

**Empirische Sonderpädagogik**, 2020, Nr. 4, S. 320-346  
ISSN 1869-4845 (Print) · ISSN 1869-4934 (Internet)

## **Verbesserung des Leseverständnisses auch bei schwachen Grundschulern der vierten Klasse: Evaluation einer Fördermaßnahme auf der Basis des Programms „Wir werden Textdetektive“ im Rahmen einer Felduntersuchung**

*Jürgen Walter*

*Europa-Universität Flensburg*

### **Zusammenfassung**

Im Rahmen der internationalen Grundschul-Leseuntersuchung (IGLU 2016) wurden eher ernüchternde Befunde ermittelt sowie eine kritische Analyse strukturell-inhaltlicher Merkmale des Deutschunterrichts an deutschen Grundschulen vorgenommen. Vor diesem Hintergrund und den häufig eher schwach ausfallenden Effektstärken strategieorientierter Förderkonzepte wurde im Rahmen einer Felduntersuchung eine weitere Evaluation des Unterrichtsprogramms „Wir werden Textdetektive“ (Gold et al., 2006) realisiert. Es zeigte sich ein signifikanter kurz- und mittelfristiger Transfer-Effekt ( $d = 0.47$  bzw.  $d = 0.30$ ) auf standardisiert erfasste Leseverständnisleistungen insgesamt bei Viertklässlern sowie ein Effekt von  $d = 0.65$  bzw.  $d = 0.56$  bei unterdurchschnittlichen Lesern dieser Klassenstufe. Die Analyse der allgemeinen Forschungslage sowie die Erfahrungen im Rahmen dieser Untersuchung legen für die Zukunft ein gezielteres und hybridorientiertes Vorgehen zur Verbesserung von Leistungen im Leseverstehen nahe.

*Schlüsselwörter:* Leseförderung, Leseverständnis, Lesestrategien, Grundschüler

## Improvement of reading comprehension even with below-average fourth grade readers: Evaluation of the program “Wir werden Textdetektive” as part of a field study

### Abstract

As a result of the Progress in International Reading Literacy Study (PIRLS 2016), sobering findings were made and a critical analysis of the structural and content-related features concerning the reading instruction at German primary schools was carried out. Against this background and the often rather weak effect sizes of strategy-oriented teaching concepts, a further evaluation of the teaching program “Wir werden Textdetektive” (Gold et al., 2006) was realized. There was a significant short- and medium-term transfer effect ( $d = 0.47$  and  $d = 0.30$ ) in a standardized reading comprehension test overall in fourth graders as well as an effect of  $d = 0.65$  and  $d = 0.56$  for below-average readers of this grade. The analysis of the general research situation as well as the experiences in the context of this study suggest a more targeted and hybrid-oriented approach to improve performance in reading comprehension in the future.

*Keywords:* reading promotion, reading comprehension, reading strategies, elementary school students

### Einleitung

Die Befunde der aktuellen internationalen Grundschul-Leseuntersuchung IGLU 2016 (Hußmann et al., 2017) werfen ein eher ernüchterndes Schlaglicht auf den Leistungsstand deutscher Schülerinnen und Schüler am Ende der Grundschulzeit sowie auf strukturell-inhaltliche Merkmale des Leseunterrichts, auch wenn Deutschland mit 537 Leistungspunkten insgesamt einen Rangplatz im Mittelfeld einnimmt.

Betrachtet man nämlich den Leistungsstand deutscher Viertklässler differenziert (Bos, Valtin, Hußmann, Wendt & Goy, 2017), so fällt auf, dass 11% von ihnen die Aufgaben lösen können, die am obersten Ende der fünfstufigen Kompetenzskala (Stufe V) gestellt werden (höchste Kompetenzstufe). Dies entspricht zwar dem EU/OECD-Mittelwert, kann aber nicht darüber hinwegtäuschen, dass ein deutlich höherer Prozentsatz von Kindern, nämlich ca. 20%, in einer Reihe von europäischen Nachbarstaaten diese Stufe erreichen (Bulgarien, England, Finnland, Irland, Nordirland und Polen). Nimmt man das untere Ende der

Leistungsskala in den Blick (Stufe I und II, d.h. eher rudimentäres Leseverständnis), so zeigt sich, dass in Deutschland hier eine recht hohe Rate von ca. 19% der Kinder anzusiedeln ist. Bei diesen Schülern ist damit zu rechnen, dass sie im Rahmen ihrer weiteren Beschulung aufgrund ihrer unzureichenden Lesekompetenz in allen Fächern Schwierigkeiten beim Lernen bekommen werden, wenn sie nicht grundlegend und nachhaltig gefördert werden. Bremerich-Vos, Wendt und Bos (2017) stellen in diesem Zusammenhang fest, dass sich bezüglich dieser Befundlage keine statistisch signifikante Veränderung im Vergleich zum Jahr 2001 feststellen lässt, und zwar auch dann, wenn man Veränderungen innerhalb der Schülerschaft berücksichtigt.

IGLU 2016 stellt nicht nur Leistungsvergleiche an, sondern beleuchtet auch strukturelle und inhaltliche Aspekte des Deutschunterrichts im Vergleich zu anderen Ländern (Bremerich-Vos, Stahns, Hußmann, & Schurig, 2017; Mullis, Martin, Foy & Hoooper, 2017). So entfallen auf deutsche Viertklässler bezüglich des expliziten Leseunterrichts (auch über Fächergrenzen hinweg) im

Schuljahr ca. 90 Stunden. Der internationale Mittelwert liegt bei ca. 160 Stunden. Auch inhaltlich zeigen sich Unterschiede. So stellen Bremerich-Vos et al. (2017, S. 293-295) fest, „dass nur eine Minderheit der Kinder von Lehrkräften unterrichtet wird, die in nennenswertem Umfang explizit lehren, wie und wann Lesestrategien einzusetzen sind. 56 beziehungsweise 62 Prozent der Schülerinnen und Schüler werden von Lehrkräften unterrichtet, die die Optionen „ein- bis zweimal pro Monat“ beziehungsweise „nie“ angekreuzt haben... Es liegt nahe, diesen Befund mit dem vergleichsweise schwachen Abschneiden der Viertklässlerinnen und Viertklässler in Deutschland im Rahmen von IGLU 2016 in Verbindung zu bringen.“

Vor dem skizzierten Hintergrund ist es dringend erforderlich, die Entwicklung, Evaluation und Implementierung von Unterrichts- und Förderkonzepten im Grundschulbereich voranzutreiben, um eine Verbesserung der Sinnentnahme beim Lesen zu ermöglichen. Dazu macht es zunächst Sinn, sich die Wirksamkeit unterschiedlicher im Grundschulbereich relevanter Förderkonzepte anzuschauen.

Die Befundlage bezüglich des Einflusses der Vermittlung hierarchieniedriger Teilkompetenzen (phonologische Bewusstheit, Wortlesekompetenz, Leseflüssigkeit) auf das Leseverstehen stellt sich auf einem geringen bis moderaten Level (Cohen, 1988) als vergleichsweise stabil dar. Meta-Analysen aus dem englischsprachigen Raum kamen bezüglich des Einflusses der phonologischen Bewusstheit auf das Leseverstehen mit  $d = 0.34$  (Ehri, Nunes, Willows, Schuster, Yaghoub-Zadeh & Shanahan, 2001) zu einem etwas stärkeren Effekt als diejenigen im deutschsprachigen Raum. So ermittelten Fischer und Pfof (2015) bezüglich eines kurzfristigen Effekts ( $\leq 1$  Jahr) ein  $d = 0.26$  und langfristig ( $> 1$  Jahr) eine Effektstärke von  $d = 0.16$ . Wolf, Schroeders und Kriegbaum (2016) fanden für kitabasierte Interventionen mit  $d = 0.18$  einen geringen und

für schulbasierte Fördermaßnahmen mit  $d = 0.39$  einen moderaten Zusammenhang.

Bezüglich des Einflusses der Förderung der Wortlesekompetenz auf der Basis eines Unterrichtskonzepts, das automatisiertes Abrufen von Graphem-Phonem-Korrespondenzen (phonics instruction) in den Vordergrund stellt, kam der National Reading Panel (NICHD, 2000) zu dem Befund, dass bei jüngeren Kindern diese Vorgehensweise mit  $d = 0.51$  einen deutlich stärkeren Effekt auf das Leseverstehen ausübt als andere Konzepte (non-phonics instruction). In Hinblick auf einen direkten Transfer auf sinnverstehendes Lesen konnten brauchbare Effekte bei schwachen Lesern ( $d = 0.32$ ), jedoch nicht bei älteren Kindern ( $d = .12$ ) beobachtet werden, die bereits über eine Wortlesekompetenz verfügten.

Der Leseflüssigkeit wird inzwischen eine Brückenfunktion zwischen hierarchieniedrigen und hierarchiehöheren Teilkompetenzen zugeschrieben (NICHD, 2000; Rasinski, 2004; Samuels, 2006; Klauda & Guthrie, 2008; Rosebrock, Nix, Rieckmann & Gold, 2013). So konnte Therrien (2004) meta-analytisch ermitteln, dass Maßnahmen zur Verbesserung der Leseflüssigkeit, wenn sie von Lehrern gesteuert werden, einen substanziellen Transfer von  $d = 0.71$  auf das Leseverstehen nach sich ziehen können im Gegensatz zu peergestützten Vorgehensweisen mit einer Effektstärke von  $d = 0.22$ . Darüber hinaus zeigten Rosebrock, Rieckmann, Nix und Gold (2010) bei deutschen Zwölfjährigen in einer Felduntersuchung, dass sich peergestützt durch eine Verbesserung der Leseflüssigkeit mit Hilfe des wiederholten lauten Lesens kurzfristig ein Transfer von  $d = 0.36$  und ca. vier Monate später im Follow-Up von  $d = 0.30$  auf das Leseverstehen erzielen lässt.

Während an Grundschulen im Leseunterricht häufig die soeben angesprochenen Konzepte zum Einsatz kommen, die hierarchieniedrige Prozesse beinhalten, wird das beobachtete eher selten vorkommende explizite Lehren und Einüben von Lesestrategien, also hierarchiehöherer Prozesse, of-

fensichtlich eindeutig als Manko ausgemacht (Bremerich-Vos et al., 2017; Mullis et al., 2017) und schließlich deutlich mit den Defiziten deutscher Kinder in Verbindung gebracht.

In diesem Zusammenhang muss die Frage gestellt werden, ob im deutschsprachigen Raum auf dem Markt befindliche strategieorientierte Unterrichts- und Förderkonzepte tatsächlich die Basis dafür bieten, Lehrkräften ein Instrumentarium an die Hand zu geben, um eine nachhaltige Wende bezüglich der Verbesserung der sinnerfassenden Lesekompetenz bei Kindern im Grundschulalter und darüber hinaus zu ermöglichen.

Betrachtet man nämlich einschlägige Meta-Analysen (Rosenshine & Meister, 1994; Davis, 2010; Mayer & Marks, 2019) sowie neuere breit angelegte Untersuchungen (Droop, van Elsäcker, Voeten, & Verhoeven, 2016; Muijselaar, Swart, Steenbeek-Planting, Droop, Verhoeven & de Jong, 2018), so zeigen sich durchgängig sehr unterschiedliche Effektstärken (zwischen  $d = 0.21$  und  $d = 0.36$ ) und auch keine signifikanten Verbesserungen. Damit ist die empirische Befundlage bei Weitem nicht so eindeutig, wie manchmal der Eindruck erweckt worden ist (vgl. z.B. Gold, 2010).

Das Ziel des vorliegenden Beitrags besteht darin, nach kritischer Durcharbeitung des Forschungsstands die Planung und Realisierung einer Felduntersuchung darzustellen, die mit dem Ziel durchgeführt wurde, die empirische Basis auf diesem Gebiet zu verbreitern und zu eruieren, ob und in welchem Maße sich durch das Einüben hierarchiehöherer Kompetenzen (Textlesestrategien) ein echter Transfer (Klauer, 2011) auf das Leseverstehen in diesem Fall von Viertklässlern erreichen lässt. Dazu wurde als Grundlage auf das im deutschen Sprachraum weit verbreitete Unterrichtsprogramm „Wir werden Textdetektive“ (Gold, Mokhlesgerami, Rühl, Schreblowski & Souvignier, 2006) zurückgegriffen.

## Forschungsstand

Die Befundlage bezüglich der Verbesserung des Textverständnisses mithilfe von Lese-strategie-Trainings im deutschsprachigen Bereich kann auf den ersten Blick insgesamt als vergleichsweise ernüchternd beschrieben werden. Tabelle 1 enthält zusammengefasst Befunde aus empirischen Studien zur Wirksamkeit verschiedener Versionen des Programms „Wir werden Textdetektive“ bezüglich eines Transfers auf die Lesekompetenz.

Wie ersichtlich (vorletzte Spalte) zeigen sich in fünf der sieben Untersuchungen im Vergleich zur konventionell unterrichteten Kontrollgruppe keine statistisch signifikanten Transfereffekte unmittelbar nach dem Ende der Unterrichtsphase.

Eine Ausnahme stellen hier die Ergebnisse von Mokhlesgerami (2004) dar, und zwar in einer Programmvariante, bei der im Gegensatz zum kompletten Programm der Programmteil „Motivationale Selbstregulation“ weggelassen wurde und immerhin ein schwacher Effekt ( $d = 0.19$ ) erzielt werden konnte. Des Weiteren finden sich bei Trenk-Hinterberger (2006) in einer Form des Programms ohne zusätzliche Wiederholungsstunden interpretierbare Verbesserungen der Lesekompetenz ausschließlich bei Sachtexten, interessanterweise jedoch doppelt so starke Verbesserungen in einer Programm-Variante ohne Leistungsmotiv-Förderung im Vergleich zur Variante mit Leistungsmotiv-Förderung ( $d = 0.31$  vs.  $d = 0.14$ ).

