AUS DER KLINIK UND POLIKLINIK FÜR KINDER- UND JUGENDPSYCHIATRIE, PSYCHOSOMATIK UND PSYCHOTHERAPIE

KLINIK DER LUDWIG-MAXIMILIANS UNIVERSITÄT MÜNCHEN

DIREKTOR: PROF. DR. MED. GERD SCHULTE-KÖRNE

Zur Entwicklung digitaler, evidenzbasierter Instrumente für die Erfassung und Förderung von Lese-Rechtschreibfähigkeiten bei Grundschulkindern

Dissertation
zum Erwerb des Doktorgrades der Humanbiologie
an der Medizinischen Fakultät der
Ludwig-Maximilians-Universität zu München

vorgelegt von

Ruth Maria Görgen (M.A.)

aus

Düren

Jahr

2021

Mit Genehmigung der Medizinischen Fakultät
der Universität München

Berichterstatter:	Prof. Dr. med. Gerd Schulte-Körne
Mitberichterstatter:	Prof. Dr. Gent Michele Noterdaeme
	Prof. Dr. Beate Sodian
Mitbetreuung durch den promovierten Mitarbeiter:	PD Dr. Kristina Moll
Dekan:	Prof. Dr. med. Thomas Gudermann
Tag der mündlichen Prüfung:	01.12.2021

INHALTSVERZEICHNIS

ZUSAMMENFASSUNG

Lese- und Rechtschreibstörungen (LRS) gehören mit einer Prävalenz von 6-8 % zu den häufigsten Entwicklungsstörungen bei Kindern. Bei hoher Persistenz des Störungsbildes ergeben sich häufig weitreichende Folgen für die psychische, schulische und berufliche Entwicklung der Betroffenen. Eine frühzeitige Identifikation von Schriftsprachdefiziten sowie eine daran anschließende effektive Förderung des Lesens und Rechtschreibens ist daher erforderlich. Hierfür bieten sich insbesondere digitale Instrumente an, die neben motivationalen Aspekten auch eine bessere Erreichbarkeit versprechen und ein flexibles und selbstständiges Training zu Hause ermöglichen. Allerdings mangelt es derzeit an entsprechenden Programmen und wissenschaftlichen Belegen für deren Wirksamkeit. Um wirksam Anwendung finden zu können, sollten digitale Förderinstrumente aber auf Basis wissenschaftlich fundierter Erkenntnisse konzipiert und evaluiert werden. Übergreifendes Ziel der vorliegenden kumulativen Dissertation ist es, die Entwicklung digitaler, evidenzbasierter Instrumente im Bereich der Diagnostik und Förderung des Lesens und Rechtschreibens voranzubringen. Zu diesem Zweck wurden relevante Basiskompetenzen für das Lesen und Rechtschreiben identifiziert und darauf basierend schrittweise digitale Anwendungen entwickelt und evaluiert.

Im Rahmen dieser Arbeit wurde zunächst eine bereits vorhandene Übersichtsarbeit zur Leseförderung durch eine Übersichtsarbeit zur Rechtschreibförderung *(Fachartikel eins)* ergänzt. Diese ermöglichte die Identifikation wirksamer Maßnahmen nicht nur für das Lesen, sondern auch für das Rechtschreiben. Neben einer Förderung der Phonem-Graphem-Korrespondenz (d. h. Buchstabe-Laut-Zuordnung) und des orthografischen Regelwissens (d. h. explizites Wissen über orthografische Gesetzmäßigkeiten, wie z. B. die Mitlautverdopplung), wurde in der durchgeführten Metaanalyse vor allem die Wirksamkeit morphembasierter Trainings (d. h. Vermittlung orthografischer Gesetzmäßigkeiten mithilfe von Wortbausteinen) zur Verbesserung der Rechtschreibfähigkeit nachgewiesen. Zusätzlich wurde das Potenzial digitaler Förderprogramme analysiert.

Mithilfe neu entwickelter, digitaler Screening-Aufgaben wurde in *Fachartikel zwei* die Rolle der morphologischen Bewusstheit für das Lesen und Rechtschreiben näher untersucht. Hierbei wurden neben non-verbalen kognitiven Fertigkeiten auch zusätzliche Basiskompetenzen wie die phonologische Bewusstheit (d. h. Bewusstheit über die Lautstruktur der gesprochenen Sprache) und das sublexikalische orthografische Wissen (d. h. implizites und explizites Wissen über häufige und (un)zulässige Buchstabenkombinationen) miteinbezogen. Die Ergebnisse legen nahe, dass die morphologische Bewusstheit einen wichtigen Prädiktor für die Schriftsprachfähigkeiten und insbesondere für die Rechtschreibfähigkeit darstellt. Eine weitere wichtige Erkenntnis ist, dass sich die Lese-

Rechtschreibfähigkeiten sowie damit zusammenhängende Basiskompetenzen mithilfe einer digitalen Anwendung erfassen ließen.

Der *dritte Fachartikel* beinhaltet die Evaluation des digitalen Lese- und Rechtschreibtrainings *Meister Cody-Namagi*. Im Rahmen dieser Arbeit wurde zur Entwicklung des Programms beigetragen und die ersten drei Module des Programms, die primär auf die Förderung der Lesefertigkeiten abzielen, evaluiert. Es konnte gezeigt werden, dass das digitale Training die Lesefähigkeiten von Kindern mit Lesestörung signifikant verbessern kann. Das Training erwies sich darüber hinaus als besonders motivierend für die teilnehmenden Kinder. Vielversprechend ist zudem die Erkenntnis, dass eine alleinige Durchführung des Trainings im häuslichen Umfeld und ohne externe Anleitung gelang. Ausgehend von diesen positiven Ergebnissen, wird derzeit das vierte Modul entwickelt, das auf die systematische digitale Förderung der Rechtschreibfähigkeit abzielt.

SUMMARY

With a prevalence of 6-8 %, reading and spelling disorders are among the most frequent developmental disorders in children. For those affected, the persistence of the disorder often has serious consequences for socio-emotional well-being and academic achievement. Therefore, identification of reading and spelling deficits and effective support of literacy skills is crucial. This can be achieved by using digital tools, which, apart from motivational aspects, also offer better accessibility as well as flexible and independent usage at home. However, there is a lack of evaluated digital programs to support literacy skills making it difficult to prove effectiveness of digital tools. In order to be used effectively digital tools should be evidence-based and evaluated.

The overall aim of this thesis is to foster the development of digital evidence-based tools to assess and support literacy skills. For this purpose, relevant precursors of literacy skills as well as effective approaches to support those skills were identified, and subsequently digital tools for the assessment and support of reading and spelling skills were developed and evaluated.

As part of this work, an existing review regarding the effectiveness of reading intervention programs was complemented by a systematic review examining the effectiveness of spelling approaches *(article one).* This allowed the identification of effective approaches for the remediation of spelling deficits. Results demonstrated the effectiveness of phonics intervention (i. e. training of letter-sound mappings), orthographic instruction (i. e. graphotactic or orthographic phonological spelling rules such as consonant doubling) and especially morphological instruction (i. e. analyzing and manipulating morphemes). Moreover, the potential of digital tools was analysed.

In *article two* the role of morphological awareness in reading and spelling was examined after controlling for non-verbal cognitive abilities, age and grade as well as phoneme awareness and sublexical orthographic sensitivity. Therefore, novel digital tasks were developed. The results showed that morphological awareness is an important predictor of literacy skills in the German orthography and is especially important for spelling. Another key finding is that reading and spelling, as well as other precursor skills, could be assessed using a digital application.

The *third article* contains the evaluation of a novel digital game-based reading training. Within this thesis, parts of the training program were developed and the first three modules aiming to support reading skills were evaluated. Results confirmed that the digital reading training can significantly improve reading skills of children with reading disorders. The training was also found to be highly motivating for the participating children. Another promising finding was that the training seemed feasible for the usage in home environment without external

instruction. Based on these positive results, the fourth training module is currently being developed, which aims to systematically support spelling skills.

EINLEITUNG ZUR KUMULATIVEN DISSERTATION

Hintergrund

Lese- und Rechtschreibfähigkeiten sind grundlegend für den schulischen und beruflichen Werdegang und unerlässlich für die gesellschaftliche Teilhabe. Der komplexen Anforderung des Schriftspracherwerbs werden jedoch nicht alle Kinder gerecht. Einige entwickeln sogar massive Schriftsprachschwierigkeiten, die sowohl die schulische Entwicklung als auch das psychische Wohlbefinden gefährden. Dabei wird bei ca. 6 % der Kinder eine Lesestörung und bei bis zu 7 % eine Rechtschreibstörung festgestellt, während ca. 8 % der Kinder unterdurchschnittliche Leistungen sowohl im Lesen als auch im Rechtschreiben erbringen (Lese- Rechtschreibstörung, LRS; Moll, Bruder, Kunze, Neuhoff, & Schulte-Körne, 2014). Diese Defizite sind nicht auf eine unzureichende Beschulung, neurologische, visuelle oder auditive Beeinträchtigungen zurückzuführen, sondern sind durch genetische-, neurobiologische- und Umweltfaktoren bedingt (American Psychiatric Association, 2013; Fisher & DeFries, 2002; Richlan, Kronbichler, & Wimmer, 2013; Shaywitz & Shaywitz, 2008). Aufgrund der hohen Persistenz und negativen Auswirkungen der Defizite ist eine frühzeitige Erfassung und wirksame Förderung der Lese- und Rechtschreibfähigkeiten dringend erforderlich.

Das häufige Auftreten isolierter Schriftsprachdefizite unterstreicht die mögliche Dissoziation von Lese- und Rechtschreibproblemen (Fayol, Zorman, & Lété, 2009; Moll & Landerl, 2009; Wimmer & Mayringer, 2002) und legt eine separate Betrachtung der beiden Schriftsprachkomponenten und ihrer zugrundliegenden Defizite nahe. Ein Großteil der bisherigen Forschung bezieht sich jedoch auf die Leseentwicklung, während zur Rechtschreibentwicklung vergleichsweise wenig Studien existieren (Treiman & Kessler, 2005).

