

Originalartikel

# Computerunterstützte Leseverständnisförderung: Die Effekte automatisch generierter Rückmeldungen

Wolfgang Lenhard<sup>1</sup>, Herbert Baier<sup>1</sup>, Darius Endlich<sup>1</sup>, Alexandra Lenhard<sup>2</sup>,  
Wolfgang Schneider<sup>1</sup> und Joachim Hoffmann<sup>1</sup>

<sup>1</sup>Institut für Psychologie, Universität Würzburg

<sup>2</sup>Psychometrica, Dettelbach

**Zusammenfassung:** In einer Gruppe von  $N = 210$  Hauptschülerinnen und -schülern der 6. Klasse verglichen wir die Effekte der computerunterstützten Leseverständnisförderung *conText* mit einer etablierten Leseverständnisintervention («Wir werden Lesedetektive») und einer unbehandelten Kontrollgruppe im Laufe eines Schuljahres. *conText* ist ein intelligentes tutorielles System, das die Qualität von Textzusammenfassungen mithilfe von Verfahren aus dem Bereich der automatischen Sprachverarbeitung bewertet und an die Schüler zurückmeldet. Im Vergleich zu einer unbehandelten Kontrollgruppe verbesserten beide Interventionen das Lesestrategiewissen, jedoch erwies sich nur *conText* im Hinblick auf die Verbesserung des Leseverständnisses als erfolgreich. Die Ergebnisse werden im Hinblick auf die Rahmenbedingungen von Leseverständnisinterventionen diskutiert.

**Schlüsselwörter:** Lesestrategiewissen, intelligente tutorielle Systeme, latente semantische Analyse, Lesekompetenz, Leseförderung

## Computer Assisted Reading Comprehension Instruction: Effects of Automatically Generated Content Feedback

**Abstract:** We compared the effects of computer assisted instruction (“*conText*”) with an explicit strategy instruction training (“Let’s Become Reading Detectives”) and a non-treatment control group in a group of  $N = 210$  sixth graders over one school year. *conText* is an intelligent tutoring system that uses algorithms from automatic speech processing to assess the quality of summaries and generates automatic and instant feedback. Compared to the non-treatment group, students from the intervention group considerably improved reading strategy knowledge. However, only *conText* improved reading comprehension. We discuss these results with regard to the general requirements of reading comprehension trainings.

**Keywords:** reading strategy knowledge, intelligent tutoring systems, latent semantic analysis, reading comprehension, intervention

## 1 Ansätze zur Förderung des Leseverständnisses

Schrift kommt praktisch in allen Lebensbezügen vor und gewinnt durch die fortlaufende Verlagerung von Lebensbereichen in digitale Welten nicht weniger, sondern immer mehr an Bedeutung. Die Fähigkeit, Texte zu entschlüsseln und deren Inhalte zu rekonstruieren, hat dabei unter den Kulturtechniken eine exponierte Stellung: Wir lesen er-

heblich häufiger, als wir schreiben oder rechnen. Im schulischen Bereich ist Leseverständnis eine wichtige Grundfähigkeit in fast allen Unterrichtsfächern. Lesekompetenz ist aber nicht nur eine zentrale Voraussetzung für die schulische und akademische Laufbahn: Fast kein Beruf ist ohne grundlegende Lesefähigkeit möglich. Auch außerhalb von Schule und Beruf erschließen sich zahlreiche Lebensbereiche nur mittels der Fähigkeit zum verstehenden Lesen. Für eine Teilhabe am gesellschaftlichen, kulturellen und politischen Leben ist Leseverständnis essenziell (vgl. Artelt

et al., 2007, S. 4). Die Vermittlung von Lesekompetenz ist aus diesem Grund sowohl ein wichtiges Ziel als auch die Voraussetzung schulischer Bildungsprozesse.

## 1.1 Befunde zur Effektivität von Fördermaßnahmen

Aufgrund der enormen Bedeutung der Lesefähigkeit werden effektive Fördermaßnahmen und Interventionsansätze dringend benötigt. Betrachtet man die Interventionsforschung zur Verbesserung der Lesefertigkeiten, so kann zumindest im internationalen Bereich auf eine reichhaltige Forschungstradition zurückgeblieben werden. Aktuelle Metaanalysen erlauben einen dezidierten Blick auf die Effekte unterschiedlicher Herangehensweisen und Arbeitsformen, sowie auf Effekte bei leistungsschwachen Kindern und Jugendlichen (siehe z.B. Antoniou & Souvignier, 2007; Berkeley, Scruggs & Mastropieri, 2009; Hattie, 2009, S. 129 ff.; Edmonds et al., 2009; Slavin, Cheung, Groff & Lake, 2008). Es gibt zudem zahlreiche Empfehlungen von staatlichen Stellen, die eine Fülle von Ansatzpunkten zur Lesekompetenzförderung aufzeigen (Artelt et al., 2007; National Reading Panel, 2000; RAND Reading Study Group, 2002). Die fünf wichtigsten Aspekte der Leseförderung umfassen phonologische Bewusstheit, Buchstaben-Laut-Zuordnung («Phonics»), Leseflüssigkeit, Wortschatz und Leseverständnis (vgl. Hattie, 2009, S. 140).

In Bezug auf die Förderung des Leseverständnisses spielt zumindest ab der Sekundarstufe die Vermittlung von Lesestrategien eine herausragende Rolle. Eine ganze Reihe an Förderprogrammen stützt sich deshalb in unterschiedlich starker Form auf die Vermittlung von Strategien. Deren Anwendung wird dabei entweder explizit im Unterrichtsgespräch vermittelt, implizit eingeübt oder im Rahmen kooperativer Arbeitsformen demonstriert. Hierzu gehören die Verfahren *Reciprocal Teaching* (Palincsar & Brown, 1984), *Informed Strategies Instruction* (Paris, Cross & Lipson, 1984), *Direct Explanation of Strategies* (Duffy, Roehler & Herrmann, 1988), *Transactional Strategies Instruction* (Pressely et al., 1992), *Concept Oriented Reading Instruction* (Guthrie, Wigfield & Perencevich, 2004), *Collaborative Strategic Reading* (Kim et al., 2006) und *Peer Assisted Learning Strategies* (Fuchs, Fuchs, Mathes & Simmons, 1997; für einen Überblick siehe van Kraayenoord, 2010). Auch im deutschsprachigen Raum wurde insbesondere in den letzten Jahren eine Reihe strategiebasierter Programme publiziert, darunter *Wir werden Textdetektive* (Gold, Mokhlesgerami, Rühl, Schreblowski & Souvignier, 2004), *Wir werden Lesedetektive* (Rühl & Souvignier, 2006; siehe Kap. 1.4), *Lesen. Das Training* (Bertschi-Kaufmann et al., 2007), *Programm zur Förderung der Lesekompetenz und Lesemotivation (LEKOLEMO)*; Streblov, Holodynski & Schiefele, 2007), *Lesen: Texte verarbeiten, verstehen und nutzen* (Druschky, Meier & Stadler, 2006) und das *Berliner Eltern-Kind-*

*Leseprogramm* (McElvany & Artelt, 2007), um nur einige zu nennen.

Die Bewertung der Effektivität dieser Ansätze fällt indes nicht leicht, da diese stark vom Untersuchungssetting abhängt. Die Höhe der Effekte wird unter anderem von den folgenden Faktoren beeinflusst (Berkeley et al., 2009; Souvignier & Antoniou, 2007): Größe der Interventionsgruppe, Interventionskontext (Labor, Schule oder Therapie), angewandte Sozialformen, Trainer (Wissenschaftler versus Lehrkraft oder Eltern), Leistungsniveau der Kinder und Jugendlichen, Dauer der Interventionen und die verwendeten Testverfahren (selbst konstruiert versus standardisiert). Besonders bei älteren Studien, bei denen verstärkt selbst konstruierte Maße zur Anwendung kamen, fallen die Effekte im Schnitt höher aus ( $d = .85$  versus  $d = .56$  bei selbst konstruierten versus standardisierten Tests; Souvignier, 2009). Verschiedene neuere Untersuchungen berichten bei schwachen Lesern in der Sekundarstufe bei Verwendung standardisierter Verfahren mittelhohe Effekte. Edmonds et al. (2009) ermitteln in ihrer Metaanalyse zu «Older Struggling Readers» bei Studien im Zeitraum zwischen 1994 und 2004 beispielsweise einen Effekt von  $d = .47$ . Große Effekte wurden insbesondere für spezifische Leseverständnisinterventionen sowie für Multi-Komponenten-Ansätze gefunden, wohingegen Maßnahmen auf Wortebene oder zur Steigerung der Leseflüssigkeit keine oder geringe Effekte hatten. Berkeley et al. (2009; «Students with learning Disabilities», Zeitraum 1995–2006) dokumentieren bei standardisierten Verfahren mit  $d = .52$  einen Effekt in vergleichbarer Höhe.

Betrachtet man nicht ausschließlich den leistungsschwachen Bereich, so ergibt sich ein anderes Bild. Slavin et al. (2008) untersuchten die Effekte von Lese-Lehrgängen, gemischten Methoden, computerunterstütztem Unterricht und «Instructional-Process Programs» (Programme zur Implementierung spezifischer Instruktionmethoden und der Professionalisierung der Lehrkräfte) und bezogen insgesamt 33 Studien mit etwa 39 000 Schülerinnen und Schülern ein. Es stellte sich heraus, dass keine einzige Studie zu Leselehrgängen die erforderlichen Qualitätskriterien erfüllte. Computerunterstützter Unterricht – Slavin et al. (2008) zählen hierzu computerverwaltete integrierte Lehr-Lernsysteme älteren Datums (s.u.) und am PC ergänzend dargebotene Inhalte – erwies sich mit einer durchschnittlichen Effektstärke von  $d = .10$  als wenig effektiv. Wurde individualisierte computerunterstützte Instruktion mit Arbeit in Kleingruppen und im Klassenkontext kombiniert, so zeigt sich mit  $d = .23$  immer noch ein relativ geringer Effekt, der aber deutlich höher lag als bei ausschließlichem Einsatz des Computers. In einer ähnlichen Höhe rangierten Programme, die auf die Verbesserung des Unterrichts und der eingesetzten Lehr-Lernmethoden setzen ( $d = .21$ ). Am günstigsten schnitten Programme ab, die sich reziproken Lernens bedienen ( $d = .32$ ). Isolierte Strategieprogramme, die auf die Vermittlung kognitiver und metakognitiver Strategien abzielen, erbrachten dagegen mit einer gewichteten Effektstärke von  $d = .09$  nur geringe Verbesserungen.