Bezüglich der längerfristigen Effekte (Follow-Up, letzte Spalte in Tabelle 1) ca. drei Monate nach Beendigung der Lernphase konnten in drei von fünf relevanten Untersuchungen (Mokhlesgerami, 2004; Trenk-Hinterberger, 2006; Antoniou & Souvignier, 2007) signifikante Verbesserungen bei der Erfolgsvariablen Leseverständnis beobachtet werden. Hierbei zeigten sich bei Mokhlesgerami (2004) mit Gymnasiasten bei dem reinen Strategie-Training mit  $d = 0.52$  deutlich stärkere Effekte als bei dem

Tabelle 1: Befunde aus empirischen Studien zur Wirksamkeit verschiedener Versionen des Programms „Wir werden Textdetektive“ bezüglich eines Transfers auf die Lesekompetenz (verändert nach Gold, 2010, S. 126-127)

Autoren Programm	Schulform (Unterrichts- stunden)	Stichprobe $N_{Ex}/N_{Kon}$ (Anz. Strategien)	d-Werte LV Vorher (T1) – Nachher (T2)	d-Werte LV Vorher (T1) – Follow-Up (T3)
Mokhlesgerami (2004) Textdetektive Vollständig	Gymnasium (17)	241 / 263 (7)	d = 0.18 (n.s.)	d = 0.34*
Textdetektive Reines Strategie-Training	Gymnasium (17)	89 / 263	d = 0.19*	d = 0.52*
Trenk-Hinterberger (2006) Textdetektive ohne Wiederholungsstunden	Gymnasium (20)	352 / 139 (7)	d = 0.14* (Sachtext, mit Leistungsmotiv- Förderung) d = 0.31* (Sachtext, ohne Leistungsmotiv- Förderung)	d = 0.14 (n.s.)
Textdetektive mit Wiederholungsstunden	Gymnasium (20 + 6)	288 / 139		d = 0.53* (Sachtext)
Mokhlesgerami, Souvignier & Gentsch (2006) Textdetektive	Haupt-, Real- und Gesamtschulen (17)	376 / 85 (7)	d = 0.11 (n.s.)	d = 0.04 (n.s.)
Mokhlesgerami, Souvignier, Rühl & Gold (2007) Textdetektive	Gymnasium, Gesamtschule (21)	387 / 230 (7)	d = 0.13 (n.s.)	
Rühl (2006) Textdetektive	Haupt-, Real-, Gesamtschulen (23)	404 / 159 (7)	d = 0.05 (n.s.)	d = 0.01 (n.s.)
Souvignier & Rühl (2005) Lesedetektive	Lernhilfe (25)	78 / 52 (4)	d = 0.29 (n.s.)	
Antoniou & Souvignier (2007) Lesedetektive	Lernhilfe (29)	45 / 28 (4)	d = 0.45 (n.s.)	d = 0.80*

Anmerkungen.  $N_{Ex}$  = N-Trainingsgruppe;  $N_{Kon}$  = N-Kontrollgruppe; d-Werte = Effektstärke d; LV = Leseverstehen; \* =  $p < .05$ ; n.s. = nicht signifikanter Effekt.

vollständig durchgeführten Programm (inkl. motivationale Selbststeuerung) mit  $d = 0.34$ . Trenk-Hinterberger (2006) fand einen interpretierbaren längerfristigen Effekt ( $d = 0.53$ ), jedoch nur auf der Basis von Sachtexten und wenn das Programm um zusätzliche sechs Wiederholungsstunden verlängert wurde (vgl. Tabelle 1, zweite Spalte). In der etwas ausgedünnten Version (Lesedetektive), in der nur vier anstelle von sieben Strategien vermittelt werden (Tabelle 1, dritte Spalte), konnten Antoniou und Souvignier

(2007) bei Kindern mit Lernschwierigkeiten zwar keinen kurzfristigen Lernfortschritt (Tabelle 1, letzte Zeile), dafür aber einen starken längerfristigen Effekt ( $d = 0.80$ ) nachweisen.

Versucht man aus Tabelle 1 ein Befundmuster zu extrahieren, so kann festgestellt werden, dass sich auf der Basis des Unterrichtsprogramms „Wir werden Textdetektive“ im Rahmen des Deutschunterrichts durch Lehrkräfte, die über die praktischen Umsetzungsmöglichkeiten informiert und

mit Stundenabläufen versorgt wurden, per se keineswegs ohne weiteres Verbesserungen der Leseverständnisleistungen einstellen (vgl. auch Trenk-Hinterberger, 2006, S. 212). Auffällig ist, dass sich immer dann signifikante und halbwegs prägnante Verbesserungen erzielen ließen, wenn die Zusatzmodule weggelassen wurden (Mokhlesgerami, 2004), wenn durch Zusatzstunden die Intensität der Vermittlung gesteigert und/oder als Erfolgsvariable das Textgenre Sachtext herangezogen wurde (Trenk-Hinterberger, 2006). Auf dieser Linie liegen auch die Befunde von Antoniou und Souvignier (2007). Diese konnten zeigen, dass durch die „Entfrachtung“ des Programms auf nur vier Strategien (Vorwissen aktivieren, Textschwierigkeiten klären inkl. Wortbedeutungen, Zusammenfassen und Fragen generieren) sowie die gründliche Vermittlung der Strategien mithilfe des reziproken Lernens inkl. selbstständiges Üben selbst Kindern mit Lernschwierigkeiten zu einer starken längerfristigen Verbesserung ( $d = 0.80$ ) der sinnverstehenden Lesekompetenz verholfen werden kann. Bei der Einordnung der Befunde aus Tabelle 1 sollte jedoch beachtet werden, dass die Interventionen Alternativtreatments zu den Kontrollgruppen darstellen, die üblichen Leseunterricht bekamen (vgl. auch Muijselaar et al., 2018).

In einer neueren Meta-Analyse (Zeitraum zwischen 1990 und 2017) zur Frage der Verbesserung des Textverständnisses durch die Vermittlung von Verstehensstrategien im Primar- und Sekundarstufenbereich filterten Mayer & Marks (2019) aus insge-

samt 302 Publikationen schließlich 23 Studien heraus, die den inhaltlichen und methodologischen Ansprüchen genügen. Darunter befanden sich auch die Effektivitätsuntersuchungen zu den „Text-Detektiven“. Als Outcome-Variablen fungierten das Strategiewissen sowie das Leseverständnis, erfasst durch informelle und standardisierte Verfahren.

Wie aus Tabelle 2 zu ersehen ist, zeigt sich hinsichtlich der Verbesserung des Leseverständnisses (T1 / T2) auf der Basis standardisierter Verfahren ein kleiner Effekt (Cohen, 1988) von  $d = 0.21$ . Dabei konnte in neun Studien kein statistisch signifikanter Effekt ermittelt werden. Die Befunde fallen damit schwächer aus als in den Meta-Analysen von Rosenshine und Meister (1994) mit  $d = 0.32$  sowie Davis (2010) mit  $d = 0.31$ , die sich auf reziprokes Lernen in unterschiedlichen Kontexten und Altersgruppen beziehen.

Bei der Langfristauswertung (Tabelle 2) ergibt sich ein  $d = 0.15$ . Bei sieben der 13 Datensätze ließ sich hier kein statistisch signifikanter Effekt nachweisen (Mayer & Marks, 2019, S. 19-22).

Im Rahmen der informellen Erfassung der Verbesserung des Leseverständnisses ergeben sich erwartungsgemäß mittlere bis starke Effekte zwischen  $d = 0.66$  und  $d = 0.94$  (Tabelle 2). Bei einer solchen Vorgehensweise besteht das Prüfverfahren jedoch in der Regel aus Aufgaben, die mehr oder weniger identisch mit denjenigen sind, die auch während des Trainings verwendet wurden, womit deutlich weniger ein Trans-

Tabelle 2: Befunde (Effektstärken  $d$ ) der Meta-Analyse nach Mayer und Marks (2019)

Outcome-Variable	T1 / T2 (d)	N	95% KI	T1 / T3 (d)	N	95% KI
Leseverständnis standardisiert	0.21	24	[0.15, 0.27]	0.15	13	[0.07, 0.23]
Leseverständnis informell	0.66	18	[0.54, 0.77]	0.94	8	[0.76, 1.11]
Strategiewissen	0.51	24	[0.44, 0.57]	0.57	15	[0.48, 0.67]

Anmerkungen. N = Anzahl der Datensätze.

fer auf „normale“ Lesesituationen erfasst wird. Bei der Betrachtung der kurzfristigen Effekte liegen bei den Befunden in neun (von 18) Datensätzen die Konfidenzintervalle nicht im positiven Bereich, bei den Langfriststudien sind dies drei (von acht).

Ebenfalls erwartungsgemäß (Klauer, 2011) stellen sich die Befunde bei der Verbesserung des kurzfristigen Strategiewissens dar ( $d = 0.51$ ). In drei Studien konnte jedoch kein statistisch signifikanter Effekt gesichert werden. Bei den längerfristigen Auswirkungen des Lesetrainings auf das Strategiewissen (T1/T3) zeigten sich mit  $d = 0.57$  ähnlich starke Auswirkungen. Gemäß Tabelle 2 und abweichend von der Original-Publikation (Mayer & Marks, 2019) ergibt sich hier bei der unteren Grenze des Konfidenzintervalls ein positives Vorzeichen (persönliche Mitteilung).

### ***Einflussgrößen auf die Wirksamkeit von Lesestrategie-Trainings***

Bezüglich des Einflusses der Merkmale der Intervention bzw. Implementation lassen sich nach Mayer und Marks (2019, S. 28-20) wegen der geringen Anzahl der jeweiligen Subgruppen nur hypothetische Aussagen treffen. Es ist für die Autoren aber offensichtlich, dass Trainingsvarianten einen Vorteil besitzen, die ausführlich Zeit und Gelegenheit zur Einübung, Anwendung und Überprüfung der angemessenen Realisierung der Strategien zur Verfügung stellen (Gold et al., 2006, S. 79). Diese Aussage ist nicht nur kompatibel zu früheren Ergebnissen der Lesestrategieforschung (Rosenshine, & Meister, 1994; Rosenshine, Meister & Chapman, 1996), sondern auch zu der bereits weiter oben formulierten Schlussfolgerung aus den Befunden in Tabelle 1, dass nämlich Strategietrainings offensichtlich immer dann besonders erfolgreich sind, wenn die inhaltlich ausgerichtete Intensität gesteigert wird (Trenk-Hinterberger, 2006; Antoniou & Souvignier, 2007) und eine gründliche Vermittlung, intensives selbstständiges Üben auch des Transfers auf nicht

behandelte Texte (Klauer, 2011) sowie förderliches fachliches Feedback erfolgen (Schünemann, Spörer, Völlinger & Brunstein, 2017).

Meta-analytische Untersuchungen (Rohrbeck, Ginsburg-Block, Fantuzzo & Miller, 2003), die zur Verbesserung des Erwerbs, der Anwendung und des Transfers von Lesestrategien auf die Technik des reziproken Lehrens (reciprocal teaching) zurückgreifen, scheinen immer dann gewisse Vorteile zu besitzen, wenn die Interventionen bei jungen Grundschulern (Klasse 1 bis 3) aus sozial benachteiligten Familien stattfinden und strukturierte Peer-Interaktionen, Kontingenzmanagement, Hinweiskarten, korrekatives Feedback sowie individuelle Evaluationsprozeduren zum Einsatz kommen. Maßnahmen, bei denen mehr als 50% bestimmter Elemente, nämlich Zielsetzung und Verstärkerauswahl, die eigenständig von den Schülern selber kontrolliert werden, zeigen höhere Effekte als primär lehrerkontrollierte Vorgehensweisen. Die Autoren beklagen jedoch die insgesamt unzureichende Datenlage sowie die in den Einzelarbeiten vorkommende ungenügende Darstellung des konkreten Vorgehens im Detail (Rohrbeck et al., 2003, S. 252).

Diese Befunde werden im deutschsprachigen Raum durch eine Reihe von Untersuchungen auf der Basis von Kleingruppenförderungen mit speziell geschulten Trainern im Wesentlichen bestätigt (Spörer, Seuring, Schünemann & Brunstein, 2008; Spörer, Brunstein & Kieschke, 2009; Seuring & Spörer, 2010; Schünemann, Spörer & Brunstein, 2013; Koch & Spörer, 2016).

Ein weiterer hier hervorzuhebender Einflussfaktor auf die Effekte von strategieorientierten Maßnahmen zur Verbesserung des Textverständnisses bezieht sich auf die Flexibilität der Lehrkräfte sowie die Qualität des Unterrichts. So kommen Duffy et al. (1987) zu der Überzeugung, dass Lehrkräfte durch das Angebot komplett vorgegebener Trainingsprogramme diese eher unflexibel handhaben und an wörtlichen Formulierungen kleben. Mokhlesgerami et al. (2006, S.

85) weisen in diesem Zusammenhang auf das Problem hin, dass in ihrer Untersuchung in fünf von 23 Experimentalklassen (Haupt-, Real- und Gesamtschüler) die Implementation des Programms abgebrochen werden musste, weil die Texte zu schwierig waren und bei der Vermittlung einzelner Strategien Schwierigkeiten auftraten, die bei Gymnasiasten nicht zu beobachten waren. Vorgebeugt werden kann nach Gold (2010, S. 95) solchen enttäuschenden Ergebnissen nur durch eine Anpassung des Schwierigkeitsniveaus der Texte und einer flexiblen Anpassung der Lernmaterialien. Andererseits konnten Hacker und Tenent (2002) auf der Basis einer Videoanalyse von 17 Grundschullehrkräften mehr oder minder große (eher unerwünschte) Abweichungen bezüglich der in Handbüchern vorgegebenen empirisch nachgewiesenen optimalen Vorgehensweise hinsichtlich der Strategie-Vermittlung, der Dialoggestaltung sowie des Scaffolding bei der Vermittlung von Lesestrategien beobachten. Ähnliche Befunde haben Seymour und Osana (2003) zusammengetragen.

Dass das Wissen über Lernstrategien bzw. die Verbesserung des Leseverständnisses der Schüler mit der allgemeinen Unterrichtsqualität (Strukturiertheit des Vorgehens, Motivierungsqualität) sowie der Intensität der Auseinandersetzung der Lehrkräfte mit der Thematik zusammenhängen, zeigen sowohl Souvignier und Mokhlesgerami (2005) als auch Souvignier und Behrmann (2016).