Angesichts zahlreicher Methoden, die für die Behandlung von Schriftsprachstörungen zur Verfügung stehen, wird folglich nicht immer auf symptomspezifische und damit wirkungsvolle Methoden zurückgegriffen. Zudem werden Diagnostik- und Fördermaßnahmen nicht selten zu spät oder gar nicht eingeleitet, sodass sich Schwierigkeiten verhärten und Misserfolgserlebnisse zunehmen. Um mehr Kinder mit Förderbedarf im Lesen und/oder Rechtschreiben unterstützen zu können, bietet sich die digitale Erfassung und Förderung der Schriftsprachfähigkeiten an. Digitale Instrumente sollten jedoch auf Basis wissenschaftlich fundierter Erkenntnisse konzipiert werden und sich dabei an der typischen Schriftsprachentwicklung orientieren und symptomspezifisch fördern.

Die vorliegende kumulative Dissertation schließt mit der Metaanalyse zur Rechtschreibförderung zunächst die Forschungslücke im Bereich effektiver Rechtschreibförderung. Anschließend werden mithilfe neu entwickelter Screening-Aufgaben wesentliche Vorläuferfertigkeiten des Schriftspracherwerbs spezifisch für das Lesen und Rechtschreiben untersucht. Der Fokus liegt hierbei auf der Rolle der morphologischen

Verarbeitung, deren Einfluss auf das Lesen und Rechtschreiben in der deutschen Orthografie noch wenig untersucht ist. Schließlich wird ein neu entwickeltes, digitales Förderprogramm evaluiert, das auf die evidenzbasierte, symptomspezifische Förderung des Lesens und Rechtschreibens abzielt. Im Rahmen dieser Dissertation wurde an der Entwicklung des Programms mitgewirkt und das Lesemodul evaluiert. Hieraus gewonnene Erkenntnisse werden für das sich in Entwicklung befindende Rechtschreibmodul genutzt.

Nachfolgend wird zunächst der typische Schriftspracherwerb skizziert und anschließend auf die Identifikation von Schriftsprachschwierigkeiten und Förderaspekten eingegangen. Hierbei werden Gemeinsamkeiten und Unterschiede für das Lesen und Rechtscheiben herausgestellt. Schließlich werden Erkenntnisse aus der digitalen Lernforschung vorgestellt, bevor die drei Fachartikel zusammengefasst werden, die Gegenstand der vorliegenden kumulativen Dissertation sind.

Typischer Schriftspracherwerb

Während das Erlernen der gesprochenen Sprache zu großen Teilen implizit erfolgt, erfordert die geschriebene Sprache explizites Lernen. Dieses wird in der Regel mit dem Schuleintritt initiiert. Allerdings werden bereits im Kindergarten die Grundlagen hierfür gelegt. So werden bereits im Kindergartenalter Gedächtniseinträge für prägnante visuelle Merkmale, wie bestimmte Buchstaben oder Logos, aufgebaut oder buchstabenähnliche Kritzeleien hervorgebracht. Nach Frith‘ Modell zum Schriftspracherwerb (1985) ist dieses frühe Schriftsprachverständnis der sogenannten logographischen Phase zuzuordnen.

Einer der wichtigsten Meilensteine für die Entwicklung des Lesens und Rechtschreibens ist die Einsicht in das alphabetische Prinzip, wobei Kinder erkennen, dass die gesprochene Sprache aus Lautsegmenten besteht, die wiederum durch Schriftzeichen repräsentiert werden (Küspert & Schneider, 2018). Diese Erkenntnis baut auf der Ausbildung der phonologischen Bewusstheit auf, die es Kindern bereits im Kindergartenalter ermöglicht, sich auf formale Aspekte der Sprache zu fokussieren (Küspert & Schneider, 2018). Die Aufmerksamkeit richtet sich dabei anfänglich auf größere Spracheinheiten, wie Sätze, Wörter und Silben, bis hin zu kleinsten sprachlichen Einheiten, den Lauten. Ausgehend von der Fähigkeit die lautliche Struktur von Sprache analysieren und manipulieren zu können (d. h. phonologische Bewusstheit), werden Kinder mithilfe expliziter Anleitung in die Lage versetzt, Buchstaben und Laute einander zuordnen zu können. Durch die Zuordnung von Buchstaben und Lauten in der sogenannten alphabetischen Phase (Frith, 1985) gelingt es Kindern, lautgetreu zu verschriften und synthetisierend, also zusammenlautend, zu lesen. Da die Zuordnung von Buchstaben und Lauten in den meisten Orthografien nicht immer eindeutig ist, d. h. die Lauttreue teilweise begrenzt ist (z. B. kann der Buchstabe V als /f/ wie in *Vater* oder als /v/ wie in *Vase* gesprochen werden), reicht die alphabetische Strategie nicht aus, um eine angemessene Lese- und Rechtschreibkompetenz zu entwickeln (Treiman & Kessler, 2014).

Dies gilt im Besonderen für das Rechtschreiben, da die Phonem-Graphem-Korrespondenz (d. h. die Laut-Buchstabe Zuordnung) nicht automatisch zu einer korrekten Verschriftung von Wörtern führt (z. B. „Kint“ statt *Kind* oder „Schturm“ statt *Sturm*). Aber auch das synthetisierende Lesen erweist sich als sehr mühsam und wenig effizient. Folglich gewinnt in der anschließenden orthografischen Phase (Frith, 1985) die Beachtung größerer Schriftspracheinheiten und orthografischer Regelmäßigkeiten stetig an Bedeutung. Durch wiederkehrende Verknüpfungen zwischen Sprech- und Schriftwort bzw. Wortteilen werden zunehmend orthografische Repräsentationen im Gedächtnis aufgebaut, die wiederum die Automatisierung der Lese- und Rechtschreibfähigkeiten vorantreiben (Ehri, 1992; Perfetti, 1992). Bekannte Wörter müssen schließlich nicht mehr einzelheitlich, also Buchstabe für Buchstabe erlesen werden, sondern können mithilfe größerer schriftsprachlicher Einheiten, wie Silben und Wortbausteinen, effizienter, wenn nicht ganzheitlich, gelesen werden. Darüber hinaus können Schreibweisen mithilfe größerer Schriftspracheinheiten abgerufen oder hergeleitet werden.

Fortschreitende automatisierte Lese- und Rechtschreibfähigkeiten ermöglichen schließlich, dass mehr kognitive Ressourcen für den Ausbau wortübergreifenden Wissens (z. B. Satzgrammatik) und hierarchiehöherer Prozesse (z. B. Texte verstehen und produzieren) zur Verfügung stehen (Graham et al., 2012; Graham & Santangelo, 2014; Perfetti & Hogaboam, 1975).

Identifikation von Schriftsprachschwierigkeiten

Die Überprüfung der Schriftsprachkompetenz sollte, insbesondere zu Beginn des Schriftspracherwerbs, nicht nur die Erfassung der Lese- Rechtschreibfertigkeiten selbst, sondern auch die Erfassung der Basiskompetenzen für das Lesen und Rechtschreiben beinhalten. Zu den für die Schriftsprachentwicklung relevanten Basiskompetenzen zählen die Verarbeitungsfähigkeiten in der Phonologie, Morphologie und Orthografie, die jedoch nicht gleichermaßen bedeutend für das Lesen und Rechtschreiben sind und deren Einfluss auf die Schriftsprachfähigkeiten sich zudem von Orthografie zu Orthografie unterscheiden (Desrochers, Manolitsis, Gaudreau, & Georgiou, 2018; Georgiou, Parrila, & Papadopoulos, 2008; Landerl, Wimmer, & Frith, 1997).

Im Bereich der Phonologie wird der phonologischen Bewusstheit sowie der Graphem-Phonem- und Phonem-Graphem-Korrespondenz besondere Bedeutung zuteil. Ein Großteil der Kinder mit Schriftsprachschwierigkeiten zeigt im Vergleich zu Kindern ohne Schriftsprachprobleme geringere Fähigkeiten in der phonologischen Bewusstheit und so wurde die phonologische Bewusstheit vielfach als wesentlicher Prädiktor für die Lese- und Rechtschreibfähigkeiten herausgestellt (Bishop & Snowling, 2004; Snider, 1997; Wagner & Torgesen, 1987). Aufgrund ihrer Bedeutung für den Schriftspracherwerb, ist die phonologische Bewusstheit häufig bereits fester Bestandteil bei der Diagnostik und Überprüfung der

Schriftsprachkompetenz. Hierbei wird zumeist auf Wortbearbeitungsaufgaben zurückgegriffen, die mitunter das Analysieren und Manipulieren von Silben oder Lauten beinhalten (z. B. Wie viele Silben hat *Elefant*? Aus welchen Lauten besteht *Tisch*?). Im Deutschen gilt die phonologische Bewusstheit als besonders wichtiger Marker für die Rechtschreibleistung (Moll, Fussenegger, Willburger, & Landerl, 2009; Moll, Wallner, & Landerl, 2012). Dies ist auf die Asymmetrie der deutschen Orthografie zurückzuführen, die höhere Anforderungen an das phonologische System beim Schreiben als beim Lesen stellt. Beispielsweise kann das geschriebene Wort *Hund (hʊnt)* auf Basis der Graphem-Phonem-Korrespondenz korrekt als /hʊnt/ gelesen werden, während für das Schreiben auf Basis der Phonem-Graphem-Korrespondenz die Schreibweisen „Hunt" und „Hund" in Frage kommen.