## 1.2. Computerbasierte Leseverständnisförderung

Es lässt sich somit festhalten, dass bei einer breiten Anwendung von Programmen zur Leseverständnisförderung in Schulen positive, aber nicht durchschlagende Effekte zu erwarten sind – das gilt sowohl für computerbasierte als auch für konventionelle Programme. Betrachtet man die von Slavin et al. (2008) untersuchten computerunterstützten Interventionsmaßnahmen genauer, so fällt auf, dass die meisten berücksichtigten Computerprogramme bereits sehr alt sind und im Zeitraum zwischen Ende der 1970er-Jahre und der ersten Hälfte der 1990er-Jahre entwickelt wurden. In einigen der Programme – wie beispielsweise bei *Accelerated Reader* – wird außerdem kaum mit dem Computer gearbeitet. Der PC dient hier vorwiegend als Diagnoseinstrument und stellt auf der Basis der Testergebnisse Inhalte zusammen, die dann in Form ausgedruckter Materialien oder Bücher bearbeitet werden. Am Ende wird lediglich erneut eine Erfolgskontrolle am Computer durchgeführt. Die Vorgehensweise ähnelt verschiedenen deutschsprachigen Angeboten wie dem «Münsteraner Lernserver» oder den internetbasierten Diagnose- und Förderangeboten der deutschen Schulbuchverlage. Die untersuchten Programme schöpfen also bei Weitem nicht die Möglichkeiten heutiger Computersysteme aus, sodass die Ergebnisse der Metaanalyse in diesem Punkt als veraltet betrachtet werden müssen. Darüber hinaus ist die Zuordnung verschiedener Programme zur Rubrik «gemischte Methoden» fragwürdig, da sie genauso gut als computer-unterstützter Unterricht hätten klassifiziert werden können (Greenleaf & Petrosino, 2009). Und schließlich stellt sich auch die Frage, ob es angesichts der großen Unterschiede zwischen verschiedenen Computerprogrammen überhaupt sinnvoll ist, ihre Effektivität metaanalytisch zu betrachten oder ob nicht geeignete Unterkategorien innerhalb des computerunterstützten Unterrichts notwendig wären, um diesen angemessen beurteilen zu können. Schließlich werden die Effekte von Unterricht ohne Computerunterstützung auch nicht bewertet, ohne dabei verschiedene Lehr-Lernformen zu differenzieren.

Trotz dieser eher ernüchternden Ergebnisse älterer Programme gibt es eine Reihe an Vorteilen, die der Computer potenziell bietet (vgl. Lenhard, Lenhard, Schug & Kowalski, 2011): Zum einen stellt er für viele Kinder und Jugendliche ein sehr attraktives Lernmedium dar. Zum anderen sind Rechner stark mit Spielen und Unterhaltung assoziiert und werden deshalb häufig als spielerische Abwechslung zum normalen Schulalltag erlebt. Darüber hinaus lässt sich eine Reihe elementarer Lernprinzipien umsetzen, wie z. B. eine unmittelbare und individualisierte Rückmeldung. Und schließlich scheint eine vom Computer automatisch generierte Rückmeldung bei niedrigen Leistungen weniger angstinduzierend zu sein, als das bei Bewertungen durch Lehrkräfte der Fall ist.

Das Problem der Verarbeitung natürlicher Sprache setzte dem Einsatz des Rechners in der Leseverständnisförde-

rung allerdings lange Zeit deutliche Grenzen. Aus diesem Grund wurden Rechner bislang hauptsächlich als Basis für kollaborative Lernformen, als Verteilungsmöglichkeit für Texte und multimediale Inhalte, sowie für die Verwendung geschlossener Antwortformate eingesetzt. Während kollaborative Lernformen dazu angetan sind, eine intensive Auseinandersetzung mit Textinhalten zu fördern, verleiten geschlossene Antwortformate u. U. dazu, sich nur oberflächlich mit Texten auseinanderzusetzen (z. B. Bangert-Downs, Kulik, Kulik & Morgan, 1991). Damit ein Computerprogramm auf individueller Ebene effektiv ist, müsste es also in der Lage sein, eine intensive Auseinandersetzung mit den Inhalten anzuregen und unmittelbare, individualisierte Rückmeldungen zu geben. Wir haben versucht, ein Programm (*conText*) zu entwickeln, das diesen Forderungen nahe kommt. Um die Effekte dieses Programms zu überprüfen, wurde in der vorliegenden Untersuchung die computerbasierte Intervention mit einem bereits evaluierten Leseverständnistraining und einer Wartelistenkontrollgruppe kontrastiv verglichen. Das Computerprogramm wird im Folgenden vorgestellt:

## 1.3 conText

*conText* (Lenhard, Baier, Lenhard, Schneider & Hoffmann, 2012) ist ein intelligentes tutorielles System, das auf den *impliziten* Erwerb prozeduraler Fertigkeiten vermittelt über das computerunterstützte Erarbeiten und Zusammenfassen von Texten abzielt. Hieraus leitet sich auch das Akronym ab (*conText* = *mit Texten* arbeiten, aber auch aus dem *Zusammenhang* erschließen). Es gibt Nutzern Rückmeldungen zur Inhaltsabdeckung von Textzusammenfassungen, verwickelt sie auf diese Weise in eine intensive Auseinandersetzung mit den Textinhalten und leitet dazu an, den Inhalt eines Originaltextes vollständiger in der Zusammenfassung zu repräsentieren. Von zentraler Bedeutung für die Generierung der Rückmeldungen ist hierbei ein Verfahren aus der automatischen Sprachverarbeitung, die sog. *Latente Semantische Analyse* (LSA; Deerwester, Dumais, Furnas, Landauer & Harshman, 1990). Das Verfahren ermöglicht es, den semantischen Gehalt von Wörtern und Texten numerisch in Form von Vektorräumen zu repräsentieren und diesen somit mathematisch abbilden zu können.

Ausgangspunkt der LSA sind große Textsammlungen, die in einzelne Dokumente (z. B. Absätze) aufgeteilt werden. Für die weitere Analyse wird eine Frequenzmatrix erstellt, deren Spalten die einzelnen Abschnitte und deren Zeilen die einzelnen Wörter repräsentieren. Die Zellen der Matrix enthalten die Auftretenshäufigkeit eines Wortes in einem bestimmten Dokument. Nach einer Filterung und Gewichtung der Zellen wird diese Matrix einer Singulärwertzerlegung unterzogen. Die resultierende Wissensbasis – üblicherweise als semantischer Raum bezeichnet – stellt nun eine generalisierte, vektorielle Repräsentation des semantischen Gehaltes von Wörtern, sowie eine kompri-

mierte Form des in den Texten gespeicherten Wissens dar. Einzelne Wörter, Sätze oder Texte lassen sich durch einen Vektor in diesem semantischen Raum abbilden. Dadurch wird es möglich, sehr performant den semantischen Gehalt zweier Texte quantitativ miteinander zu vergleichen.

Eine faszinierende Anwendungsmöglichkeit der LSA liegt in der inhaltlichen Auswertung von Aufsätzen zur automatischen und unmittelbaren Bewertung und der Generierung von Feedback im Rahmen von Tutorssystemen. Ein solches Tutorssystem ist Summary Street (Kintsch, Caccamise, Franzke, Johnson & Dooley, 2007), von dem die Entwicklung von *conText* inspiriert wurde. Bei *conText* werden die folgenden Programmschritte sequenziell durchlaufen: 1. Die Nutzer bekommen einige kurze Hinweise zur Bedienung des Programms und zum Vorgehen beim Zusammenfassen von Texten. 2. Der Originaltext wird dargeboten. 3. Die Nutzer beginnen mit dem Schreiben der Zusammenfassung. Währenddessen erhalten sie eine simultane Rückmeldung über die Länge der Zusammenfassung. Gleichzeitig muss die Zusammenfassung in eigenen Worten erfolgen. Wie in einer Textverarbeitung werden unbekannte Wörter als potenzielle Falschreibungen markiert. Außerdem werden plagiierte Textstellen gekennzeichnet. 4. Es findet eine grundlegende stilistische Überprüfung statt (z. B. Überprüfung auf obszöne Begriffe, Anzahl unterschiedlicher Wörter und durchschnittliche Satzlänge). 5. Es werden Sätze markiert, die potenziell irrelevant sind, da sie inhaltlich nur eine geringe Übereinstimmung mit Sätzen des Originaltextes aufweisen. Auch gibt es eine Überprüfung der inhaltlichen Überlappung der Sätze eines Textes mit dem Ziel, redundante Inhalte zu entfernen. 6. Zum Schluss erhalten die Nutzer eine Rückmeldung über die inhaltliche Abdeckung der einzelnen Teilkapitel des Originaltextes in Form von Balkendiagrammen. Sie können nun entscheiden, ob sie ihren Entwurf überarbeiten möchten oder ob sie mit dem Ergebnis zufrieden sind.

Im Laufe der letzten Jahre gab es zahlreiche Laborstudien mit Studierenden der Psychologie sowie einzelne Studien an Schulen (Lenhard, 2008). Dabei stellte sich heraus, dass *conText* Nutzer in eine intensive Auseinandersetzung mit Texten verwickelt. Im Vergleich zu Studierenden, die einen Text ohne Feedback zusammenfassten, arbeiteten Studenten, die inhaltliche Rückmeldungen erhielten, länger an den Texten, überarbeiteten ihre Zusammenfassung häufiger, erzielten eine höhere Qualität der Zusammenfassung (erfasst mittels menschlicher Bewertung) und machten einen stärkeren Fortschritt vom ersten zum letzten Entwurf. In der gegenwärtigen Studie betteten wir *conText* in den regulären Unterricht ein, indem wir Schulbuchtexte des normalen Unterrichts in das System einspeisten. Die Texte waren ungefähr 150 Wörter lang.