Schließlich sollte nicht unerwähnt bleiben, dass die Auswahl der Textsorte, nämlich Sachtext vs. narrativer Text, die als Erfolgsvariable herangezogen wird, den Fördereffekt von Trainingsmaßnahmen zur Verbesserung des Textverstehens mit beeinflussen kann. In der IGLU 2016 (Hußmann et al., 2017) wurde nämlich festgestellt, dass bei deutschen Grundschulern zwischen den Leistungen bezüglich der beiden Textsorten eine Differenz von signifikanten neun Punkten zugunsten literarischer Texte zu beobachten war. Umgekehrt zeigten sich

bei der Überprüfung der Wirksamkeit des Textdetektive-Programms eher stärkere Effekte auf der Basis von Sachtexten und nicht bei narrativen Texten (Trenk-Hinterberger, 2006).

## Ableitungen für die Planung der Untersuchung

Für die Planung der vorliegenden Untersuchung wurde davon ausgegangen, dass nur eine Verbesserung des Leseverstehens als eine Verbesserung schulischer Leistungen zu werten ist (Lenhard et al., 2012) und nicht primär das Verständnis und die Anwendung von Lesestrategien, die eine vergleichsweise geringe gemeinsame Varianz mit konkreten Verständnisleistungen aufweisen (Rühl, 2006; Spörer, Brunstein & Kieschke, 2009, S. 280; Seuring & Spörer, 2010; Muijselaar et al., 2018). Aus diesem Grund sollte als Erfolgsvariable primär die Verbesserung der Leseleistung auf der Basis unbekannter Texte herangezogen werden.

Im Gegensatz zu den weiter oben geschilderten Arbeiten (vgl. Tabelle 1) und in Anlehnung an die Untersuchungen zum reziproken Lernen (Spörer et al., 2008, 2009; Seuring & Spörer, 2010) sollten in der aktuellen Studie als Trainer speziell instruierte Master-Studierende (Lehramt Sonderpädagogik) zum Einsatz kommen. Diese besaßen sowohl Erfahrung in der Förderung von leseschwachen Schülern, d.h. ein Bewusstsein bezüglich der Notwendigkeit eines adaptiven Unterrichtens hinsichtlich des Schwierigkeitsniveaus der Texte, der zusätzlichen Berücksichtigung von Dekodier-Strategien und der Leseflüssigkeit im Rahmen des Programm-Einsatzes, einer möglichen Interferenz einzelner Strategien (Klauer, 2010) sowie der Notwendigkeit des Erstellens neuer Übungsbögen und Materialien (Gold, 2010, S. 95).

Darüber hinaus wurde von den Trainern erwartet, dass sie allgemeine förderliche Unterrichtsmerkmale realisieren, nämlich



motivationalen Support und Strukturiertheit des Unterrichts (Souvignier & Mokhlesgerami, 2005), Peer-Interaktionen, Einsatz von Kontingenzmanagement, Hinweiskarten, korrekatives Feedback sowie individuelle Evaluationsprozeduren (Schünemann et al., 2013, 2017). Dies sollte jedoch der Vergleichbarkeit mit früheren Untersuchungen wegen nicht streng nach dem Konzept des reziproken Lernens, sondern nach den Vorgaben der Direkten Instruktion (Gold, 2010) angereichert mit Elementen reziproken Lernens umgesetzt werden.

Als schriftliche Unterlagen sollten sowohl die Materialien des Textdetektive-Programms dienen (Gold et al., 2006; Rühl & Souvignier, 2006; Trenk-Hinterberger & Souvignier, 2006) als auch *zusätzlich* Stundenabläufe und Vorgehensweisen, die im Rahmen einer Voruntersuchung (Pein, 2016) nach den in diesem Abschnitt genannten Vorgaben entwickelt wurden.

Aufgrund der aktuellen Befundlage ergeben sich für die vorliegende Untersuchung folgende zwei Forschungsfragen: Können erstens auf der Basis von geschulten Lehrkräften und vor dem Hintergrund der Realisierung der bisher beschriebenen als günstig erkannten Durchführungsbedingungen (siehe auch weiter unten) signifikante kurz- und mittelfristige Fördereffekte bezüglich des Leseverständnisses auf der Basis des Unterrichtsprogramms „Wir werden Textdetektive“ bei Viertklässlern repliziert werden? Speziell aus sonderpädagogischer Sicht wird zweitens gefragt, ob sich praktisch relevante Fördereffekte auch für Schüler dieser Klassenstufe mit unterdurchschnittlicher Kompetenz im Leseverständnis erzielen lassen.

## Methode

### *Stichprobenerhebung, Untersuchungsdesign und Stichprobensbeschreibung*

Die vorliegende Untersuchung basiert auf vier Teiluntersuchungen (TU1 bis TU4), die von vier unterschiedlichen Förderteams durchgeführt wurden. Dementsprechend setzt sich die Gesamtstichprobe aus vier Teilstichproben zusammen. Die Vorgehensweise bei der Stichprobenerhebung war jedoch in jeder TU vergleichbar.

So nahmen z.B. in der TU1 aus zwei Grundschulen jeweils zwei 4. Klassen teil. Alle Kinder (N = 88) wurden mit dem FLVT 5-6 (Souvignier, Trenk-Hinterberger, Adam-Schwebe & Gold, 2008) im Rahmen des Vortests untersucht und dann zu Zwillingspaaren gemäß der Merkmale Leistung, Alter, Geschlecht und Migrationshintergrund zusammengestellt. Da es sich bei der vorliegenden Stichprobe um Viertklässler handelt, die Normen des FLVT 5-6 sich jedoch mindestens auf Fünftklässler beziehen, die ein Jahr länger beschult worden sind (Souvignier et al., 2008, S. 29), war es wichtig zu verhindern, Kinder zu rekrutieren, deren Leistungsvermögen vor dem Hintergrund der FLVT-Normen schwach, jedoch aufgrund ihrer Altersklasse tatsächlich eher überdurchschnittlich ausfällt. Dies wurde realisiert, indem sichergestellt wurde, dass grundsätzlich nur jeweils durchschnittliche und möglichst viele schwache Schüler (gemäß FLVT-Norm) an der Untersuchung teilnahmen unter der Maßgabe, dass der durchschnittliche T-Wert der rekrutierten Gruppe insgesamt kleiner als 45 sein sollte. Diese Vorgehensweise erklärt die teilweise deutliche Heterogenität zwischen und innerhalb der TU. In einem zweiten Schritt wurden dann letztendlich 46 Probanden per Zufall der Experimentalgruppe (EG) und 42 der Kontrollgruppe (KG) zugewiesen (vgl. Tabelle 3). Auf diese Weise wurde ein randomisiertes Matched-Pairs-Design (Bortz & Döhning, 2016) realisiert. Analog wurde so bei allen TU vorgegangen, die dann zur

Tabelle 3: Mittelwerte und Standardabweichungen (Roh- und Gesamt-T-Werte) der Leistungsmerkmale (FLVT 5-6) der Experimental- und Kontrollgruppen (EG und KG) der vier Teiluntersuchungen (TU1 bis TU4) sowie der Gesamtgruppe, Altersangaben in Monaten, Geschlechterverteilung, N der Experimental- und Kontrollgruppe, Anzahl der Förderstunden und Förderwochen (ohne Ferien) in den jeweiligen TU

	EG (FLVT-RW)	KG (FLVT-RW)	Alter (Monate)	N W/M	N EG/KG	Anzahl Förder- stunden	Förder- Wochen
TU1	17.4 (7.4)	17.1 (7.1)	127.4 (6.8)	41/47	46/42	20	8
TU2	8,2 (3.6)	8.0 (3.5)	131.1 (6.9)	10/10	10/10	20	9
TU3	10.2 (4.6)	11.1 (4.5)	127.1 (5.4)	14/14	14/14	27	8
TU4	18.0 (5.5)	17.9 (4.9)	123.2 (6.8)	12/12	12/12	16	8
Gesamt	15.3 (7.3)	15.0 (6.9)	127.2 (6.8)	77/83	82/78	20.8	8.25
T-Wert	42.3 (8.9)	42.1 (8.3)					

Gesamtstichprobe zusammengezogen wurden. Dropouts ergaben sich wegen längerer Krankheit einzelner Kinder.

Während bei der TU1 das Gruppen-Auswahlkriterium ( $T < 45$ ) auf der Basis aller getesteten  $N = 88$  Schüler eingehalten werden konnte (vgl. Tabelle 3), wurden in TU2 insgesamt  $N = 43$  Kinder aus zwei 4. Klassen einer Schule vorgetestet, von denen dann insgesamt  $N = 20$  an der Untersuchung teilnahmen (vgl. Tabelle 3). In TU3 wurde der Gruppen-Grenzwert ( $T < 45$ ) zur Teilnahme vor dem Hintergrund von  $N = 28$  der insgesamt  $N = 55$  Kinder an zwei Grundschulen realisiert. Dies traf in TU4 in einer Grundschule bei  $N = 24$  von insgesamt  $N = 44$  untersuchten Schülerinnen und Schülern zu.

Insgesamt wurden damit in sechs Grundschulen in Schleswig-Holstein insgesamt  $N = 230$  Kinder vorgetestet, von denen dann  $N = 160$  Schüler unmittelbar in die Untersuchung einbezogen wurden (EG und KG, vgl. Tabelle 3). Die Viertklässler wurden pro Schule jeweils als EG und KG aufgeteilt. Im Ganzen befanden sich damit  $N = 82$  Kinder in der EG und  $N = 78$  Kinder in der KG.

Wie aus Tabelle 3 hervorgeht, zeigen sich innerhalb der jeweiligen TU schon rein numerisch kaum Unterschiede zwischen den FLVT-Werten der jeweiligen Experi-

mental- und Kontrollgruppen. Eine inferenzstatistische Überprüfung des Gesamt-Mittelwertunterschieds sowohl bezüglich der abhängigen Variable Leseverständnis ( $t = 0.88$ ;  $df = 158$ ;  $p > 0.05$ ) als auch des Alters in Monaten ( $t = 1.04$ ;  $df = 158$ ;  $p > 0.05$ ) der Schüler erweist sich als nicht signifikant.

Wie aus Tabelle 3 ebenfalls hervorgeht (vorletzte Zeile), nahmen an der Untersuchung insgesamt 77 Mädchen und 83 Jungen teil, deren Verteilung über die Experimental- und Kontrollgruppe gleich ist (Fischer's exakter Test). Dasselbe gilt für die in Tabelle 3 nicht dargestellten Häufigkeiten der Merkmale Migrationshintergrund (EG:  $N = 22$ ; KG:  $N = 21$ ) und sonderpädagogischer Förderbedarf (EG:  $N = 8$ ; KG:  $N = 4$ ).

Zur expliziten Beantwortung der zweiten Forschungsfrage, ob sich praktisch relevante Fördereffekte auch für Schüler mit deutlich unterdurchschnittlicher Kompetenz im Leseverständnis erzielen lassen, musste sichergestellt werden, dass es sich hier um ausnahmslos schwach lesende Kinder handelt. Aus diesem Grund wurde die Gesamtstichprobe bezüglich der FLVT-Gesamttestwerte im Vortest medianhalbiert. Als Ergebnis zeigt sich in der so gewonnenen Teilstichprobe ein Ausgangswert (FLVT-Rohwert) von  $M = 9.20$  ( $SD = 3.7$ ) für die Experimental- und  $M = 9.03$  ( $SD = 3.7$ )

für die Kontrollgruppe (N = 37 bzw. N = 41). Dies entspricht gemäß der FLVT-Norm für Fünftklässler (Gesamtstichprobe) einem  $T = 34.5$ , was bedeutet, dass es sich bei diesen Viertklässlern in beiden Gruppen ausschließlich um sehr schwache Leser handelt (Min.:  $T = 19$ ; Max.:  $T = 41.5$ ).

### **Messinstrument**

Zur Überprüfung der Leseverständnisleistung der Kinder wurde, wie bereits erwähnt, der Frankfurter Leseverständnistest für 5. und 6. Klassen (FLVT 5-6) (Souvignier, Trenk-Hinterberger, Adam-Schwebe & Gold, 2008) herangezogen. Dieses Verfahren, das in zwei parallelen Versionen vorliegt, fand auch im Rahmen der einschlägigen Untersuchungen zur Verbesserung des Leseverständnisses (vgl. Tabelle 1) seinen Einsatz. Aus diesem Grund wurde der Test auch hier benutzt. Des Weiteren besitzt das Verfahren den Vorteil, dass neben dem Gesamt-Testwert das Leseverständnis differenziert nach Textsorten (narrative Texte und Sachtexte) ermittelt werden kann. Die interne Konsistenz (Cronbachs Alpha) wird mit  $\alpha = .88$  (Form A) sowie  $\alpha = .86$  (Form B) angegeben. Die (minderungskorrigierte) Paralleltestreliabilität der beiden Testformen liegt bei  $r_{tt} = .80$ . Bei der Verrechnung der Daten für die vorliegenden Viertklässler wurden die Rohwerte zugrunde gelegt.

### **Vorgehensweise**

Da die Diagnostikphasen, vor allem aber die Förderung selbst, von Art und Umfang einen nicht unerheblichen Eingriff in den Schulalltag der teilnehmenden Klassen darstellten, mussten bei der Durchführung in den einzelnen Schulen Kompromisse geschlossen werden, die sich allgemein auf den zeitlichen Ablauf, die Anzahl der Förderstunden, die wöchentliche Förderfrequenz sowie die Art der Behandlung der Kontrollgruppe auswirkten.

Im Einzelnen lassen sich die Gemeinsamkeiten und Unterschiede bezüglich der

Vorgehensweise an den einzelnen Schulstandorten wie folgt beschreiben. Die Eingangsdagnostik (Vortests) wurde in allen TU innerhalb der ersten drei Monate des Jahres 2019 durchgeführt. Die Interventionen begannen teils kurz vor oder direkt nach den Osterferien. Eine Ausnahme stellt hier TU4 dar, bei der die Intervention und der Nachttest (NT) bereits vor den Osterferien stattfanden.

