgerungen basierende Verknüpfungsprozesse (Kendeou, 2015). Inferenzen sind ein mehrdimensionales Konstrukt, welches eint, dass es sich

um Schlussfolgerungen handelt, bei denen das Vorwissen eine entscheidende Rolle spielt (Philipp, 2020). Personen müssen beim Lesen selbstständig die Bezüge zwischen Informationen innerhalb von Texten und über Texte hinweg aktiv herstellen, indem sie übergreifende Themen als solche erkennen und ein integriertes mentales Modell aufbauen, welches inferenzbasierte Inhalte enthält. Dieses integrierte mentale Modell ist laut dem zentralen kognitionspsychologischen Modell der Leseforschung zum Verstehen multipler Dokumente, dem „Dokumentenmodell“ (Rouet et al., 2019), eine zentrale Ebene textbasierter Repräsentation und stellt die kognitive Verschmelzung von Informationen dar, insbesondere zu geteilten, intertextuellen Inhalten. Solche Verknüpfungsprozesse können in Diskurssynthesen rekonstruiert werden. Ein Beispiel für eine abstrahierte Makroproposition im studentischen Text ist die Proposition [29] „diesen Punkt“, welche die Propositionen [11] „Bildungszugang“ und [20] „Verteilung“ subsummiert.

Zum Verknüpfen zeigen *lesebezogene* Studien als Muster, dass es Leserinnen und Leser schwerer fällt, intertextuelle Verknüpfungen vorzunehmen als intratextuelle (Philipp, 2020) und dass komplexe Verknüpfungen wie das Feststellen von Konflikten oder komplexere Inferenzen weniger gut geleistet werden (List et al., 2020; Philipp, 2020).

In *schreibbezogenen* Studien spielen Verknüpfungen ebenfalls eine Rolle, sind allerdings noch kein prominenter Gegenstand. Segev-Miller (2007) spricht beispielsweise von „Superpropositionen“, womit sie eigenständig gebildete intertextuelle Propositionen meint, die in Diskurssynthesen vorkommen – also wenn Propositionen aus mehreren Bezugstexten durch eine einzige Makroproposition ersetzt wurden. Wenn solche eigenständig gebildeten Inferenzen in eigenen Texten vorkommen, ist dies günstig für die Qualität von Diskurssynthesen (Spivey, 1991).

Die Frage, welche Arten von Inferenzen Personen bilden und wie sich Verknüpfungsleistungen in den Texten niederschlagen, rückt allmählich in den Fokus der Forschung und verdient noch mehr Beachtung (List et al., 2020). Deshalb lautet die *zweite Fragestellung*: Welchen Anteil an möglichen Inferenzen aus Inhalten beider Bezugstexte realisieren Studierende in ihren Synthesen? Der Anteil an gebildeten Inferenzen wird analog zur ersten Fragestellung wiederum nach explizitem/implizitem Vorhandensein sowie nach Korrektheit ausgewertet.

2.2.3 Organisieren

Organisieren – bzw. auch Strukturieren – bezeichnet den dritten Prozess, der bei der Diskurssynthese beteiligt ist. Dabei geht es darum, die Makrostruktur zu klären. Beim *Lesen* ist dies sowohl in einzelnen Texten als auch über die Texte hinweg wichtig. Das heißt, dass die intertextuellen Zusammenhänge von Dis-

kursen zu erkennen sind, damit keine unverbundenen, also nicht-integrierten Modelle einzelner Bezugstexte entstehen, sondern tatsächlich integrierte mentale Modelle (List & Alexander, 2019). Hierbei nutzen Personen lesend und schreibend intratextuelle Strukturen, um intertextuelle Sinnstrukturen zu entwickeln (Barzilai & Eshet-Alkalai, 2015), sie sind aber auch gefragt, ganz eigene Strukturen zu entwickeln (etwa textsortenkonforme Strukturen bei Schreibaufträgen), was selbst Erwachsenen noch schwerfällt (Escorcía et al., 2017; Segev-Miller, 2007; Spivey, 1991). Denn dazu müssen sie teils erhebliche Umstrukturierungen vornehmen, was beim Schreiben dem Planen entspricht.

Zum Organisieren mehren sich *lesebezogene Studien*, die in Typologien oder ordinalen Skalen erfassen, welche Art von mentaler Modellbildung rekonstruierbar ist (s. z. B. List et al., 2020) und auch inwiefern die thematische Struktur die Textsortenspezifika erfüllt (Barzilai & Eshet-Alkalai, 2015; Segev-Miller, 2007). Auch hier zeigen sich relativ durchgängig Probleme dergestalt, dass eine günstige Strukturierung in aller Regel nur einem Teil der Testpersonen glückt. Wir können dies am Beispiel der studentischen Synthese aus Kasten 2 illustrieren. Die Textstruktur folgt primär der Strukturierung des ersten Bezugstexts. Die Inhalte aus dem zweiten Bezugstext werden darin untergebracht, was ein typisches Vorgehen ist (Segev-Miller, 2007).

Die Frage nach der Organisation verdient noch mehr Aufmerksamkeit, da sie eigene schwierigkeitsgenierende Merkmale hat. Deshalb widmen wir uns dem Organisieren mit der *dritten Fragestellung* unserer Studie: Welcher Grad an Integration zeigt sich in der Textstruktur der studentischen Synthesen zu beiden Messzeitpunkten? Diese Fragestellung fokussiert die Unterschiede in der Textorganisation und kategorisiert die Organisationsmuster entlang der realisierten intertextuellen Integration.

In der Forschung ist bei Strukturierungsleistungen allerdings häufig unklar, ob das intertextuelle Integrieren ein stabiles Merkmal ist oder ob es sich um eine variable Fähigkeit handelt. Dies wollen wir in unserer Studie genauer ergründen, indem wir die Struktur der Studierendentexte zu t1 und t2 mit Blick auf den Grad an Integration hinsichtlich ihrer Stabilität betrachten. Zwischen beiden Messzeitpunkten liegt eine dreimonatige Vorlesungszeit. Von Interesse ist also, inwiefern sich Integrationsleistungen ohne gezielte Intervention im Abstand eines Semesters unterscheiden. Die *vierte Fragestellung* lautet: Wie unterscheiden sich personenbezogen die Textstrukturen in Bezug auf die geleistete Integration zu t1 und t2?

2.2.4 Zusammenhänge zwischen den drei Prozessen

Die vorliegende Studie fokussiert zusätzlich eine bislang empirisch unterbeachtete Frage zu Zusammenhängen zwischen den drei Prozessen. Auswählen, Verknüpfen und Organisieren wirken aus theoretischer Sicht aufs Engste mit-

einander verbunden (Spivey, 1990). Aus empirischer Sicht ist das aber noch ungeklärt. Den Zusammenhängen zwischen Prozessen wird bislang wenig Beachtung geschenkt. Escorcia et al. (2017) etwa erheben und unterscheiden zwar Auswahl- und Verknüpfungshandlungen in den Variablen „Ideen finden“ und „Ideen generieren“, korrelieren diese aber nicht miteinander. Auch Mateos und Solé (2009) unterscheiden das Auswählen von Ideen vom Integrieren und Verknüpfen von Informationen. In einem Ratingverfahren werden dann aber Auswahl-, Verknüpfungs- und Organisationsmaße in einer einzigen Variable („types of synthesis“) vermengt. So ist zum einen unklar, welcher der drei Prozesse in welchem Maße zur Zuordnung einer Synthese zu einem Synthesetyp beiträgt, und zum andern bleibt unexpliziert, wie die Prozesse zusammenhängen. Den blinden Fleck in Bezug auf die Zusammenhänge zwischen den Prozessen wollen wir mit der *fünften Fragestellung* adressieren: Welche produktbezogenen Zusammenhänge bestehen zwischen den Indikatoren des Auswählens und Verknüpfens und des Organisierens?

2.2.5 Zusammenhang zwischen Textqualität und den drei Prozessen

Da Diskurssynthesen zuvorderst Texte sind, ist für die Beforschung des Synthetisierens auch die Perspektive der Schreibforschung bedeutsam. Die Textqualität ist in der Schreibforschung eine wichtige Variable, um Schreibleistungen zu erfassen. Eine im Kontext des Schreibens von Synthesen bislang wenig beachtete Frage betrifft den Zusammenhang zwischen den drei oben skizzierten Prozessen und der Textqualität. Aus Interventionsstudien ist bekannt, dass unterschiedliche Fördermaßnahmen wie textvergleichendes Lesen oder das Analysieren von Textstrukturen zu besseren Synthesetexten führen. Bisherige Befunde zum Verhältnis der Textqualität und der drei Prozesse sind bislang uneindeutig (van Ockenburg et al., 2019). Die Studien von Anmarkrud et al. (2013) sowie List und Du (2021) liefern erste empirisch Erkenntnisse zu diesen Zusammenhängen. Gemäß Anmarkrud et al. (2013) korreliert die Fähigkeit, relevante und nicht-relevante Inhalte korrekt zu erkennen, positiv mit der Textqualität. Auch List und Du (2021) berichten in ihrer Studie von (schwachen) positiven Korrelationen zwischen Textqualität und Auswahl- und Verknüpfungsprozessen.

Die Frage, wie die drei Prozesse des Synthetisierens mit der Textqualität empirisch zusammenhängen und wie jeder der drei Prozesse für sich mit der Textqualität zusammenhängt, ist bislang nicht eindeutig geklärt. Diesen Zusammenhang untersuchen wir mit der *sechsten Fragestellung*. Sie lautet: Inwiefern korreliert die Textqualität mit a) dem Anteil und der Korrektheit übernommener Propositionen (Auswählen), b) dem Anteil und der Korrektheit der realisierten intertextuellen Verknüpfungen (Verknüpfen) und c) der Art der Textstrukturierung (Organisieren)?

3 Methodisches Vorgehen

In unserer Untersuchung rekonstruieren wir aus Produktdaten die drei kognitiven Prozesse resp. deren Ergebnisse in den Texten. Wir analysieren, welchen Anteil an Propositionen Studierende aus Bezugstexten auswählen, wie sie ihre Diskurssynthesen strukturieren und welchen Anteil an möglichen Verknüpfungen sie realisieren.

3.1 Sample

Gegenstand der Untersuchung sind für die Diskurssynthese relevante Lese- und Schreibleistungen von Lehramtsstudierenden im dritten Semester. 25 Studierende der Pädagogischen Hochschule Zürich haben zu zwei Zeitpunkten im Abstand von drei Monaten eine Diskurssynthese verfasst. Von den Studierenden waren 21 (84 %) weiblich und vier (16 %) männlich. Zwei (8 %) Studierende streben das Lehramt für die Kindergartenstufe an, 19 (76 %) für die Primarstufe und vier (16 %) für die Sekundarstufe 1. Alle Studierenden hatten mehrere Fächer in ihrem Studienprofil.

3.2 Auftrag für die Studierenden

Die Schreibaufgabe und die zwei Bezugstexte haben wir aus Escorcía et al. (2017) übernommen und auf Deutsch übersetzt. Die Studierenden haben die Schreibaufgabe zweimal im Abstand von drei Monaten (Oktober 2017 und Januar 2018) gelöst. Die Texte wurden am Computer auf einer Online-Lernplattform verfasst. Dafür stand jeweils eine Stunde Zeit zur Verfügung. Der komplette Arbeitsauftrag nebst den nötigen beiden Bezugstexten ist in Kasten 1 im Teilkapitel 2.1 dargestellt.

3.3 Überblick über die abhängigen Variablen

Zur Erfassung des Leistungsstands stützen wir uns auf produktbezogene Daten. Sämtliche Ratings wurden verblindet durchgeführt. Kein Rater wusste, welcher Text von welchem Messzeitpunkt und welcher Person stammte. Alle Rater haben vor dem Raten einen Schulungs- und Eichungsprozess durchlaufen. Tabelle 1 enthält die Instrumente, die wir entlang von Auswahl-, Verknüpfungs- und Organisationsprozessen nach Spivey (1990) konzipiert haben.

Tabelle 1: Übersicht über die eingesetzten Instrumente bzw. der durchgeführten Analysen

Analysefokus	Instrument	Rating
a) <i>Auswählen</i> (Leseleistung) – Fragestellung 1: Anteil der Übernahme von Propositionen aus Bezugstexten	<ul style="list-style-type: none"> – Rating: Anteil der Übernahme von Propositionen aus beiden Bezugstexten gemäß Musterlösung (s. Abbildung 1) – 22 Variablen: 12 Variablen zu Bezugstext 1, 10 Variablen zu Bezugstext 2 – Ratings von Explizitheit und Korrektheit 	<ul style="list-style-type: none"> – 2 Personen (Erst- und Zweitautor) – 4.400 Ratings – Übereinstimmungen: 87 % bei Bezugstext 1, 83 % bei Bezugstext 2 (Korrektur der Abweichungen durch Zweitautor)
b) <i>Verknüpfen</i> (Leseleistung) – Fragestellung 2: Anteil der intertextuellen Inferenzbildung	<ul style="list-style-type: none"> – Rating: Anteil gebildeter Inferenzen aus dem integrierten mentalen Modell (IMM) (s. Abbildung 1) – 16 Variablen: 7 Variablen zu abstrahierten Propositionen aus beiden Bezugstexten und 9 Variablen zu Verknüpfungen zwischen den abstrahierten Propositionen im IMM und den thematisch entsprechenden Propositionen in den Bezugstexten – Rating nach Explizitheit und Korrektheit 	<ul style="list-style-type: none"> – 2 Personen (Erst- und Zweitautor) – Total 3.200 Ratings – Übereinstimmung: 84 % (Korrektur der Abweichungen durch Zweitautor)
c) <i>Organisieren</i> (Schreibleistung) – Fragestellung 3: Grad der Integration in der Textstruktur und Fragestellung 4: personenbezogene Unterschiede t1 und t2	<ul style="list-style-type: none"> – Rating nach Art der Textstrukturierung anhand eines Kategoriensystems (s. Tabelle 2) – 8 graduell unterschiedliche Strukturierungstypen 	<ul style="list-style-type: none"> – 2 studentische Rater – Total: 100 Ratings – Übereinstimmungen: 32 % perfekte Übereinstimmung, 42 % Differenz mit Unsicherheit eines Raters, 26 % absolute Differenzen (Korrektur der Differenzen durch dritten Rater (Erstautor))
d) <i>Korrelationen zwischen Prozessen</i> – Fragestellung 5: Zusammenhänge zwischen den drei Prozessmaßen	<ul style="list-style-type: none"> – Berechnung von Korrelationskoeffizienten (Pearsons r resp. Spearmans ρ) 	
e) <i>Textqualität</i> – Fragestellung 6: Korrelation zwischen Textqualität und den drei Prozessmaßen	<ul style="list-style-type: none"> – Holistisches Rating der Textqualität: Rangreihenfolge von 1 (min.) bis 10 (max.) – Berechnung von Korrelationskoeffizienten (Pearsons r resp. Spearmans ρ) 	<ul style="list-style-type: none"> – 2 studentische Rater – Total 100 Ratings (bei Abweichungen von über 2 Punkten drittes Rating durch dritten Rater in 8 Fällen)

3.3.1 Analytisches Rating von Leseleistungen zum Auswählen (Frage 1) und Verknüpfen (Frage 2)

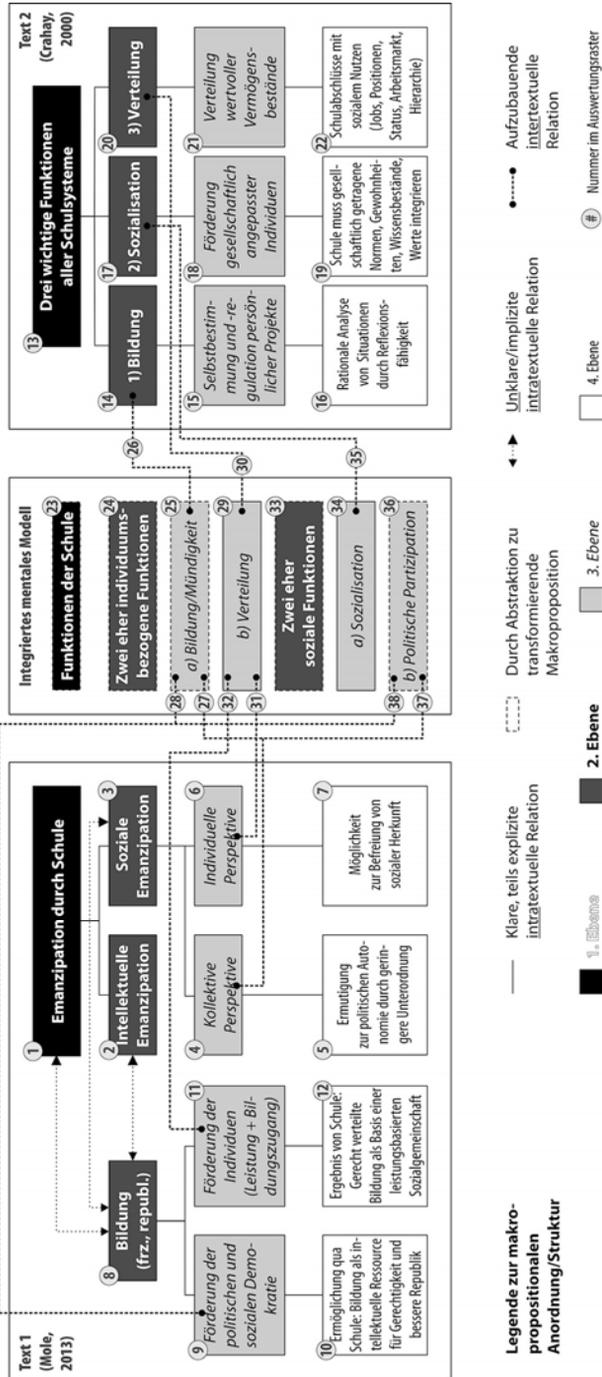
Zur Analyse von Auswahl- und Verknüpfungsprozessen haben wir ein Lesemodell erstellt (s. Abbildung 1). Das Modell bildet zu beiden Seiten in den großen Vierecken die propositionale Struktur der beiden Bezugstexte ab. Die Variablen 1–12 entsprechen den Propositionen des ersten Bezugstexts, die Variablen 13–22 jenen des zweiten. Das mittlere Viereck stellt das integrierte mentale Modell (Perfetti et al., 1999) dar. Darin sind einerseits integrierte Propositionen enthalten, die sich aus den Propositionen der beiden Bezugstexte speisen. Eine integrierte Proposition ist eine Makroproposition, also eine inferenzbasierte Abstraktion respektive Generalisierung von thematisch gleichartigen Propositionen aus beiden Bezugstexten. Das sind die Variablen 23–25, 29, 33, 34 und 36. Andererseits enthält das IMM Verknüpfungen zwischen den abstrahierten Makropropositionen und den Propositionen der Bezugstexte. Das sind die Variablen 26–28, 30–32, 35, 37 und 38. Diese Verknüpfungen sind in der Abbildung 1 als gestrichelte Linien dargestellt. Das IMM beinhaltet insgesamt 16 Variablen.