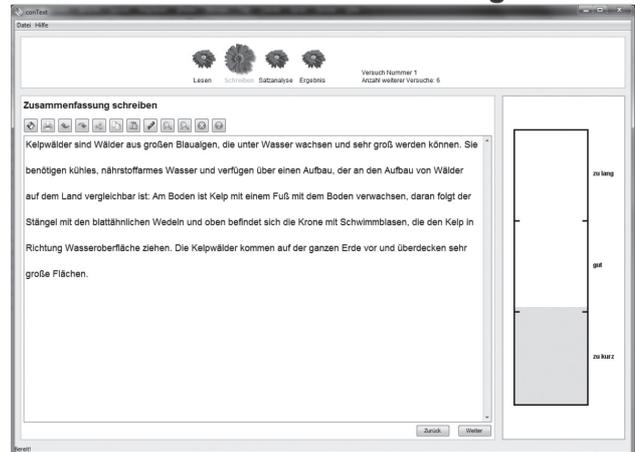
### 1.4 Wir werden Lesedetektive

Als Alternativbehandlung wurde das etablierte Förderprogramm «Wir werden Lesedetektive» (Rühl & Souvignier,

### Darstellung des Originaltextes



### Schreiben der Zusammenfassung



### Inhaltliche Rückmeldung

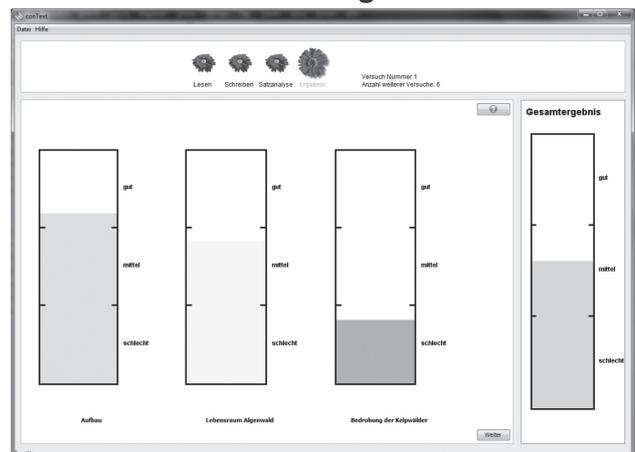


Abbildung 1. Die Schülerinnen und Schüler lesen einen Text, den sie in der Folge zusammenfassen. Nach verschiedenen Überprüfungsschritten erhalten sie eine Rückmeldung über die Inhaltsabdeckung der verschiedenen Passagen des Originaltextes sowie eine Gesamtbewertung.

2006; im Folgenden kurz *Lesedetektive*) eingesetzt, das das Ziel hat, explizites Lesestrategiewissen zu vermitteln. Die *Lesedetektive* stellen eine reduzierte Form des umfangreicheren Strategietrainings *Wir werden Textdetektive* (Gold et al., 2004) dar, das speziell an leistungsschwächere Schüler angepasst wurde. Es wird insbesondere für Schüler mit Lernbehinderungen und für Hauptschüler der 5. und 6. Klasse empfohlen, wobei jeweils Auffrischungstunden für das an die Intervention anschließende Schuljahr empfohlen werden. Beide Verfahren basieren auf dem Konzept von Paris und Jacobs (1984), das von Hasselhorn und Körkel (1986) für den deutschsprachigen Raum adaptiert wurde. Schreblowski (2004) überarbeitete den Ansatz für den Einsatz in adaptiven Kleingruppen in Schulen und ergänzte es um eine motivationale Komponente. Diese Vorlage wurde schließlich von Gold et al. (2004) im Rahmen einer formativen Evaluation zum strategieorientierten Unterrichtsprogramm *Wir werden Textdetektive* weiterentwickelt und von Rühl und Souvignier (2006) für leistungsschwache Schüler und Schülerinnen angepasst.

Bei den *Lesedetektiven* stehen vier zentrale Strategien im Mittelpunkt: a. Überschrift beachten, b. Umgang mit schwierigen Textstellen, c. Zusammenfassen von Erzähltexten und d. Zusammenfassen von Sachtexten. Die Strategie des Zusammenfassens umfasst also etwa die Hälfte der Instruktion des Programms. Die Inhalte sind in fünf Lerneinheiten gegliedert, die verteilt auf insgesamt ca. 25 Schulstunden vermittelt werden. Die *Lesedetektive* bestehen aus einem Lehrermanual (72 Seiten), das die Stundenkonzepte beinhaltet, sowie einem Arbeitsheft für die Schüler (40 Seiten) mit den konkreten Übungen und Materialien. Jeder Schüler und jede Schülerin arbeitet in einem eigenen Arbeitsheft. Die Lerneinheiten umfassen kurze Texte mit Fragen, lehrerzentrierte Instruktionen, Advance Organizers, Wiederholungen und Checklisten für die Überprüfung des eigenen Wissens. Jede Lerneinheit beginnt mit einer expliziten Erklärung zu einer bestimmten Strategie. Diese wird von der Lehrkraft im Klassenkontext demonstriert. In einer zweiten Phase arbeiten die Schüler und Schülerinnen mit ihrem Arbeitsheft. Dabei kommen Klassendiskussionen, Partnerarbeit und auch Einzelarbeit zum Einsatz. Zum Schluss diskutiert die Klasse Anwendungsprobleme und Rahmenbedingungen des Einsatzes einer bestimmten Strategie. Insgesamt umfasst die Intervention sechs Texte – drei narrative und drei expository – mit einer Länge von jeweils etwa 150 Wörtern, die von den Schülern im Laufe des Trainings mehrfach bearbeitet werden.

Die Textdetektive wurden in beispielhafter Art und Weise in unterschiedlichen Schulformen in den Klassenstufen 5 und 6 evaluiert (siehe Gold, Trenk-Hinterberger & Souvignier, 2009). An diesen Studien nahmen über 4000 Schüler und Schülerinnen teil. Im Hinblick auf das Strategiewissen zeigten sich moderate Effekte von  $d = .46$  im Posttest

und  $d = .36$  im Follow-Up nach sechs Monaten in Bezug auf die Gesamtstudie. Am stärksten profitieren Kinder aus Lernhilfeschulen ( $d = .74$ , bzw.  $d = .59$ ), wohingegen die schwächsten Effekte in Haupt-, Real- und Gesamtschule zu verzeichnen waren. Bezüglich des Leseverständnisses ergab sich ein Effekt  $d = .16$  im Posttest und  $d = .27$  im Follow-Up, wiederum mit ausgeprägteren Effekten in der Lernhilfeschule, aber Nulleffekten in Haupt-, Real- und Gesamtschule. Neuere Untersuchungen in Gymnasien (Souvignier & Trenk-Hinterberger, 2010) zeigen, dass für eine Nachhaltigkeit der Effekte des Programms und insbesondere auch für eine stärkere Verbesserung des Leseverständnisses Auffrischungssitzungen notwendig sind.

Die Wirksamkeit des Programms *Wir werden Lesedetektive* wurde an Schülern mit Lernbehinderung untersucht (Antonou & Souvignier, 2007). Direkt im Anschluss an die Förderung zeigte sich zwischen Treatment- und Kontrollgruppe lediglich ein Effekt für das Lesestrategiewissen, aber nicht für das Leseverständnis. Jedoch nahmen nach der Untersuchung die Effekte weiter zu, sodass beim Follow-Up vier Monate später auch ein positiver Effekt für das Leseverständnis zu beobachten war. Untersuchungen der *Lesedetektive* an Hauptschulen gab es unseres Wissens bislang nicht.

Vergleicht man die *Lesedetektive* mit *conText*, so gibt es eine Reihe an Gemeinsamkeiten und eine Reihe an Unterschieden (Tab. 1): Beide Verfahren fokussieren hauptsächlich auf die Strategie des Zusammenfassens. Diese umfasst bei den *Lesedetektiven* 50% des Inhalts. Bei *conText* geht es ausschließlich um das Training des Zusammenfassens. Beide Verfahren verwenden Texte der gleichen Länge, unterscheiden sich aber in der Anzahl der Texte (sechs Texte in den *Lesedetektiven* versus 11 bis 13 Texte bei *conText*). In den *Lesedetektiven* werden Strategien direkt vermittelt, wohingegen sie in *conText* implizit durch gelenktes Üben trainiert werden. In den *Lesedetektiven* gibt es als Arbeitsformen das Unterrichtsgespräch, Partner- sowie Einzelarbeit, wohingegen in *conText* Schüler und Schülerinnen einzeln am Rechner arbeiten und individualisiertes Feedback erhalten. Es lässt sich also feststellen, dass beide Interventionen in Bezug auf die Zielsetzungen und Inhalte Ähnlichkeiten aufweisen, sich aber in der Herangehensweise und der Form deutlich unterscheiden.

## 2 Fragestellung

Ziel der Studie war die Klärung der Frage, ob sich das computerbasierte Training *conText* positiv auf das Leseverständnis und auf verschiedene Determinanten des Leseverständnisses (verbale Intelligenz, Leseflüssigkeit und Lesestrategiewissen) auswirkt<sup>1</sup>. Neben dem Vergleich mit einer unbehandelten Kontrollgruppe interessierte uns der Vergleich mit den Effekten etablierter Leseverständnisin-

1 Bei der Wahl der Determinanten wurde auf das Modell von Schaffner (2009) zurückgegriffen und ein Teil der von ihr beschriebenen individuellen Lernermerkmale berücksichtigt. Nicht einbezogen wurden das themenspezifische Vorwissen, das Leseinteresse und die Lesehäufigkeit.

Tabelle 1  
Vergleich der Untersuchungsbedingungen

	Kontrollgruppe	Lesedetektive	conText
Stichprobe	6. Klasse Hauptschule	6. Klasse Hauptschule	6. Klasse Hauptschule
Intervention	Regulärer Unterricht	Direkte Vermittlung expliziten Lesestrategiewissens	Training von Textzusammenfassungen mittels automatisch generierter Rückmeldungen
Vermittelte Strategien		(1) Überschrift beachten (2) Umgang mit schwierigen Textstellen, (3, 4) Zusammenfassen von Erzähl- und Sachtexten	Zusammenfassen von Texten
Anzahl investierter Unterrichtseinheiten		20.0	17.8
Textarten		Erzähl- und Sachtexte	Sachtexte
Textquelle		Übungstexte des Lesedetektive-Programms	Reguläre Schulbuchtexte der teilnehmenden Klassen
Länge der Texte		etwa 150 Wörter pro Text	etwa 150 Wörter pro Text
Anzahl an Texten		6	11 bis 13
Arbeitsformen		Unterrichtsgespräch, Einzel- und Partnerarbeit	Einzelarbeit am Computer mit individualisiertem Feedback

erventionen. Wir entschieden uns trotz der prinzipiellen Unterschiede für den Einsatz der Lesedetektive als Alternativbehandlung, da es sich um ein etabliertes, geschlossenes Programm handelt, das für die Anwendung im realen Unterricht für die Zielgruppe zur Verfügung stand. Da die Interventionen im Rahmen des regulären Schulunterrichts von den Lehrkräften selbst appliziert wurden, handelt es sich um eine Evaluation der Programme in einem ökologisch validen Setting. Die Analyse einzelner Wirkfaktoren ist eingedenk der prinzipiellen Unterschiede der Programme auf diese Weise allerdings nicht möglich.