Alle Nachttests wurden direkt nach Ende des Förderzeitraums realisiert. Die Follow-Up-Untersuchungen (FU) der TU1 und TU2 fanden ca. vier bis fünf Wochen nach dem Nachttest statt; bei der TU3 waren es zwei Wochen und bei der TU4 zehn Wochen. Diese Unterschiede machen eine statistische Kontrolle bei der Berechnung des Treatment-Effekts notwendig.

Wie aus Tabelle 3 hervorgeht, fanden in TU1 und TU2 je 20 Fördersitzungen (Schulstunden) statt, in TU3 waren es 27 und in TU4 mit 16 am wenigsten. Auch die Effekte dieser Unterschiede sollten statistisch kontrolliert werden. Die Sitzungen erstreckten sich auf einen bezüglich der verschiedenen TU fast identischen reinen Förderzeitraum zwischen acht und neun Wochen (vgl. Tabelle 3). Die zeitliche Verteilung der Intervention innerhalb einer Woche schwankte je nach TU zwischen zwei und vier Sitzungen.

Die Anzahl der Schülerinnen und Schüler pro Fördergruppe betrug je nach TU zwischen sieben und 12 Kinder. Eine Ausnahme bildete TU1, in der eine der drei Experimentalgruppen eine Stärke von N = 23 aufwies.

Die geplante Kontrolle eines Zuwendungseffekts per Zuwendungskontrollgruppe fand ausnahmslos in allen TU statt und wurde im Kontrast zur EG je nach Möglichkeit und Bedarf vor Ort in TU1 als Rechtschreibtraining (mit identischer Stundenzahl), in TU2 als Mathematikunterricht (eine Stunde wöchentlich) und in TU3 und TU4 als Unterstützung der jeweiligen Lehrkraft im Sachunterricht (jeweils zwei Stunden wöchentlich) realisiert.

Schließlich ist zu bemerken, dass eine zielgenaue Programm-Implementation in den regulären Deutschunterricht aufgrund der Rahmenbedingungen in den einzelnen Schulen (Einbettung in den Stundenplan, Zuständigkeit der Lehrkräfte) kaum möglich war, so dass die Förderung in der Regel eher einem Zusatzprogramm (Kompaktkurs) für die Schüler entsprach.

### **Realisierte Förderprinzipien und Inhalte**

In dem Programm „Wir werden Textdetektive“ werden sieben Detektivmethoden (Strategien und Techniken) in 14 Lerneinheiten angeboten, und zwar die Organisationsstrategien „*Wichtiges unterstreichen*“ und „*Wichtiges zusammenfassen*“, die elaborativen Strategien „*Überschrift beachten*“ und „*Bildlich vorstellen*“ sowie die metakognitiven Strategien „*Umgang mit Textschwierigkeiten*“, „*Verstehen überprüfen*“ und „*Behalten überprüfen*“. Hinzu kommen noch zwei Bausteine, die das Ziel haben, bei den Schülern die motivationale und kognitive Selbstregulation zu verbessern (Gold et al., 2006, S. 7).

Die im Rahmen der Stichprobenerhebung und -beschreibung genannten TU wurden von Master-Studierenden der Sonderpädagogik mit Ausnahme von TU2 als Zweierteam durchgeführt. Vor Beginn der Förderung wurden sie sowohl in die Original-Programminhalte eingearbeitet als auch explizit auf die durch die Autoren der „Textdetektive“ angeregte mögliche Modifikation bestimmter Inhalte sowie auf die Einhaltung bestimmter Vermittlungsprinzipien hin instruiert (Gold et al., 2006; Rühl & Souvignier, 2006; Trenk-Hinterberger & Souvignier, 2006).

Des Weiteren wurde auf eine Unterrichtsplanung von Pein (2016) zurückgegriffen, die Adaptationen des Programms bereits praktisch erprobt und umgesetzt sowie Schwierigkeiten bei der Durchführung aufgedeckt hat. Diese bestanden u.a. darin, dass sehr schwache Leser auf jeden Fall Hil-

fen zur Verbesserung des Wortschatzes, der Dekodierfähigkeit und der Leseflüssigkeit benötigten, um den Anforderungen des Programms gerecht zu werden. Dies geschah z.B. durch mehrfaches Lesen der durch das Programm zur Verfügung gestellten Texte sowie individuelles Feedback.

Es wurde weiterhin großer Wert darauf gelegt, dass den Kindern möglichst viel Zeit zum Üben und Festigen der Detektivmethoden zur Verfügung gestellt werden sollte (Gold et al., 2006, S. 8). Aus diesem Grund wurde in der Regel der separate Baustein zur Motivationsförderung weggelassen (Moklesgerami, 2004). Dessen Inhalte wurden jedoch direkt im Vermittlungsprozess umgesetzt.

Eine weitere Veränderung im Vergleich zum Basisprogramm betrifft die explizite Behandlung der Unterschiede zwischen narrativen Texten und Sachtexten, die im Zusammenhang mit dem Zusammenfassen von Texten ausführlich thematisiert wurden.

Etwas systematischer dargestellt orientierten sich die Durchführenden an den in Walter (2007) dargestellten neun Handlungsanweisungen, die sich nach Rosenshine et al. (1996) als die Effektivität von Förderprogrammen steigernden Bestandteilen aus 23 erfolgreichen Studien zur Verbesserung der sinnverstehenden Lesens herausfiltern lassen:

- (1) *Liefere aufgabenspezifische verbale Lösungsprozeduren*: Bei einigen Strategien wurden in der vorliegenden Untersuchung die verbalen Lösungsprozeduren auch als fertiges Fragegerüst (DM 3 – Umgang mit Textschwierigkeiten und DM 4 – Verstehen überprüfen) auf den Arbeitsbögen vorgegeben.
- (2) *Liefere Modellverhalten für angemessenes Aufgabenlösen*: Wurde eine Strategie neu eingeführt oder wiederholt, so verbalisierte die Lehrerin die kognitiven Denkopoperationen laut, so dass sich bei den einzelnen Detektivmethoden jeweils eine *strategiespezifische* Instruktion ergab. Dadurch wurde den Kindern die Möglichkeit eröffnet, die Angemes-

senheit ihres Strategieverhaltens im Sinne eines Soll-Ist-Vergleichs abzugleichen.

- (3) *Antizipiere mögliche Schwierigkeiten und (4) Verändere die Schwierigkeit der Übungstexte:* Aufgrund der Durchführung des Programms in einer Grundschulklasse war absehbar, dass einige Schüler Schwierigkeiten bei der Bearbeitung der vorgegebenen Texte haben würden. Aus diesem Grund wurden einige Texte im Voraus gekürzt oder verändert. So wurde zudem je nach Entwicklungsstand der Kinder die Methode „Behalten überprüfen“ aufgrund der hohen Komplexität nicht behandelt. Es wurde in Teilen der Text des Märchens „Der Froschkönig“ hinzugenommen, um die anspruchsvolle Strategie „Verstehen überprüfen“ an einem bereits bekannten Text einführen zu können. Eine weitere Vorgehensweise war die, dass beispielsweise zuerst der Text „Der Hirtenjunge“ aus dem Förderschulmanual und nachfolgend der Text „Achmed der Narr“ aus dem Regelschulmanual eingesetzt wurde, um die Textschwierigkeit zu steigern.
- (5) *Liefere eine Hinweiskarte (cue card):* Diese existieren im Textdetektive-Manual und wurden auch in dieser Untersuchung entsprechend angewandt.
- (6) *Leite die Schüler aktiv an:* Im Rahmen der vorliegenden Untersuchung wurde hauptsächlich ein lehrerzentriertes Üben realisiert und durch Kleingruppenarbeit ergänzt. Durch kognitive Aktivitäten wie Zusammenfassen, Erinnern, Verbessern und Aufzeigen von Sachverhalten wurde der Lernprozess der Kinder aktiv und modellierend von der Lehrkraft stimuliert.
- (7) *Liefere Feedback und Korrekturen:* Rückmeldungen bezüglich der Qualität des Gelernten bekamen die Schüler sowohl durch einen Selbstüberprüfungstest (Förderschulmanual) als auch durch permanentes Feedback der Förderlehrkräfte. Des Weiteren wurde den Kindern

z.T. die Möglichkeit gegeben, ihre Lösungsansätze untereinander zu vergleichen (Tafel, Rundgang).

- (8) *Liefere eine Prüfliste:* Der Zweck einer Prüfliste (checklist) besteht darin, den Schülern Anhaltspunkte zur Überwachung (monitoring) des Lösungsprozesses anzubieten. Dabei werden die Schüler angehalten, gegen Ende der Lösungsprozedur zu fragen: „Was habe ich noch nicht verstanden?“ oder „Habe ich die richtigen Fragen gestellt?“ oder „Habe ich wirklich das Wichtigste verstanden?“ Diese Vorgehensweise ist bereits in dem standardisierten Textdetektive-Manual berücksichtigt und dementsprechend auch in dieser Untersuchung umgesetzt worden.
- (9) *Überprüfe, ob die Schüler die Strategie wirklich beherrschen:* Durch das Lesespiel, das im Regelschulmanual der Textdetektive ein Baustein ist, wurde dies umgesetzt. Alternativ konnte aber auch aus Zeitgründen so vorgegangen werden, dass der Fokus weniger stark auf das Lesespiel, sondern sehr stark auf das Einüben der Strategien gelegt wurde und dafür während der Arbeitsphasen und der Aufgabenbesprechungen explizit darauf geachtet wurde, ob die Detektiv-Methoden korrekt Anwendung fanden.

Vor dem Hintergrund der Überlegungen von Klauer (2011), dass die Effektivität des Programms dadurch gesteigert werden könnte, dass mehr Zeit für das Erlernen und Automatisieren der Detektivstrategien zu besseren Erfolgen führen sollte, wurden folgende Modifikationen bzw. Schwerpunktverschiebungen des Programms vorgenommen:

- *Weglassen des Ringwurfspiels.* Abweichend vom Regelschulmanual wurde in der vorliegenden Umsetzung des Trainingsprogramms das Ringwurfspiel nicht oder nur zum Teil realisiert. Die dort angestrebten Ziele wurden im Zusammenhang mit dem Buchstabenspiel

und durch die Thematisierung motivationsförderlicher Kausalattributionen im Zusammenhang mit dem Erlernen der einzelnen Detektiv-Strategien *aufgabenspezifisch* umgesetzt.

- *Konsequenter Einsatz von Kontingenzmanagement zur Verbesserung des Arbeitsverhaltens und der Motivation.* Je nach Bedarf wurden nach dem Prinzip des Response-Cost gezielt konstruktivere Verhaltensweisen, z.B. leises Sprechen bei Wortmeldungen, Unterlassen von unaufgeforderten Gesprächen mit dem Nachbarn, Unterlassen respektlosen Verhaltens gegenüber der Lehrperson angebahnt. Als Token dienten der Rahmenhandlung entsprechend z.B. Detektivlupen. Das Kontingenzmanagement wurde aber auch eingesetzt, wenn Schüler selbstgesetzte Lernziele im Zusammenhang mit der Beschäftigung mit dem Trainingsmaterial erreichten.
- *Modifikation von Arbeitsblättern.* Bei der Durchführung des Programms zeichneten sich bei den Schülern große Schwierigkeiten während des Ausfüllens der Arbeitsblätter aus dem Regelschulmanual ab. Einerseits wurde auf den Arbeitsblättern zu wenig Platz für die kindliche Schrift gelassen und andererseits wurde bei komplexen Aufgabenstellungen freies Formulieren verlangt, welches von den Grundschulern häufig nicht realisiert werden konnte. Deshalb wurde während der Vermittlung der Trainingsinhalte mehrmals auf Arbeitsblätter aus dem Förderschulmanual zurückgegriffen. Außerdem wurden neue Arbeitsblätter aus den Vorlagen der Regel- (bzw. Förderschulmanualen) konzipiert, welche den Bedürfnissen der Grundschul Kinder eher gerecht wurden. Des Weiteren wurden die Bilder der Strategiekarten teilweise neugestaltet und auf den jeweiligen Arbeitsblättern abgebildet, um einen stärkeren Wiedererkennungseffekt herbeizuführen.
- *Besonderheiten bei der Strategie „bildlich Vorstellen“.* Diese Strategie wurde je nach Interessen und Möglichkeiten der Schüler unterschiedlich umgesetzt. So wurde einerseits zur Durchführung eine optionale Traumreise mit den Kindern durchgeführt. Es wurde eine Geschichte vorgelesen, wozu Musik abgespielt wurde. Es stellte sich heraus, dass die Kinder diese kurze Entspannungseinheit sehr gut annahmen. Eine weitere Option bestand darin, dass die Schüler ihre Vorstellungen zu dem Text malten. Das Aufzeichnen von Narrativtexten erfolgte dabei nicht im Darstellen einer einzelnen Abbildung, sondern die Schüler malten mehrere kleine Bilder, ähnlich wie bei einem ablaufenden Film. Dazu wurde eine neue Hinweiskarte entwickelt.
- *Modifikationen bei den Strategien „Wichtiges unterstreichen“, „Wichtiges zusammenfassen“, „Verstehen überprüfen“.* Erwartungsgemäß hat eine Reihe von Schülern Schwierigkeiten, zwischen wichtigen und unwichtigen Informationen im Text zu unterscheiden (Mokhlesgerami, 2004). Komplette Textpassagen bis hin zum ganzen Text werden von den Schülerinnen und Schülern während des Übens und Bearbeitens unterstrichen. Dies hat Konsequenzen für das Zusammenfassen und das Überprüfen des Verstehens. Aus diesem Grund wurde sehr stark die effektive Strategie des Stellens von W-Fragen (vgl. Walter, 2007) in den Fokus gerückt, die nicht nur für die Überprüfung des eigenen Textverständnisses wichtig ist (Gold et al., 2006, S. 43), sondern auch für die Schüler eine Technik darstellen kann, überhaupt Wichtiges in einem Text zu identifizieren. Aus diesem Grund wurde mithilfe eines im Plenum erarbeiteten Frageleitfadens die Unterscheidung zwischen wichtigen und unwichtigen Informationen erarbeitet. Der erarbeitete Leitfaden enthielt transferierbare Fragen, durch deren Beantwortung eine Hilfestellung zur erfolgreichen Zusammenfassung von narrativen Texten sowie

Sachtexten geschaffen wurde (Scaffolding).