Das Lesemodell berücksichtigt die ersten vier Hierarchieebenen der Bezugstexte. Dieses Modell haben wir konsensuell erstellt, indem drei Personen, darunter Erst- und Zweitautor, die beiden Bezugstexte unabhängig voneinander lasen und eine Makrostruktur für jeden Bezugstext sowie für das IMM erstellten. Diese drei Musterlösungen wurden diskutiert und die entstehende Musterlösung aus Abbildung 1 abgeglichen und justiert.

Der genaue Blick auf die Bezugstexte offenbart erhebliche Unterschiede zwischen den Texten, was deren Schwierigkeitsgrad betrifft. Während der zweite Bezugstext (s. Kasten 1, Text 2) auf der Textoberfläche explizite Hinweise auf Gliederung (z. B. „drei wichtige Funktionen“), Thematisierung (z. B. „Die erste Funktion – *Bildung*“) und Verknüpfung (z. B. „für dieses Ziel“) aufweist, zeichnet sich Text 1 gerade durch Abwesenheit expliziter Merkmale aus. Die makrostrukturelle Kohärenz wird in Text 1 vornehmlich durch semantische Bezüge hergestellt. Die Konstruktion einer (mental)en Makrostruktur wirkt dadurch im Vergleich mit Text 2 anforderungsreicher.

Das gesamte Modell besteht aus 38 Variablen und dient als Benchmarklösung zum Rating der studentischen Synthesen. Alle 38 Variablen wurden doppelt beurteilt, einerseits nach explizitem/implizitem Vorkommen und andererseits nach korrektem Vorhandensein. Zwei Rater (Erst- und Zweitautor) haben Texte aus dem Korpus, die nur für einen Messzeitpunkt vorlagen und nicht in die Studie einfließen, für Schulungszwecke genutzt. In mehreren Durchgängen wurden Texte zuerst individuell geratet, danach wurden Abweichungen gemeinsam diskutiert. In diesem mehrstufigen Eichungsprozess wurde beurteilt,

Abbildung 1: Lesemodell mit propositionaler Struktur beider Bezugstexte und integriertem mentalem Modell als Musterlösung



welchen Anteil an Propositionen aus den Bezugstexten Studierende in ihren Texten übernehmen und welchen Anteil des IMM (Makropropositionen und Verknüpfungen von Propositionen) Studierende in ihren Texten integrieren. Die Übereinstimmung der Ratings betrug insgesamt 85 %. Beim ersten Bezugstext betrug sie 87 %, beim zweiten Bezugstext 83 % und beim integrierten mentalen Modell 84 %. Alle Abweichungen wurden vom Zweitautor überprüft und nach Konsultation von beiden Raterurteilen final festgelegt.

3.3.2 Analytisches Rating von Schreibleistungen zum Organisieren (Fragen 3 und 4)

Die Analyse der Textstrukturierung haben wir auf Basis des Kategoriensystems von Segev-Miller (2007) vorgenommen. Die Kategorien bilden unterschiedliche Textstrukturmuster ab. Zwei studentische Rater wurden an zwei Halbtagen mit Synthesen, die nicht in die Untersuchung aufgenommen werden konnten, geschult und haben daraufhin alle Texte zu t1 und t2 hinsichtlich der Strukturmuster beurteilt. Die acht Strukturmuster enthält Tabelle 2.

Tabelle 2: Kategorien zur Analyse der Textstrukturmuster

Kat.-Nr.	Bezeichnung	Erläuterung
0	Keine Bezugnahme	Es wird auf keinen der beiden Bezugstexte Bezug genommen.
1	Zusammenfassen eines Texts	Nur Inhalte eines Bezugstexts werden zusammengefasst.
2	Auflisten	Inhalte beider Texte sind separat entlang der Bezugstexte gebündelt und „en bloc“ nacheinander aufgeführt.
3	Hybride Strukturierung ohne Integration	Texte weisen unterschiedliche Strukturmuster auf, jedoch ohne Integration von Inhalten aus beiden Texten.
4	Hybride Strukturierung mit Integration	Texten weisen unterschiedliche Strukturmuster auf, wobei mindestens ein Textteil eine integrierende Struktur aufweist.
5	Ein-Text-Skelett	Ein Bezugstext dient als Gerüst für den strukturellen Rahmen. Die Textstruktur folgt der Struktur eines Bezugstexts.
6	Strukturierung nach Ideen	Auftrennen und Neukombination von Inhalten beider Bezugstexte. Die Inhalte sind nach Ideen gebündelt. Die Verknüpfung der Inhalte und die Textstruktur insgesamt sind eher additiv.
7	Synthese	Eine eigene Struktur ist erkennbar. Die Verknüpfung der Inhalte ist eher vergleichend. Gemeinsamkeiten und Unterschiede der Positionen werden deutlich.

Die Kategorien sind graduell zu verstehen, wobei die Kategorie 0 von keiner Integrationsleistung zeugt und Kategorie 7 von sehr hoher Integrationsleistung und der Struktur einer Diskurssynthese entspricht. Die studentische Synthese (s. Kasten 2) entspräche der Kategorie Nr. 5. Unser Kategoriensystem enthält im Vergleich zur Kategorisierung von Segev-Miller (2007) zwei zusätzliche Kategorien, da sich nicht alle Texte eindeutig einer Kategorie zuweisen ließen. Manche Texte wiesen hybride Textstrukturierungen auf. Es kam vor, dass der erste Textteil nach dem Muster des Auflistens strukturiert war und der zweite eine Strukturierung nach Ideen aufwies. Deshalb fügten wir die zwei neue Kategorien *Hybride ohne Integration* (3) und *Hybride mit Integration* (4) hinzu. Die Rater stimmten in 32 % der Urteile perfekt überein, in 42 % aller Ratings markierte einer der beiden Rater einen Zweifelsfall und in 26 % gab es unterschiedliche Urteile. Abweichungen wurden von einem dritten Rater (dem Erstautor) korrigiert.

Zusätzlich zur Strukturmusteranalyse führten wir eine *personenzentrierte Analyse* durch. Wir haben untersucht, inwiefern sich die Art der Strukturierung personenbezogen zwischen beiden Messzeitpunkten unterscheidet. Dafür haben wir die acht Kategorien auf zwei Kategoriengruppen reduziert, um zu unterscheiden, ob eine Person die Textstruktur additiv und damit eher unangemessen oder integrativ und damit eher angemessen realisiert hat. Die Kategorien 0–3 aggregierten wir zur Variable *nicht-integrierenden Struktur*. Die Kategorien 4–7 gruppieren wir zur Variable *integrierende Struktur*. Mit diesem Verfahren lässt sich feststellen, welche Personen zu welchem Messzeitpunkt eine integrierende respektive eine nicht-integrierende Struktur realisiert haben. Hierbei lassen sich vier Typen charakterisieren:

- Konstant Integrierende: Personen realisieren zu beiden Messzeitpunkten eine integrierende Textstruktur.
- Aufwärts-Integrierende: Personen realisieren zu t1 keine integrierende Struktur, zu t2 aber schon.
- Abwärts-Integrierende: Personen realisieren zu t1 eine integrierende Struktur, zu t2 aber nicht.
- Konstant Nichtintegrierende: Personen realisieren zu beiden Messzeitpunkten keine integrierende Struktur.

Für jede dieser personenbezogenen Typen untersuchten wir außerdem, inwiefern sie sich beim Auswählen und Verknüpfen von Inhalten unterschieden. Die Strukturmaße werden in einem Diagramm mit Auswahl- und Verknüpfungsmaßen in Zusammenhang gebracht (s. Abbildung 3 in Teilkapitel 4.2.2).

3.3.3 Holistische Analyse der Textqualität

Zwei Rater haben nach einer Schulung die studentischen Texte holistische beurteilt, indem sie die Texte in eine Rangreihenfolge von 1 bis 10 brachten. Das Ratingverfahren haben wir aus Kirkpatrick und Klein (2009) übernommen. Je höher die zugewiesene Zahl, als desto besser wurden die Texte beurteilt. Die Rater kannten die Aufgabenstellung, die beiden Bezugstexte jedoch nicht. Die Interraterreliabilität betrug $r = ,81$. Falls die Abweichung zwischen zwei Ratings mehr als zwei Punkte betrug, gab es ein drittes Rating. Das war bei 8 von 50 Texten nötig. Bei zwei Ratings wurde der Mittelwert gebildet, bei drei der Median verwendet.

3.3.4 Korrelationsanalysen (Fragen 5 und 6)

Wie in Teilkapitel 2.2.4 und 2.2.5 ausgeführt, berechnen wir Korrelationen zwischen Auswahl-, Verknüpfungs- und Organisationsmaßen, um zu eruieren, inwiefern diese Prozesse miteinander zusammenhängen. Zusätzlich führen wir Korrelationsanalysen zwischen der Schreibleistungen (Textqualität) und den prozessbezogenen Maßen des Auswählens, Verknüpfens und Organisierens durch. Je nach Skalenniveau verwenden wir unterschiedliche Korrelationskoeffizienten. Für Korrelationen zwischen den intervallskalierten Daten (Maße des Auswählens, Verknüpfens und der Textqualität) verwenden wir Pearsons r , bei ordinalskalierten Daten (Organisationsmaße) Spearmans Rho.

4 Ergebnisse

4.1 Auswählen und Verknüpfen (Fragestellungen 1 und 2)

Die Ergebnisse zu den ersten beiden Fragestellungen werden im Diagramm in Abbildung 2 zusammen dargestellt. Die erste Fragestellung lautete: Welchen Anteil an Propositionen übernehmen Studierende aus welcher hierarchischen Ebene der beiden Bezugstexte? Die zweite: Welchen Anteil an möglichen Inferenzen aus Inhalten beider Bezugstexte realisieren Studierende in ihren Synthesen? Die Abbildung 2 zeigt die gesamthafte prozentuale Verteilung aller Variablen, die auf das Auswählen (Übernahme von Propositionen) und Verknüpfen (Inferenzbildung) bezogen sind. Für alle Variablen A–I werden Ausprägungen an Explizitheit und Korrektheit mit separaten Balken im Diagramm ausgewiesen. Die Variable A *Gesamt* ist eine aggregierte Variable, die alle Variablen des Auswählens und Verknüpfens – also B–I einschließt (s. Teilkapitel 3.3.1). Die Variablen, die das Auswählen betreffen, sind B (*Text 1*), C (*Text 2*) sowie E–H (*Ebenen 1–4*). Jene, die das Verknüpfen betreffen, sind D (*IMM*; integriertes mentales Modell) und I (*Verknüpfungen*). Die Variable I *Verknüpfun-*

gen weist ausschließlich textuelle Verbindungen zwischen dem IMM und Propositionen aus Bezugstexten aus.

Abbildung 2: Verteilung der abhängigen Variablen im Auswählen und Verknüpfen im Gesamtsample

		Messzeitpunkt 1 (t1) (N = 25)			Messzeitpunkt 2 (t2) (N = 25)			
A)	Gesamt	Explizitheit	25	20	55	22	22	56
		Korrektheit	20		80	24		76
B)	Text 1	Explizitheit	22	16	62	22	21	57
		Korrektheit	13		87	19		81
C)	Text 2	Explizitheit	49	18	33	40	23	37
		Korrektheit	38		62	40		60
D)	IMM	Explizitheit	12	24	64	12	21	67
		Korrektheit	15		85	18		82
E)	Ebene 1	Explizitheit	32	15	53	36	15	49
		Korrektheit	33		67	45		55
F)	Ebene 2	Explizitheit	27	16	57	33	13	54
		Korrektheit	23		77	33		67
G)	Ebene 3	Explizitheit	32	22	46	22	26	52
		Korrektheit	25		75	24		76
H)	Ebene 4	Explizitheit	25	26	49	22	24	54
		Korrektheit	11		89	15		85
I)	Verknüpfungen	Explizitheit	10	18	72	10	24	66
		Korrektheit	15		85	15		85

Grad an Explizitheit ■ explizit vorhanden ■ implizit vorhanden □ nicht vorhanden

Korrektheit □ korrekt vorhanden □ nicht korrekt vorhanden

(Angaben in Prozent, Abweichungen von 100 Prozent basieren auf Rundungsfehlern; IMM = integriertes mentales Modell, Angaben zur Ebene 4 beziehen sich nur auf die Texte 1 und 2, Angaben zu den Verknüpfungen nur auf das IMM)

4.1.1 Aggregierte Auswahl- und Verknüpfungsmaße: Variable A

Die aggregierte Variable A *Gesamt* weist folgende Ausprägungen auf: Gesamthaf ist der Anteil an übernommenen Propositionen und Inferenzbildung gering (45 %_{t1} und 44 %_{t2}). Die beiden Messzeitpunkte unterscheiden sich kaum in Bezug auf Übernahme und Inferenzen (Unterschied 0–6 % zwischen t1 und t2 für alle Variablen). Der Anteil an expliziter und impliziter Übernahme und Inferenzbildung fällt für beide Messzeitpunkte ähnlich aus (t1: 25 %_{expl.} und 20 %_{impl.}; t2: 22 %_{expl.} und 22 %_{impl.}).

Auch der Anteil an korrekten Übernahmen und korrekter Inferenzbildung ist für beide Messzeitpunkte gering, die Korrektheit nimmt zu t2 aber leicht zu (20 %_{t1} und 24 %_{t2}). Vergleicht man die Korrektheit der Variablen B–I einzeln, ist der Unterschied zwischen den Messzeitpunkte gering (Unterschied 0–12 % zwischen t1 und t2).

4.1.2 Auswählen: Variablen B, C, E–H

Die Betrachtung der auswahlbezogenen Variablen lässt folgende Muster erkennen: Die Übernahme von Propositionen aus beiden Bezugstexten ist unvollständig. Das zeigt der Durchschnitt aus den Variablen B *Text 1* und C *Text 2*. Aus Variable C *Text 2* wurde zu beiden Messzeitpunkten ein deutlich größerer Anteil an Propositionen übernommen als aus B *Text 1* (Text 2: 67 %_{t1} und 63 %_{t2}; Text 1: 38 %_{t1} und 43%_{t2}). Die Hierarchieebenen (Variablen E–I) unterscheiden sich wenig, was Übernahmen betrifft (min. 43 % E_{2,t1}; max. 54 % E_{3,t1} und min. 46 % E₂ und E_{4,t2}; max. 51 % E_{3,t1}). Tendenziell lassen sich mehr explizite Übernahmen aus Bezugstexten in höheren Hierarchieebenen und mehr implizite Übernahmen in tieferen Hierarchieebenen feststellen.

Es gibt deutlich mehr korrekte Übernahmen aus C *Text 2* als aus B *Text 1* (Text 2: 38 %_{t1} und 40 %_{t2}; Text 2: 13 %_{t1} und 19 %_{t2}) sowie markant mehr korrekte Übernahmen aus höheren Hierarchieebenen als aus tieferen zu beiden Messzeitpunkten.

4.1.3 Verknüpfen: Variablen D (IMM) und I (Verknüpfungen)

Die Verknüpfungsvariablen weisen folgende Spezifika auf: Bezogen auf die aggregierte Verknüpfungsvariable D *IMM* gibt es wenig Inferenzbildung über beide Bezugstexte zu beiden Messzeitpunkten (IMM: 32 %_{t1}; 33 %_{t2}). Ebenso ist der Anteil an Verknüpfungen (Variable I – Verbindungen zwischen inferierten Makropropositionen im IMM und Propositionen aus den Bezugstexten) tief (28 %_{t1}; 34 %_{t2}). Sowohl beim integrierten mentalen Modell (Variable D) als auch bei den Verknüpfungen I zeigt sich zu beiden Messzeitpunkten eine Tendenz zu impliziter Inferenzbildung (ca. zwei Drittel sind implizit). Der Anteil an korrekten Inferenzen (Variable D) und korrekten Verknüpfungen (Variable I) ist zu beiden Messzeitpunkten klein (15 %_{t1} und 15–18 %_{t2}).

Exkurs: Korrelation zwischen Explizitheit und Korrektheit

Um zu eruieren, inwiefern Explizitheit und Korrektheit statistisch zusammenhängen, haben wir eine zusätzliche Analyse durchgeführt. Die Korrelationsanalysen zwischen Explizitheits- und Korrektheitsmaßen ergab Folgendes: Explizites Vorhandensein korreliert signifikant mit korrektem Vorhandensein bei allen neun Variablen des Auswählens und Verknüpfens: $r = ,67$ – $,90$ ($p < ,001$). Zwischen implizitem Vorhandensein und Korrektheit gibt es bei den beiden Variablen C *Text 2* und E *Ebene 1* signifikante negative Korrelationen: C_{t1}: $r = -,48$ ($p = ,015$) und E_{t1}: $r = -,55$ ($p = ,004$). Alle anderen Korrelationsanalysen waren nicht signifikant. Demnach sind nur Explizitheit und Korrektheit stark miteinander assoziiert.

4.2 Organisieren (Fragestellungen 3 und 4)

Die Ergebnisse zum Prozess des Organisierens beziehen sich auf die Fragestellungen 3 und 4. Zuerst wird die Verteilung der Strukturierungsmuster über beide Messzeitpunkte dargestellt. Im Anschluss daran werden Resultate zu personenzentrierten Strukturanalysen und die Zusammenhänge dieser Ergebnisse mit Auswahl- und Verknüpfungsvariablen präsentiert.

4.2.1 Strukturierungsmuster

Die Tabelle 3 enthält Ergebnisse zur dritten Fragestellung, die folgendermaßen lautet: Welcher Grad an Integration zeigt sich in der Textstruktur der studentischen Synthesen zu beiden Messzeitpunkten? Insgesamt realisieren zu t1 56 % der Studierenden eine integrierende Textstruktur (Kat. 4–7), 44 % eine nicht-integrierende Textstruktur (Kat. 0–3). Zu t2 beträgt das Verhältnis Integration zu Nicht-Integration 60 zu 40 %. Die beiden Messzeitpunkte unterscheiden sich in Bezug auf Strukturierungsleistungen in den zusammengefassten Kategorien damit nur geringfügig. Gleiches gilt für Unterschiede zwischen Messzeitpunkten innerhalb der einzelnen Kategorien. Es ist zu beiden Messzeitpunkten eine Dominanz der Strukturmuster Auflisten (2) und Hybride mit Integration (4) beobachtbar.

Tabelle 3: Prozentuale Verteilung der Textstrukturierung nach Kategorien

	0) Keine Bezugnahme ^A	1) Zusammenfassen eines Texts ^A	2) Auflisten ^A	3) Hybride ohne Integration ^A	4) Hybride mit Integration ^B	5) Ein-Text-Skelett ^B	6) Strukturierung nach Ideen ^B	7) Synthese ^B
t1	8 %	–	32 %	4 %	24 %	4 %	16 %	12 %
t2	–	–	36 %	4 %	24 %	12 %	16 %	8 %

(Legende: ^A = später zusammengefasst zu nicht-integrierender Textstruktur;

^B = später zusammengefasst zu integrierender Textstruktur)

4.2.2 Strukturtypen

Die nachfolgenden Ergebnisse beziehen sich auf die vierte Fragestellung, die folgendermaßen lautet: Wie unterscheiden sich personenbezogen die Textstrukturen in Bezug auf Integration zu t1 und t2? Generell verteilen sich die Studierenden auf die vier im Methodenteil (3.3.2) genannten Strukturtypen wie folgt (s. Abbildung 3): Von 25 Studierenden haben 8 konstant eine integrierende Struktur realisiert, 7 haben zu t2 die Struktur ihrer Texte verbessert, 6 haben zu t2 die Textstruktur verschlechtert und 4 haben konstant eine nicht-integrierende Struktur realisiert. Der Anteil von Aufwärts- (28 %) und Abwärts-Integrierenden (24 %) ist vergleichbar groß.