Wir erwarteten, dass sich *conText* durch das intensive Üben mit Texten auf die Leseflüssigkeit und das Leseverständnis auswirkt, verbale Intelligenz und Lesestrategiewissen jedoch nicht profitieren. Wir vermuteten eine Verbesserung des Lesestrategiewissens und des Leseverständnisses durch das Training mit den *Lesedetektiven*. Im direkten Vergleich zwischen beiden Interventionen gingen wir von einem Vorteil der *Lesedetektive* in Bezug auf die Verbesserung des Lesestrategiewissens aus. Aufgrund des höheren Anteils an Textarbeit nahmen wir einen Vorteil von *conText* in Bezug auf die Leseflüssigkeit an.

### 3 Methode

#### 3.1 Stichprobe

Unser Ziel war es, leistungsschwächere Jugendliche zu fördern, da hier eine höhere praktische Bedeutsamkeit besteht als bei leistungsstarken Jugendlichen. Aus der PISA-Studie ist bekannt, dass insbesondere in der Hauptschule der Anteil an schwachen Lesern und Lesern stark erhöht ist. In der Kohorte von 2009 erreichten beispielsweise

81,3% der 15-jährigen Jugendlichen an Hauptschulen auf der sechsstufigen Kompetenzskala ein Niveau von II oder darunter, wohingegen dieser Anteil an Gymnasien lediglich 6,7% betrug (nach Naumann, Artelt, Schneider & Stanat, 2010, S. 57, Abb. 2.11). Ein Niveau von IV oder höher erreichten unter den Hauptschülern dagegen lediglich 2,6%, im Gegensatz zu 64,9% an den Gymnasien. Aus diesem Grund führten wir die Studie in Hauptschulen aus und wählten als Klassenstufe den 6. Jahrgang.

Im Zuge der Rekrutierung fiel auf, dass die Schülerzahlen an Hauptschulen stark im Zurückgehen begriffen waren, sodass zum einen nur kleine Klassen auffindbar waren und zum anderen lediglich zwei Schulen zwei parallele Klassen pro Klassenstufe hatten, alle anderen dagegen nur jeweils eine Klasse. Im Schnitt befanden sich in den Klassen 15.29 Jugendliche ( $SD = 3.45$ ). Trotz der Gesamtzahl von 14 Klassen in der Untersuchung betrug die Stichprobengröße deshalb insgesamt nur 226 Schüler und Schülerinnen. Die Datenerhebung fand in Nordbayern statt. Es nahmen ausschließlich ländliche Hauptschulen teil.

#### 3.2 Genehmigungen

Vor der Untersuchung wurde die Genehmigung der zuständigen Schulbehörden eingeholt. Die Teilnahme erfolgte für Schulen, Klassen und Jugendliche auf freiwilliger Basis und Schüler und Eltern wurden vorweg über die Ziele und Inhalte informiert. In der Kontrollgruppe erhielten zwei Jugendliche nicht die Zustimmung der Eltern. Von diesen Jugendlichen wurden keine Daten erhoben. Davon abgesehen erteilten alle anderen Eltern ihre schriftliche Teilnahmegenehmigung.

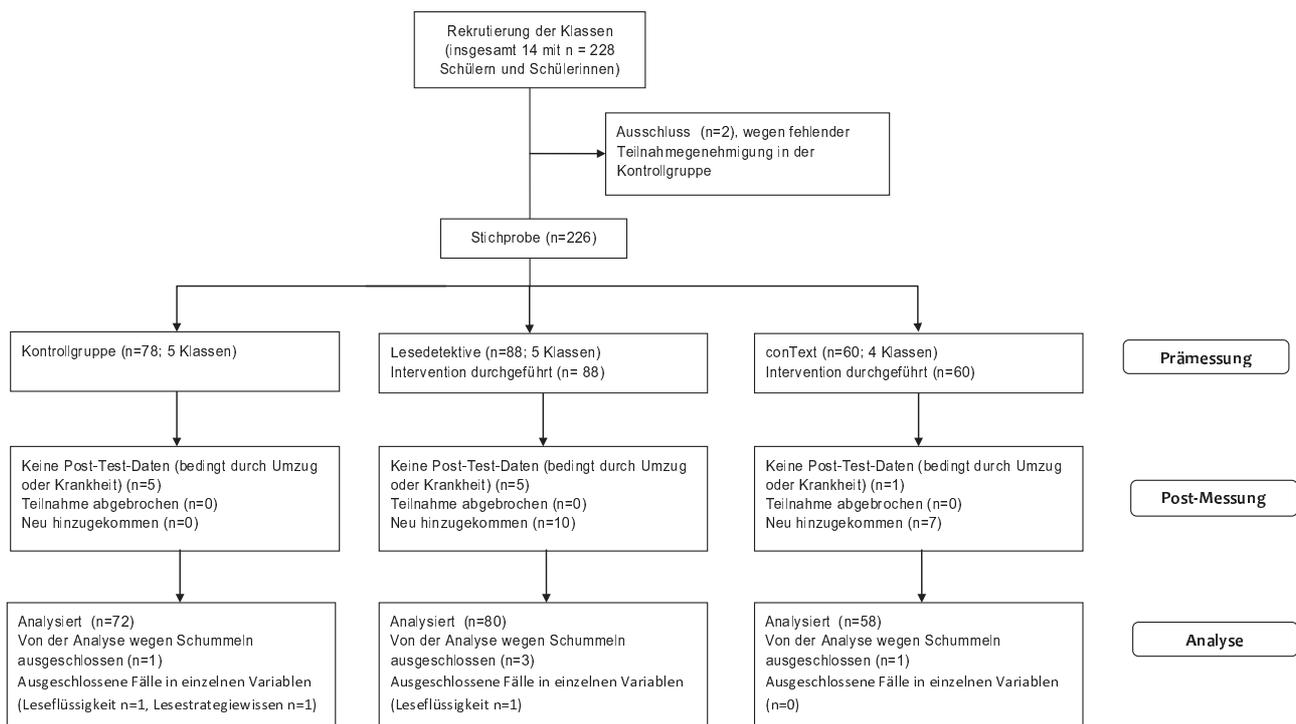


Abbildung 2. Consort Diagramm des Untersuchungsverlaufs. Von den anfänglichen 228 Schülerinnen und Schülern nahmen 210 an Prä- und Posttest und der Intervention teil.

### 3.3 Studiendesign und Ablauf

In der Untersuchung gab es zwei Interventionsgruppen (*Lesedetektive* und *conText*) und eine unbehandelte Kontrollgruppe. Die Rekrutierung der beiden Interventionsgruppen erfolgte zu Beginn des Schuljahres 2009/2010, die der Kontrollgruppe zu Beginn des Schuljahres 2010/2011. Jeweils im Oktober fand ein Prätest statt. Daran schloss sich bei den Interventionsgruppen im Zeitraum von Anfang November bis Ende Juni die Förderung an. Ende Juli wurde in allen Gruppen die Postmessung erhoben. Die Erhebung der Kontrollgruppe war also zwar um ein Jahr versetzt, wies ansonsten aber den gleichen Zeitraum zwischen Prä- und Postmessung auf wie die Interventionsgruppen.

Während des Schuljahres durchliefen die Kinder entweder die explizite Strategieinstruktion *Wir werden Lesedetektive* oder das implizite computerbasierte Training *conText*. Die Lehrkräfte der *Lesedetektive*-Klassen führten das Programm entsprechend des Handbuchs durch und wurden dazu angehalten, die vermittelten Strategien im Unterricht weiter anzuwenden. Für das Training mit *conText* griffen wir auf reguläre Schulbuchtexte aus den Bereichen Erdkunde, Biologie und Geschichte zurück und bereiteten diese für den Einsatz im Rahmen der webbasierten Version von *conText* auf. Beide Interventionen wurden im Rahmen des Deutschunterrichts durchgeführt. Die Kontrollgruppe nahm am regulären Unterricht teil. Eine Lehrerin der Kontrollgruppe berichtete, zusätzlich ein alter-

natives Lesetraining angewandt zu haben (*STIKK*; Sigel & Feneberg, 2007), das angibt, Lesetechniken, Leseflüssigkeit, Augenbewegungen und Konzentration zu fördern. Davon abgesehen erhielten die Kinder der Kontrollgruppe nur die regulär vorgesehene Leseförderung im Rahmen des Schulunterrichts.

Zur Untersuchung der Effekte der Interventionen wandten wir ein Prä-Post-Design an. Von den 226 Schülern der Prämessung nahmen 215 ebenfalls an der Postmessung teil, die anderen 11 Jugendlichen waren entweder am Posttest krank oder hatten die Schule verlassen. Dies entspricht einer Drop-Out-Quote von 5,1 %. Kein Schüler hatte im Laufe des Jahres seine Teilnahme zurückgezogen. 5 Datensätze mussten aus der Analyse entfernt werden, weil die betreffenden Jugendlichen beim Schummeln erwischt wurden. In den Interventionsgruppen kamen über das Jahr hinweg 17 Schülerinnen und Schüler hinzu. Sie nahmen an den Interventionen teil, wurden aufgrund des Fehlens des Prätests jedoch nicht in der Analyse berücksichtigt. Insgesamt standen für die Analyse 210 Datensätze zur Verfügung (Tab. 2). Die unterschiedliche Anzahl der Schüler pro Untersuchungsbedingung ist auf die unterschiedliche Anzahl an Klassen sowie auf unterschiedlich große Klassenstärken zurückzuführen. Signifikante Unterschiede in Abhängigkeit von der Untersuchungsbedingung gab es weder hinsichtlich des Drop-Out noch hinsichtlich der Klassengröße.

Tabelle 2  
Stichprobenkennwerte

Gruppe	N	Alter M (SD)	Geschlecht (Jungen/Mädchen)	Migrations- hintergrund (%)	Lese- Rechtschreibstörung
Kontrollgruppe	72	12.0 (.6)	38/34	12	14
Lesedetektive	80	12.0 (.6)	43/37	3	15
conText	58	11.9 (.6)	34/24	3	11

Anmerkungen: Der Altersunterschied wurde mittels einer einfaktoriellem ANOVA geprüft. Bei der Untersuchung der anderen Variablen wurden  $\chi^2$ -Tests angewandt. Mit Ausnahme des Anteils an Schülerinnen und Schülern mit Migrationshintergrund waren die Unterschiede nicht signifikant.