Im Gegensatz zur phonologischen Bewusstheit findet die morphologische Bewusstheit in der Erfassung der Schriftsprachkompetenz bislang nur wenig Berücksichtigung, obwohl immer wieder auf ihre für das Lesen und Rechtschreiben bedeutende Rolle hingewiesen wird (Goodwin & Ahn, 2010; Mann, 2000; Nagy, Carlisle, & Goodwin, 2013). Die morphologische Bewusstheit bezieht sich auf die Kenntnis über Morpheme (Wortbausteine) und die Fähigkeit, diese analysieren und manipulieren zu können. Dieser Fähigkeit wird mit fortschreitender Schriftspracherfahrung wachsende Bedeutung beigemessen (Berninger, Abbott, Nagy, & Carlisle, 2010). Bei Kindern mit Defiziten im Lesen und Rechtschreiben werden geringere Fähigkeiten in der morphologischen Bewusstheit beschrieben (Ben-Dror, Bentin, & Frost, 1995; Casalis, Colé, & Sopo, 2004; Siegel, 2008). Befunde zum Einfluss der morphologischen Bewusstheit auf die Schriftsprachleistungen stammen überwiegend aus dem angloamerikanischen Raum und fokussieren oft auf die Rolle der morphologischen Bewusstheit im Bereich des Leseverständnisses und der Wortlesegenauigkeit (z. B. Casalis, Quémart, & Duncan, 2015; Kirby et al., 2012; Levesque, Kieffer, & Deacon, 2017). Die wenigen vorhandenen orthografieübergreifenden Studien sprechen dafür, dass die morphologische Bewusstheit die Schriftsprachfähigkeiten in Abhängigkeit von der jeweiligen Orthografie in unterschiedlichem Ausmaß prädiziert (Casalis et al., 2015; Desrochers et al., 2018; Manolitsis, Georgiou, Inoue, & Parrila, 2019). Im Hinblick auf die deutsche Orthografie liegen neben Zusammenhängen zwischen der morphologischen Bewusstheit und dem Leseverständnis vor allem Zusammenhänge mit der Lesegeschwindigkeit und der Rechtschreibleistung nahe. Aufgrund der Konsistenz der deutschen Orthografie in Leserichtung (hohe Graphem-Phonem-Korrespondenz), wird im Deutschen vergleichsweise schnell ein adäquates Niveau im Bereich der Lesegenauigkeit erreicht. So machen Kinder nach einem Jahr der Beschulung nur noch wenige Lesefehler. Schwierigkeiten im Lesen zeigen sich dann vor allem in einer geringen Lesegeschwindigkeit, was ein Indiz dafür ist, dass die ganzheitliche Worterkennung bzw. Erkennung größerer Schriftspracheinheiten, wie Wortbausteine, unzureichend entwickelt ist. Im Umkehrschluss scheint es logisch, dass die Fähigkeiten im Bereich der morphologischen

Bewusstheit als Risikomarker für defizitäre Wortlesefähigkeiten (vor allem Lesegeschwindigkeit) herangezogen werden können. Aufgrund der Asymmetrie der deutschen Orthografie und der Inkonsistenz der Phonem-Graphem-Korrespondenz (Schreibrichtung) scheint es ferner vielversprechend, die morphologische Bewusstheit als Risikomarker für Rechtschreibdefizite heranzuziehen. Bei erneuter Betrachtung des vorgenannten Beispiels zur Schreibweise des gesprochenen Wortes *hʊnt (Hund),* lässt sich verdeutlichen, dass sich die korrekte Endung (/t/ vs. /d/) mithilfe der Nominalflexion zu *hʊndə (Hunde)* erschließen lässt. Daher ist es nur folgerichtig, die morphologische Bewusstheit als Prädiktor von Rechtschreibkompetenzen stärker in den Blick zu nehmen. Ungeachtet dessen, dass sich die morphologische Bewusstheit über Aufgaben zur Wortbildung (d. h. Derivation, Komposition und Konversion) und Aufgaben zur Flexion (d. h. Konjugation und Deklination) erfassen lässt, gibt es allerdings immer noch wenige Bemühungen den Zusammenhang zwischen der morphologischen Bewusstheit und den Lese-Rechtschreibfertigkeiten im Deutschen näher zu ergründen, geschweige denn, die morphologische Bewusstheit in diagnostische Verfahren zum Lesen und Rechtschreiben aufzunehmen.

Zusätzlich zur morphologischen Bewusstheit kann orthografisches Wissen auf sublexikalischer und lexikalischer Ebene Aufschluss über abgespeicherte orthografische Repräsentationen geben, die mit voranschreitender Schriftspracherfahrung an Relevanz gewinnen (Conrad, Harris, & Williams, 2012; Ise, Arnoldi, & Schulte-Körne, 2012; Rothe, Schulte-Körne, & Ise, 2013). Verglichen mit lexikalischen Aufgaben, die ähnlich wie Rechtschreibtests wortspezifisches orthografisches Wissen erfassen, kann mit sublexikalischen Aufgaben die implizite Kenntnis orthografischer Muster unabhängig von wortspezifischem Wissen gemessen werden (Burt, 2006). Erkennt ein Kind also ein Pseudowort mit einem häufigen Doppelkonsonanten (z. B <ll> im Deutschen) eher als plausible deutsche Schreibweise als ein Pseudowort mit einem seltenen oder illegalen Doppelkonsonanten (z. B <dd> oder <vv> im Deutschen), deutet dies auf implizites orthografisches Wissen hin, welches schon bei Vorschulkindern nachweisbar ist und besonders bedeutend für die Prädiktion der Rechtschreibleistung sein kann (Cassar & Treiman, 1997; Pacton, Michel, & Perruchet, 2005; Rothe et al., 2013).

Die Erfassung der Basiskompetenzen im Bereich der Phonologie, Morphologie und Orthografie ermöglicht eine frühe Risikoidentifikation von Schriftsprachschwierigkeiten und kann darüber hinaus auf zugrundeliegende Defizite hinweisen. Maßgeblich ist jedoch auch die Berücksichtigung der Symptomebene, also die Erfassung der Lese- und Rechtschreibleistung selbst. Im Bereich des Lesens sollte diese vor allem eine Überprüfung der Wortlesefähigkeit (Lesegeschwindigkeit und Lesegenauigkeit) und des Leseverständnisses beinhalten und im Bereich des Rechtschreibens qualitative und quantitative Fehleranalysen ermöglichen

(Galuschka, Rothe, & Schulte-Körne, 2015; Moll & Landerl, 2014). Eine differenzierte Diagnostik ist grundlegend für die Auswahl geeigneter Förderansätze.

Förderung des Lesens und Rechtschreibens

Die S3-Leitlinie zur Behandlung von Lese- und/oder Rechtschreibstörungen (Galuschka & Schulte-Körne, 2015) empfiehlt, dass eine wirksame Förderung des Lesens und Rechtschreibens symptomorientiert erfolgen und somit direkt an den Lese- und/oder Rechtschreibschwierigkeiten inklusive damit assoziierter Basiskompetenzen ansetzen sollte.

Aufgrund ihrer Relevanz für den Schriftspracherwerb, ist die phonologische Bewusstheit häufig Gegenstand von Lese- und Rechtschreibinterventionen. Die Wirksamkeit kann jedoch durch die Transparenz der jeweiligen Orthografie beeinflusst werden (Landerl et al., 2013). In der relativ transparenten deutschen Orthografie nimmt die Wirksamkeit von Trainings zur phonologischen Bewusstheit mit beginnender Beschulung ab (Galuschka, Ise, Krick, & Schulte-Körne, 2014; Wolf, Schroeders, & Kriegbaum, 2016). Daher sollte eine phonologische Förderung spätestens zu Schuleintritt mit einer Förderung der Graphem-Phonem- und Phonem-Graphem-Korrespondenz kombiniert werden (Bus & Ijzendoorn, 1999; Galuschka et al., 2014). Da die Kenntnis der Graphem-Phonem- und Phonem-Graphem-Korrespondenz eine wesentliche Grundlage für den Schriftspracherwerb darstellt, fokussiert sich ein Großteil der Forschung zur Schriftsprachförderung auf diese Fähigkeit (z. B., Ehri, 2001; McArthur et al., 2012; Weiser & Mathes, 2011). Die Effektivität konnte auch in einer Metaanalyse zur Wirksamkeit von Leseförderansätzen (Galuschka et al., 2014) bestätigt werden.

Für automatisiertes und damit effizientes Lesen und Rechtschreiben ist die Kenntnis über Graphem-Phonem- und Phonem-Graphem-Korrespondenzen allerdings nicht ausreichend (Treiman & Kessler, 2014). Besonders vielversprechend erscheint daher die Förderung des Morphemwissens, da größere schriftsprachliche Einheiten sowohl die direkte Worterkennung als auch das Verschriften eines Wortes erleichtern. So wird in einschlägiger Literatur darauf hingewiesen, dass sich eine Förderung der morphologischen Bewusstheit positiv auf die Lese- und Rechtschreibleistung auswirkt (Apel & Diehm, 2013; Arnbak & Elbro, 2000; Bowers, Kirby, & Deacon, 2010; Goodwin & Ahn, 2010). Dabei ist im Hinblick auf das Wortlesen anzunehmen, dass das Erkennen (un)bekannter Wörter durch die Beachtung von Morphemen beschleunigt werden kann, insbesondere bei langen, zusammengesetzten Wörtern, wie sie im Deutschen sehr häufig sind (z. B. *Straßenbahnhaltestelle*). Zusätzlich kann das Rechtschreiben dadurch erleichtert werden, dass sich Schriftbilder über wiederkehrende Morpheme herleiten lassen, vor allem wenn dies nicht über simple Phonem-Graphem-Zuordnungen gelingt, z. B. beim Verschriften des Wortes *fɛɐ̯ˈliːpt (verliebt)*. Hier erleichtert das Wissen über das häufige Derivationsmorphem und Präfix *ver* (und nicht etwa *fer)*, über den

Wortstamm *lieb* (von *lieben* und nicht etwa *liep)* sowie das häufige Flexionsmorphem und Suffix -*t* die Wortschreibung.