### Verrechnung der Daten

Zur inferenzstatistischen Überprüfung des Interventionseffektes (Treatment-Effekts) zum Nachtest- bzw. Follow-Up-Zeitpunkt wurden hierarchische Regressionsanalysen gerechnet, und zwar sowohl mit den FLVT-Gesamt-Rohwerten als auch getrennt nach den Textsorten (Sachtext, narrativer Text) jeweils als abhängige Variablen. Da davon auszugehen ist, dass der Treatment-Effekt möglicherweise mit Variablen wie Geschlecht, Migrations- und Förderstatus (Sonderpädagogischer Förderbedarf) eines Schülers kovariiert (Hußmann et al., 2017), wurden diese Variablen der untersten Hierarchie-Ebene (Schritt 1) zugewiesen. Wie aus Tabelle 3 hervorgeht, unterscheiden sich Experimental- und Kontrollgruppe in ihrer Gänze im Vortest zwar nicht signifikant voneinander, jedoch zeigt sich eine deutliche Heterogenität im Vorwissen (FLVT-Vortest), so dass es Sinn macht, deren Einfluss auf der zweiten Hierarchie-Ebene zu kontrollieren (Schritt 2). Des Weiteren

ergeben sich in den TU Unterschiede bezüglich einiger Merkmale der Förderung, die in Schritt 3 zusammengefasst und kontrolliert werden (Schritt 3: Förderdauer bzw. Follow-Up-Zeitpunkt). Der hier vor allem interessierende Treatment-Effekt (Schritt 4) sollte dann Auskunft darüber geben, ob und in welchem Maße unter Kontrolle aller übrigen Variablen die Gruppenzugehörigkeit der Kinder zur Varianzaufklärung beiträgt.

### Ergebnisse

In Abbildung 1 findet sich die deskriptive Visualisierung der Befunde als Entwicklungsverlauf von Experimental- und Kontrollgruppe inkl. Angabe der Mittelwerte über die drei Messzeitpunkte auf der Basis der FLVT-Rohwerte im Gesamttest. Die dazu gehörigen SD liegen bei 7.3, 6.3 und 5.9 bzw. 7.0, 6.8 und 6.7 für die Kontrollgruppe.

Wie ersichtlich starten beide Gruppen mit fast identischen Ausgangswerten im Vortest. Eine Veränderung von ca. drei Punkten zugunsten der Experimentalgruppe

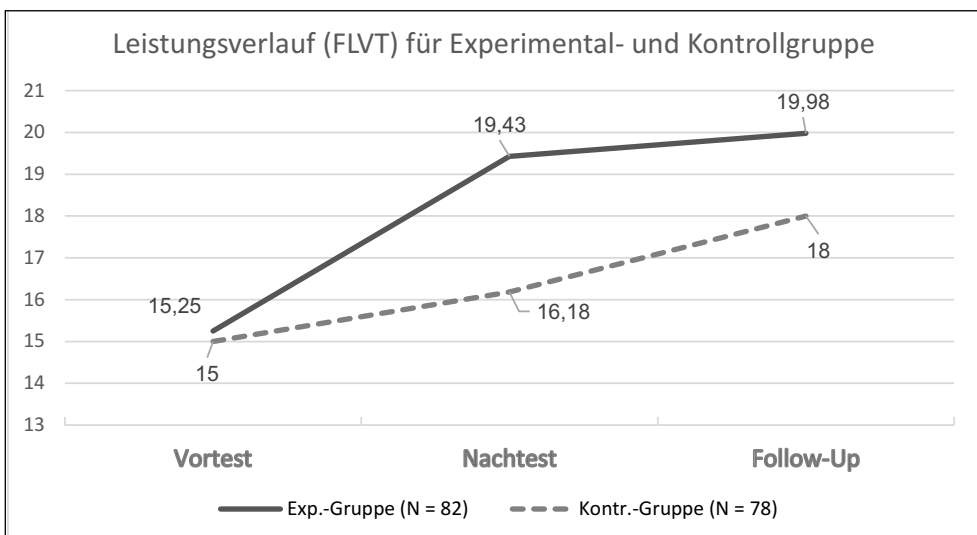


Abbildung 1. Visualisierung der FLVT-Gesamtwerte (Rohwerte) über drei Messzeitpunkte.

ergibt sich dann im Nachtest. Bei insgesamt leicht ansteigenden Werten zeigt sich dann im Follow-Up eine etwas geringere Differenz von knapp zwei Punkten ebenfalls zugunsten der Trainingsgruppe. Auffällig ist, dass sich auch die Kontrollgruppe im Nachtest und Follow-Up steigert. Ob und in welchem Maße der Leistungsverlauf der Experimentalgruppe im Kontrast zur Kontrollgruppe als Treatment-Effekt interpretiert werden darf, wird im Folgenden durch die inferenzstatistische Analyse der Daten ermittelt.

Tabelle 4 enthält die Befunde der hierarchischen Regressionsanalyse, die Aufschluss über den Einfluss möglicher intervenierender Variablen auf die hier zunächst

betrachtete abhängige Variable (FLVT) unmittelbar nach der Intervention gibt.

In der ersten Tabellenzeile sind als wesentliche Parameter das nicht standardisierte Regressionsgewicht B mit seinem Standardfehler SE(B), das standardisierte Beta-Gewicht ( $\beta$ ) zusammen mit seiner Irrtumswahrscheinlichkeit ( $p$ ) sowie die jeweils durch den entsprechenden Schritt (= Modell) aufgeklärten Varianzanteile in ihrer Summe ( $\Sigma R^2$ ), in ihrer jeweiligen Veränderung ( $\Delta R^2$ ) sowie bezüglich der Signifikanz ihrer Veränderung ( $p(\Delta)$ ) aufgeführt.

Wie aus Tabelle 4 zu ersehen ist (zweite Zeile, letzte Spalte,  $p(\Delta)$ ), bilden die Variablen Geschlecht, Migrationsstatus und För-

Tabelle 4: Hierarchische Regressionsanalyse zum Einfluss von Umgebungsvariablen (Schritt 1: Geschlecht, Migration, Förderbedarf), des Vorwissens (Schritt 2), Merkmalen der Förderung (Schritt 3: Förderdauer) und der Gruppenzugehörigkeit (Schritt 4: Treatment) auf die abhängige Variable sinnerfassendes Lesen (FLVT-Rohwert) unmittelbar nach der Intervention

	B	SE (B)	$\beta$	p	$\Sigma R^2$	$\Delta R^2$	p ( $\Delta$ )
<b>Schritt 1</b>					.080	.080	.005*
Geschlecht	-0.15	1.04	-.11	.883			
Migration	-2.19	1.17	-.14	.063			
Förderbedarf	-6.24	1.96	-.25	.002*			
<b>Schritt 2</b>					.553	.473	.000*
Geschlecht	0.54	0.73	.04	.461			
Migration	-0.24	0.83	-.02	.773			
Förderbedarf	0.50	1.47	.02	.732			
Vorwissen	0.71	0.06	.75	.000*			
<b>Schritt 3</b>					.555	.001	.513
Geschlecht	0.56	0.73	.04	.446			
Migration	-0.10	0.86	-.01	.904			
Förderbedarf	0.52	1.47	.02	.727			
Vorwissen	0.70	0.06	.74	.000*			
Förderdauer	-0.08	0.12	-.04	.513			
<b>Schritt 4</b>					.606	.052	.000*
Geschlecht	0.58	0.69	.04	.401			
Migration	-0.18	0.81	-.01	.829			
Förderbedarf	-0.08	1.40	-.00	.956			
Vorwissen	0.69	0.05	.73	.000*			
Förderdauer	-0.08	0.11	-.04	.507			
<b>Treatment</b>	<b>3.06</b>	<b>0.69</b>	<b>.23</b>	<b>.000*</b>			

Anmerkungen. Abhängige Variable = FLVT-Nachtest; Vorwissen = FLVT-Vortest; Förderdauer = Anzahl der Fördersitzungen;  $\Sigma$  = Summierung;  $\Delta$  = Veränderung; \* = hoch signifikanter Effekt ( $p < .01$ ). Kodierung Geschlecht: 0 (männlich) bzw. 1 (weiblich); Migration: 0 (nein) bzw. 1 (ja); Förderbedarf: 0 (nein) bzw. 1 (ja); Treatment: 2 (Experimental-Gruppe) bzw. 1 (Kontrollgruppe).



derbedarf in ihrer Gesamtheit einen hoch signifikanten Effekt auf die abhängige Variable ab (Schritt 1;  $p(\Delta) = .005$ ), der im Wesentlichen dem nachvollziehbaren Einfluss des Umstands zuzuschreiben ist, ob ein Kind einen offiziellen Förderstatus besitzt oder nicht (fünfte Zeile,  $\beta = -.25$ ;  $p = .002$ ). Der Wert  $B = -6.24$  zeigt in diesem Zusammenhang an, dass Kinder mit Förderbedarf (bei Konstanthaltung von Geschlecht und Migrationsstatus) am Ende der Förderung einen im Schnitt um 6.24 Punkte niedrigeren Rohwert im FLVT aufweisen als Kinder ohne Förderbedarf.

Das zusätzliche Einfügen des Vorwissens durch die Erweiterung des Modells (Schritt 2) signalisiert zusätzlich eine signifikante und prägnante Kovariation ( $\Delta R^2 = .473$ ;  $p(\Delta) = .000$ ) desselben mit der abhängigen Variablen und führt dazu, dass der Einfluss des Förderbedarfs als Einzelvariable verschwindet ( $\beta = .02$ ;  $p = .732$ ).

Das Erweitern des Modells um die Variable Förderdauer (Schritt 3) hat keinerlei zusätzliche Varianzaufklärung ( $p(\Delta) = .513$ ) zur Folge.

Das abermalige Erweitern des Modells (Schritt 4) um die hier zu allererst interessierende Einflussgröße, nämlich die der Zugehörigkeit zur Experimental- oder Kontrollgruppe (Treatment, fett), bewirkt unter Konstanthaltung aller anderen hier erhobenen Kontrollvariablen (Geschlecht, Migrationsstatus, Förderbedarf, Vorwissen) einen *zusätzlichen* signifikanten varianzaufklärenden Effekt (Schritt 4;  $\Delta R^2 = .052$ ;  $p(\Delta) = .000$ ) von 5.2%. Das nicht standardisierte Regressionsgewicht von  $B = 3.06$  für die Einflussgröße Treatment signalisiert, dass die Kinder der Trainingsgruppe (Kodierung: 1 = KG; 2 = EG) unter Konstanthaltung (Eliminierung) des Einflusses der anderen Variablen im Schnitt um 3.06 Punkte besser abschneiden als die Kinder der Kontrollgruppe (Tabelle 4, letzte Zeile,  $p = .000$ ). Insgesamt klärt das Gesamtmodell (Schritt 4) damit 60.6 % der Unterschiede im Nachtest auf ( $\Sigma R^2 = 0.606$ ), was insgesamt zu einem höchst signifikant von Null verschiedenen

Gesamt-Modell führt (ANOVA:  $F = 38.98$ ;  $df = 6/152$ ;  $p = .000$ ).

Auf der Basis der standardisierten  $\beta$ -Gewichte (Tabelle 4, vierte Spalte) lassen sich die relativen Gewichte der Prädiktorvariablen direkt miteinander vergleichen. So wird sichtbar, dass das Vorwissen mit  $\beta = .73$  ( $p = .000$ ) einen ca. dreimal so starken Einfluss auf die Nachtestergebnisse ausübt wie das Treatment mit  $\beta = .23$  ( $p = .000$ ).

Tabelle 5 enthält Befunde der hierarchischen Regressionsanalyse mit den FLVT-Follow-Up-Werten als abhängige Variable.

Bis auf die Tatsache, dass neben dem Förderbedarfs-Status auch noch der Migrations-Status im Modell 1 (= Schritt 1) als jeweilige Einzelgrößen zu signifikant niedrigeren Werten bei den Kindern führen, ergibt sich bei Schritt 2 (Tabelle 5) das gleiche Bild wie die Datenanalyse im Zusammenhang mit den Nachtestwerten in Tabelle 4: Bei Einführung des starken Faktors Vorwissen, d.h. Vortestwerte, ( $\beta = .71$ ;  $p = .000$ ) verschwindet der Einfluss der Umgebungsvariablen.

Die Merkmale der Förderung, also Förderdauer und FU-Abstand (zeitlicher Abstand des Follow-Up-Tests zum Nachtest in Wochen), kovariieren als ganze dann in Schritt 3 bei der Analyse der Follow-Up-Werte mit der abhängigen Variable ( $\Delta R^2 = .036$ ;  $p(\Delta) = .001$ ). Dies kommt dadurch zustande, dass sowohl die Förderdauer ( $\beta = -.23$ ;  $p = .007$ ) als auch der FU-Abstand ( $\beta = -.31$ ;  $p = .000$ ) negativ mit den Follow-Up-Werten korrelieren.

Schließlich leistet die hier vor allem interessierende Treatment-Variable (fett) in Schritt 4 zusätzlich einen signifikanten Beitrag zur Varianzaufklärung ( $\beta = .15$ ;  $\Delta R^2 = .023$ ;  $p(\Delta) = .003$ ), wobei die Beiträge des Vorwissens ( $\beta = .71$ ;  $p = .000$ ), der Förderdauer ( $\beta = -.23$ ;  $p = .006$ ) und des FU-Abstands ( $\beta = -.31$ ;  $p = .000$ ) signifikant bleiben.