### 3.4 Durchführung der Untersuchungen

Alle teilnehmenden Klassen bekamen die verwendeten Materialien kostenlos zur Verfügung gestellt. Vor dem Beginn der Untersuchung erhielt jede Lehrkraft eine individuelle Einführung in die jeweils angewandten Interventionsverfahren sowie das Angebot für umfangreichere Fortbildungen und tutorielle Hilfe bei der Umsetzung. Die beteiligten Lehrkräfte wurden in regelmäßigen Abständen kontaktiert, um zu fragen, ob zusätzliche Hilfe benötigt würde. Eine Schule der *Lesedetektive*-Gruppe nahm daraufhin einen vierstündigen Workshop zum Thema Leseverständnisdiagnostik und -förderung in Anspruch, ansonsten arbeiteten die Lehrkräfte mit den vorbereiteten Stundenentwürfen des *Lesedetektive*-Manuals. In der *conText*-Gruppe nutzten die Lehrkräfte keine Weiterbildungsangebote, forderten aber im Schnitt vier Mal tutorielle Hilfe bei der Anwendung des Programms in der Klasse an. Die Hilfe bezog sich ausschließlich auf Anwendungsaspekte (Login in der Webapplikation, Starten der Aufgabe etc.).

Nach Abschluss der Untersuchung fand eine Nachbefragung der Lehrkräfte statt, die sich auf die Anwendbarkeit der Programme bezog. Alle Lehrkräfte der *conText*-Gruppe gaben an, dass – abgesehen von anfänglichen Problemen beim Einloggen – das Arbeiten mit dem Programm gut funktioniert habe. Zwei Lehrkräfte der *Lesedetektive*-Gruppe empfanden die verwendeten Texte als sehr leicht, aber prinzipiell für die Hauptschule geeignet. Eine Lehrkraft gab an, dass die Detektivkärtchen eine gute Stütze beim Erarbeiten von Texten seien. Alle Lehrkräfte der *Lesedetektive*-Gruppe bestätigten, das Verfahren komplett angewandt zu haben.

### 3.5 Abhängige Maße

Neben den personenbezogenen Daten wurden die verbale Intelligenz, die Leseflüssigkeit, das Lesestrategiewissen und das Leseverständnis mittels standardisierter Verfahren erfasst.

Die Erfassung der *verbalen Intelligenz* erfolgte mit dem Subtest «Analogien» aus dem *KFT 4-12+* (Heller & Perleth, 2000). Bei diesem Verfahren muss die Beziehung zwischen den beiden Wörtern eines Paares erfasst werden,

und auf ein weiteres Paar übertragen werden. Die Reliabilität dieses Subtests beträgt  $r_{tt} = .81$ . Der Test enthält 20 Items. Die Retestreliabilität lag in der unbehandelten Kontrollgruppe dieser Untersuchung bei einem Intervall von 8 Monaten bei  $r_{tt} = .56$ . Die folgenden Angaben der Retestreliabilität beziehen sich ebenfalls auf die Kontrollgruppe.

Die *Leseflüssigkeit* wurde mit dem *Salzburger Lesescreening 5-8 (SLS 5-8; Auer, Gruber, Mayringer & Wimmer, 2005)* erhoben. Die Aufgabe besteht bei diesem Test darin, 67 Aussagen danach zu bewerten, ob sie richtig oder falsch sind (z.B. «Wenn jemand einen Brief in eine Flasche steckt, sie verschließt und ins Meer wirft, so ist das eine Frischhaltefolie»). Der Test misst also de facto nicht allein Leseflüssigkeit, sondern erhebt auch basale Verständnisleistungen. Die Aussagen sind nach Schwierigkeit angeordnet und die Reliabilität wird mit  $r_{tt} = .89$  angegeben. In dieser Untersuchung lag sie nach 8 Monaten bei  $r_{tt} = .76$ .

Der *Lesestrategie-Wissenstest* (Lingel, Neuenhaus, Artelt & Schneider, 2010) schildert fünf Lernsituationen, zu denen jeweils fünf bzw. sechs Vorschläge zur deren Gestaltung gemacht werden. Aufgabe ist es zu bewerten, wie sinnvoll diese Vorschläge sind. Die Einschätzungen werden mit Expertenbewertungen verglichen. Es sind maximal 36 Punkte möglich. Die interne Konsistenz dieses Verfahrens wird mit  $\alpha_{cr} = .89$  angegeben. In dieser Untersuchung lag die Retestreliabilität im Abstand von 8 Monaten bei  $r_{tt} = .34$ .

Zur Messung des Leseverständnisses setzten wir den *Frankfurter Leseverständnistest 5-6 (FLVT 5-6; Souvignier, Trenk-Hinterberger, Adam-Schwebe & Gold, 2008)* ein. Dieser besteht aus zwei Texten mit je 18 Multiple-Choice-Fragen. Insgesamt ist folglich ein maximaler Punktwert von 36 möglich. Der Gesamttest hat eine interne Konsistenz von  $\alpha_{cr} = .87$  und eine Paralleltestreliabilität von  $r_{tt} = .71$ . In der Prämessung wurde Version A des FLVT 5-6 verwendet, in der Postmessung Version B. Da Hauptschüler bei Version B des Tests im Schnitt 2.93 Punkte weniger erreichen (E. Souvignier, persönliche Mitteilung, 07.10.2010), wurde das Ergebnis im Leseverständnis in der Postmessung um diese Konstante korrigiert. Die Retestreliabilität in dieser Untersuchung lag im Abstand von 8 Monaten bei  $r_{tt} = .58$ .

Mit Ausnahme des SLS 5-8 hatte kein Untertest eine strenge Zeitbegrenzung. Die gesamte Testung nahm ca. 70 Minuten in Anspruch. Die Testungen fanden vorzugsweise in der 3. und 4. Stunde statt.

### 3.6 Statistische Analyse der Daten

Da es sich um eine Interventionsstudie handelt, entscheiden wir uns für einen konservativen Umgang mit fehlenden Werten. Fälle mit fehlender Postmessung wurden komplett aus der Analyse genommen (Angaben zu fehlenden Werten siehe Abb. 2). Im Falle fehlender Werte in einzelnen abhängigen Variablen wurden die betreffenden Fälle aus den jeweiligen Analysen ausgeschlossen. Dies betraf in der Kontrollgruppe je einen Fall bei der Leseflüssigkeit und beim Lesestrategiewissen und in der *Lesedetektive*-Gruppe einen Fall bei der Leseflüssigkeit. Die Analyse der Daten erfolgte mit  $\chi^2$ -Tests (Geschlecht, Migrationsstatus und Auftreten von Lese-Rechtschreibstörungen) und ANOVAs (Alter) zur Ermittlung von Unterschieden zwischen den Gruppen. Unterschiede zwischen Prä- und Postmessung in den abhängigen Leistungsmaßen wurde mit ANCOVAs untersucht (einseitige Testung). Post-Hoc-Tests für paarweise Vergleiche zwischen den Untersuchungsgruppen wurden ebenfalls mittels ANCOVAs durchgeführt, wobei jeweils Zweier-Kombinationen von Gruppen in die Analyse eingingen. Auf diese Weise war es möglich, über die Schätzung der aufgeklärten Varianz auf die Effektgröße zu schließen. Die Umrechnung von  $\eta^2$  in  $d$  erfolgte nach Cohen (1988, S. 281 f.). In den Post-Hoc-Tests wurde zur Verhinderung der Alpha-Inflation das Alpha-Niveau mittels der Bonferroni-Holm-Prozedur adjustiert.

### 3.7 Nutzungshäufigkeit und Nachbefragung der Lehrkräfte

Detaillierte Nutzungshäufigkeiten in der *conText*-Gruppe standen über die Log-Dateien des Webservers zur Verfügung. Drei Klassen in der *conText*-Gruppe bearbeiteten zwischen 11 und 13 verschiedene Texte. Eine kleinere Klasse, bestehend aus 14 Schülerinnen und Schülern, bearbeitete lediglich drei Texte. Wir schlossen aus den Nutzerstatistiken alle Überarbeitungszyklen aus, bei denen ersichtlich war, dass die Schülerinnen und Schüler sich nicht intensiv mit dem Text befasst hatten (weniger als 10 geänderte Buchstaben oder Bearbeitungszeiten unter 30 Sekunden). Im Schnitt überarbeiteten die Schüler der *conText*-Gruppe ihre Zusammenfassungen  $M = 5.7$  Mal ( $SD = 4.9$ ) und investierten pro Text  $M = 31.0$  Minuten reine Bearbeitungszeit ( $SD = 9.3$ ; ohne abgebrochene Zyklen).

Weitere Nutzungsdaten wurden durch Befragung der Lehrkräfte im Anschluss an die Untersuchung erhoben. Die Lehrkräfte der *Lesedetektive*-Gruppe gaben an, durchschnittlich 20.0 Unterrichtseinheiten á 45 Minuten auf das Training verwendet zu haben. In der *conText*-Gruppe belief sich die investierte Anzahl an Unterrichtseinheiten im Mittel auf 15.8 (reine Arbeitszeit) bzw. 17.5 (real aufgewendete Unterrichtsstunden). Im Schnitt investierten die Klassen im Untersuchungszeitraum alle zwei Wochen eine Unterrichtseinheit auf die Trainings. Beide Interventions-

gruppen unterscheiden sich bezüglich der Anzahl aufgewendeter Unterrichtsstunden nicht signifikant.

## 4 Ergebnisse

### 4.1 Analyse von Prätestunterschieden

Die verschiedenen Untersuchungsgruppen unterschieden sich nicht bezüglich der Geschlechtsverteilung ( $p = .73$ ), des Alters zum Zeitpunkt der Prämessung ( $p = .46$ ) und der Häufigkeit des Auftretens einer Lese-Rechtschreibstörung ( $p = .50$ ). Das Geschlechtsverhältnis war in allen Gruppen leicht zugunsten der Jungen ausgeprägt (54.2 %) und der Anteil an Jugendlichen mit einer selbst berichteten Lese-Rechtschreib-Störung (LRS) war mit 19.0 % deutlich erhöht. Dieser Wert ist erwartungskonform, da es sich um eine selektierte Stichprobe aus dem leistungsschwachen Bereich handelt.