Um der deutschen Orthografie gerecht zu werden, müssen zusätzlich orthografische Regeln erlernt werden. Förderprogramme zum orthografischen Regelwissen, wie das *Marburger Rechtschreibtraining* (Schulte-Körne & Mathwig, 2001) oder das *Würzburger orthografisches Training* (Berger et al., 2009), berücksichtigen häufig die Erarbeitung phonographischer und graphotaktischer Regelmäßigkeiten (z. B. die Verschriftung von *st/sp*) sowie phonologisch-orthografische Regelmäßigkeiten (z. B. Vokallänge und Mitlautverdopplung). Die Förderung des orthografischen Regelwissens gilt als sehr ergiebig für die Verbesserung der Rechtschreibfähigkeit (Squires & Wolter, 2016).

Besonders im Bereich der Rechtschreibförderung fehlt es an systematischen Übersichtsarbeiten, die die Effektivität der verschiedenen Förderansätze untersuchen und dabei Einflussfaktoren wie die Transparenz der Orthografie und den Entwicklungsstand des Kindes berücksichtigen. Sowohl für das Lesen als auch für das Rechtschreiben scheint es allerdings vielversprechend, den Fokus stärker auf eine ganzheitliche Förderung zu legen. Dies schließt zur Verbesserung der basalen Wortlesefähigkeit vor allem die Förderung der phonologischen Bewusstheit, der Graphem-Phonem-Korrespondenz und der morphologischen Bewusstheit ein und zur Verbesserung der Rechtschreibfähigkeit zusätzlich die Vermittlung orthografischen Regelwissens. Systematische Untersuchungen sollten dies insbesondere für die Rechtschreibfähigkeit näher ergründen.

Potenziale digitaler Schriftsprachförderung

Neuere Technologien wie Smartphones und Tablets können kindliche Lernprozesse unterstützen und somit besonders vorteilhaft bei Kindern mit Lese-Rechtschreibstörung eingesetzt werden (Chauhan, 2016; Jamshidifarsani, Garbaya, Lim, Blazevic, & Ritchie, 2019). Während Lese- und Schreibaktivitäten in der Schule oder in der Freizeit bei wiederkehrendem Misserfolg zunehmend gemieden werden, bieten digitale Technologien Potenziale, die motivierend für Kinder mit Schriftsprachschwierigkeiten sein können (Morgan, Fuchs, Compton, Cordray, & Fuchs, 2008; Ronimus, Eklund, Pesu, & Lyytinen, 2019). So kommen digitale Förderprogramme dem kindlichen Lernen allein durch ihre Interaktivität entgegen und bieten ferner die Möglichkeit, sich adaptiv auf das individuelle Leistungsniveau des Kindes einzustellen (Huemer, Pointner, Schöfl, & Landerl, 2019). Hierdurch können wiederkehrende Misserfolgserlebnisse und damit einhergehende Frustrationen reduziert und stattdessen mehr positive Lernerlebnisse geschaffen werden. Die Bedeutung adaptiver und individualisiert aufbereiteter Inhalte wurde im Kontext digitaler Schriftsprachförderung vielfach hervorgehoben (Blok, Oostdam, Otter, & Overmaat, 2002; Jamshidifarsani et al., 2019; MacArthur, Ferretti, Okolo, & Cavalier, 2001; Perelmutter, McGregor, & Gordon, 2017; Santoro & Bishop, 2010). Auch die Einbettung der Lernaufgaben in eine Rahmengeschichte, sowie die Integration von

Herausforderungen, Belohnungssystemen und sichtbaren Lernfortschritten sind essenziell und werden von Kindern mit Lernstörungen bei digitalen Förderprogrammen besonders geschätzt (Ke & Abras, 2013; Ronimus, Kujala, Tolvanen, & Lyytinen, 2014). Ferner kann die Tatsache, dass Lernsoftware häufig selbstständig ohne Betreuungsperson und im häuslichen Umfeld durchgeführt werden kann, dazu beitragen, die Akzeptanz einer entsprechenden Förderung zu steigern. Hierbei sollte jedoch sichergestellt werden, dass die Anwendung für Kinder intuitiv handhabbar ist. Digitale Anwendungen bieten somit auch zeitökonomische Vorteile und können zudem durch eine breitere Verfügbarkeit (z. B. im Appstore) den Zugang zu Diagnostik- und Fördermaßnahmen erleichtern. Nicht zuletzt vor dem Hintergrund der weltweiten, zeitweiligen Schulschließungen während der COVID-19-Pandemie können digitale Förderinstrumente wertvolle Unterstützung bieten.

Obwohl die überwiegende Mehrheit der Haushalte in Deutschland bereits über Smartgeräte verfügt (Smartphone: 77,9 %; Tablet: 47,5 %; Statistisches Bundesamt (Destatis), 2018) und sogar mehr als 50 % der 8-11-jährigen Kinder ein eigenes Smartphone besitzen (Medienpädagogischer Forschungsverbund Südwest, 2019), werden die Potenziale innovativer Technologien für die Lese- und Rechtschreibförderung bislang noch nicht ausreichend genutzt. So lassen die wenigen bestehenden computerisierten Förderprogramme innovative Merkmale wie Adaptivität und Interaktivität meist außer Acht (Jamshidifarsani et al., 2019), was die Notwendigkeit der Weiterentwicklung in diesem Bereich unterstreicht.

Übersicht der Fachartikel

Übergeordnetes Ziel und Ergebniszusammenfassung

Da neuere Technologien wie Smartphones und Tablets das Potenzial bieten, den Zugang zur Erfassung und Förderung des Lesens und Rechtschreibens zur erleichtern, war es Ziel der vorliegenden kumulativen Dissertation, zur Entwicklung digitaler, evidenzbasierter Diagnostik- und Förderinstrumente beizutragen. Hierfür wurden für die Lese- und Rechtschreibentwicklung wesentliche Fertigkeiten identifiziert und darauf basierend digitale Screening-Aufgaben und ein Förderprogramm entwickelt und evaluiert.

Ausgehend von einer bereits vorhandenen Übersichtsarbeit zur effektiven Leseförderung (Galuschka et al., 2014), sollten zunächst wirksame Methoden der Rechtschreibförderung anhand einer weiteren Übersichtsarbeit identifiziert werden *(Fachartikel eins).* Um möglichst alle relevanten Studien berücksichtigen zu können, wurde eine umfassende und systematische Literaturrecherche in PSYNDEX, ERIC, PsycINFO und Medline durchgeführt. Zudem wurde die Datenbank ProQuest nach relevanten unpublizierten Arbeiten (z. B. Dissertationen und Masterabschlussarbeiten) durchsucht. Darüber hinaus wurden die Literaturverzeichnisse verschiedener Übersichtsartikel auf relevante Artikel hin überprüft und Experten über den Mailverteiler der *Society for the Scientific Studies of Reading* kontaktiert. Schließlich konnten 34 kontrollierte Studien, in denen Fördermaßnahmen zur Verbesserung der Rechtschreibfähigkeit bei Kindern mit schwachen Schriftsprachfähigkeiten durchgeführt wurden, in die Übersichtsarbeit eingeschlossen werden. Von den 34 Studien wurden 28 Studien in die Metaanalyse (d. h. quantitative Analyse) eingeschlossen. Sechs Studien wurden ausschließlich qualitativ analysiert, da sie das Einschlusskriterium der Kontrollgruppe nicht erfüllten (d. h. Vorhandensein einer untrainierten Kontrollgruppe, einer Placebo-Kontrollintervention, eines unspezifischen Kontrolltrainings oder regulären Schulunterrichts als Kontrollintervention). Die in den Studien durchgeführten Fördermaßnahmen wurden folgendermaßen kategorisiert: „phonological awareness intervention“, „phonics intervention“, „memorization intervention“, „morphological intervention“, „orthographic intervention“, „audio-visual cue intervention“, „assisted writing“ und „supportive software“. Orthografieübergreifend stellten schließlich die Förderung der Phonem-Graphem Zuordnung („phonics intervention“), des orthografischen Regelwissens („orthographic intervention“) und der morphologischen Bewusstheit („morphological intervention“) die effektivsten Fördermaßnahmen zur Verbesserung der Rechtschreibfähigkeit von rechtschreibschwachen Schülerinnen und Schülern des Primar- und Sekundarbereichs dar, wobei die Förderung der morphologischen Bewusstheit die größte Effektstärke aufwies (g' = 0.80). Moderatoranalysen zeigten zudem, dass der Fördererfolg durch Faktoren wie das Setting, die Klassenstufe, sowie die Expertise des durchführenden Personals beeinflusst werden kann. So wurden beispielsweise deutlich größere Effektstärken für eine

Rechtschreibförderung im Individualsetting als im Klassensetting gefunden (*g'* = 0.81 vs. *g'* = 0.28). Zudem wiesen computergestützte und computerbasierte Förderungen geringere Effektstärken auf als Förderungen im analogen Format, die dennoch im mittleren Bereich lagen (*g'* = 0.54 und *g'* = 0.44 vs. *g'* = 0.68). Demnach hat eine digitale Rechtschreibförderung durchaus das Potenzial eine Alternative oder zusätzliche Unterstützung zur individuellen Förderung darzustellen, auch wenn eindeutig Optimierungsbedarf besteht, zumal es sich bei den computerisierten Programmen in der Regel um ältere Technologien handelte (d. h. keine Verwendung intuitiver und adaptiver Lernsoftware über Smartphones oder Tablets) und diese meist nur begleitend eingesetzt wurden.

Eine ausführliche Beschreibung der Methode und Analysen ist dem ersten Fachartikel der vorliegenden kumulativen Dissertation zu entnehmen.