Abbildung 3: Verteilung der abhängigen Variablen im Auswählen und Verknüpfen in vier Gruppen zur Stabilität bei den Integrationsleistungen in einer 2-x-2-Gruppenzugehörigkeit

		Konstant Integrierende (32 %; N = 8)			Aufwärts-Integrierende (28 %; N = 7)			Abwärts-Integrierende (24 %; N = 6)			Konstant Nichtintegrierende (16 %; N = 4)			
A)	Gesamt	Explizitheit t1	25	18	57	22	23	55	28	19	53	24	20	56
		Explizitheit t2	23	19	58	22	26	52	18	21	61	29	19	52
		Korrektheit t1	22		78	20		80	19		81	21		79
		Korrektheit t2	23		77	27		73	21		79	26		74
B)	Text 1	Explizitheit t1	20	15	65	21	15	64	26	19	55	19	15	66
		Explizitheit t2	18	18	64	18	26	56	25	24	51	33	17	50
		Korrektheit t1	37		63	39		61	33		67	45		55
		Korrektheit t2	37		63	38		62	35		65	56		44
C)	Text 2	Explizitheit t1	48	18	34	44	27	29	48	12	40	60	13	27
		Explizitheit t2	40	20	40	37	29	34	33	20	47	58	23	19
		Korrektheit t1	37		63	39		61	33		67	45		55
		Korrektheit t2	37		63	38		62	35		65	56		44
D)	Integriertes mentales Modell (IMM)	Explizitheit t1	14	21	65	8	26	66	16	24	60	6	28	66
		Explizitheit t2	16	20	64	16	24	60	3	21	76	8	19	73
		Korrektheit t1	18		82	13		87	17		83	13		87
		Korrektheit t2	16		84	29		71	10		90	11		89
E)	Ebene 1	Explizitheit t1	25	21	54	38	14	48	33	6	61	33	17	50
		Explizitheit t2	33	17	50	19	29	52	44	6	50	58		42
		Korrektheit t1	25		75	43		57	33		67	33		67
		Korrektheit t2	46		54	33		67	50		50	58		42
F)	Ebene 2	Explizitheit t1	23	14	63	27	23	50	23	17	60	38	6	56
		Explizitheit t2	31	11	58	34	16	50	27	4	69	41	22	37
		Korrektheit t1	20		80	27		73	19		81	31		69
		Korrektheit t2	33		67	34		66	27		73	39		61
G)	Ebene 3	Explizitheit t1	35	23	42	26	30	44	36	20	44	32	11	57
		Explizitheit t2	24	20	56	21	34	45	18	30	52	25	18	57
		Korrektheit t1	32		68	21		79	23		77	23		77
		Korrektheit t2	22		78	32		68	23		77	20		80
H)	Ebene 4 (ohne IMM)	Explizitheit t1	21	23	56	22	27	51	38	26	36	18	29	53
		Explizitheit t2	16	20	64	22	29	49	17	26	57	39	21	40
		Korrektheit t1	11		89	10		90	14		86	7		93
		Korrektheit t2	11		89	12		88	14		86	29		71
I)	Verknüpfungen (nur IMM)	Explizitheit t1	15	11	74	6	14	80	11	20	69	6	36	58
		Explizitheit t2	15	25	60	14	22	64	2	28	70	6	22	72
		Korrektheit t1	18		82	11		89	15		85	17		83
		Korrektheit t2	17		83	23		77	7		93	8		92

Grad an Explizitheit ■ explizit vorhanden ■ implizit vorhanden □ nicht vorhanden

Korrektheit □ korrekt vorhanden □ nicht korrekt vorhanden

(Angaben in Prozent, Abweichungen von 100 Prozent basieren auf Rundungsfehlern)

In Abbildung 3 werden für jeden Strukturtyp Auswahl- und Verknüpfungsmaße aufgeführt, um zu eruieren, inwiefern sich die Strukturtypen hinsichtlich Auswahl- und Verknüpfungsprozessen unterscheiden. Damit wird Bezug genommen auf einen Aspekt der fünften Fragestellung, welche die produktbezo-

genen Zusammenhänge zwischen Indikatoren des Auswählens und Verknüpfens und des Organisierens fokussiert.

Die Ergebnisse in Abbildung 3 verdeutlichen, dass sich die Strukturtypen global bei Auswahl- und Verknüpfungsmaßen nicht stark voneinander unterscheiden. Sowohl bezüglich Explizitheit und auch bezüglich Korrektheit zeigt sich bei allen Variablen nur ein geringer Unterschied zwischen den Strukturtypen für beide Messzeitpunkte. Auffällig ist, dass alle Strukturtypen zu beiden Messzeitpunkten mehr Propositionen aus Variable C *Text 2* als aus Variable B *Text 1* übernehmen. Der Anteil korrekter Übernahmen und korrekter Inferenzbildung (Variable A *Gesamt*) ist bei allen Strukturtypen zu beiden Messzeitpunkten tief (20–27 %). Der Anteil an korrekten Inferenzen (Variable D *IMM*) und Verknüpfungen (Variable I) ist bei Aufwärts-Integrierenden zu t2 höher als zu t1, nämlich mindestens doppelt so hoch (IMM: 13 %_{t1} und 29 %_{t2}; Verknüpfungen: 11 %_{t1} und 23 %_{t2}). Die Gruppe der konstant Nichtintegrierenden hat zu beiden Messzeitpunkten einen höheren Anteil korrekter Übernahmen aus beiden Bezugstexten (Variablen B und C) als andere Strukturgruppen (45 %_{t1} und 56 %_{t2}).

4.3 Textqualität (Fragestellung 6)

Die Textqualität ist insbesondere für die Korrelationsanalysen relevant. Im Durchschnitt war die Textqualität zu t1 ($M = 6.06$, $SD = 1,87$) tiefer als zu t2 ($M = 6.32$, $SD = 1,63$). Dieser Unterschied ist nicht signifikant, $t(24) = -,56$, $p = ,58$.

4.4 Korrelationsanalysen (Fragestellungen 5 und 6)

Die fünfte Fragestellung lautete: Welche produktbezogenen Zusammenhänge bestehen zwischen den Indikatoren des Auswählens, Verknüpfens und Organisierens? Die Analyse anhand der Korrelationskoeffizienten – jeweils für beide Messzeitpunkte ausgewiesen – in Tabelle 4 brachte folgende Ergebnisse hervor: Insgesamt lassen sich keine klaren Korrelationsmuster zwischen den drei Indikatoren feststellen. Es treten nur vereinzelt signifikante Korrelationen zwischen Variablen in Erscheinung. Ein systematischer Zusammenhang lässt sich jedoch nicht erkennen.

Tabelle 4: Ergebnisse zu Korrelationen zwischen Indikatoren des Auswählens, Verknüpfens und Organisierens sowie zwischen Textqualität und den drei genannten Indikatoren

	t1	t2	t1	t2	t1	t2	t1	t2
	<i>Text 1_{expl.+impl.}</i>		<i>Text 1_{expl.}</i>		<i>Text 1_{impl.}</i>		<i>Text 1_{Korrektheit}</i>	
Text 2 ^A	,28*	,39	,53**	,50*	-,01	,59**	,27	,50*
IMM ^A	,29	,39	,28	-,12	,13	,55**	,23	,03
Struktur ^B	,16	-,07	,00	-,27	,15	,18	,08	-,08
Textqualität ^C	,40*	,29	,30	,24	,28	,18	,35	,25
	<i>Text 2_{expl.+impl.}</i>		<i>Text 2_{expl.}</i>		<i>Text 2_{impl.}</i>		<i>Text 2_{Korrektheit}</i>	
IMM ^A	,40*	,32	,07	,08	,43*	,29	,22	,15
Struktur ^B	-,25	-,15	-,13	-,14	-,15	,18	-,31	-,08
Textqualität ^C	,25	,34	,46*	,06	-,44*	,42	,28	,14
	<i>IMM_{expl.+impl.}</i>		<i>IMM_{expl.}</i>		<i>IMM_{impl.}</i>		<i>IMM_{Korrektheit}</i>	
Struktur ^B	,01	,52**	,29	,36	-,19	,33	,21	,25
Textqualität ^C	,49*	,51**	,51**	,35	,19	,25	,51**	,22

^A = Gleiche Variable wie in der Horizontale, nur bezogen auf jeweils anderen Gesamtscore aus dem Lesemodell, Pearsons *r*

^B = Bezogen auf prozentuale Verteilung nach Kategorien (s. Tabelle 3), Spearmans Rho

^C = Pearsons *r*

(* $p < ,05$; ** $p < ,01$)

Auffällig ist besonders die Abwesenheit von Korrelationen zwischen Struktur- und Auswahlmaßen. Das Organisieren scheint mit anderen Prozessen kaum verbunden zu sein, da die Strukturmaße von insgesamt 32 Analysen nur an einer einzigen Stelle signifikant mit einer anderen Variable korrelieren (IMM_{expl.+impl.} zu t2). Ein zweiter auffälliger Punkt aus den Korrelationsanalysen ist, dass Auswahlprozesse zwischen Text 1 und Text 2 zusammenhängen, vor allem zum zweiten Messzeitpunkt: Je expliziter und korrekter Propositionen aus dem einem Bezugstext ausgewählt werden, desto expliziter und korrekter werden sie auch aus dem zweiten Bezugstext ausgewählt.

Die sechste Fragestellung lautete: Inwiefern korreliert die Textqualität mit a) dem Anteil und der Korrektheit übernommener Propositionen (Auswählen), b) dem Anteil und der Korrektheit der realisierten intertextuellen Verknüpfungen (Verknüpfen) und c) der Art der Textstrukturierung (Organisieren)? Insgesamt ergibt sich, dass die Textqualität allein mit den *Verknüpfungsmaßen* aus dem integrierten mentalen Modell in mittlerem Maß korreliert. Dieser Effekt zeigt sich auf der Ebene der aggregierten Verknüpfungen (IMM_{expl.+impl.}) für beide Messzeitpunkte und zusätzlich bei explizit übernommenen Verknüpfungen aus Text 1. Ebenfalls zeigt sich eine mittlere Korrelation zwischen Textqualität und der Korrektheit der Verknüpfungen. Indikatoren der Auswahlprozesse hängen hingegen nicht systematisch mit der Textqualität zusammen. Zwischen den

Variablen *Strukturierung* und Textqualität gibt es zu beiden Messzeitpunkten keine signifikante Korrelation: $r = ,19$ ($p = ,38$) zu t1 bzw. $,21$ ($p = ,32$) zu t2.

5 Diskussion

Unsere Studie hatte zum Ziel, die kognitiven Prozesse des Auswählens, Organisierens und Verknüpfens, die beim Verfassen von Diskurssynthesen beteiligt sind, produktbezogen mit diversen geschulten Ratern zu rekonstruieren. Wir haben aufgrund von Schreibleistungen von Studierenden Rückschlüsse auf Prozesse gezogen. Dieser Ansatz ist natürlich methodisch kritisierbar, weil Produkte und Prozesse nicht dasselbe sind, wir also nur begrenzt valide Auskünfte über faktisch abgelaufene Prozesse geben können. Wir haben also deren rekonstruierbare Spuren fokussiert.

Drei Erkenntnisinteressen standen dabei im Vordergrund. Das erste betraf die Leistungen von Studierenden (Fragestellungen 1–4). Wir haben anhand einer Musterlösung genau kartiert, welche Auswahl-, Verknüpfungs- und Organisationsleistungen erbracht werden können, und haben die studentischen Diskurssynthesen daran gemessen. Ein zweiter Interessenfokus lag auf den Zusammenhängen zwischen den drei Prozessen (Fragestellung 5). Wir eruierten, inwiefern die drei Leistungsmaße (und damit auch die drei Prozesse) miteinander zusammenhängen. Das dritte Forschungsinteresse bestand darin, zu erheben, welches der Leistungsmaße wie stark mit der Textqualität korreliert (Fragestellung 6). Die folgenden Ausführungen sind entlang dieser drei Erkenntnisinteressen strukturiert.

5.1 Fokus 1: Auswahl-, Verknüpfungs- und Organisationsleistungen

Die vorliegende Untersuchung brachte generische Schwierigkeiten von Studierenden bei allen drei Prozessen hervor. Besonders deutlich offenbarten sich die Schwierigkeiten bei Verknüpfungs- und Organisationsleistungen, jenen Prozessen also, die Inferenzleistungen erfordern. Diese lagen in unserer Studie auf niedrigem Niveau. Die Studierenden realisierten nur ca. ein Drittel aller möglichen Verknüpfungen. Das Produzieren einer integrativen Textstruktur gelang zwei Fünftel der Studierenden nicht. Diese Probleme zeigten sich zu beiden Messzeitpunkten in vergleichbarem Ausmaß. Es war keine Verbesserung der Leistungen aufgrund des Durchlaufens eines Studiensemesters auszumachen. Die Befunde decken sich mit anderen Studien, in denen Inferenzprozesse untersucht und als Schwierigkeit benannt wurden (Barzilai & Eshet-Alkalai, 2015; Braasch et al., 2013; Gil et al., 2010; List & Du, 2021; Mateos & Solé, 2009; Primor & Katzir, 2018; Segev-Miller, 2007; Wiley & Voss, 1999).

Das Auftreten von Schwierigkeiten beim Inferieren in unserer Studie war demnach erwartbar, daher haben wir genauer geprüft, worin sich die schwierigkeitsgenerierenden Merkmale zeigten, nämlich bei einem bislang blinden Fleck in der Forschung zur multiplen Textintegration: den Bezugstexten. Aus der Leseforschung zu Einzeltexten ist bekannt, dass kohärente Texte mit salienter Struktur leichter mental repräsentiert werden können und weniger kognitive Ressourcen fürs Verstehen beanspruchen (Kintsch, 1990; McNamara et al., 1996). Die von uns verwendeten Bezugstexte unterscheiden sich stark hinsichtlich des Schwierigkeitsgrads (s. Teilkapitel 3.3.1). Der erste Text weist im Gegensatz zum zweiten eine implizite Makro- und Mikrostruktur auf und ist damit insgesamt anspruchsvoller. Diese textseitigen Faktoren schlugen sich in den Auswahl- und Verknüpfungsleistungen nieder: Die Studierenden haben aus dem schwierigeren Text markant weniger Propositionen ausgewählt und integriert. Paradoxiertweise zeigen Untersuchungen aus der Leseforschung aber, dass solche anforderungsreichen Texte größeres Potenzial fürs Lernen mit und durch Texte beinhalten (Kintsch, 1990; McNamara et al., 1996), vermutlich weil die Implizitheit Leserinnen und Leser zu intensiveren Verarbeitungs- und Inferenzleistungen drängt (Primor & Katzir, 2018). Personen mit hoher Lesekompetenz und viel Vorwissen füllen inkohärente Stellen und fehlende Informationen unter Zuhilfenahme des Vorwissens auf und generieren so Verknüpfungen. Die Textinhalte werden dadurch aktiver verarbeitet und sind dem Verständnis zuträglicher (Kintsch, 1990; McNamara et al., 1996). Für das Schreiben von Diskurssynthesen ist aber weithin ungeklärt, wie der Schwierigkeitsgrad der Bezugstexte das Integrieren von Texten und somit das Lernen via multiple Textintegration beeinflusst. Untersuchungen zum Verhältnis von Textschwierigkeit und Verstehensleistungen über mehrere Texte stellen ein dringendes Desiderat dar.

5.2 Fokus 2: Zusammenhänge zwischen Auswahl-, Verknüpfungs- und Organisationsleistungen

Der zentrale Befund lautet, dass produktbezogene Auswahl-, Verknüpfungs- und Organisationsmaße in dieser Studie nicht miteinander korrelieren. Es scheint im Gegenteil so zu sein, dass das Auswählen, Verknüpfen und Organisieren – zumindest aus einer Produktperspektive – relativ unabhängig voneinander in Erscheinung treten. Auch den vier Strukturierungstypen, die wir im Längsschnitt erfasst haben, zeigen sich keine musterhaften Zusammenhänge zwischen den drei Maßen. Dieser Befund liegt quer zur Theorie (s. Teilkapitel 2.2.4), denn die drei Prozesse scheinen eng miteinander verbunden: Das Verknüpfen bedingt das Auswählen von Propositionen. Ein integriertes mentales Modell setzt das Verstehen der einzelnen Texte voraus. Auch das Organisieren

ist an Prozesse des Auswählens und Verknüpfens von Inhalten gebunden (Spivey, 1990; van Ockenburg et al., 2019).

Nimmt man andere empirische Studien in den Blick, die strukturanalog zu unserer Untersuchung Korrelationen zwischen Auswahl-, Verknüpfungs- und Organisationsvariablen ausgewiesen haben, zeigen sich teils Zusammenhänge. List und Du (2021) berichten eine Korrelation zwischen einer Auswahl- und einer Verknüpfungsvariable. Auch Barzilai und Eshet-Alkalai (2015) fanden eine Korrelation zwischen einer Variable zum Auswählen (Verstehen wichtiger Propositionen) und Inferenzleistungen. Die Korrelationen in diesen beiden Studien treten jedoch vereinzelt auf, eine Systematik ist nicht erkennbar. Andere Studien lassen keine Rückschlüsse auf Korrelationen zu (Anmarkrud et al., 2013; Escorcía et al., 2017). Insgesamt ergibt sich aus empirischer Sicht kein einheitliches Bild zu diesen Zusammenhängen (Philipp, 2018). Weitere Forschung zum Verhältnis und zum Zusammenhang zwischen den drei Prozessen ist also nötig und wünschenswert.