Einen signifikanten Unterschied gab es im Hinblick auf den Anteil an Jugendlichen mit Migrationshintergrund, der in der Kontrollgruppe höher ausfiel als in den Interventionsgruppen,  $\chi^2(2) = 9.7$ ,  $p < .05$ . Dies war auf einen unterdurchschnittlichen Anteil an Kindern mit Migrationshintergrund in den beiden Interventionsgruppen zurückzuführen, der dort nur bei 4.3 % lag. In der Kontrollgruppe betrug der Anteil dagegen 16.7 %. Dass der Anteil in den beiden Interventionsgruppen extrem niedrig ausfällt, könnte gegebenenfalls darauf zurückzuführen sein, dass in diesen beiden Gruppen nicht alle Schüler mit Migrationshintergrund die Familiensprache korrekt berichteten. Nach Aussagen der Lehrkräfte gab es keine bedeutsame Veränderung der Zusammensetzung der Schülerschaft in beiden Schuljahren. Zwischen den beiden Interventionsgruppen gab es keinen signifikanten Unterschied. Um zu entscheiden, ob die Variable «Migrationshintergrund» als Kovariate in die Hauptanalyse einfließen sollte, wurde für jede abhängige Variable eine ANCOVA mit der Muttersprache als unabhängige Variable berechnet. Die Postmessung ging jeweils als abhängige Variable ein und der Wert der Prämessung als Kovariate. In keiner abhängigen Variablen zeigte sich bei Kontrolle des Ausgangswertes ein signifikanter Gruppenunterschied, sodass auf die weitere Einbeziehung des Migrationshintergrundes verzichtet wurde. Auch änderte die Einbeziehung des Migrationshintergrundes als Kovariate die Interpretation der Ergebnisse aus Abschnitt 4.2 nicht.

Zur Ermittlung signifikanter Gruppenunterschiede in den Prätest-Ergebnissen von verbaler Intelligenz, Leseflüssigkeit, Lesestrategiewissen und Leseverständnis wurde eine MANOVA mit dem Faktor Untersuchungsgruppe durchgeführt. Es zeigte sich kein signifikanter Gruppenunterschied, weder im Gesamttest, noch in den einzelnen Variablen, sodass von einer initialen Äquivalenz der Gruppen im Hinblick auf die erhobenen Leistungsmaße ausgegangen werden kann. Bei allen varianzanalytischen Tests wurde auf Homogenität geprüft. Es lagen in keiner

Variable Hinweise für Verletzungen der Homogenitätsannahmen vor. In allen abhängigen Variablen lagen die Intraklassenkorrelationen in der Prämessung bei  $\rho \leq .10$ , was als Hinweis für eine Auswertbarkeit der Daten auf Individualebene gewertet werden kann.

### 4.2 Effekte der Lesetrainings

Die *verbale Intelligenz* veränderte sich während des Förderzeitraums in den drei untersuchten Gruppen nicht unterschiedlich.

Im Hinblick auf Veränderungen der *Leseflüssigkeit* zeigte sich ein deutlicher Haupteffekt der Gruppe,  $F(2, 204) = 6.03, p < .01$ . Sowohl die *conText*-Gruppe,  $F(1, 126) = 11.24, p = .001, \eta^2 = .082, d = .60$ , als auch die *Lesedetektive*-Gruppe,  $F(1, 147) = 4.24, p = .024, \eta^2 = .028, d = .34$ , schnitten signifikant besser ab als die Kontrollgruppe. Außerdem zeigte sich nach der Korrektur nach Bonferroni-Holm ( $\alpha = .025$ ) eine Überlegenheit von *conText* gegenüber den *Lesedetektiven*,  $F(1, 134) = 2.86, p = .045, \eta^2 = .021, d = .29$ .

Auch bei der Veränderung des Lesestrategiewissens lag ein signifikanter Gruppenunterschied vor,  $F(2, 205) = 6.47, p < .01$ , wobei sowohl die *Lesedetektive*-Gruppe,  $F(1, 147) = 4.24, p = .001, \eta^2 = .064, d = .52$ , als auch die *conText*-Gruppe,  $F(1, 126) = 7.60, p = .003, \eta^2 = .057, d = .49$ , sich stärker verbesserten als die Kontrollgruppe. Die Veränderungen in der *Lesedetektive*- und der *conText*-Gruppe unterschieden sich dagegen nicht voneinander ( $p = .75$ ).

Beim Leseverständnis zeigte sich ebenfalls ein signifikanter Effekt der Gruppe,  $F(2, 206) = 5.19, p < .01$ , der vor allem auf eine Überlegenheit der *conText*-Gruppe gegenüber der Kontrollgruppe,  $F(1, 127) = 10.86, p < .001, \eta^2 = .079, d = .59$  zurückzuführen war. Numerisch bestand außerdem eine Überlegenheit der *conText* – gegenüber der *Lesedetektive*-Gruppe, die allerdings aufgrund der Bonferroni-Holm-Korrektur ( $\alpha_{\text{Bonf.-Holm}} = .025$ ) nur marginale Signifikanz erreichte,  $F(1, 135) = 3.48, p = .032, \eta^2 = .025, d = .32$ . Die *Lesedetektive*-Gruppe unterschied sich von der Kontrollgruppe nur deskriptiv. Die deskriptiven Statistiken und Effektstärken finden sich in Tabelle 3 und 4.

## 5 Diskussion

Ziel der Untersuchung war es, die Effektivität der computerbasierten Leseverständnisintervention *conText* zu überprüfen und die Effekte mit einer Alternativbehandlung (*Lesedetektive*) und einer unbehandelten Kontrollgruppe zu vergleichen.

Bei Förderung mithilfe des computerbasierten Trainings *conText*, bei dem Schüler das Zusammenfassen von Texten üben und dabei Rückmeldung über die Qualität der Zusammenfassung vom Computer erhalten, verbes-

Tabelle 3  
Ergebnisse im Prä- und Posttest nach Untersuchungsgruppe

	Kontrollgruppe (N = 72)			Lesedetektive (N = 80)			conText (N = 58)		
	Prä M(SD)	Post M(SD)	Diff M(SD)	Prä M(SD)	Post M(SD)	Diff M(SD)	Prä M(SD)	Post M(SD)	Diff M(SD)
Verbale Intelligenz	9.7 (3.8)	10.3 (4.1)	.6 (3.7)	9.2 (3.7)	10.7 (3.5)	1.5 (3.3)	9.9 (3.2)	10.4 (3.3)	.5 (3.0)
Leseflüssigkeit	31.5 (7.1)	33.7 (7.3)	2.2 (5.0)	32.0 (6.0)	35.7 (6.9)	3.6 (4.2)	30.3 (6.3)	35.3 (6.6)	5.0 (3.6)
Lesestrategiewissen	21.0 (4.8)	21.5 (6.1)	.3 (6.3)	21.7 (5.1)	24.5 (5.2)	2.7 (6.5)	21.6 (5.4)	24.2 (4.8)	2.6 (5.4)
Leseverständnis	22.1 (5.6)	23.3 (5.3)	1.2 (5.0)	22.2 (5.9)	24.4 (5.3)	2.2 (5.3)	23.0 (5.5)	26.2 (4.6)	3.1 (4.9)

Anmerkungen: Die Tabelle berichtet die Prä- und Posttestergebnisse sowie die Differenzwerte der drei Untersuchungsgruppen in den abhängigen Maßen verbale Intelligenz, Leseflüssigkeit, Lesestrategiewissen und Leseverständnis.

Tabelle 4  
Effektstärken der Treatments im Vergleich

	Kontrollgruppe (Retest)	Lesedetektive (Retest)	conText (Retest)	Lesedetektive versus Kontrollgruppe $\eta^2/d$	conText versus Kontrollgruppe $\eta^2/d$	conText versus Lesedetektive $\eta^2/d$
Verbale Intelligenz		.45				
Leseflüssigkeit	.44	.86	1.39	.028/.34	.082/.60	.021/.29
Lesestrategiewissen		.42	.48	.064/.52	.057/.49	
Leseverständnis	.24	.42	.63		.079/.59	.025/.32

Die Tabelle zeigt den Anteil aufgeklärter Varianz  $\eta^2$  und die Effektstärke  $d$  der Unterschiede zwischen Prä- und Postmessung für signifikante Unterschiede innerhalb der Gruppen, sowie für signifikante und marginal signifikante Unterschiede zwischen den Untersuchungsgruppen. Die Berechnung der Effektstärken erfolgte nach Cohen (1988, S. 281 f.) auf der Basis des aufgeklärten Anteils an Varianz, bzw. innerhalb der Gruppen anhand der Folgedifferenzen aus Tabelle 2. Die ersten drei Spalten beziehen sich auf den Vergleich der Ergebnisse innerhalb der einzelnen Gruppen und sind z. T. auf Retesteffekte zurückzuführen.

serte sich das Leseverständnis während des Förderzeitraums am stärksten. Darüber hinaus zeigte sich auch ein bedeutsamer Zuwachs an Leseflüssigkeit und – entgegen den Erwartungen – auch an Lesestrategiewissen. In der *Lesedetektive*-Gruppe wurde explizit Lesestrategiewissen vermittelt. Es zeigte sich im Vergleich zur unbehandelten Kontrollgruppe eine deutliche Verbesserung des Lesestrategiewissens und der Leseflüssigkeit. Im Hinblick auf das Leseverständnis profitierten die Schülerinnen und Schüler hingegen gegenüber der Kontrollgruppe nicht bedeutsam. Im direkten Vergleich zwischen den beiden Interventionsmaßnahmen erwies sich *conText* bezüglich der Leseflüssigkeit gegenüber «Wir werden Lesedetektive» überlegen, das Lesestrategiewissen wurde von beiden Programmen gleichermaßen verbessert. Im Hinblick auf das Leseverständnis zeigte sich eine numerische Überlegenheit von *conText*, die allerdings nur marginale Signifikanz erreicht. *conText* verbessert das Leseverständnis folglich mindestens so stark wie «Wir werden Lesedetektive», eventuell sogar stärker.

## 5.1 Mögliche Wirkmechanismen von conText

Unter den Lesestrategien gelten insbesondere das schriftliche Zusammenfassen oder das mündliche Nacherzählen als zentrale Ansätze. Das Wiedergeben eines Textinhalts in eigenen Worten erfordert die Verknüpfung mit Hintergrundwissen und die Unterscheidung zwischen zentralen Inhalten und Details. Das Zusammenfassen zwingt den Leser dazu, den eigenen Leseprozess aufmerksamer zu überwachen und den Text besser zu organisieren. Es basiert im Wesentlichen auf denselben Prozessen, die Leser für den Aufbau eines Situationsmodells benötigen. Auf diese Weise entsteht ein vollständigeres, strukturiertes und stärker verdichtetes Abbild des Textinhaltes, das besser mit der eigenen Vorwissensbasis verknüpft ist (vgl.

Kintsch et al., 2007; McNamara, Ozuru, Best & O'Reilly, 2007, S. 489 f.).