Insgesamt klärt das Gesamtmodell damit 62.1 % der Unterschiede zum Follow-Up-Zeitpunkt auf (Schritt 4:  $\Sigma R^2 = 0.621$ ), was insgesamt zu einem höchst signifikant

Tabelle 5: Hierarchische Regressionsanalyse zum Einfluss von Umgebungsvariablen (Schritt 1: Geschlecht, Migration, Förderbedarf), des Vorwissens (Schritt 2), Merkmalen der Förderung (Schritt 3: Förderdauer und FU-Zeitpunkt) und der Gruppenzugehörigkeit (Schritt 4: Treatment) auf die abhängige Variable sinnerfassendes Lesen im FLVT-Follow-Up-Test

	<b>B</b>	<b>SE (B)</b>	<b>β</b>	<b>p</b>	<b>ΣR<sup>2</sup></b>	<b>ΔR<sup>2</sup></b>	<b>p (Δ)</b>
<b>Schritt 1</b>					.143	.143	.000*
Geschlecht	-0.87	0.95	-.07	.361			
Migration	-3.15	1.06	-.22	.004*			
Förderbedarf	-7.42	1.79	-.31	.000*			
<b>Schritt 2</b>					.562	.419	.000*
Geschlecht	-0.15	0.68	-.01	.831			
Migration	-1.32	0.78	-.09	.092			
Förderbedarf	-1.29	1.38	-.05	.353			
Vorwissen	0.63	0.05	.71	.000*			
<b>Schritt 3</b>					.598	.036	.001*
Geschlecht	-0.39	0.66	-.03	.554			
Migration	-1.28	0.77	-.09	.097			
Förderbedarf	-1.63	1.34	-.07	.223			
Vorwissen	0.64	0.05	.72	.000*			
Förderdauer	-0.45	0.17	-.23	.007*			
FU-Abstand	-0.81	0.22	-.31	.000*			
<b>Schritt 4</b>					.621	.023	.003*
Geschlecht	-0.38	0.64	-.03	.555			
Migration	-1.32	0.75	-.09	.081			
Förderbedarf	-2.04	1.31	-.09	.122			
Vorwissen	0.64	0.05	.71	.000*			
Förderdauer	-0.45	0.16	-.23	.006*			
FU-Abstand	-0.80	0.21	-.31	.000*			
<b>Treatment</b>	<b>1.91</b>	<b>0.64</b>	<b>.15</b>	<b>.003*</b>			

Anmerkungen. Abhängige Variable = FLVT-Follow-Up; Förderdauer = Anzahl der Fördersitzungen; Vorwissen = FLVT-Vortest; FU-Abstand = Follow-Up-Zeitpunkt; Σ = Summierung; Δ = Veränderung; \* = hoch signifikanter Effekt (p < .01). Kodierung Geschlecht: 0 (männlich) bzw. 1 (weiblich); Migration: 0 (nein) bzw. 1 (ja); Förderbedarf: 0 (nein) bzw. 1 (ja); Treatment: 2 (Experimental-Gruppe) bzw. 1 (Kontrollgruppe).

von Null verschiedenen Gesamt-Modell führt (ANOVA: F = 35.35; df = 7/151; p = .000).

Bezüglich der Beantwortung der zweiten Forschungsfrage zeigt sich in der sehr schwach lesenden Teilstichprobe ein Ausgangswert (FLVT-Rohwert) von M = 9.20 (SD = 3.7) für Kinder aus der Experimental- und M = 9.03 (SD = 3.7) für diejenigen aus der Kontrollgruppe (N = 37 bzw. N = 41). Der Verlauf der Veränderung nach der Intervention (M = 15.4; SD = 5.0 bzw. M = 12.03; SD = 5.5) und im Follow-Up

(M = 16.4; SD = 4.8 bzw. M = 13.7; SD = 5.5) für Experimental- und Kontrollgruppe entspricht demjenigen der Gesamt-Stichprobe (Abbildung 1), jedoch auf deutlich niedrigerem Niveau, aber steilerem Anstieg. Die Signifikanzprüfung der Treatment-Effekte stellt sich hier wie folgt dar: Die Treatment-Variable leistet sowohl zum Nachtestzeitpunkt einen zusätzlichen signifikanten Beitrag zur Varianzaufklärung (β = .31; ΔR<sup>2</sup> = .093; p (Δ) = .002) als auch zum Follow-Up-Zeitpunkt (β = .28; ΔR<sup>2</sup> = .074; p (Δ) = .006).

### **Befunde auf der Basis der Textsorten**

Analog zur Auswertung der FLVT-Gesamtestwerte im Nachttest und Follow-Up wurden Analysen getrennt für die Unterskalen Sachtext und narrativer Text durchgeführt. Dabei wird aus Platzgründen und der Übersichtlichkeit halber hier jeweils auf den 4. Schritt eingegangen. Beim Sachtext zeigt sich für den Nachttest analog zu Tabelle 4 eine grundsätzlich vergleichbare Befundstruktur mit einem Gewicht für das Vorwissen von  $\beta = .45$  ( $p = .000$ ) und für das Treatment von  $\beta = .25$  ( $p = .000$ ). Für das Follow-Up-Maß ergibt sich analog zu Tabelle 5 für das Vorwissen ein  $\beta = .55$  ( $p = .000$ ) und für das Treatment ein  $\beta = .15$  ( $p = .021$ ). Darüber hinaus stellt sich heraus, dass der Effekt Förderbedarf ( $\beta = -.17$ ;  $p = .012$ ) und der Follow-Up-Abstand ( $\beta = -.30$ ;  $p = .004$ ) signifikante Varianzanteile aufklären, jedoch nicht die Trainingsdauer.

Bei der Analyse der Werte für den Bereich narrativer Texte zeigen sich analog zu den bisherigen Befunden zum Nachttest (Tabelle 4) signifikante Effekte für das Vorwissen ( $\beta = .51$  ( $p = .000$ )) und für das Treatment von  $\beta = .16$  ( $p = .012$ ). Für das Follow-Up-Maß ergibt sich analog zu Tabelle 5 für das Vorwissen ein  $\beta = .66$  ( $p = .000$ ), für das Treatment ein  $\beta = .12$  ( $p = .040$ ), für die Trainingsdauer ein  $\beta = -.22$  ( $p = .018$ ) und für den FU-Abstand ein  $\beta = -.21$  ( $p = .019$ ).

Im Zusammenhang mit der zweiten Forschungsfrage ergeben sich bezüglich der hier vor allen interessierenden Interventionseffekte vergleichbare Befunde. Bezüglich der Sachtexte zeigt sich sowohl zum Nachtzeitpunkt ( $\beta = .31$ ,  $p = .005$ ) als auch zum Follow-Up ( $\beta = .24$ ,  $p = .022$ ) ein signifikanter Treatment-Effekt. Dasselbe gilt für die Leistungen auf der Basis narrativer Texte ( $\beta = .22$ ,  $p = .046$ ;  $\beta = .21$ ,  $p = .048$ ).

### **Bestimmung der Effektstärken**

Da bei der inferenzstatistischen Datenanalyse die Treatment-Effekte allesamt statistisch signifikant ausgefallen sind, liegt die Voraussetzung dafür vor, Aussagen über die Prägnanz, also die praktische Bedeutsamkeit der Fördereffekte, zu machen.

Zu diesem Zweck wurden die jeweiligen Semipartialkorrelationen zwischen dem Kriterium (FLVT-Werte) und dem Prädiktor (Treatment) herangezogen. Die quadrierte semipartielle Korrelation stellt den von  $X_1$  (Gruppenzugehörigkeit) erklärten Anteil der (Einheits-)Varianz relativ zur Gesamtvarianz von  $Y$  dar und ist daher ein angemessener Indikator für die praktische Bedeutung eines Einflussfaktors als die partielle Korrelation. Der besseren Interpretierbarkeit wegen empfiehlt es sich, die Semipartialkorrelation in den gewohnten  $d$ -Wert (Cohen, 1988) umzurechnen (Lenhard & Lenhard, 2016).

Tabelle 6 enthält die auf der Basis der Semipartialkorrelationen zwischen dem Treatment und den abhängigen Variablen FLVT-Nachttest sowie FLVT-Follow-Up umgerechnete Treatment-Effektstärken, und zwar aufgeschlüsselt nach der Gesamtstichprobe und der deutlich leseschwachen Teilstichprobe.

Wie aus Tabelle 6 zu ersehen ist, ergeben sich erwartungsgemäß direkt nach Trainingsende für alle drei Erfolgsmaße im Vergleich zum Follow-Up die jeweils größeren Effektstärken.

Der Fördereffekt bezogen auf den FLVT-Gesamttest ( $d = 0.47$  bzw.  $d = 0.65$  für die Teilstichprobe) direkt nach der Förderung speist sich offensichtlich stark aus dem Effekt für den FLVT-Sachtext ( $d = 0.51$  bzw.  $d = 0.65$  für die Teilstichprobe). Im Vergleich dazu fällt die Befund-Prägnanz für das narrative Textverständnis mit  $d = 0.32$  bzw.  $d = 0.45$  deutlich ab. Sichtbar enger beieinander liegen die jeweiligen Fördereffekte zum jeweiligen Follow-Up-Zeitpunkt bei der Gesamtstichprobe mit  $d = 0.30$  (FLVT-Gesamt),  $d = 0.29$  (Sachtext) sowie

Tabelle 6: Auf der Basis der Semipartialkorrelation zwischen dem Treatment und den abhängigen Variablen FLVT-Nachtest sowie FLVT-Follow-Up (aufgeschlüsselt nach dem FLVT-Gesamtwert, FLVT-Sachtext sowie FLVT-Narrativ) umgerechnete Treatment-Effektstärken nach Cohen-d (vgl. Lenhard & Lenhard, 2016)

Gesamtstichprobe	Nachtest	Follow-Up
FLVT-Gesamt	0.47	0.30
FLVT-Sachtext	0.51	0.29
FLVT-Narrativ	0.32	0.23
Schwache Teilstichprobe		
FLVT-Gesamt	0.65	0.56
FLVT-Sachtext	0.65	0.47
FLVT-Narrativ	0.45	0.43

d = 0.23 (narrativer Text) und bei der Teilstichprobe mit d = 0.56 (FLVT-Gesamt), d = 0.47 (Sachtext) sowie d = 0.43 (narrativer Text).

## Diskussion

In der aktuellen Studie konnten im Gegensatz zu den meisten in Tabelle 1 dargestellten Befunden mit dem Programm „Wir werden Textdetektive“ unmittelbar nach dem Ende der Intervention (Nachtest) bei einer vergleichbaren Anzahl von Unterrichtsstunden statistisch signifikante und prägnante Verbesserungen der sinnentnehmenden Lesekompetenz von etwa einer halben Standardabweichung ( $d = 0.47$ ) bei der Gesamtstichprobe und ein noch stärkerer Effekt von  $d = 0.65$  für die leseschwache Teilstichprobe (zweite Forschungsfrage) beobachtet werden. Dabei liegt die Prägnanz der Intervention beim Verständnis von Sachtexten in etwa gleichauf ( $d = 0.51$  bzw.  $d = 0.65$ ).

Auch mittelfristig konnte der Interventionseffekt inferenzstatistisch abgesichert werden. Bezüglich der Follow-Up-Auswertung liegen die vorliegenden Befunde bei der Gesamtstichprobe mit  $d = 0.30$  in einem Bereich, der mit denjenigen an Gymnasiasten gewonnenen Befunden von Mokhlesgerami (2004) vergleichbar ist. Die Autorin ermittelte bei der vollständigen Programm Durchführung ein  $d = 0.34$ , das je-

doch unter der Bedingung des reinen Strategietrainings (abgespeckte Version des Programms) mit  $d = 0.52$  deutlich die hier an der Gesamtstichprobe gewonnene Effektstärke übertrifft, jedoch nicht diejenige der schwachen Leser ( $d = 0.56$ ). Ähnliche Befunde konnte Trenk-Hinterberger (2006) mit  $d = 0.53$  ebenfalls bei Gymnasiasten bei einer Vorgehensweise ermitteln, die an die 20 regulären Förderstunden noch zusätzlich sechs Stunden zur Wiederholung und Einübung der Strategien angehängt hat.

Damit können beide Fragestellungen klar bejaht werden: Wenn die Förderung vor dem Hintergrund der bisher beschriebenen als günstig erkannten Durchführungsbedingungen von geschulten Lehrkräften durchgeführt wird und die weiter oben beschriebenen inhaltlichen Veränderungen im Vorgehen realisiert werden, können in der Tat signifikante kurz- und mittelfristige Fördereffekte bezüglich des Leseverständnisses erzielt werden, die außerdem bei leseschwachen Viertklässlern sogar noch stärker ausfallen. Dies ist umso bemerkenswerter, als das Textdetektive-Programm für Kinder ab der fünften Klasse empfohlen wird.

Die explizite Behandlung der Unterschiede zwischen narrativen Texten und Sachtexten, die im Zusammenhang mit dem Zusammenfassen von Texten im Rahmen des durchgeführten Förderunterrichts ausführlich thematisiert wurden, führte im Vergleich beider Textsorten für den Nachtest zu einer differenziellen Verbesserung

bezüglich der Verständnisleistungen zugunsten der Sachtexte. Dies bestätigt die Befunde von Trenk-Hinterberger (2006; vgl. auch Tabelle 1). Im Gegensatz zu diesen konnte jedoch in der vorliegenden Untersuchung darüber hinaus auch ein signifikanter Effekt auf der Basis der Textsorte „narrativer Text“ erzielt werden. Nach Zwaan (1996) dürfte der Grund für die stärkeren Effekte beim Verstehen von Sachtexten darin zu sehen sein, dass es bei diesen im Vergleich zu narrativen Texten (Erzählungen) für den Leser deutlich einfacher ist, ein vollständiges Situationsmodell zu generieren.

Insgesamt fallen die Fördereffekte in dieser Untersuchung höher aus als die mittleren Effektstärken in einschlägigen Meta-Analysen von Rosenshine & Meister (1994) mit  $d = 0.32$ , Davis (2010) mit  $d = 0.31$  sowie Mayer & Marks (2019) mit  $d = 0.21$  bzw.  $d = 0.15$ . Die Befunde deuten darauf hin, dass die Vorgaben einer flexiblen, aber stets zielführenden Vorgehensweise gut umgesetzt wurden. Die Abschwächung des Fördereffekts im Follow-Up ist nachvollziehbar, weil das gezielte Fördern des Leseverständnisses im regulären Unterricht nicht fortgeführt wurde.