5.3 Fokus 3: Zusammenhänge zwischen Textqualität und Auswahl-, Verknüpfungs- und Organisationsleistungen

Die Textqualität wurde in unserer Untersuchung holistisch erfasst. Den Ratern waren die Rahmenbedingungen und die Aufgabenstellung bekannt, die Bezugstexte jedoch nicht. In den Korrelationsanalysen trat das Verknüpfen als einziger Indikator für Textqualität musterhaft hervor. Die Auswahl- und Organisationsmaße korrelierten nicht mit der Textqualität. Die Korrelation zwischen Verknüpfungsleistungen und der Textqualität deckt sich mit Resultaten anderer Studien, bei denen aber teils auch Auswahl- und Organisationsleistungen mit Textqualität zusammenhängen (Ackerman, 1991; Anmarkrud et al., 2013; List & Du, 2021; Mateos & Solé, 2009; Spivey & King, 1989). Inferenzleistungen scheinen für die Textqualität von Diskurssynthesen also hochgradig relevant. Das zeigt sich deutlich auch bei Spivey (1991). Dort war die Fähigkeit, Inhalte aus zwei Texten zu verknüpfen und diese als Makropropositionen im Text zu präsentieren, der stärkste Prädiktor für die holistisch beurteilte Textqualität. Mit dem Verknüpfen scheint der Prozess des Strukturierens eng verwandt und auch relevant für die Textqualität zu sein, da beide Prozesse dazu führen, dass aus der Lektüre der Bezugstexte ein kohärenter Text resultiert. In unserer Untersuchung fanden wir aber keine Korrelationsmuster zwischen Organisationsmaßen und Textqualität, obwohl ein Zusammenhang theoretisch nahe liegt. Bislang existieren wenig Untersuchungen zum Zusammenwirken von Strukturierungsprozessen und Textqualität beim Schreiben über mehrere Texte. Es bedarf diesbezüglich weiterführender Forschung. Wie oben erwähnt, hatte auch ein hoher Anteil ausgewählter Inhalte aus Bezugstexten keinen Einfluss auf die Textqualität. Spivey (1991) zeigte in ihrer Studie, dass sich ein

hoher Anteil übernommener Propositionen im Gegenteil sogar negativ auf die Textqualität auswirkte. Das losgelöste Wiedergeben von Inhalten aus Bezugstexten allein reicht nicht zu einer guten Synthese. Es bedarf dafür eines aktiven Stiftens von Kohärenz, indem Inhalte durch Inferenzbildung und Verknüpfung mit Vorwissen miteinander in Beziehung gebracht werden.

5.4 Abschlussbemerkung

Alles in allem lohnt sich (nicht nur) im Licht der Befunde unserer Studie ein genauer Blick in der Leseforschung darauf, welche textseitigen Faktoren sich auf die kognitiven Prozesse des Lesens und Schreibens auswirken. Wir denken, dass insbesondere die Eigenheiten der eingesetzten Texte und Textsorten hinsichtlich ihrer schwierigkeitsgenierenden Merkmale jenseits von Lesbarkeitsindizes und Umfangsmaßen mehr Beachtung verdienen. So ließe sich auch einholen, was beispielsweise Modelle wie das MD-TRACE-Modell (Rouet & Britt, 2011) postulieren: ein kontextsensitives Vorgehen. Dies in Verbindung mit den individuumsbezogenen Merkmalen (Barzilai & Strømsø, 2018) multimethodisch in puncto Prozess- (Mason & Florit, 2018) und Produktdaten herauszuarbeiten (Primor & Katzir, 2018), ist ein dringend zu bestellendes Feld der Schreib- und Leseforschung.

Literaturverzeichnis

- Ackerman, J. M. (1991): Reading, Writing, and Knowing: The Role of Disciplinary Knowledge in Comprehension and Composing. *Research in the Teaching of English*, 25(2), 133–178.
- Anmarkrud, Ø., McCrudden, M. T., Bråten, I. & Strømsø, H. I. (2013): Task-Oriented Reading of Multiple Documents: Online Comprehension Processes and Offline Products. *Instructional Science*, 41(5), 873–894. <https://doi.org/10.1007/s11251-013-9263-8>
- Barzilai, S. & Eshet-Alkalai, Y. (2015): The Role of Epistemic Perspectives in Comprehension of Multiple Author Viewpoints. *Learning and Instruction*, 36, 86–103. <https://doi.org/10.1016/j.learninstruc.2014.12.003>
- Barzilai, S. & Strømsø, H. I. (2018): Individual Differences in Multiple Document Comprehension. In J. L. G. Braasch, I. Bråten & M. T. McCrudden (Eds.), *Handbook of Multiple Source Use* (pp. 99–116). New York: Routledge.
- Braasch, J. L. G., Bråten, I., Strømsø, H. I., Anmarkrud, Ø. & Ferguson, L. E. (2013): Promoting Secondary School Students' Evaluation of Source Features of Multiple Documents. *Contemporary Educational Psychology*, 38(3), 180–195. <https://doi.org/10.1016/j.cedpsych.2013.03.003>
- Escorcia, D., Passerault, J.-M., Ros, C. & Pylouster, J. (2017): Profiling Writers: Analysis of Writing Dynamics Among College Students. *Metacognition and Learning*, 12(2), 233–273. <https://doi.org/10.1007/s11409-016-9166-6>
- Gil, L., Bråten, I., Vidal-Abarca, E. & Strømsø, H. I. (2010): Summary Versus Argument Tasks When Working with Multiple Documents: Which is Better for Whom? *Contemporary Educational Psychology*, 35(3), 157–173. <https://doi.org/10.1016/j.cedpsych.2009.11.002>

- Kendeou, P. (2015): A General Inference Skill. In E. J. O'Brien, A. E. Cook & R. F. Lorch (Eds.), *Inferences during Reading* (pp. 160–181). Cambridge University Press. <https://doi.org/10.1017/CBO9781107279186.009>
- Kintsch, E. (1990): Macroprocesses and Microprocesses in the Development of Summarization Skill. *Cognition and Instruction*, 7(3), 161–195. https://doi.org/10.1207/s1532690xci0703_1
- Kintsch, W. & van Dijk, T. A. (1978): Toward a Model of Text Comprehension and Production. *Psychological Review*, 85(5), 363–394. <https://doi.org/10.1037/0033-295X.85.5.363>
- Kirkpatrick, L. C. & Klein, P. D. (2009): Planning Text Structure as a Way to Improve Students' Writing from Sources in the Compare-Contrast Genre. *Learning and Instruction*, 19(4), 309–321. <https://doi.org/10.1016/j.learninstruc.2008.06.001>
- List, A. & Alexander, P. A. (2019): Toward an Integrated Framework of Multiple Text Use. *Educational Psychologist*, 54(1), 20–39. <https://doi.org/10.1080/00461520.2018.1505514>
- List, A. & Du, H. (2021): Reasoning Beyond History: Examining Students' Strategy Use When Completing a Multiple Text Task Addressing a Controversial Topic in Education. *Reading and Writing*, 34(4), 1003–1048. <https://doi.org/10.1007/s11145-020-10095-5>
- List, A., Du, H. & Lee, H. Y. (2020): How do Students Integrate Multiple Texts? An Investigation of Top-down Processing. *European Journal of Psychology of Education*. <https://doi.org/10.1007/s10212-020-00497-y>
- Mason, L. & Florit, E. (2018): Complementary Methods for Assessing Online Processing of Multiple Sources. In J. L. G. Braasch, I. Bråten & M. T. McCrudden (Eds.), *Handbook of Multiple Source Use* (pp. 425–446). New York: Routledge.
- Mateos, M. & Solé, I. (2009): Synthesising Information from Various Texts: A Study of Procedures and Products at Different Educational Levels. *European Journal of Psychology of Education*, 24(4), 435–451. <https://doi.org/10.1007/BF03178760>
- McCrudden, M. T. (2018): Text Relevance and Multiple-Source Use. In J. L. G. Braasch, I. Bråten, & M. T. McCrudden (Eds.), *Handbook of Multiple Source Use* (pp. 168–183). New York: Routledge.
- McNamara, D. S., Kintsch, E., Songer, N. B. & Kintsch, W. (1996): Are Good Texts Always Better? Interactions of Text Coherence, Background Knowledge, and Levels of Understanding in Learning from Text. *Cognition and Instruction*, 14(1), 1–43. https://doi.org/10.1207/s1532690xci1401_1
- Nelson, N. (2001): Writing to Learn. In P. Tynjälä, L. Mason, & K. Lonka (Eds.), *Writing as a Learning Tool: Integrating Theory and Practice* (pp. 23–36). Dordrecht: Springer.
- Perfetti, C. A., Rouet, J.-F. & Britt, M. A. (1999): Toward a Theory of Documents Representation. In H. van Oosterdorp & S. R. Goldman (Eds.), *The Construction of Mental Representations during Reading* (pp. 99–122). Mahwah: Erlbaum.
- Philipp, M. (2018): *Lesekompetenz bei multiplen Texten: Grundlagen, Prozesse, Didaktik*. Tübingen: Francke.
- Philipp, M. (2020): *Multiple Dokumente verstehen. Theoretische und empirische Perspektiven auf Prozesse und Produkte des Lesens mehrerer Dokumente*. Weinheim: Beltz Juventa.
- Primor, L. & Katzir, T. (2018): Measuring Multiple Text Integration: A Review. *Frontiers in Psychology*, 9, 1–16. <https://doi.org/10.3389/fpsyg.2018.02294>
- Rouet, J.-F. & Britt, M. A. (2011): Relevance Processes in Multiple Document Comprehension. In M. T. McCrudden, J. P. Magliano, & G. Schraw (Eds.), *Text Relevance and Learning from Text* (pp. 19–52). Charlotte: Information Age.

- Rouet, J.-F., Britt, M. A. & Potocki, A. (2019): Multiple-Text Comprehension. In J. Dunlosky & K. A. Rawson (Eds.), *The Cambridge Handbook of Cognition and Education* (pp. 356–380). Cambridge: Cambridge University Press.
- Segev-Miller, R. (2007): Cognitive Processes in Discourse Synthesis: The Case of Intertextual Processing Strategies. In M. Torrance, L. van Waes, & D. Galbraith (Eds.), *Writing and Cognition: Research and Applications* (pp. 231–250). Amsterdam: Elsevier. [https://doi.org/10.1108/S1572-6304\(2007\)0000020016](https://doi.org/10.1108/S1572-6304(2007)0000020016)
- Spivey, N. N. (1990): Transforming Texts: Constructive Processes in Reading and Writing. *Written Communication*, 7(2), 256–287. <https://doi.org/10.1177/0741088390007002004>
- Spivey, N. N. (1991): The Shaping of Meaning: Options in Writing the Comparison. *Research in the Teaching of English*, 25(4), 390–418.
- Spivey, N. N. & King, J. R. (1989): Readers as Writers Composing from Sources. *Reading Research Quarterly*, 24(1), 7–26. <https://doi.org/10.1598/RRQ.24.1.1>
- van Ockenburg, L., van Weijen, D. & Rijlaarsdam, G. (2019): Learning to Write Synthesis Texts: A Review of Intervention Studies. *Journal of Writing Research*, 10(3), 299–326. <https://doi.org/10.17239/jowr-2019.10.03.01>
- Wiley, J. & Voss, J. F. (1999): Constructing Arguments from Multiple Sources: Tasks That Promote Understanding and Not Just Memory for Text. *Journal of Educational Psychology*, 91(2), 301–311. <https://doi.org/10.1037/0022-0663.91.2.301>

Ist argumentatives materialgestütztes Schreiben dem informierenden überlegen?

Eine Sekundäranalyse experimenteller Studien zu den Effekten der Zieltextsorte auf Verstehensleistungen

Maik Philipp

Zusammenfassung: Dem materialgestützten Schreiben (MGS) wird ein hohes epistemisches Potenzial für das tiefe Verstehen von multiplen Dokumenten zugestanden. Dabei scheinen strategische Prozesse der Informationstransformation den Ausschlag zu geben, welche ihrerseits von der Zieltextsorte des MGS abhängen. Dieser Beitrag widmet sich der empirischen Frage, inwieweit differenzielle Effekte bei Paarvergleichen in 18 experimentellen Studien tatsächlich festgestellt wurden. Mittels Vote-Countings wurden für drei besonders häufig untersuchte Zieltextsorten – Argumentation, Beschreibung und Zusammenfassung – und mit einem theoriebasierten Kodieraster jeweils für sechs Gruppen von abhängigen Variablen in expressiven und rezeptiven Tests Effekte zusammengefasst. Dabei zeigte sich keine durchgängige Überlegenheit einer Zieltextsorte, dafür aber Muster in Form relativer Überlegenheiten in einzelnen Gruppen von Variablen. Hinzu kamen Effekte des intertextuellen Verhältnisses der Dokumente, über die die Testpersonen schrieben.

Schlüsselwörter: materialgestütztes Schreiben, Leseverstehen, Deep Reading, Lese-strategien, Textsorte

Abstract: Writing-from-sources (WFS) is regarded as a cognitively demanding, yet fruitful activity when it comes to learning from multiple documents. Transformational processes seem to be the key activities of comprehension and they depend on the text genre of WFS. This chapter summarizes the empirical evidence of studies that used statistical pair comparisons from 18 experiments to shed light on differential outcomes of different genres used in WFS. By means of vote-counting procedures comprehensive codings and analyses were conducted for the three most studied genres (argumentation, narration/description and summary). Theoretically derived codes were used for 3 groups of outcomes and 3 groups of processes as well for expressive or receptive outcomes and for intertextually complementary or conflicting documents. Results indi-

cate no clear-cut superiority of any of the three genres, but there emerge patterns depending on intertextual relationships and kind of tasks in each genre.

Keywords: writing-from-sources, discourse synthesis, deep reading comprehension, reading strategies, genre

1 Einleitung

Wer heute liest, tut dies zunehmend digital und auf der Basis verschiedener multimodaler Dokumente. Gerade strittige oder ambige Themen – von politischen Fragen bis hin zu (natur-)wissenschaftlichen Kontroversen – sind von einer Vielstimmigkeit im Diskurs und zunehmend auch von gezielten Desinformationen (Fake News) gekennzeichnet, die von Leserinnen und Lesern einiges abverlangen, wenn sie kritisch partizipieren und sich informieren wollen. Hinzu kommt, dass Aufgabenformate wie das «materialgestützte Schreiben» sich dessen bedienen und zusätzlich zur erfolgreichen Rezeption auch die glückende Produktion von Texten einfordern. Als ein authentisches Beispiel hierfür kann der folgende Schreibauftrag aus Tabelle 1 gelten, den Studierende mithilfe von fünf digitalen Dokumenten bearbeiteten, die sich teils widersprachen, teils ergänzten.

Tabelle 1: Zwei Varianten eines Auftrags des materialgestützten Schreibens, Variante 1 mit der Zieltextsorte Zusammenfassung, Variante 2 mit der Zieltextsorte Argumentation

<i>Auftrag Variante 1: Stellen Sie sich vor, Sie müssen anderen Studierenden einen kurzen Bericht schreiben, in dem die Ursachen des Klimawandels zusammengefasst sind.</i>	<i>Auftrag Variante 2: Stellen Sie sich vor, Sie müssen anderen Studierenden einen kurzen Bericht schreiben, in dem Sie Ihre persönliche Meinung zu den Ursachen des Klimawandels äußern und begründen.</i>
<i>Stützen Sie Ihren Bericht auf Informationen, die in den folgenden fünf Texten enthalten sind. Sie finden Informationen, die für das Schreiben Ihres Berichts relevant sind, und Informationen, die für das Schreiben des Berichts nicht relevant sind. Verwenden Sie die relevantesten Informationen und versuchen Sie, sich klar auszudrücken und die Informationen auszuarbeiten – vorzugsweise in Ihren eigenen Worten.</i>	

(Quelle: Übersetzung von Gil et al., 2010a, S. 163; der untere Teil des Arbeitsauftrags galt für beide Varianten)

Wenn Leserinnen und Leser solche Aufträge erfolgreich bearbeiten wollen, müssen sie diverse lese- und schreibstrategische Handlungen absolvieren, um dem geforderten Produkt zu entsprechen. Die Frage ist, ob hierbei verschiedene Schreibaufträge mit divergierenden Zieltextsorten empirisch differenzielle Ergebnisse nach sich ziehen. Lesedidaktisch ist das deshalb eine wichtige Frage, weil die Aufgabenstellungen selbst schon als eine Intervention betrachtet wer-

den können – und sie gelten im Vergleich zu logistisch aufwändigeren Fördermaßnahmen wie der Vermittlung von Strategien als niederschwelliger und leichter implementierbar (Stadtler, Bromme & Rouet, 2018). Hinzu kommt, dass bereits metaanalytisch belegt ist, dass generell verschiedene Schreibfördermaßnahmen mit variierenden Zieltextsorten differenzielle Wirkungen auf das Verstehen von einzelnen Sachtexten haben (Graham & Hebert, 2011). Für das materialgestützte Schreiben ist das bislang nur wenig ausgeschildert (Philipp, 2020) und gilt dennoch als aussichtsreicher Weg (Rouet, Britt & Potocki, 2019).

An diesem Punkt setzt der Beitrag an. Er widmet sich der empirischen Frage, ob die Variation der Zieltextsorte – so wie im Beispiel der zwei Varianten von Schreibaufträgen aus Tabelle 1 – sich in Tests statistisch signifikant niederschlägt, also ob es spezifische Vor- oder Nachteile gibt. Die Basis hierfür ist eine kriterienbasierte Sichtung von experimentellen Studien, in denen gezielt die Zieltextsorte – und zwar nur die Zieltextsorte – zwischen Teilnehmenden variiert wurde. Diese Studienbefunde systematisch nach einem feiner differenzierten Auswertungsschema auszuwerten und zu quantifizieren, ist damit Ziel des Beitrags.

2 Materialgestütztes Schreiben – ein Hybrid strategischer Lese- und Schreibprozesse

Das materialgestützte Schreiben – international auch als «Writing from Sources» oder «Discourse Synthesis» deklariert – bezeichnet das Schreiben eines eigenen (Sach-)Textes auf der Basis mehrerer gelesener Dokumente in Hinblick auf einen mehr oder minder expliziten Auftrag hin (Philipp, 2017). Entscheidend ist hierbei, dass kein einzelner Text, sondern mehrere (durchaus multimodale) Texte nebst zugehörigen Metadaten als zusätzlich bedeutungstragende Informationseinheiten von der lesenden Person zu verstehen und in Hinblick auf ein gefordertes schriftliches Aufgabenprodukt zu verarbeiten sind. Dabei sind diverse strategische Lese- und Schreibprozesse nötig, die im Folgenden entfaltet werden. Zunächst geht es um tiefe Verstehensleistungen im rezeptiven Teil des materialgestützten Schreibens (2.1), während schreibnahe (Planungs- und Transformations-)Prozesse in ihrem Verbund danach beleuchtet werden (2.2).

2.1 Leseverstehen multipler Dokumente und das Konzept «Deep Reading»

2.1.1 *Deep Reading und einzelne Texte*

Das verstehende Lesen wird in der Kognitionspsychologie als Prozessverbund modelliert, bei dem automatisiert ablaufende sowie direkt steuerbare Prozesse interagieren. Das produktbezogene Ergebnis dieser Prozesse sind verschiedene Formen der kognitiven Repräsentation. Dabei werden gemäß dem einflussreichen «Construction-Integration-Modell» drei Varianten des Leseverstehens unterschieden: a) die wortwörtliche Repräsentation, b) die propositionale Repräsentation (verstanden als vorwissensbasierte Verknüpfung von sprachlichen Hauptinhalten und ihren Verbindungen untereinander) sowie c) das mentale Modell als nicht-sprachliche, umfassende Repräsentation von Textinhalten (Richter & Schnotz, 2018). Für das Erlangen eines mentalen Modells müssen lesende Personen diverse kognitive Reduktions- und Anreicherungsprozesse bei den Inhalten erfolgreich absolviert haben, während diese bei der wortwörtlichen Repräsentation nicht nötig sind. Prototypisch lässt sich an diesen beiden Repräsentationsarten das Kontinuum von oberflächlicher und Tiefenverarbeitung, von «Shallow Reading» und «Deep Reading» aufspannen, da die wortwörtliche Repräsentation kein zwingendes Verstehen des Inhalts voraussetzt, das mentale Modell hingegen sehr deutlich.