In dem *zweiten Fachartikel* der vorliegenden Arbeit wurde der Einfluss der morphologischen Bewusstheit auf das Lesen und Rechtschreiben im Deutschen unter Berücksichtigung weiterer wesentlicher Prädiktoren (IQ, phonologische Bewusstheit und sublexikalisches orthografisches Wissen) analysiert. Obwohl die Relevanz der morphologischen Bewusstheit für die Schriftsprachentwicklung unumstritten ist, beruhen die meisten Erkenntnisse hierzu auf Studien aus dem englischsprachigen Raum. Es ist daher unklar, inwiefern diese Erkenntnisse auf andere Orthografien übertragen werden können. So unterscheiden sich beispielsweise die englische und deutsche Orthografie im Hinblick auf ihre Konsistenz. Im Gegensatz zum Englischen, das als sehr inkonsistente Orthografie gilt, liegt im Deutschen eine Asymmetrie zwischen der Konsistenz im Lesen und Schreiben vor: Während die Graphem-Phonem-Korrespondenz (Leserichtung) durch eine hohe Konsistenz gekennzeichnet ist, ist die Phonem-Graphem-Korrespondenz (Schreibrichtung) relativ inkonsistent. Die korrekte Schreibung von Wörtern kann also nicht zuverlässig anhand der Phonem-Graphem-Korrespondenz erfolgen, weshalb es naheliegt, zusätzlich auf größere Einheiten, wie Morpheme, zurückzugreifen. Da es im deutschsprachigen Raum an publizierten Testverfahren zur morphologischen Bewusstheit und zum sublexikalischen orthografischen Wissen mangelt, wurden digitale Screening-Aufgaben entwickelt, die neben der Erfassung der phonologischen Bewusstheit auch die Erfassung der morphologischen Bewusstheit und des sublexikalischen orthografischen Wissens ermöglichten. Um die Fähigkeiten in diesen Bereichen unabhängig vom Wortschatz der Kinder überprüfen zu können, wurden für alle Aufgaben Pseudowörter verwendet. Eingebettet wurden die Aufgaben in eine umfassende App, die der Erfassung des psychischen Wohlbefindens und der schulischen Leistungen diente. Die App beinhaltete neben den experimentellen Screening-Aufgaben diverse standardisierte Tests (u. a. Intelligenztest und Lese- und Rechtschreibtests), die auf ein digitales Format übertragen wurden. Auf Grundlage der Daten von 3.122 Kindern der dritten und vierten Klassenstufen wurde überprüft, welchen Einfluss die phonologische Bewusstheit, das sublexikalische orthografische Wissen und vor allem die morphologische Bewusstheit auf das Lesen und Rechtschreiben haben, nachdem für nonverbale kognitive Fähigkeiten und Klassenstufe bzw. Alter kontrolliert wurde. Der spezifische Einfluss der einzelnen Prädiktoren wurde mithilfe von schrittweisen Regressionsanalysen berechnet. Die Ergebnisse zeigten, dass die morphologische Bewusstheit einen größeren Beitrag zur Varianzaufklärung der Lese- und Rechtschreibfähigkeiten leistete als die phonologische Bewusstheit und das sublexikalische orthografische Wissen. Hierbei war die Varianzaufklärung durch die morphologische Bewusstheit größer für das Rechtschreiben als für das Lesen, was darauf hinweist, dass die Rolle der morphologischen Bewusstheit von der Struktur der Orthografie abhängt und die Annahme bestätigt, dass sie in der asymmetrischen deutschen Orthografie besonders wichtig für die Rechtschreibfähigkeit ist. Trotzdem erwies sich die morphologische Bewusstheit auch als bedeutsamer Prädiktor für die Wortlesefähigkeit (d. h. Leseflüssigkeit). Nicht zuletzt konnte mit der Studie gezeigt werden, dass Lese- und Rechtschreibfähigkeiten sowie damit assoziierte Basiskompetenzen ökonomisch über digitale Geräte wie Smartphones und Tablets erfasst werden können.

Bei dem *dritten Fachartikel* handelt es sich um eine Evaluationsstudie, deren Ziel es war zu untersuchen, ob ein Training mit *Meister Cody-Namagi* die Lesefertigkeiten von Kindern mit Lesestörung verbessern kann. Meister Cody-Namagi ist ein digitales Förderprogramm, das in der Forschungsabteilung der Klinik und Poliklinik für Kinder- und Jugendpsychiatrie, Psychosomatik und Psychotherapie des Klinikums der Universität München in Zusammenarbeit mit der Softwarefirma Meister Cody GmbH entwickelt wurde und auf die systematische, evidenzbasierte Förderung der Lese- und Rechtschreibfähigkeiten von Grundschulkindern abzielt. Im Rahmen der Evaluationsstudie wurden die ersten drei Module des Förderprogramms evaluiert, mit denen vor allem die Förderung der Wortlesefertigkeiten und der Basisfertigkeiten des Lesens und Rechtschreibens intendiert wird. In Modul eins werden mit Übungen zur Silbenerkennung, Vokallängenunterscheidung, Lautanalyse und - synthese grundlegende phonologische Fähigkeiten gefördert. Darauf aufbauend wird in Modul zwei die Phonem-Graphem- und Graphem-Phonem-Korrespondenz trainiert, die insbesondere zu Beginn des Schriftspracherwerbs eine essenzielle Förderkomponente darstellt. Das dritte Modul trainiert die Wortlesefähigkeiten. Hierbei wird durch Übungen wie lexikalisches Entscheiden, das Erkennen von Wortbausteinen und Blitzlesen, sowohl die Lesegenauigkeit als auch die Lesegeschwindigkeit trainiert. Ein besonderes Merkmal des Förderprogramms ist dessen Adaptivität, wodurch es sich individuell dem Leistungsstand des jeweiligen Kindes anpasst. Die adaptive Umsetzung und die ins Programm integrierte Begleitung durch den Zauberer Meister Cody ermöglichen es, dass Kinder das Training selbstständig im häuslichen Umfeld durchführen können und auf dem für sie adäquaten Schwierigkeitslevel trainieren. Zur Überprüfung der Wirksamkeit des Förderprogramms wurde eine randomisiert-kontrollierte Studie im Wartekontrollgruppendesign mit aktiver Wartekontrollgruppe durchgeführt. In die Studie eingeschlossen wurden 50 Kinder der zweiten und dritten Klassenstufen mit Lesestörung (d. h. ≤ 16. Perzentil in mindestens einem der beiden Lesesubtests des Salzburger Lese- und Rechtschreibtests; SLRT-II Weiterentwicklung; Moll & Landerl, 2014), die zufällig der Experimental- oder Wartekontrollgruppe zugeteilt wurden. Die Experimentalgruppe trainierte selbstständig über einen Zeitraum von acht bis elf Wochen mit Meister Cody-Namagi, während die Wartekontrollgruppe im gleichen Zeitraum zunächst digitale Spiele zum logischen Denken durchführte, welche keinerlei schriftsprachliche Komponenten enthielten. Anschließend erhielt auch die Wartekontrollgruppe eine Förderung mit Meister Cody-Namagi. Um die Fördereffekte zu evaluieren, wurden die Kinder vor und nach dem Training getestet. Hierbei wurden mithilfe verschiedener Tests die non-verbalen kognitiven Leistungen, die Vorläuferfertigkeiten (wie z. B. die phonologische Verarbeitung und Fähigkeiten im Bereich der Phonem-Graphem-Korrespondenz) sowie die Leistungen im Lesen und Rechtschreiben erhoben. Darüber hinaus wurden Fragebögen zu den Trainingserfahrungen mit dem Kind und einem Elternteil durchgeführt. Es zeigte sich, dass Kinder mit Lesestörung von einem Training mit Meister Cody-Namagi profitieren. Verglichen mit der Kontrollgruppe, zeigten die Kinder der Experimentalgruppe signifikante Verbesserungen im Hinblick auf trainiertes Wortmaterial und eine Tendenz zur Generalisierung auf untrainiertes Wortmaterial. Da es sich um einen relativ kurzen Förderzeitraum handelte, ist anzunehmen, dass ein längerer Förderzeitraum Generalisierungseffekte verstärkt hätte. Darüber hinaus erwies sich das Training als besonders motivierend für die Kinder, was gemessen daran, dass Kinder mit Leseschwierigkeiten häufig eine geringe Lesemotivation aufweisen, ein sehr wertvolles Ergebnis ist. Dabei spielten für die Kinder vor allem die Einbettung in die Rahmenhandlung um den Zauberer Meister Cody und das integrierte Belohnungssystem eine wichtige Rolle. Schließlich konnte mit der Studie auch gezeigt werden, dass das digitale Förderprogramm selbstständig von Kindern der zweiten und dritten Schulklassen durchgeführt werden kann. Das Programm scheint somit vielversprechend vor allem für diejenigen Kinder zu sein, die keinen Zugang zu einer Therapie haben oder bei denen eine höhere Förderfrequenz erforderlich ist. Eine detaillierte Beschreibung des Trainings, des zugrundeliegenden Konzepts und der Evaluationsstudie ist dem dritten Fachartikel der vorliegenden Arbeit zu entnehmen.

Zusammenfassende Diskussion und Ausblick

In der vorliegenden kumulativen Dissertation wurden wesentliche Komponenten für die Entwicklung und Förderung des Lesens und Rechtschreibens identifiziert und mit aktuellen Erkenntnissen der digitalen Lernforschung kombiniert. Mithilfe einer App war es möglich, die Leistungen von Grundschulkindern im Lesen und Rechtschreiben sowie damit assoziierten Basiskompetenzen zu erfassen. Ausgehend von dieser Erkenntnis sollte die Entwicklung digitaler Screening-Verfahren im Bereich des Lesens und Rechtschreibens weiter verfolgt werden, um mögliche Defizite in der Schriftsprachentwicklung effizient und hierdurch möglichst frühzeitig identifizieren und folglich entsprechende Fördermaßnahmen einleiten zu können. Dabei sollte der Einschluss der morphologischen Bewusstheit bedacht werden, deren bedeutende Rolle für die Schriftsprachentwicklung im Deutschen im Rahmen dieser Dissertation herausgestellt wurde. Ferner liefert die vorliegende Arbeit wichtige Erkenntnisse für die digitale Förderung von Schriftsprachfähigkeiten. Es wurde gezeigt, dass ein digitales und adaptives Lesetraining, das gleichzeitig evidenzbasierten Ansätzen folgt, die Lesefertigkeiten von Grundschulkindern verbessern kann. Zudem erwies sich die digitale Förderung mit dem Programm Meister Cody-Namagi als besonders motivierend für die Kinder, was häufig auftretendem Vermeidungsverhalten von Lese- und Schreibaktivitäten entgegenwirken könnte. Diese positiven Ergebnisse der digitalen Leseförderung sprechen dafür, die Entwicklung des Programms weiterzuführen und um ein spezifisches Rechtschreibförderprogramm zu ergänzen. Basierend auf den Ergebnissen der in dieser Arbeit enthaltenen Metaanalyse zur Rechtschreibförderung sollte das, sich derzeit in Entwicklung befindende, digitale Rechtschreibförderprogramm eine Förderung der Phonem-Graphem-Korrespondenz, der morphologischen Bewusstheit und des orthografischen Regelwissens enthalten.