Einen deutlich größeren Follow-Up-Effekt als den in dieser Studie konnten Antoniou & Souvignier (2007) in einer Untersuchung mit lernschwachen Schülern (nicht-sprachlicher IQ = 94.4) aus einer Gesamtschule beobachten. Während sich beim Posttest-Ergebnis nach 29 Programm-Sitzungen keine signifikante Verbesserung bei den Schülern ergab, wurde im Follow-Up, ca. drei Monate später, ein starker Effekt von  $d = 0.80$  ermittelt (vgl. Tabelle 1). Die Autoren erklären sich diesen Befund damit, dass Kinder mit Lernschwierigkeiten mehr Zeit benötigen, um das deklarative Strategiewissen in handlungswirksames automatisiertes Verhalten umzusetzen. Das mag zutreffen. Beachtenswert ist aber auch, dass die Lehrkräfte sich im Rahmen des Projekts verpflichtet haben, auch nach Beendigung des Programms (Posttest im April) die Strategien

bis zum Follow-Up (im Juli) weiter im Unterricht zu implementieren (S. 48).

Damit scheint sich auch vor dem Hintergrund der vorliegenden eigenen Befunde insgesamt zu bestätigen, dass strategieorientierte Programme zur Verbesserung des Leseverständnisses umso erfolgreicher sind, je mehr den Schülern Zeit und Gelegenheit gegeben wird, deklaratives in prozedurales Strategie-Wissen weiterzuentwickeln.

Den alleinigen Fokus auf diese Art der Interpretation der Ergebnisse zu richten, ist jedoch aus zweierlei Gründen kritisch zu sehen: erstens wegen der häufig zu beobachtenden vergleichsweise geringen gemeinsamen Varianz zwischen dem Strategie-Wissen/-Nutzen und dem Leseverständnis (Rühl, 2006; Spörer et al., 2009; Muijselaar et al., 2018) sowie zweitens aufgrund der Tatsache, dass im Rahmen von Strategietrainings offensichtlich auch andere das Leseverständnis fördernde Kompetenzen erworben werden bzw. erworben werden müssen. Letzterer Gesichtspunkt resultiert aus der in dieser Untersuchung notwendig gewordenen Unterstützung der schwachen Leser im Bereich des Wortlesens und der Leseflüssigkeit. Dies war die Voraussetzung dafür, dass sie sich überhaupt mit den Detektivstrategien aktiv auseinandersetzen konnten.

Des Weiteren fand Rühl (2006), dass die Varianz im Leseverstehen von Haupt-, Real- und Gesamtschülern ohne Migrationshintergrund bei einer schrittweisen Regression am besten durch den Wortschatz ( $\beta = .45$ ), dann erst durch das Verständnis von Lesestrategien ( $\beta = .14$ ) und schließlich durch die Metakognition ( $\beta = .12$ ) erklärt werden kann. Bei Probanden mit Migrationshintergrund spielte neben dem Wortschatz ( $\beta = .19$ ) auch noch die Dekodiergeschwindigkeit ( $\beta = .20$ ) und erst dann die Metakognition ( $\beta = .17$ ) eine Rolle. Mayer (2020) kommt auf Basis einer Stichprobe von  $N = 184$  Viertklässlern aus Sonderpädagogischen Förderzentren zur Erklärung des Textverstehens zu ähnlichen Befunden (Worterkennung:  $\beta = .53$ ; Strategien:  $\beta = .17$ ; Wort-

schatz:  $\beta = .17$ ). In diesem Zusammenhang sei auch auf das pfadanalytisch ermittelte Modell zur Vorhersage der (sinnerfassenden) Lesekompetenz von Artelt, Schiefele, Schneider & Stanat (2002) verwiesen, dessen Parameter mit diesen Befunden übereinstimmen.

Darüber hinaus ermittelten Spörer et al. (2009, S. 280) sowie Seuring und Spörer (2010) gemeinsame Varianzanteile zwischen standardisiert erfasstem Leseverständnis und vier Lesestrategien (Klären, Fragen, Vorhersagen, Zusammenfassen), die deutlich unter 10% liegen und z.T. sogar negativ ausfallen.

All diese Befunde werfen die Frage auf, was in strategieorientierten Programmen zur Verbesserung des Leseverstehens neben der reinen Strategie-Anwendung von den Schülern tatsächlich sonst noch gelernt wird bzw. gelernt werden muss, und wie sich diese Einflussgrößen mit welchem Gewicht auf das Leseverstehen auswirken.

Dass im Rahmen von Strategietrainings offensichtlich auch andere für das Leseverständnis förderliche Kompetenzen aufgebaut werden, wird besonders bei Antoniou & Souvignier (2007) deutlich. Von den dort insgesamt  $N = 73$  Kindern der Gesamtstichprobe sprachen die wenigsten ( $N = 19$ ) zu Hause ausschließlich deutsch. Bei all den Aktivitäten während der Programmdurchführung kamen die Kinder selbstverständlich nicht darum herum, unterschiedliche Texte mehrfach laut oder leise zu lesen, was im Sinne des Konzepts des wiederholten Lautlesens (Rosebrock et al., 2013) auch zur Verbesserung der Leseflüssigkeit (Walter, 2018) und als Konsequenz nachweislich auch des Leseverstehens führt (Kuhn & Stahl, 2003; Therrien, 2004; Klaua & Guthrie, 2008). Bei der Einübung der Strategie „Umgang mit Textschwierigkeiten“ geht es außerdem sehr stark darum, die Bedeutung unbekannter Wörter zu ergründen, was mit hoher Wahrscheinlichkeit eine Verbesserung des Wortschatzes zur Konsequenz hat.

Diese Komponenten finden sich jedoch, ähnlich wie in der vorliegenden Untersu-

chung, implizit oder explizit und auch mehr oder weniger in vielen strategieorientierten Förderprogrammen zum Leseverstehen wie z.B. bei Spörer und Brunstein (2009), wo etwa ein Drittel der gesamten Trainings-Aktivitäten im Rahmen von PALS (Peer-Assisted Learning Strategies) durch ein begleitetes Lautlesen (Partner-Lesen) abgedeckt wird. So ist es nicht verwunderlich, dass Spörer et al. (2008, S. 247) zu dem Befund kommen, dass sich die Schüler „... insbesondere im flüssigen Lesen und im Zusammenfassen von Textabschnitten“ verbessern. In der Interventionsstudie von Seuring und Spörer (2010) steigerten sich Fünftklässler vom Prätest zum Follow-Up-Zeitpunkt durch das Drei-Lesestrategien-Training nicht nur im Leseverständnis ( $d_{\text{Repeated Measures}} = 0,67$ ), sondern auch in der Leseflüssigkeit ( $d_{\text{Repeated Measures}} = 0,50$ ). Umgekehrt verbesserten sich Probanden, die ein reines Leseflüssigkeitstraining durchliefen, erwartungsgemäß in der Leseflüssigkeit ( $d_{\text{Repeated Measures}} = 0,96$ ), jedoch auch im Leseverständnis ( $d_{\text{Repeated Measures}} = 0,71$ ) und darüber hinaus in der Strategie „Fragen“ ( $d_{\text{Repeated Measures}} = 0,45$ ) und „Vorhersagen“ ( $d_{\text{Repeated Measures}} = 0,81$ ). Die genannten Effektstärken wurden vom Autor dieses Beitrags unter der Annahme einer Korrelation von  $r = 0.70$  zwischen den Messzeitpunkten nach Lenhard & Lenhard (2016) berechnet.

Im aktuellen Zusammenhang besonders relevant ist die Tatsache, dass auch Lenhard et al. (2012) in einer Vergleichsstudie feststellten, dass sich bei Hauptschülern der 6. Klasse als Konsequenz sowohl der Durchführung eines strategieorientierten computergestützten Programms zur Leseverständnisförderung als auch durch die „Lesedetektive“ die Leseflüssigkeit signifikant und prägnant verbesserte ( $d = 0.60$  bzw.  $d = 0.34$ ).

In eine vergleichbare Richtung gehen die auf der Basis von  $N = 1105$  Fünftklässlern pfadanalytisch gewonnenen Ergebnisse von Völlinger, Spörer, Lubbe und Brunstein (2018), die mit Hilfe ihres ReStMe-Modells (Reading Strategies Mediation Modell) zei-

gen, dass Leseflüssigkeit und Wortschatz sowohl einen direkten Einfluss auf das Leseverstehen ausüben als auch einen indirekten über die Lesestrategien. Die Kompetenz bezüglich der Qualität des Gebrauchs von Lesestrategien hatte auch einen direkten Einfluss auf die Verstehensleistungen, dieser war jedoch insofern nicht perfekt, als die direkten Effekte von Wortschatz und Leseflüssigkeit auch unter Auspartialisierung der Qualität des Gebrauchs von Lesestrategien statistisch signifikant blieben.

### Fazit

Auch wenn dies in der vorliegenden Untersuchung nicht explizit überprüft wurde, dürfte man im Kontext der *Gesamtforschungslage* sowie der in dieser Untersuchung realisierten Leseverständnisförderung zu der Schlussfolgerung kommen, dass sich in reinen Strategie-Trainings oder in Variationen davon, bei denen eine Verbesserung z.B. der Dekodierfähigkeit und der Leseflüssigkeit – aus welchen Gründen auch immer – zu kurz kommt, deutlich schwerer Effekte auf das Leseverstehen nachweisen lassen. Was konkret in vielen Interventionsstudien methodisch passiert, bleibt jedoch häufig intransparent (Rohrbeck et al. 2003, S. 252). Dies erschwert die Interpretation und den Vergleich von Befunden nicht unerheblich.

Vermutlich ist es so, dass die aktive Auseinandersetzung mit Texten (vgl. auch Hacker & Tenet, 2002) zu einer Verbesserung des Wortschatzes, der Dekodierfähigkeit und der Leseflüssigkeit führt, was genügend Informationsverarbeitungskapazitäten zur Konstruktion eines Situationsmodells im Sinne der Cognitive-Load-Theorie (Seufert, 2018) frei werden lässt, so dass im Zweifel das zusätzliche Lehren der Strategienutzung nur wenige oder keine Effekte zeigt.

Dies müsste jedoch im Rahmen von hybridmäßig (vgl. Klauer, 2011, S. 117; Völlinger et al., 2018) angelegten Interventionsstudien noch genauer untersucht werden. Für die Förderpraxis bedeutet die aktuelle

Befundlage aber schon jetzt, dass längerfristig angewandte hybrid ausgelegte Verfahren die größte Gewähr auf nachhaltige Fördererfolge erwarten lassen oder anders ausgedrückt: Die Förderung des Textverständnisses sollte bei schwachen Lesern auf jeden Fall mehrdimensional angelegt werden.

Die Summe der vorliegenden Analyse-Ergebnisse kann sich auch als Erklärungsgrundlage für sowohl erwartungskonforme als auch erwartungswidrige Befunde der älteren (vgl. Tabelle 1) und auch neueren Lesestrategieforschung eignen, wie zum Beispiel für diejenigen von Droop et al. (2016) sowie Muijselaar et al. (2018). In beiden Untersuchungen (Grundschüler) bekamen die Kontrollgruppen einen im Vergleich zur Experimentalgruppe ähnlich ausgerichteten Leseunterricht, ohne dass jedoch explizit kognitive und metakognitive Kompetenzen (wie in der Fördergruppe) gelehrt wurden. In ersterer, die sich über zwei Schuljahre erstreckte, zeigte sich nach einem Jahr kein signifikanter Effekt der Lesestrategie-Gruppe auf das Leseverstehen und nach zwei Jahren ein schwacher von  $d = 0.20$ . Bei Muijselaar et al. (2018) konnte trotz einer signifikanten Verbesserung des Wissens um die Anwendung von Strategien kein bedeutsamer Anstieg der Lesekompetenz im Vergleich zur Kontrollgruppe erreicht werden, was die Autoren zu der Frage führt, ob es überhaupt Sinn macht, ein reines Strategie-training zur Verbesserung der Lesekompetenz bei Viertklässlern heranzuziehen. Diese Befunde verdeutlichen, dass es bei der Bewertung von Fördereffekten in Interventionsstudien zu berücksichtigen gilt, welche Aktivitäten in den jeweiligen Kontrollgruppen stattfinden, gegen die ja statistisch getestet wird. Ähnliche Gesichtspunkte müssten auch bei der Bewertung der „Frankfurter“ Untersuchungen (Tabelle 1) Berücksichtigung finden, bei denen der Unterricht der Deutschlehrer als „Alternativtraining“ gewertet werden müsste.

### Limitationen

Die vorgelegten Befunde sind in einem Rahmen erzeugt worden, der sowohl Stärken als auch Schwächen aufweist. Die Untersuchung füllt insofern eine Lücke, als es kaum Studien zum Einsatz von strategieorientierten Verfahren bei schwachen Viertklässlern gibt. Dabei wurde der Versuch unternommen, unter Bedingungen des Schulalltags einen echten Transfer auf Leseverständnisleistungen zu erzielen. Im Gegensatz zu den meisten anderen Förderexperimenten wurde hier gegen eine Zuwendungs-Kontrollgruppe getestet, was sich möglicherweise dämpfend auf die Effektivitätsmaßstäbe ausgewirkt hat.

Als eindeutiges Manko dieser Untersuchung muss das Fehlen von Kontrollvariablen wie Wortschatz, Dekodierfähigkeit und Leseflüssigkeit festgehalten werden. Möglicherweise sind hier differenzielle Fördereffekte insofern verborgen geblieben, als sich bei einigen Schülern tatsächlich Probleme beim Strategie-Erwerb aufgrund ihrer mangelnden Leseflüssigkeit ergaben. Auch wenn das Trainerteam gut vorbereitet in die Förderung ging, wurde nicht kontrolliert, in welchem Maße die Vorgaben umgesetzt wurden und sich ggf. Unterschiede in den TU zeigten. Abweichend von den Vorgaben im Manual der „Textdetektive“ wurde hier ein zeitlich befristeter Kompaktkurs realisiert, was sich vor dem Hintergrund einschlägiger Befunde eher negativ auf die Entfaltung der Lesekompetenz der Schüler ausgewirkt haben könnte.