Das tiefe oder oberflächliche Verstehen von Textinhalten wird meist als Oppositionspaar dargestellt. Ein Beispiel dafür stammt von Graesser, Léon und Otero (2002, S. 6), die die gelesenen Inhalte von (wissenschaftlichen) Texten als oberflächliches bzw. tiefes Wissen kontrastieren:

«Oberflächliches Wissen besteht aus explizit genannten Inhalten in Texten, die sich beziehen auf Listen von Konzepten, eine Handvoll einfacher Fakten oder Merkmale jedes Konzepts, einfache Definitionen von Hauptbegriffen und Hauptschritte bei Prozeduren (nicht detaillierten Schritten). Tiefes Wissen besteht aus kohärenten Erklärungen des Inhaltes, welche die lesende Person ermutigen, Inferenzen zu bilden, Probleme zu lösen, Entscheidungen zu treffen, neue Ideen zu synthetisieren, Inhalte in Teile zu zerlegen, zukünftige Ereignisse in einem System zu prognostizieren und Wissen in praktischen Situationen anzuwenden.»

Nach dieser Gegenüberstellung zeichnet sich tiefes Leseverstehen vor allem dadurch aus, dass die lesende Person auf der Basis gelesener Textinhalte und der Anwendung von eigenen Wissensbeständen Schlussfolgerungen zieht, um das Gelesene für variable Zwecke anzuwenden. Dies deckt sich mit dem Katalog der sieben Merkmale tiefen Lesens, den O'Reilly, Sabatini und Wang (2018) mit Blick auf eine sogenannte szenariobasierte Variante des Messens von tiefem Leseverstehen benennen, um dadurch das Konzept «Deep Reading» abzusteu-

cken (s. Kasten «Tiefes Leseverstehen»). Insbesondere sind die Merkmale 2, 3 und 7 deutlich anschlussfähig an die obige Definition. Der Merkmalskatalog geht allerdings noch darüber hinaus, weil es um den kritischen Umgang mit Informationen und Informationsquellen (Merkmal 5) und um die Partizipation an Diskursen und ihren Regeln (Merkmal 6) geht. Alle diese Merkmale – zumal das erste mit dem Zweck des Lesens – sind insbesondere für das Lesen multippler Dokumente im Allgemeinen (2.1.2) und das materialgestützte Schreiben im Besonderen (2.2) von eminenter Bedeutung.

Tiefes Leseverstehen ...

- 1) ... *wird zumindest teilweise vom Zweck und dem Ziel des Lesens bestimmt.* Leserinnen und Leser adaptieren je nach Aufgabenanforderungen einen sogenannten «Kohärenzstandard», sie definieren also ein Soll an Leseverstehen, welches ihre Aufmerksamkeit und ihren lesestrategischen Umgang lenkt. Je komplexer die lesebezogene Aufgabe ist, desto wahrscheinlicher ist es, dass ein tieferes Leseverstehen angestrebt und erreicht wird.
- 2) ... *beinhaltet, über das wortwörtliche Verstehen gegebener Inhalte hinauszugehen, indem die lesende Person gezwungen wird, implizite Inhalte zu erschließen.* Texte erfordern es, entfernt voneinander liegende Informationen kohärent aufeinander zu beziehen sowie implizite Informationen mit Vorwissen anzureichern. Das Bilden von Inferenzen ist dazu nötig, sich die (durchaus implizite) Bedeutung von Texten durch aktives schlussfolgerndes Denken zu erschließen.
- 3) ... *bedeutet, mehr als nur die Hauptinhalte oder -fakten zu wissen, sondern die Verbindung zwischen ihnen.* Dieses Merkmal hängt mit der Nr. 2 zusammen: Texte sind keine isolierte Sammlung von Informationen, sondern bedürfen der aktiven kognitiven Integration und Restrukturierung, um eine kohärente Repräsentation zu erzielen. Erst solche kognitiven Verknüpfungen der lesenden Person gelten als hinreichend für komplexeres Verstehen.
- 4) ... *inkludiert die Fähigkeit, bedeutungsvolle Fragen zu formulieren, die das eigene Verstehen herausfordern oder über die Aussagen der Autorin/des Autors hinausgehen und diese infrage stellen.* Mit diesem Merkmal ist gemeint, dass Leserinnen und Leser im Falle des tieferen Leseverstehens sowohl ein mentales Zwiegespräch mit dem Text führen, um ihr sich entwickelndes und veränderndes Erkenntnisinteresse entlang des Textes zu verbalisieren, als auch ein metakognitives Befragen ihrer selbst vornehmen, um die Auseinandersetzung mit dem Textinhalt zu gewährleisten.
- 5) ... *beinhaltet die Fähigkeit, die Autorinnen und Autoren sowie ihre Aussagen und ihre Belege zu evaluieren.* Leseverstehen bedeutet mehr als die (Re-)Konstruktion eines Textinhaltes. Zunehmend geht es darum, den Wahrheitsgehalt von Aussagen, die Qualitäten von Texten und die Absichten und Qualifikationen von Autorinnen und Autoren kritisch im Blick zu behalten. Dieses evaluative Moment dient als kognitiv anspruchsvolle und voraussetzungsreiche Sicherungsmaßnahme dazu, epistemische Urteile für und gegen die Aufnahme von Informationen, Perspektiven oder ganzen Texten zu fällen. Gleichzeitig ist ein solch kritisches Lesen essenziell im versierten Umgang mit digitalen Internet-Texten.
- 6) ... *ist zwar ein Merkmal mit individuellen Ursprüngen, ist allerdings auch sozial insofern, als es an den Maßstäben von Fachdisziplinen gemessen wird und sich in sozialen Diskursen vollzieht.* Lesen wird als kommunikativer Akt innerhalb sozialer Kontexte modelliert. Das bedeutet für das tiefe Leseverstehen, dass Texte in ihrem Kontext zu sehen – und zu verstehen – sind. Für das angemessene Verstehen müssen daher lesende Personen Wissensbestände aus jenen Kontexten berücksichtigen, in denen die Kommunikation stattfindet. Das gilt insbesondere für die Teilhabe an disziplinären

Diskursen, bei denen der Umgang mit Divergenz und Vorläufigkeit, mit über viele einzelne Dokumente verteilten Konstituenten geteilten Wissens und mit unterschiedlichen Perspektiven typisch ist.

- 7) ... manifestiert sich im Anwendungsbezug in variablen Situationen. Lesen dient häufig instrumentellen Zwecken, d. h. die Nutzung der Informationen aus und über Texte stellt oftmals das Motiv dar, etwas zu lesen. In der adäquaten variablen Anwendung der Textinhalte manifestiert sich ein wünschenswert tiefes Verstehen. (Quelle: Darstellung gemäß O'Reilly et al., 2018, S. 198–202)

2.1.2 Deep Reading und multiple Dokumente

Die bisherigen Ausführungen konzentrierten sich auf das Verstehen singulärer, also einzelner (Sach-)Texte und das Konzept des Deep Reading. Seit den 1990er Jahren und zuletzt mit deutlich anziehender Forschungskonjunktur widmet sich die Leseforschung einem Feld, in dem Personen nicht mehr nur einzelne Texte lesen, sondern (multimodale) Dokumente, d. h. Texte mit zugehörigen, bedeutungsrelevanten Metadaten über die Texte. Für die lesende Person ist dieses quantitative Mehr an zu verarbeitenden Texten nicht folgenlos. Sie kann nicht mit einer Kohärenzerwartung die Texte lesen, sondern muss im Leseprozess selbst erkennen, welche Informationen welcher Text zum übergeordneten Thema offeriert und wie diese Texte im Verhältnis untereinander stehen. Diese Texte können zudem verschiedenster Art sein, nämlich aus verschiedenen Zeiten stammen, verschiedene Modalitäten aufweisen, diversen Genres zugehören, ein im- oder explizites intertextuelles Verhältnis aufweisen etc. (Rouet et al., 2019).

All dies erfordert neue, vor allem lesestrategische Herangehensweisen. Diese beziehen sich im Wesentlichen auf zwei Gruppen von kognitiven Prozessen: das «intertextuelle Integrieren» und das «Sourcing» (Philipp, 2020). Beide Prozessgruppen zeichnen sich durch ihre jeweils unterschiedliche Funktion aus: Beim «intertextuellen Integrieren» geht es darum, zielbezogen Informationen aus einzelnen Dokumenten auszuwählen, zu strukturieren und inferenzbasiert stimmig zu einem kohärenten mentalen Gesamtmodell zu verschmelzen. Diese Prozesse führen im absoluten Optimalfall laut einer besonders einflussreichen theoretischen Modellierung, dem «Dokumentenmodell» (Britt & Rouet, 2012), zu einer Repräsentation der gelesenen Dokumente auf zwei miteinander verbundenen Ebenen. Inhaltlich ist zuvorderst die Ebene des «integrierten mentalen Modells» anzuführen. Auf ihr sind die separaten, aber auch geteilten Repräsentationen der Inhalte aus verschiedenen Dokumenten zu lokalisieren. D. h. Leserinnen und Leser haben hier durch informatorische Transformationsprozesse eine Vorstellung erlangt, wovon die gelesenen Dokumente gesamthaft kohärent inhaltlich handeln und welche Inhalte nur aus einzelnen Dokumenten stammen. Die zweite Ebene wird «Intertextmodell» genannt. Sie beinhaltet das (intertextuelle) Verhältnis der Dokumente, also ob sich Dokumente ergänzen, einander widersprechen, partiell oder vollständig gleiche Perspektiven auf einen

Sachverhalt einnehmen etc. Diese Form der Repräsentation basiert auf teils hochinferenten Verstehensleistungen sowohl einzelner Dokumente als auch intertextueller Inferenzen, die zu «Intertextprädikaten» verdichtet werden (wie «Dokument 1 widerspricht mehrheitlich Dokument 4»).

Auf der Ebene des Intertextmodells ist das «Sourcing» zu verorten. Denn für jedes Dokument sieht das Dokumentenmodell einen «Dokumentenknotten» vor, eine geordnete Sammlung von Metadaten über die Person(en), die das Dokument verfasst hat/haben und den Erscheinungskontext (Zeit, Publikationsort etc.). Diese Metadaten, welche Personen evaluieren, um kommunikative Absichten zu interpretieren und die Glaubwürdigkeit von Aussagen zu beurteilen, bedürfen einer vor allem vorwissensbasierten, interpretativen Verarbeitung. Sie werden dazu genutzt, die inhaltlichen Aussagen zu gewichten, abzulehnen, zu priorisieren etc.

Wer multiple Dokumente liest, muss damit nicht nur viele qualitativ unterschiedliche Informationsarten (Inhalte und Metadaten), sondern auch intra- und intertextuelle Informationen vorwissensbasiert kohärent in eine mentale Repräsentation des Gelesenen überführen. Da viele dieser Leistungen auf stark schlussfolgernden Prozessen basieren, die implizite Zusammenhänge sowohl auf der Ebene des integrierten mentalen Modells als auch des Intertextmodells betreffen, wirkt der kognitive Anspruchsgrad ausgesprochen hoch und wird auch deshalb in der Leseforschung explizit mit dem «Deep Reading» in Verbindung gebracht (Wiley & Guerrero, 2018). Mit Blick auf den Merkmalskatalog des «Deep Reading» aus dem Abschnitt 2.1.1 lässt sich damit plausibilisieren, dass das erfolgreiche Lesen multipler Dokumente viele dieser Eigenheiten aufweist. Besonders auffällig ist dies beim Erschließen impliziter Zusammenhänge (Nr. 2 in der Liste), dem inferenzbasierten Erkennen von Zusammenhängen (Nr. 3) und dem kritisch-evaluativen Lesen (Nr. 5).

2.2 Materialgestütztes Schreiben

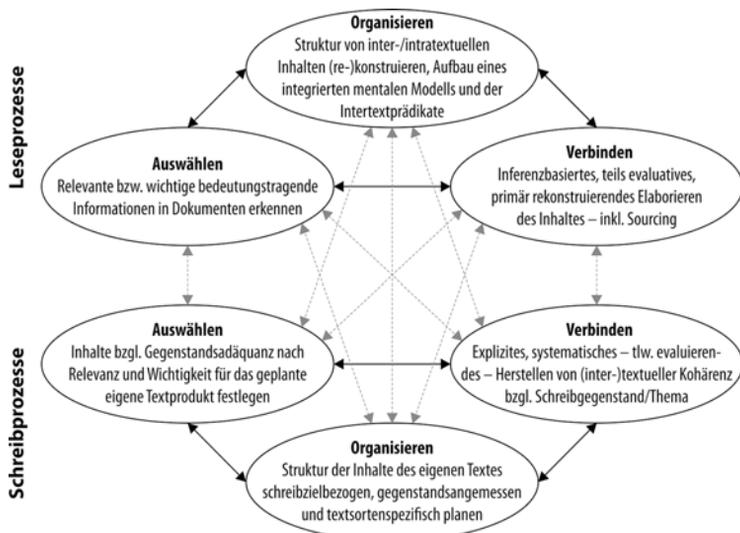
Das Leseverstehen von multiplen Dokumenten ist nur ein Bestandteil des materialgestützten Schreibens. Dem materialgestützten Schreiben ist als Format inhärent, dass das Lesen in Hinblick auf ein Schreibziel, auf ein zu erstellendes Produkt erfolgt. Dieser Zielbezug, der von Anfang an Leseprozesse beeinflusst bzw. metakognitiv beeinflussen sollte, wird zunehmend unter dem Stichwort des «Aufgabenmodells» diskutiert (Rouet & Britt, 2011). Diese metakognitive Steuerungseinheit ist auch vom Schreibauftrag abhängig, der als interpretierte und angereicherte Repräsentation Verarbeitungstiefe und Verarbeitungsprozesse reguliert. Insofern gibt es auch hier eine große Überschneidung zu Merkmalen des «Deep Reading» aus dem Merkmalskatalog in Abschnitt 2.1.1.

2.2.1 Wissenstransformation – ein Prozessverbund des strategischen Lesens und Schreibens

Das Konzept «Wissenstransformation» hat seine Popularität durch die Arbeiten von Scardamalia und Bereiter (1987) erlangt, die in den 1980er Jahren eine Kontrastierung von zwei expertisebedingten Formen des Schreibens vorgenommen haben, nämlich das Wissen-Erzählen («Knowledge Telling») und eben das Transformieren des Wissens («Knowledge Transforming»). In dieser Gegenüberstellung ist die Wissenstransformation das elaboriertere Vorgehen, welches Teile des Erzählens von Wissen aber ausdrücklich beinhaltet. Die Wissenstransformation geht davon aus, dass Schreiben als Problemlöseprozess dazu dient, beabsichtigte Inhalte des Textes (im sogenannten «inhaltlichen Problemlöseraum») mit der intendierten rhetorischen Umsetzung (im «rhetorischen Problemlöseraum») zu harmonisieren. Um die Inhalte möglichst adäquat zu kommunizieren, müssen Schreibende sie so anpassen und modifizieren – sprich: transformieren –, dass sie Publikum, Textsorte, Inhalt und kommunikative Absicht erfolgreich adressieren. Gleichzeitig wirkt der Inhalt – vor allem dessen Komplexität – darauf, welche rhetorischen Ziele sich Personen beim Schreiben stecken und diese dynamisch anpassen. Entscheidend – auch für den Zweck dieses Kapitels – ist damit, dass bei der Wissenstransformation die Erfordernisse der Zieltextsorte als Bestandteil der rhetorischen Ziele den gesamten Prozess des Schreibens leiten. Denn sie fungieren als metakognitive Richtschnur dafür, welche Inhalte auf welche Art ausgewählt, strukturiert und miteinander verbunden werden müssen, um die kommunikative Absicht einzulösen.

Das Konzept Wissenstransformation ist originär nicht spezifisch für das Schreiben über einzelne Texte oder multiple Dokumente vorgelegt worden. Gleichwohl hat es – im Diskurs über das Fachlernen mithilfe des Schreibens – bis heute Konjunktur und wird als Sammelbegriff und Heuristik für verschiedene Vorgehensweisen verwendet. Auf den allgemeinsten Nenner gebracht einen diese Vorgehensweisen planerische zielbezogene Prozesse beim Schreiben (Klein & van Dijk, 2019). Das lässt das Konzept für das materialgestützte Schreiben anschlussfähig wirken, da transformative Prozesse dort ebenfalls typisch, um nicht zu sagen *prototypisch* sind. Tatsächlich bildet das materialgestützte Schreiben eine hybride Form des Lesens und Schreibens, bei der die Grenzen zwischen Lesen und Schreiben verwischen. Das basiert auf ähnlichen kognitiven Prozessen. Diese fallen domänenspezifisch zwar anders aus und beziehen sich auf rezeptive (Lese- und Verstehens-) und expressive (Schreib-) Komponenten. Allerdings haben sie eine ähnliche domänenübergreifende Funktion und müssen zusätzlich von der lesenden und schreibenden Person konzertiert und im Verbund genutzt werden (Spivey, 1990).

Abbildung 1: Transformationsprozesse des Auswählens, Organisierens und Verbindens in ihrer Domänenspezifität im Rahmen des materialgestützten Schreibens



(Quelle der eigenen Darstellung: Philipp, in Druck, basierend vor allem auf Spivey, 1990; Legende: durchgängige Pfeile visualisieren Interaktionen zwischen den Prozessen innerhalb der Domäne, gestrichelte Doppelpfeile beziehen sich auf Interaktionen zwischen Lese- und Schreibprozessen)

Spivey (1990) hat dies herausgearbeitet, indem sie drei Transformationsprozesse expliziert, die konstitutiv im materialgestützten Schreiben sind und die in Abbildung 1 dargestellt sind. Diese Trias – bestehend aus Auswählen, Organisieren und Verbinden – dient dazu,

- dass beim *Auswählen* Personen aus Dokumenten wichtige (dokumentenintern als bedeutsam rekonstruierbare) bzw. relevante (extern als bedeutsam zugewiesene) Informationen beim Lesen erkennen und selektieren bzw. beim Schreiben für den eigenen Text planerisch vorsehen;
- dass beim *Organisieren* die makrostrukturelle Organisation zwischen den Inhalten lesend intra- und intertextuell rekonstruiert wird (etwa für das integrierte mentale Modell; 2.1.2) bzw. für das Schreiben eine eigene thematische Entfaltung (durchaus in den Konventionen der Zieltextsorte) textstrukturell gestaltet wird;
- dass im Falle des *Verbindens* die Vorwissensbestände der lesenden Person vor allem für intra- und intertextuelle Integrationsleistungen beim Lesen (Inferenzbildung) genutzt werden und beim Schreiben des eigenen Texts intertextuelle Kohärenz über die explizite Verknüpfung von Inhalten mit Vorwissen hergestellt und versprachlicht wird.