Die vorliegende Dissertation leistet einen wesentlichen Beitrag zur Entwicklung digitaler, evidenzbasierter Anwendungen, die die Erfassung und Förderung der Schriftsprachfähigkeiten und damit assoziierter Basiskompetenzen ermöglichen.

Darstellung des eigenen Beitrags

Die in dieser Arbeit als *erster Fachartikel* aufgeführte Metaanalyse entstand primär in Zusammenarbeit mit Frau Dr. Katharina Galuschka, die die Publikation initiierte. Die systematische Literaturrecherche wurde in erster Instanz von Frau Dr. Julia Kalmar und mir unter Anleitung von Frau Dr. Galuschka durchgeführt. Die Selektion nach Sichtung der Volltexte wurde zu großen Teilen von mir, mit Unterstützung von Frau Dr. Galuschka und Frau Dr. Kalmar, durchgeführt. Frau Dr. Galuschka führte die quantitativen Analysen durch, nachdem die Daten hauptsächlich von mir aufbereitet wurden. Herr Dr. Stefan Haberstroh brachte seine statistische Expertise bei den Analysen mit ein. Die qualitative Analyse wurde von mir vorgenommen. Die Analysen zur „onset entropy" stammen von Frau Dr. Xenia Schmalz. Hauptsächlich wurde der Fachartikel von Frau Dr. Galuschka und mir verfasst, unter Beteiligung von Frau Dr. Kalmar, Herrn Dr. Haberstroh, Frau Dr. Schmalz, und Herrn Prof. Dr. Schulte-Körne. Ich verfasste 50 % der Einleitung, 80 % des Methodenteils, die qualitative Analyse und 50 % der Diskussion. Den Einreich- und Reviewprozess führte ich gemeinsam und zu gleichen Teilen mit Frau Dr. Galuschka durch.

Der *zweite Fachartikel* entstammt aus einem vom Bundesministerium für Bildung und Forschung geförderten Kooperationsprojekt mit dem DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation in Frankfurt. Initiiert wurde das Projekt von Herrn Prof. Dr. Schulte-Körne und Herrn Prof. Dr. Marcus Hasselhorn (DIPF). Zu Beginn meiner Promotion lag das Studiendesign bereits vor und wurde im Zuge des Projekts von mir und den anderen Projektmitarbeiterinnen Frau Dr. Kalmar und Frau Dr. Josefine Rothe sowie von Frau Dr. Linda Visser (DIPF) und Frau Katharina Grunwald (DIPF) weiter ausgearbeitet. Der Ethikantrag inklusive der Anschreiben für die Studienteilnahme, sowie der Antrag zur Datenschutzfreigabe wurde für München hauptsächlich von Frau Dr. Kalmar und mir angefertigt, mit Unterstützung von Herrn Prof. Dr. Schulte-Körne. Die Auswahl der Testverfahren für die App erfolgte durch Frau Dr. Kalmar, Herrn Prof. Dr. Schulte-Körne und mich, gemeinsam mit dem Frankfurter Team. Die Akquise der 27.000 Adressdaten bei den bayerischen Einwohnermeldeämtern erfolgte durch Frau Dr. Kalmar, Frau Dr. Rothe und mich mit Unterstützung von Hilfskräften, die Frau Dr. Kalmar, Frau Dr. Rothe und ich einarbeiteten und supervidierten. Die Aufbereitung der Adressdaten erfolgte vorwiegend durch mich mit Unterstützung der studentischen Hilfskräfte. Die App wurde von der Softwarefirma Meister Cody GmbH programmiert, wobei die Inhalte gemeinsam von Frau Dr. Kalmar, Frau Dr. Rothe und mir sowie dem Frankfurter Team vorbereitet wurden. Die Screening-Aufgaben wurden im Rahmen einer separaten Arbeitsgruppe bestehend aus Frau Dr. Galuschka, Frau Dr. Sini Huemer, Frau Dr. Kalmar, Frau PD Dr. Kristina Moll, Frau Dr. Rothe, Herrn Prof. Dr. Schulte-Körne und mir entwickelt. Meine Verantwortlichkeit lag dabei vor allem auf der Entwicklung der Aufgabe zur morphologischen Bewusstheit, worin mich Frau PD Dr. Moll unterstützte. Die

Koordination der Studie erfolgte durch Frau Dr. Kalmar, Frau Dr. Rothe und mich sowie dem Frankfurter Projektteam, stets unter Supervision der beiden beteiligten Professoren. Die Datenaufbereitung erfolgte in Zusammenarbeit zwischen Frankfurt und München. Die Rückmeldung zu den Testergebnissen wurde inhaltlich maßgeblich von mir erstellt, mithilfe von Frau Dr. Kalmar und Frau Dr. Rothe und zu Automatisierungszwecken von Herrn Dr. Janosch Linkersdoerfer (DIPF) programmiert. Erste Auswertungen wurden vornehmlich von den Frankfurter und Münchner Projektteams durchgeführt. Die Vorbereitung der anschließenden Validierungsstudie erfolgte für München von Frau Dr. Kalmar, Frau Dr. Rothe und mir. Die Durchführung der Studie lag maßgeblich bei Frau Dr. Rothe und mir. Wir koordinierten die Studie und arbeiteten beteiligte Hilfskräfte sowie Praktikantinnen und Praktikanten ein. Testungen wurden, unter unserer Supervision, von Hilfskräften durchgeführt und Elterngespräche von Frau Dr. Rothe und mir geführt. Die Auswertung der in der Validierung erhobenen Daten erfolgte für München, unter meiner Beteiligung, hauptsächlich von Frau Dr. Rothe. Aus der Studie und den damit erhobenen Daten (online Erhebung und Validierung) sind diverse Veröffentlichungen hervorgegangen (z. B., Visser, Linkersdörfer, Kalmar, et al., 2020; Visser, Linkersdörfer, Rothe, et al., 2020) sowie auch der in dieser Arbeit enthaltene Fachartikel zwei, der von Frau PD Dr. Moll und mir initiiert wurde. Die Analysen wurden von mir durchgeführt, nach initialer Anleitung von Frau PD Dr. Moll. Gemeinsam verfassten wir die Einleitung und ich maßgeblich die Methode, die Ergebnisse und die Diskussion. Frau Elisabetta De Simone brachte sich bei der Literaturrecherche ein. Den Reviewprozess führte ich überwiegend selbstständig mit Unterstützung von Frau PD Dr. Moll durch.

Der *dritte Fachartikel* entstammt aus der Evaluationsstudie zum Förderprogramm Meister Cody-Namagi, die von Herrn Prof. Dr. Schulte-Körne, Frau PD Dr. Moll, Frau Dr. Huemer und mir initiiert wurde. Die Aufgaben der ersten drei zu evaluierenden Module lagen zu Beginn meiner Promotion größtenteils bereits vor und wurden von Herrn Prof. Dr. Schulte-Körne, Frau PD Dr. Moll und Frau Dr. Huemer konzipiert. Ich unterstützte bei der Erprobung der App während der Programmierung durch die Softwarefirma Meister Cody GmbH und brachte mich bei der finalen Konzeption der App, Aufgabeninstruktionen und notwendigen Aufgabenadaptionen inklusive Itemauswahl mit ein. Für die Studie erarbeitete ich das Studiendesign und verfasste den Antrag zur Datenschutzfreigabe, sowie den Ethikantrag inklusive der Anschreiben für die Studienteilnahme, wobei mich Frau PD Dr. Moll und Frau Dr. Huemer unterstützten. Die Rekrutierung der Teilnehmer erfolgte durch mich unter Mithilfe von Therapeuten. Die Kontrollbedingung wurde hauptsächlich von mir, unterstützt durch Frau Dr. Huemer und nach Rücksprache mit Frau PD Dr. Moll und Prof. Dr. Schulte-Körne, entwickelt. Die Studienkoordination lag maßgeblich in meiner Hand, wobei ich von Frau PD Dr. Moll unterstützt wurde. Ich führte gemeinsam mit Frau PD Dr. Moll die Elterngespräche und

- beratungen im Rahmen der Testungen durch. Die Testungen mit den Kindern wurden hauptsächlich von Studienassistentinnen, Hilfskräften sowie von Praktikantinnen und Praktikanten durchgeführt, nach Anleitung und Supervision durch Frau PD Dr. Moll und mich. Die Betreuung der Familien erfolgte telefonisch und per Mail vornehmlich durch mich, unterstützt durch die Studienassistentinnen. Die Datenanalyse erfolgte durch mich, nach Rücksprache mit Frau PD Dr. Moll, Frau Dr. Huemer und Herrn Prof. Dr. Schulte-Körne. Der aus dieser Studie entstandene Fachartikel *(Fachartikel drei)* wurde hauptsächlich von mir initiiert, verfasst und nach Rücksprache mit den Ko-Autorinnen und dem Ko-Autor überarbeitet. Den Einreich- und Reviewprozess führte ich selbstständig nach Rückmeldung der Ko-Autorinnen und des Ko-Autors durch.