### Literaturverzeichnis

- Antoniou, F. & Souvignier, E. (2007). Strategy instruction in reading comprehension: An intervention study for students with learning disabilities. *Learning Disabilities: A Contemporary Journal*, 5, 41-57.
- Artelt, C., Schiefele, U., Schneider, W. & Stanat, P. (2002). Leseleistungen deutscher Schülerinnen und Schüler im internationalen Vergleich (PISA). Ergebnisse und Erklärungsansätze. *Zeitschrift für Erziehungswissenschaft*, 5, 6-27.
- Bortz, J. & Döring, N. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften* (5. Auflage). Heidelberg: Springer.
- Bos, W., Valtin, R., Hußmann, A., Wendt, H. & Goy, M. (2017). IGLU 2016: Wichtige Ergebnisse im Überblick. In A. Hußmann, H. Wendt, W. Bos, A. Bremerich-Vos, D. Kasper, E.-M. Lankes et al. (Hrsg.), *Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 13-28). Münster: Waxmann.
- 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.-M. Lankes et al. (Hrsg.), *Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 79-142). Münster: Waxmann.
- Bremerich-Vos, A., Stahns, R., Hußmann, A. & Schurig, M. (2017). Förderung von Leseflüssigkeit und Leseverstehen im Lesunterricht. In A. Hußmann, H. Wendt, W. Bos, A. Bremerich-Vos, D. Kasper, E.-M. Lankes et al. (Hrsg.), *Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 279-296). Münster: Waxmann.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). Hillsdale, N.J.: L. Erlbaum Associates.
- Davis, D.S. (2010). *A meta-analysis of comprehension strategy instruction for upper elementary and middle school students* (Unpublished doctoral dissertation). Graduate School of Vanderbilt University, Nashville, TN.
- Droop, M., van Elsäcker, W., Voeten, M.J.M. & Verhoeven, L. (2016). Long-term effects of strategic reading instruction in the intermediate elementary grades. *Journal of Research on Educational Effectiveness*, 9, 77-102.



- Duffy, G.G., Roehler, L.R., Sivan, E., Ratcliffe, G., Book, C., Meloth, M.S., Vavrus, L.G. et al. (1987). Effects of explaining the reasoning associated with using reading strategies. *Reading Research Quarterly*, 22, 347–368.
- Ehri, L.C., Nunes, S.R., Willows, D.M., Schuster, B.V., Yaghoub-Zadeh, Z. & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel's meta-Analysis. *Reading Research Quarterly*, 36, 250-287.
- Fischer, M.Y. & Pfof, M. (2015). Wie effektiv sind Maßnahmen zur Förderung der phonologischen Bewusstheit? Eine meta-analytische Untersuchung der Auswirkungen deutschsprachiger Trainingsprogramme auf den Schriftspracherwerb. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 47, 35 – 51.
- Gold, A. (2010). *Lesen kann man lernen. Lese-strategien für das 5. und 6. Schuljahr* (2. Aufl.). Göttingen: Vandenhoeck & Ruprecht.
- Gold, A., Mokhlesgerami, J., Rühl, K., Schreblowski, S. & Souvignier, E. (2006). *Wir werden Textdetektive*. Lehrermanual (2. Auflage). Göttingen: Vandenhoeck & Ruprecht.
- Gold, A., Mokhlesgerami, J., Rühl, K., Schreblowski, S. & Souvignier, E. (2006). *Wir werden Textdetektive. Arbeitsheft* (3. Auflage). Göttingen: Vandenhoeck & Ruprecht.
- Hattie, J. (2013). *Lernen sichtbar machen*. Baltmannsweiler: Schneider Verlag Hohengehren.
- Hacker, D.J. & Tenent, A. (2002). Implementing reciprocal teaching in the classroom: Overcoming obstacles and making modifications. *Journal of Educational Psychology*, 94, 699-718.
- Hußmann, A., Wendt, H., Bos, W., Bremerich-Vos, A., Kasper, D., Lankes, E.-M. et al. (Hrsg.). (2017). *Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster: Waxmann
- Klauda, S.L. & Guthrie, J.T. (2008). Relationships of three components of reading fluency to reading comprehension. *Journal of Educational Psychology*, 100, 310–321.
- Klauer, K.J. (2010). Schädliche Interferenzen beim Training hochkomplexer Lernstrategien? *Zeitschrift für Pädagogische Psychologie*, 24, 235–239.
- Klauer, K.J. (2011). *Transfer des Lernens. Warum wir oft mehr lernen als gelehrt wird*. Stuttgart: Kohlhammer.
- Koch, H. & Spörer, N. (2016). Förderung der Lesekompetenz mittels reziproken Lehrens: Implementation und Wirksamkeit im Regelunterricht. *Zeitschrift für Pädagogische Psychologie*, 30, 213–225.
- Kuhn, M.R. & Stahl, S.A. (2003). Fluency: A review of developmental and remedial practices. *Journal of Educational Psychology*, 95, 3-21.
- Lenhard, W., Baier, H., Endlich, D., Lenhard, A., Schneider, W. & Hoffmann, J. (2012). Computerunterstützte Leseverständnisförderung: Die Effekte automatisch generierter Rückmeldungen. *Zeitschrift für Pädagogische Psychologie*, 26, 135-148.
- Lenhard, W. & Lenhard, A. (2016). *Berechnung von Effektstärken*. Abgerufen unter: <https://www.psychometrica.de/effektstaerke.html>. Dettelbach: Psychometrica. DOI: 10.13140/RG.2.2.17823.92329
- Mayer, A. (2020). Der Einfluss (schrift-)sprachlicher und kognitiv-strategischer Kompetenzen auf das Leseverständnis von Schülern mit sonderpädagogischem Förderbedarf. *Praxis Sprache*, 65, 5-18.
- Mayer, A. & Marks, D.-K. (2019). Förderung des Textverständnisses durch die Vermittlung von Verstehensstrategien – Eine Metaanalyse zur Effektivität. *Forschung Sprache*, 7, 4-36.
- Mokhlesgerami, J. (2004). *Förderung der Lesekompetenz. Implementatiom und Evaluation eines Unterrichtsprogramms in der Sekundarstufe I*. Hamburg: Kovac.
- Mokhlesgerami, J., Souvignier, E. & Gentsch, S. (2006). Förderung von Lesestrategien. Erprobung eines Unterrichtsprogramms in

- Haupt-, Real- und Gesamtschulen. *Empirische Pädagogik*, 20, 70-90.
- Mokhlesgerami, J., Souvignier, E., Rühl, K. & Gold A. (2007). Naher und weiter Transfer eines Unterrichtsprogramms zur Förderung der Lesekompetenz in der Sekundarstufe I. *Zeitschrift für pädagogische Psychologie*, 21, 169-180.
- Muijselaar, M., Swart, N.M., Steenbeek-Planting, E.G., Droop, M., Verhoeven, L. & de Jong, P.F. (2017). Developmental relations between reading comprehension and reading strategies. *Scientific Studies of Reading*, 21, 194-209. <https://doi.org/10.1080/10888438.2017.1278763>
- Muijselaar, M., Swart, N.M., Steenbeek-Planting, E.G., Droop, M., Verhoeven, L. & de Jong, P.F. (2018). The effect of a strategy training on reading comprehension in fourth-grade students. *The Journal of Educational Research*, 111, 690-703.
- Mullis, I., Martin, M.O., Foy, P. & Hooper, M. (2017). *PIRLS 2016. International results in reading*. Boston: TIMSS & PIRLS International Study Center.
- 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, DC: U.S. Government Printing Office.
- Pein, M. (2016). *Adaption des Trainingsprogramms "Wir werden Textdetektive" im Rahmen einer Feldstudie innerhalb einer vierten Grundschulklasse*. Master-Thesis, Europa-Universität Flensburg.
- Rasinski, T.V. (2004). *Assessing reading fluency*. Honolulu, HI: Pacific Resources for Education and Learning.
- Rohrbeck, C.A., Ginsburg-Block, M., Fantuzzo, J.W. & Miller, T.R. (2003). Peer-assisted learning interventions with elementary school students: A meta-analytic review. *Journal of Educational Psychology* 95, 240-257.
- Rosebrock, C., Nix, D., Rieckmann, C. & Gold, A. (2013). *Leseflüssigkeit fördern. Lautleseverfahren für die Primar- und Sekundarstufe* (2. Auflage). Seelze: Kallmeyer/Klett.
- Rosebrock, C., Rieckmann, C., Nix, D. & Gold, A. (2010). Förderung der Leseflüssigkeit bei leseschwachen Zwölfjährigen. *Didaktik Deutsch*, 28, 33-58.
- Rosenshine, B. & Meister, C. (1994). Reciprocal teaching: A review of the research. *Review of Educational Research*, 64, 479-530.
- Rosenshine, B., Meister, C. & Chapman, S. (1996). Teaching students to generate questions: A review of the intervention studies. *Review of Educational Research*, 66, 181-221.
- Rühl, K. (2006). *Förderung des Textverstehens. Prüfung der differentiellen Wirksamkeit eines strategieorientierten Unterrichtsprogramms*. Hamburg: Kovac.
- Rühl, K. & Souvignier, E. (2006). *Wir werden Lesedetektive – Lehrermanual & Arbeitsheft*. Göttingen: Vandenhoeck & Ruprecht.
- Samuels, S.J. (2006). Toward a model of reading fluency. In S.J. Samuels & A.E. Farstrup (Eds.), *What research has to say about fluency instruction* (pp. 24-46). Newark, DE: International Reading Association.
- Schünemann, N., Spörer, N. & Brunstein, J.C. (2013). Integrating self-regulation in whole-class reciprocal teaching: A moderator-mediator analysis in incremental effects on fifth graders' reading comprehension. *Contemporary Educational Psychology*, 38, 289-305.
- Schünemann, N., Spörer, N., Völlinger, V.A. & Brunstein, J.C. (2017). Peer feedback mediates the impact of self-regulation procedures on strategy use and reading comprehension in reciprocal teaching groups. *Instructional Science*, 45, 395-415.
- Seufert, T. (2018). The interplay between self-regulation in learning and cognitive load. *Educational Research Review*, 24, 116-129.
- Seuring, V.A. & Spörer, N. (2010). Reziprokes Lehren in der Schule: Förderung von Leseverständnis, Leseflüssigkeit und Strategieanwendung. *Zeitschrift für Pädagogische Psychologie*, 24, 191-205.

- Seymour, J.R. & Osana, P. (2003). Reciprocal teaching procedures and principles: two teachers' developing understanding. *Teaching and Teacher Education, 19*, 325–344.
- Souvignier, E. & Behrmann, L. (2016). Wie viel Lehrerfortbildung ist für eine nachhaltige Implementation strategieorientierten Lesunterrichts nötig? *Unterrichtswissenschaft, 44*, 391–407.
- Souvignier, E. & Mokhlesgerami, J. (2005). Implementation eines Programms zur Vermittlung von Lesestrategien im Deutschunterricht. Die Rolle der Lehrenden. *Zeitschrift für Pädagogische Psychologie, 19*, 249–261.
- Souvignier, E. & Rühl, K. (2005). Förderung des Leseverständnisses, Lesestrategiewissens und Leseinteresses von Schülern mit Lernbehinderungen durch strategieorientierten Unterricht. *Heilpädagogische Forschung, 31*, 2–11.
- Souvignier, E., Trenk-Hinterberger, I., Adam-Schwebe, St. & Gold, A. (2008). FLVT 5-6. Frankfurter Leseverständnistest für 5. und 6. Klassen. Göttingen: Hogrefe.
- Spörer, N. & Brunstein, J.C. (2009). Fostering the reading comprehension of secondary school students through peer-assisted learning: Effects on strategy knowledge, strategy use, and task performance. *Contemporary Educational Psychology, 34*, 289–297.
- Spörer, N., Brunstein, J.C. & Kieschke, U. (2009). Improving students' reading comprehension skills: Effects of strategy instruction and reciprocal teaching. *Learning and Instruction, 19*, 272–286.
- Spörer, N., Seuring, V., Schünemann, N. & Brunstein, J.C. (2008). Förderung des Leseverständnisses von Schülern der 7. Klasse. *Zeitschrift für Pädagogische Psychologie, 22*, 247–259.
- Therrien, W.J. (2004). Fluency and comprehension gains as a result of repeated reading. A meta-analysis. *Remedial and Special Education, 25*, 252–261.
- Trenk-Hinterberger, I. (2006). *Lesekompetenz im Schulunterricht. Evaluation und Optimierung eines strategieorientierten Programms in der fünften Jahrgangsstufe*. Hamburg: Kovac.
- Trenk-Hinterberger, I. & Souvignier, E. (2006). *Wir sind Textdektive. Lehrermanual mit Kopiervorlagen*. Göttingen: Vandenhoeck & Ruprecht.
- Völlinger, V.A., Spörer, N., Lubbe, D. & Brunstein, J.C. (2018). A path analytic test of the reading strategies mediation model: Relating cognitive competencies and motivational influences to individual differences in fifth-grade students' comprehension. *The Journal of Educational Research, 111*, 733–745.
- Walter, J. (2007). Sinnverstehendes Lesen. In J. Walter & F. Wember (Hrsg.), *Sonderpädagogik des Lernens* (S. 518–539). Göttingen: Hogrefe.
- Walter, J. (2018). Zur Effektivität der Förderung der Leseflüssigkeit auf der Basis von Hörbüchern in Kombination mit wiederholtem Lesen: Weitere Evidenz. *Empirische Sonderpädagogik, 3*, 248–272.
- Wolf, K.M., Schroeders, U. & Kriegbaum, K. (2016). Metaanalyse zur Wirksamkeit einer Förderung der phonologischen Bewusstheit in der deutschen Sprache. *Zeitschrift für Pädagogische Psychologie, 30*, 9–33.
- Zwaan, R.A. (1996). Toward a model of literacy comprehension. In B.K. Britton & A.C. Graesser (Eds.), *Models of understanding text* (pp. 241–255). Mahwah: Erlbaum.

**Prof. Dr. Jürgen Walter**

Europa-Universität Flensburg  
 Institut für Sonderpädagogik  
 Abteilung Sonderpädagogik des Lernens  
 Auf dem Campus 1  
 24943 Flensburg  
 E-Mail: walter@uni-flensburg.de

Erstmalig eingereicht: 29.07.2020

Überarbeitung eingereicht: 04.11.2020

Angenommen: 05.11.2020