Diese Prozesse, und das visualisieren die Doppelpfeile in der Abbildung 1, sind nicht als diskrete Phasen zu verstehen, bei denen Personen zunächst lesen und dann schreiben und dabei jeweils auswählen, organisieren und danach verbinden. Ein solches eher rigides Vorgehen ist zwar durchaus empirisch beobachtet worden, allerdings in der Regel von geringerem Erfolg gekrönt (Philipp, 2018; Vandermeulen et al., 2020). Es zeigt sich deshalb, dass ein besseres, tieferes Verstehen mit selbstregulatorisch orchestrierten kognitiven Prozessen des Transformierens für gelingendes materialgestütztes Schreiben Hand in Hand geht (s. auch Rickert & Philipp, i. d. B.).

2.2.2 *Zieltextsorten als Stellschraube potenziell differenzieller Effekte*

Textsorten sind konventionalisierte Formen der (schrift-)sprachlichen Kommunikation, die erwartbare Elemente enthalten, um eine bestimmte kommunikative Absicht zu unterstützen, ohne dass über die Textsorte selbst schon der Inhalt determiniert wäre. Solchen Konventionen zu folgen, erfordert zum einen die Transformation des Wissens vor allem im rhetorischen Problemlöseraum gemäß dem Konzept Wissenstransformation (Scardamalia & Bereiter, 1987). Zum anderen bilden im Optimalfall das Wissen darüber, welche Merkmale Textsorten haben und wie man solche Texte herstellt, eine wichtige Ressource für die metakognitive Steuerungseinheit «Aufgabenmodell» (Rouet & Britt, 2011).

Das Aufgabenmodell wird in der Leseforschung zunehmend stärker gewichtet. Es bezeichnet eine innerhalb der lesenden bzw. schreibenden Person aktiv zu konstruierende Repräsentation zum Ziel und zur Zielerreichung der lese- und schreibbezogenen Aktivität im Sinne einer Ziel-Mittel-Analyse, die ihrerseits dafür benötigt wird, nötige Aktivitäten und deren Ergebnisse zu planen, zu überwachen und zu regulieren (Philippakos, 2018; Rouet & Britt, 2011). Damit bildet das Aufgabenmodell eine durchaus veränderliche Ressource der lesenden und schreibenden Person, um das definierte Ziel zu erreichen, z. B. das Schreiben einer Argumentation oder einer Zusammenfassung über den Klimawandel wie im Falle des Eingangsbeispiels aus Tabelle 1. Für das verstehende Lesen von Dokumenten und die Nutzung dieser Dokumente für das Schreiben eines eigenen Texts sind die Zieltextsorten deshalb wichtig, weil sie – das jeweilige Wissen seitens der Leserinnen und Leser vorausgesetzt – bereits prinzipiell als Ressource dazu dienen, Relevanzentscheidungen zur Verwertbarkeit von Informationen und Dokumenten zu treffen (McCrudden, 2018).

Dafür in Anlehnung an das Eingangsbeispiel eine Kontrastierung: Wer Zusammenfassungen schreibt, sucht vermutlich eher nach übereinstimmenden Themen in multiplen Dokumenten. Die Nutzung derselben Dokumente zum Zweck des Argumentierens zielt eher darauf ab, Positionen zu einem strittigen Thema zu erkennen und sie in ihren wechselseitigen Bezügen nebst kritischer

Evaluation der Metadaten zu verwenden. Wer zusammenfasst, wird Details tilgen müssen, während sie in wissenschaftsbasierten Argumentationen der persuasiven Glaubwürdigkeit im Sinne von Beispielen dienen und damit alles andere als verzichtbar sind. Mit Blick auf die zielbezogene Verwertbarkeit von Informationen oder sogar ganzen Texten bilden geforderte Zieltextsorten als externe Vorgaben nur den Ausgangspunkt der strategischen Kognitionen. Denn Zieltextvorgaben müssen in eine interne Variante überführt werden: als Teil des Aufgabenmodells, das Kohärenzstandards und Relevanzstandards gleichermaßen beeinflusst – und damit die Art und das Maß, wie Personen strategische Transformationsprozesse bemühen (McCrudden, 2018; Philippakos, 2018).

Die Zieltextsorten beim Schreiben haben als ein potenzieller Einflussfaktor von mehreren bereits die Aufmerksamkeit der Leseforschung in Theorie und Praxis auf sich gezogen (Rouet & Britt, 2011; Wiley, Jaeger & Griffin, 2018). Dabei, so resümieren Rouet et al. (2019, S. 369), sind solche Aufgaben vielversprechend, die intra- und intertextuelle Integrationsprozesse einfordern, und als Beispiel bemühen sie explizit Argumentationen. Tatsächlich dominiert die Argumentation die untersuchten Zieltextsorten sowohl in angewandter Forschung (Philipp, 2020) als auch in Grundlagenforschung (Primor & Katzir, 2018); meist in Verbindung mit Dokumentensets, die inhaltliche Widersprüche aufweisen. Der Nutzen konfligierender Dokumente gilt im Sinne der Salienz von intertextuellen Bezügen als besonders lernförderlich (Britt & Rouet, 2012) und dürfte die Wahrscheinlichkeit erhöhen, diverse Transformationsprozesse zu absolvieren, darunter das Sourcing (Philipp, 2020), um Widersprüche nicht nur zu erkennen, sondern auch zu erklären. Allerdings ist in bisherigen Forschungsüberblicken unklar, ob diese Zieltextsorte (in Verbindung mit einer bestimmten Variante von Dokumentensets) konsistent günstiger für Versteheleistungen ist. Dies ist Anlass des vorliegenden quantitativen Reviews.

3 Methode

3.1 Fragestellungen

Dieser Beitrag hat das Ziel, den differenziellen Effekten der Zieltextsorte beim materialgestützten Schreiben auf produktbezogene Leistungsmaße nachzugehen. Drei Fragestellungen sind hierfür leitend:

- Bei welchen Zieltextsorten lassen sich studienübergreifend Vor- und Nachteile bei den abhängigen Variablen in expressiven bzw. rezeptiven Leistungstests finden?

- Welche Effekte bestehen in puncto intertextueller Zusammenstellung der Dokumentensets?
- Welche Effekte ergeben sich aus der Kreuzklassifikation von Leistungstest (expressiv/rezeptiv) und intertextuellem Verhältnis der Dokumente im verwendeten Dokumentenset (kontrastiv/komplementär)?

3.2 Zur Auswahl der Primärstudien

Um die Fragestellungen zu beantworten, wurden per Schneeballsystem ausgehend von drei frühen, zentralen Studien (Voss & Wiley, 1997; Wiley & Voss, 1996, Wiley & Voss, 1999) Experimente recherchiert. Für die Auswertung im Rahmen dieses Beitrags mussten die Studien folgende Inklusionskriterien erfüllen:

- Die Experimentalgruppen schrieben auf der Basis eines identischen Sets von mindestens zwei Dokumenten einen eigenen Text, dessen Länge nicht definiert war, aber im Schreibauftrag als Zieltextsorte eindeutig benannt wurde.
- Der Schreibauftrag sah es vor, dass jede Experimentalgruppe eine Zieltextsorte verfasste, und jede Testperson schrieb nur einen Text.
- Die Auswirkungen der Schreibaufträge wurden in einem Produktmaß erfasst. Dieses Produktmaß konnte aus expressiven (freie, nicht vorab vorgegebene Antwort in aller Regel in einem Textprodukt erfasste Leistung) und/oder rezeptiven Tests (mittels geschlossener oder nur sehr kurzer Antwortformate) stammen.
- Die Effekte der unabhängigen Variablen mit expressiven bzw. rezeptiven Maßen wurden mit einem geeigneten statistischen Verfahren als Paarvergleiche zwischen den Experimentalgruppen mit verschiedenen Zieltextsorten bestimmt, wobei das Signifikanzniveau der Vergleiche bei $p < .05$ liegen musste.
- Die Studien wurden auf Englisch in Fachzeitschriften mit Peer-Review bis Juni 2020 publiziert.

Das Ergebnis von Recherche und Anwendung der Kriterien sind 18 gefundene Experimente, die teils über mehrere Publikationen veröffentlicht wurden. Eine Übersicht über die Studien gibt – schon mit einigen Kodierungen – die Tabelle 2.

Tabelle 2: Überblick über die 18 kodierten Primärstudien

Nr.	Studie	Testpersonen		Materialien		Zieltextsorten und abhängige Variablen					
		Altersgruppe	N	Anzahl Dokumente	Intertextuelles Verhältnis	Anzahl Zieltextsorten	Argumentation	Zusammenfassung	Beschreibung	Andere	Anzahl abhängiger Variablen
1	De La Paz & Wissinger (2015)	Kl. 11	101	3	kontrastiv	2	+	+			8
2	Du & List (2020)	Studierende	32	6	kontrastiv	2	+			Bericht	3
3	Gil et al. (2010a), Experiment 1	Studierende	87	5	kontrastiv	2	+	+			5
4	Gil et al. (2010a), Experiment 2	Studierende	47	5	kontrastiv	2	+	+			4
5	Gil et al. (2010b)	Studierende	53	7	kontrastiv	2	+	+			6
6	Hemmerich & Wiley (2002), Studie 2	Studierende	40	k. A.	k. A.	2	+			Essay	1
7	Lacroix (1999), Experiment 2	Studierende	160	4	komplementär	2		+		Bericht	2
8	Le Bigot & Rouet (2007)	Studierende	52	7	komplementär	2	+	+			14
9	Lehmann et al. (2019, 2020)	Studierende	93	3	komplementär	2	+			unbestimmte Textsorte	11
10	List et al. (2019), Studie 1; Wang & List (2019)	Studierende	143	6	kontrastiv	2	+			Bericht	5
11	Naumann et al. (2009)	Studierende	119	8	kontrastiv	2	+		+		14
12	Stadtler et al. (2014)	Studierende	189	9	kontrastiv	3	+	+		Schlüsselwörter	4
13	Voss & Wiley (1997)	Studierende	96	8	komplementär	3	+		+	Geschichtliche Darstellung	7
14	Wiley (2001), Experiment 1	Studierende	30	11	komplementär	2	+		+		3
15	Wiley (2001), Experiment 2	Studierende	40	19	komplementär	2	+		+		3
16	Wiley & Voss (1996)	Studierende	60	8	komplementär	3	+		+	Geschichtliche Darstellung	9

Nr.	Studie	Testpersonen		Materialien		Zieltextsorten und abhängige Variablen					
		Altersgruppe	N	Anzahl Dokumente	Intertextuelles Verhältnis	Anzahl Zieltextsorten	Argumentation	Zusammenfassung	Beschreibung	Andere	Anzahl abhängiger Variablen
17	Wiley & Voss (1999), Experiment 1	Studierende	64	8	komplementär	4	+	+	+	Erklärung	9
18	Wiley et al. (2009), Experiment 1	Studierende	110	7	komplementär	2	+		+		8

(Legende: + = Zieltextsorte der drei am häufigsten vorkommenden Zieltextsorte wurde in jeweiligem Experiment berücksichtigt, k. A. = keine Angabe)

3.3 Kodierungen und Auswertung

Die 18 Experimente wurden nach Anzahl und Altersgruppe der Testpersonen, der Zahl eingesetzter Dokumente und dem intertextuellen Verhältnis der Dokumente im Set (komplementär vs. kontrastiv),¹ der Anzahl geschriebener Zieltextsorten und der Art der Zieltexte sowie – besonders wichtig – in puncto abhängiger Variablen kodiert. Dabei war eine entscheidende Weichenstellung, dass nur die drei häufigsten Zieltextsorten – Argumentationen (in 17 Experimenten), Zusammenfassungen (8) und Beschreibungen (7) – genauer betrachtet wurden. Die Tabelle 2 enthält die Ergebnisse der Kodierungen. Feingliedriger wurde danach kodiert, welche abhängigen Variablen genutzt wurden. Weil diese Variablen sehr heterogen waren, erfolgte zunächst eine Unterscheidung in expressive und rezeptive Tests. Innerhalb dieser beiden Gruppen wurde genauer kodiert, zum einen welche drei Transformationsprozesse die abhängigen Variablen betrafen und zum anderen welche Arten von Transformationen die Tests erfassten. Beides stützt sich auf ein Kodierschema, welches von Primor und Katzir (2018) bereits verwendet wurde.

Die drei *Prozesse des Transformierens* sind dabei deutlich anschlussfähig an das, was im Abschnitt 2.2.1 beschrieben wurde:

- 1) *Informationsauswahl*, verstanden als Maß für die Selektion von nötigen Informationen aus Dokumenten für die korrekte Nutzung in Antworten;

1 Das intertextuelle Verhältnis konnte im Falle einer Studie (Hemmerich & Wiley, 2002, Studie 2) nicht bestimmt werden. Deshalb kann diese Studie in Analysen, bei denen das intertextuelle Verhältnis zur Unterscheidung von Gruppen nötig ist, nicht berücksichtigt werden.

- 2) *Generieren von Dokumentenrelationen*: das organisierende Verknüpfen von Informationen über Dokumentengrenzen hinweg, vor allem zum Zweck des Klärens des Verhältnisses von Informationen untereinander;
- 3) *Inferenzbildung* als Indikator stark vorwissensbasierter Anreicherungen oder sonstiger hochinferenter Transformationen bzw. Verbindungen.

Ergänzend ließen sich drei *Arten von Transformationen* kodieren:

- 1) *linguistische Transformationen* – Maße, die sich auf sprachliche Veränderungen oder die sprachliche Wiedergabe von Informationen aus Dokumenten beziehen;
- 2) *rhetorische Transformationen* – Indikatoren für die strukturelle Transformation von Inhalten aus Dokumenten;
- 3) *konzeptuelle Transformationen* – Variablen, die die Inhalte selbst und ihre (sich durch die Anwendung verändernde) Nutzung in den Tests betreffen.

Jeder einzelne Paarvergleich aus den 18 Experimenten wurde den jeweiligen drei Arten und drei Prozessen der Transformation – jeweils aufgegliedert nach expressiven und rezeptiven Maßen – zugeordnet (s. Tabelle 3 für die expressiven Tests und Tabelle 4 für die rezeptiven Tests). Für jeden Paarvergleich wurde dann im Sinne eines Vote-Count-Ansatzes (Bushman & Wang, 2009) für jede der beiden verglichenen Zieltextsorten kodiert, ob es entweder a) ein statistisch signifikantes positives Ergebnis (Überlegenheit) im Paarvergleich gab, b) ein signifikantes negatives Ergebnis vorlag (Unterlegenheit) oder c) keine signifikante Differenz bestand (kein Effekt). Das bedeutet, ein Paarvergleich in einer Variable erbrachte immer zwei Ergebnisse, z. B. eine Überlegenheit für die Zieltextsorte Argumentation und eine Unterlegenheit bei der mit der Argumentation verglichenen Beschreibung.

Diese Kodierungen wurden dann studienübergreifend zusammengefasst und ermittelt, wie viele positive, neutrale/keine und negative Effekte pro Prozess und pro Art von Transformationen vorlagen, sodass sechs vertiefte Auswertungen und abstrahierte Gesamtwerte pro Zieltextsorte das Ergebnis bildeten. Dabei wurde zusätzlich noch danach differenziert, ob das intertextuelle Verhältnis komplementär oder kontrastiv war und ob es sich bei der Gruppe abhängiger Variablen um expressive oder rezeptive Maße handelte. Zwecks Vergleichbarkeit wurden diese Zählungen in Prozentangaben überführt und in Balkendiagrammen zusammengefasst. Lagen in einer Auszählung weniger als vier Paarvergleiche vor, erfolgte keine gesonderte Auswertung.

Tabelle 3: Überblick über eingesetzte Testverfahren zur Erfassung von expressiven Leistungen

Transformationsart Prozess	Linguistische Transformationen	Rhetorische Transformationen	Konzeptuelle Transformationen
Informationsauswahl (Auswählen)	<ul style="list-style-type: none"> - Paraphrasen oder wortwörtliche Übernahmen^[8, 9, 11, 16, 17] 	<ul style="list-style-type: none"> - Referenzen zu den Quellen: a) verschiedene Arten von Referenzen^[8, 11, 12], b) Anzahl Zitationen^[2, 10, 11] - Länge des eigenen Textes: a) Anzahl Sätze^[13, 16, 17], b) Wortzahl^[1, 8] - Anzahl der verwendeten Dokumente^[3, 5] 	<ul style="list-style-type: none"> - Abdeckung von Hauptaussagen bzw. Argumenten^[7, 13, 16, 18] - Verhältnis von konzeptueller und propositionaler Übereinstimmung («Balanced Semantic-Matching-Koeffizient»)^[9] - Inakkurate Konzepte (invers kodiert)^[18]
Generieren von Dokumentenrelationen (Organisieren)	<ul style="list-style-type: none"> - Anzahl einzelner Arten von Konnektoren: a) kausale Konnektoren^[8, 11, 13, 16, 17], b) 4 weitere verschiedene Konnektoren^[8] - Gesamtanzahl Konnektoren^[2, 10, 13, 16, 17] - Verknüpfungen von Informationen in Sätzen^[16] - Anzahl integrativer Aussagen im eigenen Text^[10] 	<ul style="list-style-type: none"> - Anzahl Wechsel zwischen den Dokumenten im eigenen Text^[3, 4, 5, 9] - Textstruktur: Analysestruktur vs. Liste^[11, 13, 16] - Textqualität: Ausgewogenheit in der Themenentwicklung und Integration^[2, 10] - Textstruktur über Summenscore von fünf Teildimensionen, davon drei zur Textstruktur^[1] - Übereinstimmung von Vernetzungen der Propositionen mit Musterlösungen («Gamma»-Koeffizient)^[9] - Reihenfolge der thematischen Entwicklung^[7] 	<ul style="list-style-type: none"> - Adäquanz des beschriebenen Sachverhalts als Kausalmodell^[6, 18] - Typologie mentaler Modelle^[10] - Integrative, intertextuelle Elaboration^[9] - Nicht-integrative, intratextuelle Elaboration^[9] - Zweiseitiges Berichten von intertextuellen Konflikten^[12] - Einseitiges Berichten von intertextuellen Konflikten (invers kodiert)^[12]

Transformationsart Prozess	Linguistische Transformationen	Rhetorische Transformationen	Konzeptuelle Transformationen
Inferenzen bilden (Verbinden)	– Anzahl Elaborationen/vorwissensbasierte Transformationen von Inhalten ^[8, 11, 14, 15, 16, 17]	–	– Gesamtscore Transformationen (gebildet aus mehreren Einzelindizes) ^[1, 3, 4, 5, 13] – Anzahl von Hinzufügungen ^[8, 9, 11, 16, 17] – Anzahl irrelevanter Hinzufügungen (invers kodiert) ^[11] – Fähigkeiten zum historischen Denken gemäß eigenem Text ^[1]

(Darstellung nach Primor & Katzir, 2018, S. 12; hochgestellte Zahlen in eckigen Klammern geben die Nummer der Studien aus Tabelle 2 an, in denen ein Test zum Einsatz gekommen ist;
– = keine Studien vorhanden)

Tabelle 4: Überblick über eingesetzte Testverfahren zur Erfassung von rezeptiven Leistungen

Transformationsart Prozess	Linguistische Transformationen	Rhetorische Transformationen	Konzeptuelle Transformationen
Informationsauswahl (Auswählen)	– Beurteilung der inhaltlichen Gültigkeit von paraphrasierten Sätzen (Sentence Verification Task) ^[3, 4, 5, 17, 18] – Wiedererkennen von vorhandenen bzw. nicht-vorhandenen Sätzen ^[11]	–	– Wissenstests ^[1, 9, 11] – Fragen zu mikro- und makrostrukturellen Informationen ^[8] – Beurteilung der Wichtigkeit von Aussagen in Texten ^[13] – Erinnerung an Textinhalte ^[16] – Intratextuelle Verifikationsaufgabe ^[5]
Generieren von Dokumentenrelationen (Organisieren)	–	–	– Beurteilung der Gültigkeit von Aussagen mit integriertem Inhalt (Inference Verification Task) ^[3, 4, 5, 9, 11, 14, 15, 17] – Lesetest zu mikro- und makropropositionalen Inferenzen ^[8] – Erinnerung an konfligierende Informationen aus den Dokumenten ^[12]

Transformationsart Prozess	Linguistische Transformationen	Rhetorische Transformationen	Konzeptuelle Transformationen
Inferenzen bilden (Verbinden)	–	–	<ul style="list-style-type: none"> – Beurteilung von Analogien ^[14, 15] – Bilden von Inferenzen ^[17] – Einschätzung der Glaubwürdigkeit von Websites ^[18] – Einschätzung der Korrektheit von fremden Texten ^[18]

(Darstellung nach Primor & Katzir, 2018, S. 12; hochgestellte Zahlen in eckigen Klammern geben die Nummer der Studien aus Tabelle 2 an, in denen ein Test zum Einsatz gekommen ist; – = keine Studien vorhanden)

4 Ergebnisse

In den 18 Experimenten wurden insgesamt 1.516 Testpersonen getestet, im Mittel waren das 84 Personen pro Experiment. Die Anzahl der Experimente mit komplementären (9) und kontrastiven (8) Dokumentensets war relativ ausgewogen. Die Mehrzahl der Experimente (14) setzte auf zwei Ziltextsorten. Bei der Anzahl der abhängigen Variablen war die Spanne mit einem Spektrum von 1 bis 14 hoch und lag bei einem Medianwert von 5,5.