Ausgehend von den Studienergebnissen, beteiligte ich mich außerdem an der Weiterentwicklung des Programms durch Ergänzung des vierten Moduls, das die Förderung der Rechtschreibung beinhaltet. Dessen Konzept erarbeiteten wir in einem Team bestehend aus Herrn Prof. Dr. Schulte-Körne, Frau PD Dr. Moll, Frau Dr. Huemer, Frau Dr. Galuschka und mir in Kooperation mit Frau Dr. Irene Corvacho del Toro (Goethe Universität Frankfurt). Gemeinsam entwickelten Frau PD Dr. Moll, Frau Dr. Huemer, Frau Dr. Galuschka und ich Aufgaben und Coaching-Videos und erstellten Itemlisten für die Rechtschreibaufgaben. Im Rahmen dieser Weiterentwicklung des Förderprogramms bereitete ich die zweiteilige Studie *Online Schrift-Namagi* vor, bei der zunächst ein von Prof. Dr. Schulte-Körne, Frau PD Dr. Moll und Frau Dr. Huemer entwickeltes Screening-Verfahren normiert werden sollte, das in einer von Frau PD Dr. Moll, Frau Dr. Huemer und mir durchgeführten Studie (*Ok-Leser Studie*) zuvor bereits an 116 Erst- bis Viertklässlern erprobt wurde. Dieses Screening-Verfahren wurde im Rahmen der Studie *Online Schrift-Namagi* in eine App eingebettet, die zusätzlich standardisierte Testverfahren zum Lesen und Rechtschreiben enthielt. Auf deren Basis konnten Kinder mit Schwierigkeiten im Rechtschreiben (und Lesen) identifiziert werden, die für eine Studie zur Evaluation des Rechtschreibtrainings (inklusive der Module I und II aus Meister Cody-Namagi zur Förderung der Basisfertigkeiten) in Frage kommen. Das Studiendesign dieser zweiteiligen Studie (Normierung und Evaluation) wurde gemeinsam von Frau PD Dr. Moll, Frau Dr. Huemer, Frau Dr. Galuschka und mir nach Rücksprache mit Herrn Prof. Dr. Schulte-Körne entwickelt. Ich kümmerte mich um die Freigabe der Studie durch die Ethikkommission und den Datenschutzbeauftragten der LMU München und verfasste die Anschreiben für die Studienteilnahme und die Einwohnermeldeämter. Zudem initiierte ich die Gewinnung der Adressdaten. Fortgeführt wurden meine Vorbereitungen von Herrn Björn Witzel, der schließlich gemeinsam mit Herrn Prof. Dr. Schulte-Körne, Frau PD Dr. Moll, Frau Dr. Huemer und Frau Dr. Galuschka die Studie durchführte. Aktuell läuft der zweite Teil der Studie zur Evaluation des Rechtschreibmoduls.

Literatur

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders (DSM) 5*. Washington DC: American Psychiatric Association.

Apel, K., & Diehm, E. (2013). Morphological awareness intervention with kindergartners and first and second grade students from low SES homes: a small efficacy study. *Journal of Learning Disabilities, 47*(1), 65-75. doi:10.1177/0022219413509964

Arnbak, E., & Elbro, C. (2000). The Effects of Morphological Awareness Training on the Reading and Spelling Skills of Young Dyslexics. *Scandinavian Journal of Educational Research, 44*(3), 229-251. doi:10.1080/00313830050154485

Ben-Dror, I., Bentin, S., & Frost, R. (1995). Semantic, Phonologic, and Morphologic Skills in Reading Disabled and Normal Children: Evidence from Perception and Production of Spoken Hebrew. *Reading Research Quarterly, 30*(4), 876. doi:10.2307/748202

Berger, N., Küspert, P., Lenhard, W., Marx, P., Schneider, W., & Weber, J. (2009). *Würzburger orthografisches Training. WorT*. Berlin: Cornelsen Verlag.

Berninger, V. W., Abbott, R. D., Nagy, W. E., & Carlisle, J. F. (2010). Growth in phonological, orthographic, and morphological awareness in grades 1 to 6. *Journal of Psycholinguistic Research, 39*(2), 141-163. doi:10.1007/s10936-009-913 0-6

Bishop, D. V. M., & Snowling, M. J. (2004). Developmental Dyslexia and Specific Language Impairment: Same or Different? *Psychological Bulletin, 130*(6), 858-886. doi:10.1037/0033-2909.130.6.858

Blok, H., Oostdam, R., Otter, M. E., & Overmaat, M. (2002). Computer-Assisted Instruction in Support of Beginning Reading Instruction: A Review. *Review of Educational Research, 72*(1), 101-130. doi:10.3102/00346543072001101

Bowers, P. N., Kirby, J. R., & Deacon, S. H. (2010). The Effects of Morphological Instruction on Literacy Skills: A Systematic Review of the Literature. *Review of Educational Research, 80*(2), 144-179. doi:10.3102/0034654309359353

Burt, J. S. (2006). What is orthographic processing skill and how does it relate to word identification in reading? *Journal of Research in Reading, 29*(4), 400-417. doi:10.1111/j.1467-9817.2006.00315.x

Bus, A. G., & Ijzendoorn, M. H. (1999). Phonological awareness and early reading: A meta-analysis of experimental training studies. *Journal of Educational Psychology, 91*(3), 403-414. doi:10.1037/0022-0663.91.3.403

Casalis, S., Colé, P., & Sopo, D. (2004). Morphological Awareness in Developmental Dyslexia. *Annals of Dyslexia, 54*(1), 114-138. doi:10.1007/s11881-004-0006-z

Casalis, S., Quémart, P., & Duncan, L. G. (2015). How language affects children´s use of derivational morphology in visual word and pseudoword processing: evidence from a cross-language study. *Frontiers in Psychology, 6*, 1-10. doi:10.3389/fpsyg.2015.00452

Cassar, M., & Treiman, R. (1997). The beginnings of orthographic knowledge: Children's knowledge of double letters in words. *Journal of Educational Psychology, 89*(4), 631-644. doi:10.1037/0022-0663.89.4.631.

Chauhan, S. (2016). A meta-analysis of the impact of technology on learning effectiveness of elementary students. *Computers & Education, 105*, 14-30. doi:10.1016/j.compedu.2016.11.005

Conrad, N., Harris, N., & Williams, J. (2012). Individual differences in children's literacy development: The contribution of orthographic knowledge. *Reading and Writing, 26*(8), 1223-1239. doi:10.1007/s11145-012-9415-2

Desrochers, A., Manolitsis, G., Gaudreau, P., & Georgiou, G. (2018). Early contribution of morphological awareness to literacy skills across languages varying in orthographic consistency. *Reading and Writing, 31*(8), 1695-1719. doi:10.1007/s11145-017-9772-y

Ehri, L. C. (1992). Reconceptualizing the development of sight word reading and its relationship to recoding. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 107-143). Hillsdale, NJ: Erlbaum.

Ehri, L. C. (2001). Systematic phonics instruction helps students learn to read: Evidence from the National Reading Panel's meta-analysis. *Review of Educational Research, 71*(3), 393-447. doi:10.3102/00346543071003393

Fayol, M., Zorman, M., & Lété, B. (2009). Unexpectedly good speller too. Associations and dissociations in reading and spelling French. *British Journal of Educational Psychology, 1*(1), 63-75. doi:10.1348/000709909X421973

Fisher, S. E., & DeFries, J. C. (2002). Developmental dyslexia: genetic dissection of a complex cognitive trait. *Nature Reviews Neuroscience, 3*(10), 767-780. doi:10.1038/nrn936

Frith, U. (1985). Beneath the surface of developmental dyslexia. *Surface dyslexia, 32*, 301-330.

Galuschka, K., Ise, E., Krick, K., & Schulte-Körne, G. (2014). Effectiveness of Treatment Approaches for Children and Adolescents with Reading Disabilities: A Meta-Analysis of Randomized Controlled Trials. *PLOS One, 9*(2), 1-12. doi:10.1371/journal.pone.0089900

Galuschka, K., Rothe, J., & Schulte-Körne, G. (2015). Die methodische Beurteilung und qualitative Bewertung psychometrischer Tests am Beispiel aktueller Verfahren zur Erfassung der Lese- und/oder Rechtschreibleistung. *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie, 43*(5), 317-334. doi:10.1024/1422-4917/a000378

Galuschka, K., & Schulte-Körne, G. (2015). *Lese- und/oder Rechtschreibstörung bei Kindern und Jugendlichen, Diagnostik und Behandlung.* Retrieved from https://www.kjp.med.uni-muenchen.de/download/leitlinie_lrs_kjp_langfassung.pdf [Zugriff am 08.02.2021]

Georgiou, G. K., Parrila, R., & Papadopoulos, T. C. (2008). Predictors of word decoding and reading fluency across languages varying in orthographic consistency. *Journal of Educational Psychology, 100*(3), 566-580. doi:10.1037/0022-0663.100.3.566

Goodwin, A. P., & Ahn, S. (2010). A meta-analysis of morphological interventions: effects on literacy achievement of children with literacy difficulties. *Annals of Dyslexia, 60*(2), 183-208. doi:10.1007/s11881-010-0041-x

Graham, S., Bollinger, A., Booth Olson, C., D'Aoust, C., MacArthur, C., McCutchen, D., & Olinghouse, N. (2012). *Teaching elementary school students to be effective writers: A practice guide (NCEE 2012-4058)*. Retrieved from http://ies.ed.gov/ncee/wwc/publications_reviews.aspx#pubsearch [Zugriff am 10.11.2020]

Graham, S., & Santangelo, T. (2014). Does spelling instruction make students better spellers, readers, and writers? A meta-analytic review. *Reading and Writing, 27*(9), 1703-1743. doi:10.1007/s11145-014-9517-0

Huemer, S., Pointner, A., Schöfl, M., & Landerl, K. (2019). *Evidenzbasierte LRS-Förderung, überarbeitete Neuauflage 2019*. Retrieved from https://www.schulpsychologie.at/fileadmin/user_upload/190731_Broschu__re_Evidenzbasierte_LRS-Fo__rderung_A4-BF.pdf [Zugriff am 04.11.2020]

Ise, E., Arnoldi, C. J., & Schulte-Körne, G. (2012). Development of orthographic knowledge in Germanspeaking children: A 2-year longitudinal study. *Journal of Research in Reading, 37*(3), 233-249. doi:10.1111/j.1467-9817.2012.01535.x

Jamshidifarsani, H., Garbaya, S., Lim, T., Blazevic, P., & Ritchie, J. M. (2019). Technology-based reading intervention programs for elementary grades: An analytic review. *Computers & Education, 128*, 427-451. doi:10.1016/j.compedu.2018.10.003

Ke, F., & Abras, T. (2013). Games for engaged learning of middle school children with special learning needs. *British Journal of Educational Technology, 44*(2), 225-242. doi:10.1111/j.1467-8535.2012.01326.x

Kirby, J. R., Deacon, S. H., Bowers, P. N., Izenberg, L., Wade-Woolley, L., & Parrila, R. (2012). Children's morphological awareness and reading ability. *Reading and Writing, 25*(2), 389-410. doi:10.1007/s11145-010-9276-5

Küspert, P., & Schneider, W. (2018). *Hören, lauschen, lernen. Sprachspiele für Kinder im Vorschulalter - Würzburger Trainingsprogramm zur Vorbereitung auf den Erwerb der Schriftsprache, 7. überarbeitete Auflage*. Göttingen: Vandenhoeck & Ruprecht.