Beide Interventionsgruppen vermitteln diese Strategie, allerdings auf unterschiedliche Art und Weise. Zumindest tendenziell ist die Herangehensweise bei *conText* jedoch effektiver. Wir gehen davon aus, dass hierzu vor allem zwei Faktoren beigetragen haben könnten: Erstens wurden inhaltlich relevante Texte verwendet. Die Schülerinnen und Schüler arbeiteten an regulären Texten, die aufgrund der Einbettung in andere Fächer und aufgrund des gegebenen curricularen Bezugs für die Schüler sinnhaft und bedeutsam waren. Ob Gruppenunterschiede zwischen *conText*- und *Lesedetektive*-Gruppe bei unterrichtsfremden Inhalten ebenfalls aufgetreten wären, ist aus unserer Sicht allerdings nicht auszuschließen. Unterstützung erhält diese Ansicht aus den Ergebnissen der Interventionsforschung: Insbesondere jene Programme, die einen starken Fokus auf Inhalte und Lesemotivation aufweisen und Lesestrategien lediglich als eine Facette im Rahmen der Leseerziehung einsetzen (z. B. CORI; Guthrie et al., 2004), erweisen sich als besonders effektiv. Interventionen, die nicht in Bezug zu den konkret schulisch vermittelten Inhalten stehen und isoliert auf die explizite Vermittlung von Strategiewissen abzielen, zeigen dagegen in Bezug auf die Leseverständnisleistung keine oder geringe Effekte. So war beispielsweise in der ursprünglichen Untersuchung von Paris et al. (1984, S. 1245 f.) – dem «Urvater» der Lese- und Textdetektive – kein signifikanter Effekt des Trainings auf die Leseverständnisleistung zu finden. Slavin et al. (2008, S. 308) berichten für diese Art von Intervention ebenfalls nur kleine, positive Effekte von  $d = .09$ . Ein starker Fokus auf die pointierte Vermittlung von Lesestrategien im Rahmen geschlossener Unterrichtskonzepte verfehlt deshalb womöglich das eigentliche Ziel – die Verbesserung schulischer Leistungen (siehe auch Klauer, 2010).

Wir vermuten, dass in der Entwicklung von Leseverständnisinterventionen stärker auf die Verschränkung mit schulischen Inhalten geachtet werden muss, und bei geschlossenen, vorgefertigten Interventionen deshalb prinzi-

piell nur geringe Effekte zu erwarten sind. In ihrer Studie zu inhalts- versus strategiebasierten Intervention kommen McKeown, Beck und Blake (2009, p. 246) zu einem ähnlichen Schluss:

“Strategies questions did prompt students to bring key ideas into the discussion but ... students spent as much time focusing on the strategic actions themselves as on the content of what they were reading and seemed less likely to connect the ideas. The importance of making connections among ideas is paramount. Focusing on strategies during reading may leave students less aware of the overall process of interacting with text, especially in terms of the need to connect ideas they encounter and integrate those ideas into a coherent whole.”

Der zweite Faktor, der zur Effektivität von *conText* beigetragen haben könnte, ist die unmittelbare inhaltliche Rückmeldung durch das Computerprogramm. Diese erzeugte möglicherweise eine höhere Motivation, sich intensiv mit den Textinhalten auseinanderzusetzen. Auf der einen Seite ist aus zahlreichen empirischen Arbeiten bekannt, dass die zeitliche Nähe von Rückmeldung eine zentrale Voraussetzung für Wissenserwerb darstellt (z. B. Bangert-Downs et al., 1991; Lenhard et al., 2011; Platt, 1973). Der Computer kann in dieser Hinsicht der Lehrkraft überlegen sein, da er in sehr kurzer Zeit nahezu beliebig vielen Schülern simultan Rückmeldung geben kann. Sicherlich ist für die Entstehung und Aufrechterhaltung der Motivation eine hinreichend präzise Rückmeldung notwendig. In verschiedenen Untersuchungen konnten wir bereits zeigen, dass die Qualität der Rückmeldungen von *conText* derjenigen von menschlichen Bewertern nicht unterlegen ist (z. B. Lenhard, Baier, Hoffmann & Schneider, 2007). Jedoch sind wir davon überzeugt, dass weniger die Präzision der inhaltlichen Rückmeldung selbst, sondern vielmehr die daraus resultierende Bereitschaft zur intensiven Auseinandersetzung mit den Texten den größten Teil des Trainingseffektes bewirkte.

Ein weiterer spannender Befund liegt in der Tatsache, dass das computerbasierte Üben des Zusammenfassens gleichermaßen zur Zunahme von deklarativem Lesestrategiewissen führte wie die explizite Strategievermittlung. Dies stellt die vorherrschende Meinung in Frage, dass für die Verbesserung des Leseverständnisses zuerst Lesestrategiewissen explizit vermittelt werden müsse und aus diesem deklarativen Wissen schließlich durch wiederholtes Üben prozedurale Fähigkeiten resultieren würden (vgl. Pressley, Borkowski & Schneider, 1989). Offensichtlich ist der umgekehrte Weg genauso möglich: Durch das Üben prozeduraler Fähigkeiten kann zeitgleich deklaratives Lesestrategiewissen aufgebaut werden, ohne dieses explizit vermitteln zu müssen. Ein Grund hierfür könnte darin liegen, dass die unmittelbare Rückmeldung von Verbesserungen den Schülern bewusst macht, welche Herangehensweisen besonders effektiv sind und welche nicht. Für die Vermittlung von Lesestrategiewissen scheint die direkte Instruktion von deklarativem Lesestrategiewissen folglich nicht zwangsläufig notwendig zu sein.

## 5.2 Einschränkungen

Die vorliegende Studie weist eine Reihe an Einschränkungen auf, die bei der Interpretation der Ergebnisse und der daraus gezogenen Schlussfolgerungen berücksichtigt werden müssen. Die verwendeten Interventionsprogramme weichen im Aufbau und der Zielrichtung in verschiedener Weise voneinander ab, sodass Unterschiede in den Programmeffekten nicht auf einzelne Aspekte der Interventionsansätze zurückgeführt werden können. Dieses Faktum schränkt die interne Validität der Untersuchung ein. Die Programme unterschieden sich nicht alleine im Hinblick auf explizite versus implizite Strategievermittlung, sondern auch in Bezug auf die Darstellungsform, das verwendete Textmaterial, die Arbeitsformen und den Grad an Rückmeldung. Zum Teil wirken sich diese Faktoren zugunsten von *conText* aus, insbesondere die engere Einbettung des Programms in den Unterricht durch Verwendung regulärer Schulbuchtexte und die hohe Intensität an Rückmeldung. Andere Faktoren wie die Partnerarbeit oder die lehrerzentrierte Instruktion sollten dagegen die Wirkung der Lesedetektive günstig beeinflussen. Weitere Einschränkungen dieser Untersuchung betreffen beide Interventionen gleichermaßen. So konnten aufgrund der begrenzten Zeit für die Testungen in den Schulen lediglich einige Facetten lesebezogener Fähigkeiten erhoben werden, die zudem leicht in Gruppentestungen erfassbar waren. Diesem Umstand ist das Fehlen von Maßen zu prozeduralen Fähigkeiten, wie z. B. der Anwendung von Kontrollstrategien, geschuldet. Darüber hinaus handelt es sich um eine quasiexperimentelle Studie, bei der die Gefahr besteht, dass Clustereffekte die Ergebnisse verzerren. Letztlich wäre es notwendig, die Effekte mithilfe hierarchisch linearer Modelle auf Klassenebene auszuwerten und eine erheblich größere Anzahl an Klassen zu untersuchen. Eine weitere Einschränkung besteht darin, dass die Stichprobe nur aus Hauptschülern der 6. Klassenstufe einer bestimmten Region bestand, also eingeschränkt ist. Zur Klärung der Frage, ob die Ergebnisse auf Schüler anderer Schulformen und Jahrgangsstufen übertragbar sind, wären weiterführende Studien notwendig. Auf der anderen Seite verfügt die Untersuchung über eine hohe ökologische Validität, da die Effekte im realen Unterrichtsetting nachgewiesen wurden. Auch wurden im Wesentlichen die Ergebnisse der Forschungsarbeiten zu «Wir werden Textdetektive» in Haupt-, Real- und Gesamtschule repliziert.

Insgesamt lässt sich festhalten, dass die Ergebnisse zu *conText* einen ersten, ermutigenden Schritt darstellen, letztlich aber durch weitere Untersuchungen in anderen Schulformen und Altersbereichen, sowie unter Hinzuziehung größerer Stichproben weiter untermauert werden müssten. Es ist darüber hinaus denkbar, dass auch die Schüler in der *Lesedetektive*-Gruppe dieser Untersuchung von zusätzlichen «Booster-Sessions» profitiert hätten, wie es von Souvignier und Trenk-Hinterberger (2010) bei Jugendlichen des Gymnasiums demonstriert wurde. Spannend wäre zudem eine kombinierte Anwendung beider

Programme. Möglicherweise würden beide Ansätze voneinander profitieren.

### 5.3 Abschließendes Fazit und Schlussfolgerungen für weiterführende Forschung

*conText* hat als Lernmedium das Potenzial, bei Einbettung in den regulären Unterricht die Förderung lesebezogener Fähigkeiten effektiver zu gestalten. Durch die Verwendung regulärer Unterrichtsinhalte muss hierfür keine zusätzliche Unterrichtszeit investiert werden. Die inhaltliche Relevanz der Texte für den regulären Unterricht stellt unseres Erachtens sogar einen der zentralen Wirkmechanismen der Intervention dar. Darüber hinaus wird das Leseverständnis umso stärker verbessert, je tiefer sich Schüler mit einem Inhalt auseinandersetzen. Hierbei scheint die zeitnahe Rückmeldung über die Qualität der Textzusammenfassungen, die im Schulunterricht erst mithilfe eines Computerprogramms möglich wird, von entscheidender Bedeutung zu sein. Die Ergebnisse der Untersuchung werfen deshalb die Frage auf, ob in der Unterrichtspraxis das intensive, gelenkte Arbeiten mit bedeutungsvollen Texten nicht im Vergleich zur reinen Vermittlung von Techniken stärker in den Vordergrund treten sollte. Die Verwendung von computergeneriertem Feedback kann hierbei wertvolle Dienste leisten.

## 6 Danksagung

Das Forschungsprojekt wurde von der Deutschen Forschungsgemeinschaft gefördert (Fördernummer Schn315/29-3 und HO 1301/11-2,4). Unser Dank gilt Drs. Eileen und Walter Kintsch, University of Colorado und dem ehemals dort angesiedelten SA-LSA-Lab für die Unterstützung bei der Entwicklung des LSA-basierten Tutor-systems.