Die Resultate der 313 Paarvergleiche, das Herzstück der Auswertung, sind in Form zweier Balkendiagramme zusammengetragen. Die Abbildung 2 beinhaltet – zur leichteren Kommunikation mittels Verweisen sind die Variablen alphanumerisch durchnummeriert – Prozesse und Art der Transformationen separat ausgewiesen und die aggregierten Gesamtwerte. Die Sortierung folgt dabei der Unterscheidung von Werten zu expressiven und rezeptiven Tests und danach dem intertextuellen Verhältnis der Dokumente untereinander. Die Abbildung 3 folgt diesem Darstellungsmuster prinzipiell, hat aber die Testform mit intertextuellem Verhältnis kreuzklassifiziert und erlaubt damit vertiefende Einblicke in die Zusammenhänge dieser beiden unabhängigen Variablen.

Abbildung 2: Übersicht über Verteilung der Paarvergleiche in den expressiven und rezeptiven Leistungstests beim materialgestützten Schreiben von drei Zieltextsorten, aufgeschlüsselt nach Prozess und Art der Transformationsleistungen

	Argumentation			Zusammenfassung			Beschreibung		
A) Alle Paarvergleiche (A: 152*, Z: 74, B: 87)	26	60	14	19	69	12	5	67	29
A1) davon expressiv (A: 108*, Z: 49, B: 61)	30	56	15	16	65	18	7	64	30
A2) davon rezeptiv (A: 44, Z: 25, B: 26)	18	70	11	24	76			73	27
A3) davon kontrastiv (A: 53, Z: 31, B: 14)	21	47	32	39	52	10	21	57	
A4) davon komplementär (A: 98, Z: 43, B: 73)	29	67	4	5	81	14	1	75	23
B) Informationsauswahl – alle (A: 57, Z: 26, B: 34)	14	68	18	27	69	4	6	74	21
B1) davon expressiv (A: 36, Z: 15, B: 23)	19	61	19	27	67	7	9	65	26
B2) davon rezeptiv (A: 21, Z: 11, B: 11)	5	81	14	27	73			91	9
B3) davon kontrastiv (A: 21, Z: 11, B: 8)	24	48	29	45	55	13	25	63	
B4) davon komplementär (A: 36, Z: 15, B: 26)	8	81	11	13	80	7	4	88	8
C) Dokumentenrelationen – alle (A: 59*, Z: 30, B: 27)	32	56	12	17	70	13	63	37	
C1) davon expressiv (A: 45*, Z: 19, B: 21)	36	53	11	11	68	21	62	38	
C2) davon rezeptiv (A: 14, Z: 11, B: 6)	21	64	14	27	73		67	33	
C3) davon kontrastiv (A: 22, Z: 13, B: 3)	18	50	32	38	46	15			
C4) davon komplementär (A: 36, Z: 17, B: 24)	39	61		88	12	67	33		
D) Inferenzenbildung – alle (A: 36, Z: 18, B: 26)	36	53	11	11	67	22	8	62	31
D1) davon expressiv (A: 27, Z: 15, B: 17)	33	52	15	12	65	24	12	65	24
D2) davon rezeptiv (A: 9, Z: 3, B: 9)	44	56					56	44	
D3) davon kontrastiv (A: 10, Z: 7, B: 3)	20	40	40	29	57	14			
D4) davon komplementär (A: 26, Z: 11, B: 23)	42	58		73	27	70	30		
E) Linguistische Transformationen – alle (A: 50, Z: 26, B: 35)	30	50	20	19	65	15	9	66	26
E1) davon expressiv (A: 41, Z: 20, B: 29)	34	46	20	15	65	20	10	62	28
E2) davon rezeptiv (A: 9, Z: 6, B: 6)	11	67	22	33	67		83	17	
E3) davon kontrastiv (A: 12, Z: 4, B: 5)	17	33	50		75	25	40	20	40
E4) davon komplementär (A: 38, Z: 22, B: 30)	34	55	11	9	73	18	3	73	23
F) Rhetorische Transformationen – alle (A: 33, Z: 15, B: 16)	24	64	12	20	73	7	63	38	
F1) davon expressiv (A: 33, Z: 14, B: 16)	24	64	12	21	71	7	63	38	
F2) davon rezeptiv (A: 0, Z: 1, B: 0)									
F3) davon kontrastiv (A: 18, Z: 9, B: 5)	28	50	22	33	67		100		
F4) davon komplementär (A: 15, Z: 6, B: 11)	20	80		83	17	91	9		
G) Konzeptuelle Transformationen – alle (A: 69*, Z: 33, B: 36)	25	65	10	18	70	12	3	69	28
G1) davon expressiv (A: 34*, Z: 15, B: 16)	29	59	12	13	60	27	6	69	25
G2) davon rezeptiv (A: 35, Z: 18, B: 20)	20	71	9	22	78		70	30	
G3) davon kontrastiv (A: 23, Z: 18, B: 4)	17	52	30	33	50	17	25	50	25
G4) davon komplementär (A: 45, Z: 15, B: 32)	27	73		93	7	72	28		

□ Positiver Effekt □ Kein Effekt ■ Negativer Effekt

(Anzahl der jeweiligen Paarvergleiche (PV) in der Y-Achse angegeben:

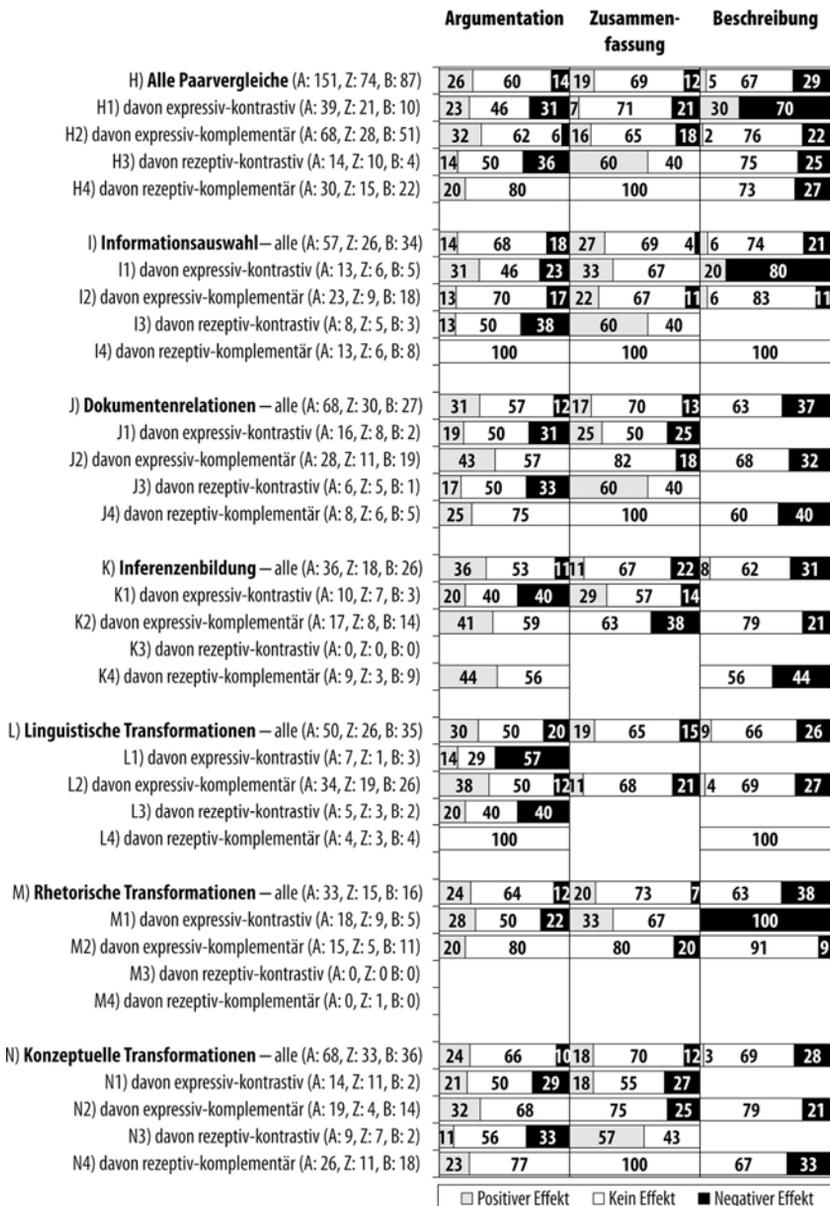
A = Anzahl PV bei Argumentationen; Z = Anzahl PV bei Zusammenfassungen;

B = Anzahl PV bei Beschreibungen; Zahlen im Balkendiagramm geben gerundete Prozente an,

Abweichungen von 100 Prozent basieren auf Rundungsfehlern;

* = Experiment von Hemmerich & Wiley, 2002, Studie 2, in Analyse berücksichtigt)

Abbildung 3: Übersicht über Verteilung der Paarvergleiche in den expressiven und rezeptiven Leistungstests beim materialgestützten Schreiben von drei Zieltextsorten, aufgeschlüsselt in der Kombination von Prozess und Art der Transformationsleistungen



(Anzahl der jeweiligen Paarvergleiche (PV) in der Y-Achse angegeben:

A = Anzahl PV bei Argumentationen; Z = Anzahl PV bei Zusammenfassungen;

B = Anzahl PV bei Beschreibungen; Zahlen im Balkendiagramm geben gerundete Prozente an, Abweichungen von 100 Prozent basieren auf Rundungsfehlern)

4.1 Argumentationen

Argumentationen haben allgemein keine klare Überlegenheit (gerade auch mit Blick auf die negativen Effekte, s. A). Dennoch deuten sich Überlegenheiten eher in den expressiven Tests (A1) und bei komplementären Dokumentensets (A4) an.

Betrachtet man die *Transformationsprozesse*, fallen die geringeren positiven Effekte vor allem bei der Informationsauswahl auf (B, C, D). Während innerhalb der Informationsauswahl keine klaren Muster erkennbar sind (B1–B4, I1–I4), sieht dies bei den Dokumentenrelationen und der Inferenzbildung anders aus. Dort zeigt sich nämlich, dass expressive Aufgaben mit komplementären Dokumentensets besonders häufig zu günstigeren Ausprägungen führten (C1, C4, D1, D4, J2, K2). Kontrastive Dokumentensets waren hingegen eher lernhinderlich (C3, D3, J1, J3, K1).

Hinsichtlich der *Transformationsarten* waren eher die expressiven Tests und Experimente mit komplementären Dokumentensets mit eindeutiger positiven Ergebnissen zugunsten des argumentierenden Schreibens assoziiert (E1, E4, F1, F4, G1, G4). Dies zeigt sich, wo die Kreuzklassifikationen genügend Paarvergleiche enthielten, auch in der Kombination (L2, M2, N2).

Insgesamt scheinen Leistungen mit höherem Anspruchsgrad hinsichtlich der intertextuellen, inhaltlichen Transformation dann besser zu sein, wenn die Aufgaben auf expressive Formate und auf komplementäre Dokumentensets setzten. Allerdings ist bei keiner der Auswertungen der Anteil der positiven Effekte höher als 43 Prozent (J2).

4.2 Zusammenfassungen

Zusammenfassungen als informierende Zieltextsorte haben allgemein positive Effekte dort in Paarvergleichen erbracht, wo die abhängigen Variablen rezeptiv erfasst wurden (A2) und bei intertextuell konfligierenden Dokumentensets (A3). Der Vorteil inhaltlich divergierender Dokumentensets zeigte sich durchgängig bei den *Transformationsprozessen* (B3, C3, D3), ist allerdings aber auf das Zusammenspiel mit rezeptiven Aufgaben zurückzuführen (B2, C2, I3, J3). Bei expressiven Aufgaben sind bei den konfligierenden Dokumentenrelationen (C1) und den Inferenzbildungen (D1) nachteilige Effekte hingegen doppelt so häufig wie vorteilhafte.

Die *Transformationsarten* ähneln dem bisherigen Muster: Kontrastive Dokumentensets (E3, F3, G3) und rezeptive Aufgaben (E2, G2) zeigen konsistentere Vorteile. Komplementäre Dokumentensets (E4, F4, G4) und expressive Tests (E1, G1) weisen teils ein gegenteiliges Muster auf. Die Kreuzkombinationen aus Abbildung 3 erbringen kein eindeutig interpretierbares Muster, zumal

häufig nur wenige oder für die Auswertung zu geringe Paarvergleichszahlen vorliegen.

Als Muster lässt sich aber für die Zusammenfassungen als Zieltextsorte festhalten, dass allgemein rezeptive Tests und kontrastive Dokumentensets recht durchgängig zu höheren Leistungen führen. Dies ist ein anderes Muster als noch bei den Argumentationen.

4.3 Beschreibungen

Die Beschreibungen als zweite informierende Textsorte fallen insofern auf, als hier die negativen Effekte in den Paarvergleichen im Verhältnis zu den positiven Effekten mit wenigen Ausnahmen (E3, G3) die Regel bilden (s. Blöcke A und H in den beiden Balkendiagrammen). Das übergeordnete Muster sowohl bei den *Prozessen* (Blöcke B–D) als auch den *Arten* der Transformationen (Blöcke E–G) zeichnet sich dadurch aus, dass Beschreibungen zu entweder überwiegend vergleichbaren oder schlechteren Testleistungen in den Experimenten führten. Die Kombination von den Merkmalen Test und Dokumentenset, die ähnlich wie bei den Zusammenfassungen von teils zu wenig auswertbaren Paarvergleichen betroffen ist, erbringt in der Tendenz das gleiche Muster wie die separate Betrachtung beider Merkmale. Damit deutet sich an, dass das Verfassen einer Beschreibung im Vergleich mit den anderen untersuchten Zieltextsorten keinen Vorteil mit sich bringt, sondern allenfalls sogar eher schlechtere Leistungen in Experimenten zeitigt.

5 Fazit

5.1 Die Ergebnisse in der Essenz

Die titelgebende Frage des Beitrags lautete, ob das argumentative materialgestützte Schreiben dem informierenden überlegen ist. Auf der Basis des Deep-Reading-Konzepts lässt sich mit Blick auf textsortenspezifisch differenziell ausfallende Transformationsprozesse annehmen, dass in Aufgabenstellungen verlangte Zieltextsorten differenzielle Effekte in Produkten nach sich ziehen. Im Rahmen einer produktbezogenen Sekundäranalyse von 18 Experimenten mit Studierenden und älteren Sekundarschuljugendlichen wurden für die drei empirisch am häufigsten untersuchten Zieltextsorten studienübergreifende Muster im Sinne eines Vote-Countings erfasst. Die Ergebnisse, die insgesamt keine eindeutige Überlegenheit einer Zieltextsorte über alle Studien hinweg demonstrierten, lauten wie folgt:

- Das *materialgestützte Argumentieren* ist dann eher überlegen, wenn die Leistungen das Schreiben und das intertextuelle Transformieren von komplementären Inhalten erfordern.
- Das *materialgestützte Informieren* – übrigens mit einer Ausnahme (Lacroix, 1999) immer auch direkt mit dem materialgestützten Argumentieren in den Experimenten direkt verglichen – weist je nach Zieltextsorte unterschiedliche Befundlagen auf. Das *Beschreiben* wirkt relativ durchgängig unterlegen, während das *Zusammenfassen* bei rezeptiven Tests und auf der Basis kontrastiver Dokumentensets tendenzielle Vorteile aufweist.

Es scheint damit, als wäre das materialgestützte Zusammenfassen zumindest zum Teil das Gegenstück zum materialgestützten Argumentieren. Die Antwort auf die im Kapiteltitel aufgeworfene Frage lautet daher ganz allgemein: Zum Teil ist das materialgestützte Argumentieren in Produktmaßen überlegen.

Etwas kontraintuitiv mutet der Befund an, dass das Argumentieren auf der Basis von komplementären Dokumentensets und das Zusammenfassen bei inhaltlich kontrastierenden Dokumenten tendenzielle Überlegenheiten demonstrierten. Möglicherweise spielen hier die in den Primärstudien nicht oder nur wenig geprüften Interaktionen zwischen Dokumentensets, Aufgabe und individuellen Merkmalen wie dem Vorwissen eine Rolle (Britt & Rouet, 2012). Allerdings lässt sich der Befund auch darüber erklären, dass es möglich ist, dass die Transformation von konvergierenden Informationen in ein argumentatives Schema bzw. von konfligierenden Informationen in die Logik von Zusammenfassungen mehr strategisches elaboratives Inferieren verlangt (van den Broek, Beker & Oudega, 2015). Es geht also um ein Transformieren der Informationen zum Zweck der Aufgabenerfüllung (Spivey, 1990), was tiefere Verarbeitungen im Sinne des Deep Readings nach sich zieht – möglicherweise auch im Sinne einer lernförderlichen «wünschenswerten Erschwernis» (Bjork & Bjork, 2011). Auch dies wäre mit Studien zu unterfüttern, die sich auf Prozesse und Zwischenprodukte wie Notizen konzentrieren (List, Du & Lee, in Druck).