## Literatur

- 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., McElvany, N., Christmann, U., Richter, T., Groeben, N., Köster, J. et al. (2007). *Förderung von Lesekompetenz*. Berlin: Bundesministerium für Bildung und Forschung (BMBF). Verfügbar unter: [http://www.bmbf.de/pub/bildungsreform\\_band\\_siebzehn.pdf](http://www.bmbf.de/pub/bildungsreform_band_siebzehn.pdf) (13.02.2012).
- Auer, M., Gruber, G., Mayringer, H. & Wimmer, H. (2005). *Salzburger Lese-Screening für die Klassenstufen 5-8 (SLS 5-8)*. Bern: Hans Huber.
- Bangert-Downs, R., Kulik, C., Kulik, J. & Morgan, M. (1991). The Instructional Effect of Feedback in Test-Like Events. *Review of Educational Research*, 61, 213–238.
- Berkeley, S., Scruggs, T. E. & Mastropieri, M. A. (2009). Reading Comprehension Instruction for Students With Learning Disabilities, 1995–2006: A Meta-Analysis. *Remedial and Special Education*, 31, 423–436.
- Bertschi-Kaufmann, A., Hagendorf, P., Kruse, G., Rank, K., Riss, M. & Sommer, T. (2007). *Lesen. Das Training*. Seelze: Friedrich.
- Cohen, J. (1988). *Statistical Power Analysis for Behavioral Sciences*. New York: Erlbaum.
- Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K. & Harshman, R. (1990). Indexing by Latent Semantic Analysis. *Journal of the American Society For Information Science*, 41, 391–407.
- Druschky, P., Meier, R. & Stadler, C. (2006). *Lesen. Texte bearbeiten, verstehen und nutzen*. Seelze: Friedrich.
- Duffy, G. D., Roehler, L. R. & Herrmann, B. A. (1988). Modeling mental processes helps poor readers become strategic readers. *The Reading Teacher*, 4, 762–767.
- Edmonds, M. S., Vaughn, S., Wexler, J., Reutebuch, C., Cable, A., Tackett, K. K. et al. (2009). A Synthesis of Reading Interventions and Effects on Reading Comprehension Outcomes for Older Struggling Readers. *Review of Educational Research*, 79, 262–300.
- Fuchs, D., Fuchs, L. S., Mathes, P. G. & Simmons, D. C. (1997). Peer-Assisted Learning Strategies: Making Classrooms More Responsive to Diversity. *American Educational Research Journal*, 34, 174–206.
- Gold, A., Mokhlesgerami, J. Rühl, K., Schreblowski, S. & Souvignier, E. (2004). *Wir werden Textdetektive – Lehrermanual & Arbeitsheft*. Göttingen: Vandenhoeck & Ruprecht.
- Gold, A., Trenk-Hinterberger, I. & Souvignier, E. (2009). Die Textdetektive. Ein strategieorientiertes Programm zur Förderung des Leseverständnisses. In W. Lenhard & W. Schneider (Hrsg.), *Diagnose und Förderung des Leseverständnisses* (S. 207–226). Göttingen: Hogrefe.
- Greenleaf, C. & Petrosino, A. (2009). Response to Slavin, Cheung, Groff, and Lake (2008). Effective Reading Programs for Middle and High Schools: A Best-Evidence Synthesis. *Reading Research Quarterly*, 44, 349–354.
- Guthrie, J. T., Wigfield, A. & Perencevich, K. C. (2004). *Motivating reading comprehension: Concept-Oriented Reading Instruction*. Mahwah, NJ: Erlbaum.
- Hasselhorn, M. & Körkel, J. (1986). Metacognitive versus traditional reading instructions: The mediating role of domain-specific knowledge on children's text processing. *Human Learning*, 5, 75–90.
- Hattie, J. (2009). *Visible Learning: A synthesis of over 800 meta-analyses relating to achievement*. New York: Routledge.
- Heller, K. A. & Perleth, Ch. (2000). *Kognitiver Fähigkeitstest für 4.-12. Klassen, Revision (KFT 4-12+R)*. Göttingen: Hogrefe.
- Kim, A.-H., Vaughn, S., Klingner, J. K., Woodruff, A. L., Klein Reutebuch, C. & Kouzekanani, K. (2006). Improving the Reading Comprehension of Middle School Students With Disabilities Through Computer-Assisted Collaborative Strategic Reading. *Remedial and Special Education*, 27, 235–249.
- Kintsch, E., Caccamise, D., Franzke, M., Johnson, N. & Doolley, S. (2007). SummaryStreet®: Computer-Guided Summary Writing. In T. K. Landauer, D. S. McNamara, S. Dennis, & W. Kintsch (Eds.), *The Handbook of Latent Semantic Analysis* (pp. 263–278). Mahwah, NJ: Erlbaum.
- Klauer, K. J. (2010). Schädliche Interferenzen beim Training hochkomplexer Lernstrategien. *Zeitschrift für Pädagogische Psychologie*, 24, 235–239.

- Lenhard, W. (2008). *Bridging the Gap to Natural Language: Latent Semantic Analysis as a Tool for the Development of Intelligent Tutoring Systems* (Research Report). Würzburg: University.
- Lenhard, W., Baier, H., Lenhard, A., Schneider, W. & Hoffmann, J. (2012). *conText*. Göttingen: Hogrefe.
- Lenhard, W., Baier, H., Hoffmann, J. & Schneider, W. (2007). Automatische Bewertung offener Antworten mittels Latenter Semantischer Analyse. *Diagnostica*, 53, 155–165.
- Lenhard, A., Lenhard, W., Schug, M. & Kowalski, A. (2011). Computerbasierte Mathematikförderung mit den «Rechenspielen mit Elfe und Mathis I». *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 43, 79–88.
- Lingel, K., Neuenhaus, N., Artelt, C. & Schneider, W. (2010). Metakognitives Wissen in der Sekundarstufe: Konstruktion und Evaluation domänenspezifischer Messverfahren. *Zeitschrift für Pädagogik*, 56, 228–238.
- McElvany, N. & Artelt, C. (2007). Das Berliner Eltern-Kind Leseprogramm: Konzeption und Effekte. *Psychologie in Unterricht und Erziehung*, 4, S. 314–332.
- McKeown, M. G., Beck, I. L. & Blake, R. G. K. (2009). Rethinking comprehension instruction: Comparing strategies and content instructional approaches. *Reading Research Quarterly*, 44, 218–253.
- McNamara, D. S., Ozuru, Y., Best, R. & O'Reilly, T. (2007). The 4-pronged comprehension strategy framework. In D. S. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies* (pp. 465–496). Mahwah, NJ: Erlbaum.
- National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Rockville, MD: National Institute of Child Health and Human Development.
- Naumann, J., Artelt, C., Schneider, W. & Stanat, P. (2010). Lesekompetenz von PISA 2000 bis PISA 2009. In E. Klieme, C. Artelt, J. Hartig, N. Jude, O. Köller, M. Prenzel, W. Schneider, & P. Stanat (Hrsg.), *PISA 2009: Bilanz nach einem Jahrzehnt* (S. 23–72). Münster: Waxmann.
- Palincsar, A. S. & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction*, 1, 117–175.
- Paris, S. G., Cross, D. R. & Lipson, M. Y. (1984). Informed strategies for learning: A program to improve children's reading awareness and comprehension. *Journal of Educational Psychology*, 76, 1239–1252.
- Paris, S. G. & Jacobs, J. E. (1984). The benefits of informed instruction for children's reading awareness and comprehension skills. *Child Development*, 55, 2083–2093.
- Platt, J. (1973). Social traps. *American Psychologist*, 28, 641–651.
- Pressley, M., Borkowski, J.G. & Schneider, W. (1989). Good information processing: What it is and how education can promote it. *International Journal of Educational Research*, 13, 857–867.
- Pressley, M., El-Dinary, P. B., Gaskins, I., Schuder, T., Bergman, J. L., Almasi, J. et al. (1992). Beyond direct explanation: Transactional instruction of reading comprehension strategies. *The Elementary School Journal*, 92, 513–555.
- RAND Reading Study Group. (2002) *Reading for understanding: toward an R&D program in reading comprehension*. Santa Monica, CA: RAND.
- Rühl, K. & Souvignier, E. (2006). *Wir werden Lesedetektive – Lehrermanual & Arbeitsheft*. Göttingen: Vandenhoeck & Ruprecht.
- Schaffner, E. (2009). Determinanten des Leseverstehens. In W. Lenhard & W. Schneider (Hrsg.), *Diagnostik des Leseverständnisses* (S. 19–44). Göttingen: Hogrefe.
- Schreblowski, S. (2004). *Training von Lesekompetenz*. Münster: Waxmann.
- Sigel, R. & Feneberg, S. (2007). *Leseförderung und Entwicklung von Lesefreude – Leitfaden für die lesekompetente Hauptschule*. München: Maiß.
- Slavin, R. E., Cheung, A., Groff, C. & Lake, C. (2008). Effective Reading Programs for Middle and High Schools: A Best-Evidence Synthesis. *Reading Research Quarterly*, 43, 290–322.
- Souvignier, E. & Antoniou, F. (2007). Förderung des Leseverständnisses bei Schülerinnen und Schülern mit Lernschwierigkeiten – eine Metaanalyse. *Vierteljahresschrift für Heilpädagogik und ihre Nachbargebiete*, 76, 46–62.
- Souvignier, E. & Trenk-Hinterberger, I. (2010). Implementati-on eines Programms zur Förderung selbstregulierten Lesens: Verbesserung der Nachhaltigkeit durch Auffrischungssitzungen. *Zeitschrift für Pädagogische Psychologie*, 24, 207–220.
- Souvignier, E. (2009). Effektivität von Interventionen zur Verbesserung des Leseverständnisses. In W. Lenhard & W. Schneider (Hrsg.), *Diagnose und Förderung des Leseverständnisses* (185–206). Göttingen: Hogrefe.
- Souvignier, E., Trenk-Hinterberger, I., Adam-Schwebe, S. & Gold, A. (2008). *Frankfurter Leseverständnistest (FLVT 5-6)*. Göttingen: Hogrefe.
- Streblow, L., Holodynski, M. & Schiefele, U. (2007). Entwicklung eines Lesekompetenztrainings (LEKOLEMO) für die siebte Klassenstufe: Bericht über zwei Evaluationsstudien. *Psychologie in Erziehung und Unterricht*, 54, 287–297.
- Van Kraayenoord, C.E. (2010). The Role of Metacognition in Reading Comprehension. In H.P. Trolldenier, W. Lenhard & P. Marx (Hrsg.), *Brennpunkte der Gedächtnisforschung* (pp. 277–304). Göttingen: Hogrefe.

Dr. Wolfgang Lenhard

Lehrstuhl für Psychologie IV  
 Universität Würzburg  
 Wittelsbacherplatz 1  
 DE-97074 Würzburg  
 wolfgang.lenhard@uni-wuerzburg.de