5.2 Forschungsperspektiven

Argumentieren und Zusammenfassen zählen im Inventar der schreibbasierten schulischen Förderung von Leseverstehen und Fachlernen zu den effektivsten Förderansätzen (Graham & Hebert, 2011; Graham, Kiuahara & MacKay, 2020), allerdings ist die relative Überlegenheit noch längst keine ausgemachte Sache, teils noch nicht einmal in Studien vergleichend untersucht (Hebert, Gillespie & Graham, 2013). Hier hat dieser Forschungsüberblick in dem methodisch limitierten Rahmen (Bushman & Wang, 2009) und mit seinen vielen, durchaus heterogenen abhängigen Variablen zumindest in der Tendenz demonstriert,

dass es sich lohnt, den differenziellen, anscheinend durchaus komplementär wirkenden Effekten der Zieltextsorten nachzuspüren.

Es wäre zu wünschen, dass eine Metaanalyse – zumal mit noch mehr Primärstudien – tatsächlich Effektstärken quantifiziert. Hinzu kommt, dass die vielen leeren Zellen aus der Tabelle 3 und der Tabelle 4 von weiteren Primärstudien gefüllt werden sollten (Primor & Katzir, 2018). Ähnliches gilt für generell mehr Studien mit rezeptiven Maßen, für Studien, die die statistische Unabhängigkeit von abhängigen Variablen nachweislich überprüfen, jüngere Altersgruppen testen und nicht zuletzt für proximale und distale Einflussfaktoren auf Testleistungen in Verbindung mit experimentell variierten unabhängigen Variablen Ergebnisse berichten (Cromley, 2020). Als wichtiges Forschungsfeld in der ohnehin diesbezüglich stärker werdenden Leseforschung dürften die Aufgabenmodelle selbst viele für die Theorieentwicklung und (Hoch-)Schulpraxis dringend nötige empirische Befunde beisteuern (List, Du & Wang, 2019; Philippakos, 2018; Schoor et al., 2021).

Literatur

- Bjork, E. L. & Bjork, R. A. (2011): Making Things Hard on Yourself, But in a Good Way. Creating Desirable Difficulties to Enhance Learning. In M. A. Gernsbacher & E. M. Pommerantz (Eds.), *Psychology and the Real World. Essays Illustrating Fundamental Contributions to Society* (pp. 56–64). New York: Worth Pub.
- Britt, M. A. & Rouet, J.-F. (2012): Learning with Multiple Documents. Component Skills and Their Acquisition. In J. R. Kirby & M. J. Lawson (Eds.), *Enhancing the Quality of Learning. Dispositions, Instruction, and Learning Processes* (pp. 276–314). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139048224.017>
- Bushman, B. J. & Wang, M. C. (2009): Vote-Counting Procedures in Meta-Analysis. In H. Cooper, L. V. Hedges & J. C. Valentine (Eds.), *The Handbook of Research Synthesis and Meta-Analysis* (2nd ed., pp. 207–220). New York: Russell Sage Foundation.
- Cromley, J. G. (2020): Learning from Multiple Representations. Roles of Task Interventions and Individual Differences. In P. N. van Meter, A. List, D. Lombardi & P. Kendeou (Eds.), *Handbook of Learning from Multiple Representations and Perspectives* (pp. 62–75). New York: Routledge. <https://doi.org/10.4324/9780429443961-6>
- La Paz, S. de & Wissinger, D. R. (2015): Effects of Genre and Content Knowledge on Historical Thinking with Academically Diverse High School Students. *The Journal of Experimental Education*, 83(1), 110–129. <https://doi.org/10.1080/00220973.2013.876228>
- Du, H. & List, A. (2020): Researching and Writing Based on Multiple Texts. *Learning and Instruction*, 66, Article 101297, 1–25. <https://doi.org/10.1016/j.learninstruc.2019.101297>
- Gil, L., Bråten, I., Vidal-Abarca, E. & Strømsø, H. I. (2010a): Summary versus Argument Tasks when Working with Multiple Documents. Which Is Better for Whom? *Contemporary Educational Psychology*, 35(3), 157–173. <https://doi.org/10.1016/j.cedpsych.2009.11.002>

- Gil, L., Bråten, I., Vidal-Abarca, E. & Strømsø, H. I. (2010b): Understanding and Integrating Multiple Science Texts. Summary Tasks Are Sometimes Better than Argument Tasks. *Reading Psychology*, 31(1), 30–68. <https://doi.org/10.1080/02702710902733600>
- Graesser, A. C., León, J. A. & Otero, J. C. (2002): Introduction to the Psychology of Science Text Comprehension. In J. C. Otero, J. A. León & A. C. Graesser (Eds.), *The Psychology of Science Text Comprehension* (pp. 1–15). New York: Routledge.
- Graham, S. & Hebert, M. (2011): Writing to Read: A Meta-Analysis of the Impact of Writing and Writing Instruction on Reading. *Harvard Educational Review*, 81(4), 710–744. <https://doi.org/10.17763/haer.81.4.t2k0m13756113566>
- Graham, S., Kiuahara, S. A. & MacKay, M. (2020): The Effects of Writing on Learning in Science, Social Studies, and Mathematics. A Meta-Analysis. *Review of Educational Research*, 90(2), 179–226. <https://doi.org/10.3102/0034654320914744>
- Hebert, M., Gillespie, A. & Graham, S. (2013): Comparing Effects of Different Writing Activities on Reading Comprehension: A Meta-Analysis. *Reading and Writing*, 26(1), 111–138. <https://doi.org/10.1007/s11145-012-9386-3>
- Hemmerich, J. A. & Wiley, J. (2002): Do Argumentation Tasks Promote Conceptual Change about Volcanoes? *Proceedings of the Annual Meeting of the Cognitive Science Society*, 24, 453–458.
- Klein, P. D. & van Dijk, A. (2019): Writing as a Learning Activity. In J. Dunlosky & K. A. Rawson (Eds.), *The Cambridge Handbook of Cognition and Education* (pp. 266–291). Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781108235631.012>
- Lacroix, N. (1999): Macrostructure Construction and Organization in the Processing of Multiple Text Passages. *Instructional Science*, 27(3), 221–233. <https://doi.org/10.1007/BF00897320>
- Le Bigot, L. & Rouet, J.-F. (2007): The Impact of Presentation Format, Task Assignment, and Prior Knowledge on Students' Comprehension of Multiple Online Documents. *Journal of Literacy Research*, 39(4), 445–470. <https://doi.org/10.1080/10862960701675317>
- Lehmann, T., Pirnay-Dummer, P. & Schmidt-Borcherding, F. (2020): Fostering Integrated Mental Models of Different Professional Knowledge Domains. Instructional Approaches and Model-Based Analyses. *Educational Technology Research and Development*, 68(3), 905–927. <https://doi.org/10.1007/s11423-019-09704-0>
- Lehmann, T., Rott, B. & Schmidt-Borcherding, F. (2019): Promoting Pre-Service Teachers' Integration of Professional Knowledge. Effects of Writing Tasks and Prompts on Learning from Multiple Documents. *Instructional Science*, 47(1), 99–126. <https://doi.org/10.1007/s11251-018-9472-2>
- List, A., Du, H. & Lee, H. Y. (in Druck): How Do Students Integrate Multiple Texts? An Investigation of Top-Down Processing. *European Journal of Psychology of Education*. <https://doi.org/10.1007/s10212-020-00497-y>
- List, A., Du, H. & Wang, Y. (2019): Understanding Students' Perceptions of Task Assignments. *Contemporary Educational Psychology*, 59, Article 101801, 1–16. <https://doi.org/10.1016/j.cedpsych.2019.101801>
- List, A., Du, H., Wang, Y. & Lee, H. Y. (2019): Toward a Typology of Integration. Examining the Documents Model Framework. *Contemporary Educational Psychology*, 58, 228–242. <https://doi.org/10.1016/j.cedpsych.2019.03.003>
- McCrudden, M. T. (2018): Text Relevance and Multiple Source Use. In J. L. G. Braasch, I. Bråten & M. T. McCrudden (Eds.), *Handbook of Multiple Source Use* (pp. 168–183). New York: Routledge. <https://doi.org/10.4324/9781315627496-10>

- Naumann, A. B., Wechsung, I. & Krems, J. F. (2009): How to Support Learning from Multiple Hypertext Sources. *Behavior Research Methods*, 41(3), 639–646. <https://doi.org/10.3758/BRM.41.3.639>
- O'Reilly, T., Sabatini, J. & Wang, Z. (2018): Using Scenario-Based Assessment to Measure Deep Learning. In K. Millis, D. L. Long, J. P. Magliano & K. Wiemer (Eds.), *Deep Comprehension. Multi-Disciplinary Approaches to Understanding, Enhancing, and Measuring Comprehension* (pp. 197–208). New York: Routledge.
- Philipp, M. (2017): *Materialgestütztes Schreiben. Anforderungen, Grundlagen, Vermittlung*. Weinheim: Beltz Juventa.
- Philipp, M. (2018): *Lesekompetenz bei multiplen Texten. Grundlagen, Prozesse, Didaktik*. Tübingen: Francke.
- Philipp, M. (2020): *Multiple Dokumente verstehen. Theoretische und empirische Perspektiven auf Prozesse und Produkte des Lesens mehrerer Dokumente*. Weinheim: Beltz Juventa.
- Philipp, M. (in Druck): Ohne Lesen nichts gewesen! Warum strategisches Lesen multipler Dokumente in den Kernbereich des wissenschaftlichen Schreibens (und damit in Schreibratgeber) gehört. In K. Burkhalter & B. Rieder (Hrsg.), *Schreibratgeber für die Hochschule: Eine Buchsorte zwischen Markt und Wissenschaft*. Frankfurt am Main: Peter Lang.
- Philippakos, Z. A. (2018): Using a Task Analysis Process for Reading and Writing Assignments. *The Reading Teacher*, 72(1), 107–114. <https://doi.org/10.1002/trtr.1690>
- Primor, L. & Katzir, T. (2018): Measuring Multiple Text Integration. A Review. *Frontiers in Psychology*, 9, Article 2294, 1–16. <https://doi.org/10.3389/fpsyg.2018.02294>
- Richter, T. & Schnotz, W. (2018): Textverstehen. In D. H. Rost, J. R. Sparfeldt & S. R. Buch (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (5., überarb. und erw. Aufl., S. 826–837). Weinheim: Beltz.
- Rouet, J.-F. & Britt, M. A. (2011): Relevance Processes in Multiple Document Comprehension. In M. T. McCrudden, J. P. Magliano & G. J. Schraw (Eds.), *Text Relevance and Learning from Text* (pp. 19–52). Charlotte: Information Age.
- Rouet, J.-F., Britt, M. A. & Potocki, A. (2019): Multiple-Text Comprehension. In J. Dunlosky & K. A. Rawson (Eds.), *The Cambridge Handbook of Cognition and Education* (pp. 356–380). Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781108235631.015>
- Scardamalia, M. & Bereiter, C. (1987): Knowledge Telling and Knowledge Transforming in Written Composition. In S. Rosenberg (Ed.), *Advances in Applied Psycholinguistics. Volume 2: Reading, Writing, and Language Learning* (pp. 142–175). Cambridge: Cambridge University Press.
- Schoor, C., Rouet, J.-F., Artelt, C., Mahlow, N., Hahnel, C., Kröhne, U. & Goldhammer, F. (2021): Readers' Perceived Task Demands and their Relation to Multiple Document Comprehension Strategies and Outcome. *Learning and Individual Differences*, 88, Article 102018, 1–13. <https://doi.org/10.1016/j.lindif.2021.102018>
- Spivey, N. N. (1990): Transforming Texts: Constructive Processes in Reading and Writing. *Written Communication*, 7(2), 256–287. <https://doi.org/10.1177/0741088390007002004>
- Stadtler, M., Bromme, R. & Rouet, J.-F. (2018): Learning from Multiple Documents. How Can We Foster Multiple Document Literacy Skills in a Sustainable Way? In E. Manalo, Y. Uesaka & C. Chinn (Eds.), *Promoting Spontaneous Use of Learning and Reasoning Strategies. Theory, Research, and Practice for Effective Transfer* (pp. 46–61). Singapore: Routledge.

- Stadtler, M., Scharrer, L., Skodzik, T. & Bromme, R. (2014): Comprehending Multiple Documents on Scientific Controversies: Effects of Reading Goals and Signaling Rhetorical Relationships. *Discourse Processes*, 51(1–2), 93–116. <https://doi.org/10.1080/0163853X.2013.855535>
- van den Broek, P., Beker, K. & Oudega, M. (2015): Inference Generation in Text Comprehension. Automatic and Strategic Processes in the Construction of a Mental Representation. In E. J. O'Brien, A. E. Cook & R. F. Lorch (Eds.), *Inferences during Reading* (pp. 94–121). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781107279186.006>
- Vandermeulen, N., van den Broek, B., van Steendam, E. & Rijlaarsdam, G. (2020): In Search of an Effective Source Use Pattern for Writing Argumentative and Informative Synthesis Texts. *Reading and Writing*, 33(2), 239–266. <https://doi.org/10.1007/s11145-019-09958-3>
- Voss, J. F. & Wiley, J. (1997): Developing Understanding while Writing Essays in History. *International Journal of Educational Research*, 27(3), 255–265. [https://doi.org/10.1016/S0883-0355\(97\)89733-9](https://doi.org/10.1016/S0883-0355(97)89733-9)
- Wang, Y. & List, A. (2019): Calibration in Multiple Text Use. *Metacognition and Learning*, 14(2), 131–166. <https://doi.org/10.1007/s11409-019-09201-y>
- Wiley, J. (2001): Supporting Understanding through Task and Browser Design. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 23, 1136–1143.
- Wiley, J., Goldman, S. R., Graesser, A. C., Sanchez, C. A., Ash, I. K. & Hemmerich, J. A. (2009): Source Evaluation, Comprehension, and Learning in Internet Science Inquiry Tasks. *American Educational Research Journal*, 46(4), 1060–1106. <https://doi.org/10.3102/0002831209333183>
- Wiley, J. & Guerrero, T. A. (2018): Prose Comprehension beyond the Page. In K. Millis, D. L. Long, J. P. Magliano & K. Wiemer (Eds.), *Deep Comprehension. Multi-Disciplinary Approaches to Understanding, Enhancing, and Measuring Comprehension* (pp. 3–15). New York: Routledge.
- Wiley, J., Jaeger, A. J. & Griffin, T. D. (2018): Effects of Instructional Conditions on Comprehension from Multiple Sources in History and Science. In J. L. G. Braasch, I. Bråten & M. T. McCrudden (Eds.), *Handbook of Multiple Source Use* (pp. 341–361). New York: Routledge.
- Wiley, J. & Voss, J. F. (1996): The Effects of 'Playing Historian' on Learning in History. *Applied Cognitive Psychology*, 10(7), 63–72. [https://doi.org/10.1002/\(SICI\)1099-0720\(199611\)10:7<63::AID-ACP438>3.0.CO;2-5](https://doi.org/10.1002/(SICI)1099-0720(199611)10:7<63::AID-ACP438>3.0.CO;2-5)
- Wiley, J. & Voss, J. F. (1999): Constructing Arguments from Multiple Sources: Tasks That Promote Understanding and Not Just Memory for Text. *Journal of Educational Psychology*, 91(2), 301–311. <https://doi.org/10.1037/0022-0663.91.2.301>

Die Autorinnen und Autoren

Mareike Ehlert ist wissenschaftliche Mitarbeiterin an der Westfälischen Wilhelms-Universität Münster. Ihre Forschungsschwerpunkte liegen in der diagnosebasierten, differenzierten Leseförderung und in der Implementation evidenzbasierter Methoden in den Leseunterricht.

Dr. **Sven Hilbert** ist Professor für Methoden der empirischen Bildungsforschung und Wissenschaftlicher Leiter des Zentrums für Hochschul- und Wissenschaftsdidaktik an der Universität Regensburg. Seine Forschungsschwerpunkte sind Statistik und Psychometrie.

Dr. **Simone Jambor-Fahlen** hat Germanistik, Allgemeine Sprachwissenschaft und Pädagogik an der Bergischen Universität Wuppertal studiert und ist derzeit wissenschaftliche Mitarbeiterin am Mercator-Institut für Sprachförderung und Deutsch als Zweitsprache der Universität zu Köln. Ihre Forschungsschwerpunkte sind Lesen und Schreiben sowie der Schriftspracherwerb.

Dr. **Britta Juska-Bacher** ist Professorin für Fachwissenschaft und Fachdidaktik Deutsch an der Pädagogischen Hochschule Bern. Ihre Forschungsschwerpunkte sind Wortschatz, Lesen und historische Erstlese-Lehrmittel.

Elisabeth Kraus ist wissenschaftliche Mitarbeiterin in den Methoden der empirischen Bildungsforschung an der Universität Regensburg. Ihre Forschungsschwerpunkte sind Psychometrie und Anwendungsgebiete der Entscheidungstheorie.

Dr. **Maik Philipp** ist Professor für Deutschdidaktik an der Pädagogischen Hochschule Zürich. Seine Forschungsschwerpunkte liegen in der Lese- und Schreibdidaktik, der Selbstregulation beim Lesen und Schreiben, dem Lernen durch Lesen und Schreiben, epistemischen Kognitionen und im digitalen Lesen.

Alex Rickert ist Dozent und Leiter des Schreibzentrums der Pädagogischen Hochschule Zürich. Seine Forschungs- und Arbeitsschwerpunkte sind Textlinguistik sowie Schreibberatung und -didaktik.

Martina Röthlisberger hat allgemeine Sprachwissenschaften studiert und ist wissenschaftliche Mitarbeiterin in der Leseforschung der Pädagogischen Hochschule Bern und an der Klinik für Kinder- und Jugendpsychiatrie der Universität Zürich.

Dr. **Anita Schilcher** ist Professorin für die Didaktik der deutschen Sprache und Literatur und ist Leiterin des Regensburger Universitätszentrums für Lehrerbildung. Ihre Forschungsschwerpunkte sind die Lese- und Schreibdidaktik, literarisches Lernen sowie professionelle Kompetenzen von Deutschlehrkräften.

Professor Dr. **Elmar Souvignier** leitet die Arbeitseinheit «Diagnostik und Evaluation im schulischen Kontext» am Institut für Psychologie in Bildung und Erziehung der Westfälischen Wilhelms-Universität Münster. Seine Forschungsschwerpunkte liegen bei der diagnosebasierten Förderung (Lesen und Mathematik), der Lernverlaufsdiagnostik sowie der Implementationsforschung.

Maria Steinert ist abgeordnete Lehrkraft am Lehrstuhl für Didaktik der deutschen Sprache und Literatur an der Universität Regensburg. Ihr Forschungsschwerpunkt ist die Lesedidaktik.

Dr. **Johannes Wild** ist Akademischer Rat am Lehrstuhl für Didaktik der deutschen Sprache und Literatur an der Universität Regensburg. Seine Forschungsschwerpunkte liegen in der Lese- und Schreibdidaktik.

Dr. **Christoph Zangger** hat Soziologie, Politikwissenschaft und mathematische Statistik studiert und ist derzeit als wissenschaftlicher Mitarbeiter am Soziologischen Institut der Universität Zürich tätig. Zu seinen Forschungsinteressen zählen u. a. soziale Ungleichheit und intergenerationale Mobilität.