Landerl, K., Ramus, F., Moll, K., Lyytinen, H., Leppanen, P. H., Lohvansuu, K., . . . Schulte-Körne, G. (2013). Predictors of developmental dyslexia in European orthographies with varying complexity. *The Journal of Child Psychology and Psychiatry, 54*(6), 686-694. doi:10.1111/jcpp.12029

Landerl, K., Wimmer, H., & Frith, U. (1997). The impact of orthographic consistency on dyslexia: a German-English comparison. *Cognition, 63*(3), 315-334.

Levesque, K. C., Kieffer, M. J., & Deacon, S. H. (2017). Morphological awareness and reading comprehension: examining mediating factors. *Journal of Experimental Child Psychology, 160*, 1-20. doi:10.1016/j.jecp.2017.02.015

MacArthur, C. A., Ferretti, R. P., Okolo, C. M., & Cavalier, A. R. (2001). Technology Applications for the Students with Literacy Problems: A critical Review. *The Elementary School Journal, 101*(3), 273-301. doi:10.1086/499669

Mann, V. A. (2000). Introduction to special issue on morphology and the acquisition of alphabetic writing systems. *Reading and Writing, 12*(3), 143-147. doi:10.1023/A:1008190908857

Manolitsis, G., Georgiou, G. K., Inoue, T., & Parrila, R. (2019). Are morphological awareness and literacy skills reciprocally related? Evidence from a cross-linguistic study. *Journal of Educational Psychology, 111*(8), 1362–1381. doi:10.1037/edu0000354

McArthur, G., Eve, P. M., Jones, K., Banales, E., Kohnen, S., Anandakumar, T., . . . Castles, A. (2012). Phonics training for English-speaking poor readers. *Cochrane Database of Systematic Reviews, 11*(1). doi:10.1002/14651858.CD009115.pub2

Medienpädagogischer Forschungsverbund Südwest. (2019). *KIM-Studie 2018. Kindheit, Internet, Medien. Basisuntersuchung zum Medienum-gang 6- bis 13-Jähriger in Deutschland*. Stuttgart: mpfs.

Moll, K., Bruder, J., Kunze, S., Neuhoff, N., & Schulte-Körne, G. (2014). Specific learning disorder: prevalence and gender differences. *PLOS One, 9*(7). doi:10.1371/journal.pone.0103537

Moll, K., Fussenegger, B., Willburger, E., & Landerl, K. (2009). RAN is not a measure of orthographic processing. Evidence from the asymmetric German orthography. *Scientific Studies of Reading, 13*(1), 1-25. doi:10.1080/10888430802631684

Moll, K., & Landerl, K. (2009). Double Dissociation Between Reading and Spelling Deficits. *Scientific Studies of Reading, 13*(5), 359-382. doi:10.1080/10888430903162878

Moll, K., & Landerl, K. (2014). *SLRT-II: Lese- und Rechtschreibtest, Weiterentwicklung*. Bern: Huber.

Moll, K., Wallner, R., & Landerl, K. (2012). Kognitive Korrelate der Lese-, Leserechtschreib- und der Rechtschreibstörung. *Lernen und Lernstörungen, 1*, 7-19. doi:10.1024/2235-0977/a000002

Morgan, P. L., Fuchs, D., Compton, D. L., Cordray, D. S., & Fuchs, L. S. (2008). Does Early Reading Failure Decrease Children's Reading Motivation? *Journal of Learning Disabilities, 41*(5), 387-404. doi:10.1177/0022219408321112

Nagy, W. E., Carlisle, J. F., & Goodwin, A. P. (2013). Morphological Knowledge and Literacy Acquisition. *Journal of Learning Disabilities, 47*(1), 3–12. doi:10.1177/0022219413509967

Pacton, S., Michel, F., & Perruchet, P. (2005). Children's Implicit Learning of Graphotactic and Morphological Regularities. *Child Dev, 76*(2), 324-339. doi:10.1111/j.1467-8624.2005.00848_a.x

Perelmutter, B., McGregor, K. K., & Gordon, K. R. (2017). Assistive technology interventions for adolescents and adults with learning disabilities: An evidence-based systematic review and meta-analysis. *Computers & Education, 114*, 139-163. doi:10.1016/j.compedu.2017.06.005

Perfetti, C. A. (1992). The presentation problem in reading acquisition. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 145-174). Hillsdale, NJ: Erlbaum.

Perfetti, C. A., & Hogaboam, T. (1975). Relationship between single word decoding and reading comprehension skill. *Journal of Educational Psychology, 67*(4), 461-469. doi:10.1037/h0077013

Richlan, F., Kronbichler, M., & Wimmer, H. (2013). Structural abnormalities in the dyslexic brain: a meta-analysis of voxel-based morphometry studies. *Human Brain Mapping, 34*(11), 3055-3065. doi:10.1002/hbm.22127

Ronimus, M., Eklund, K., Pesu, L., & Lyytinen, H. (2019). Supporting struggling readers with digital game-based learning. *Educational Technology Research and Development, 67*(3), 639–663. doi:10.1007/s11423-019-09658-3

Ronimus, M., Kujala, J., Tolvanen, A., & Lyytinen, H. (2014). Children's engagement during digital game-based learning of reading: The effects of time, rewards, and challenge. *Computers & Education, 71*, 237–246. doi:10.1016/j.compedu.2013.10.008

Rothe, J., Schulte-Körne, G., & Ise, E. (2013). Does sensitivity to orthographic regularities influence reading and spelling acquisition? A 1-year prospective study. *Reading and Writing, 27*(7), 1141–1161. doi:10.1007/s11145-013-9479-7

Santoro, L. E., & Bishop, M. J. (2010). Selecting Software with Caution: An Empirical Evaluation of Popular Beginning Reading Software for Children with Early Literacy Difficulties. *Computers in the Schools, 27*(2), 99-120. doi:10.1080/07380561003801582

Schulte-Körne, G., & Mathwig, F. (2001). *Das Marburger Rechtschreibtraining. Ein regelgeleitetes Förderprogramm für rechtschreibschwache Kinder.* Bochum: Winkler.

Shaywitz, S. E., & Shaywitz, B. A. (2008). Paying attention to reading: the neurobiology of reading and dyslexia. *Development and Psychopathology, 20*(4), 1329-1349. doi:10.1017/S0954579408000631

Siegel, L. S. (2008). Morphological awareness skills of English language learners and children with dyslexia. *Topics in Language Disorders, 28*(1), 15-27. doi:10.1097/01.adt.0000311413.75804.60

Snider, V. E. (1997). The relationship between phonemic awareness and later reading achievement. *Journal of Educational Research, 90*(4), 203-211.

Squires, K. E., & Wolter, J. A. (2016). The effects of orthographic pattern intervention on spelling performance of students with reading disabilities: A best evidence synthesis. *Remedial and Special Education, 37*(6), 357-369. doi:10.1177/0741932516631115

Statistisches Bundesamt (Destatis). (2018). Ausstattung privater Haushalte mit Informations- und Kommunikationstechnik Deutschland. Retrieved from https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Einkommen-Konsum-Lebensbedingungen/Ausstattung-Gebrauchsgueter/Tabellen/a-evs-infotechnik-d.html [Zugriff am 14.01.2021]

Treiman, R., & Kessler, B. (2005). Writing systems and spelling development. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 120-134). Malden: Blackwell Publishing.

Treiman, R., & Kessler, B. (2014). *How children learn to write words*. New York, NY: Oxford University Press.

Visser, L., Linkersdörfer, J., Kalmar, J., Görgen, R., Rothe, J., Hasselhorn, M., & Schulte-Körne, G. (2020). Comorbidities between specific learning disorders and psychopathology in elementary school children in Germany. *Frontiers in Psychiatry, 11*(292), 1-12. doi:10.3389/fpsyt.2020.00292

Visser, L., Linkersdörfer, J., Rothe, J., Görgen, R., Hasselhorn, M., & Schulte-Körne, G. (2020). The Role of ADHD Symptoms in the Relationship Between Academic Achievement and Psychopathological Symptoms. *Research in Developmental Disabilities, 97*, 1-11. doi:10.1016/j.ridd.2019.103552

Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin, 102*(2), 192-212. doi:10.1037/0033-2909.101.2.192

Weiser, B., & Mathes, P. (2011). Using encoding instruction to improve the reading and spelling performance of elementary students at risk for literacy difficulties: A best-evidence synthesis. *Review of Educational Research, 81*(2), 170-200. doi:10.3102/0034654310396719

Wimmer, H., & Mayringer, H. (2002). Dysfluent reading in the absence of spelling difficulties: A specific disability in regular orthographies. *Journal of Educational Psychology, 94*(2), 272-277. doi:10.1037/0022-0663.94.2.272

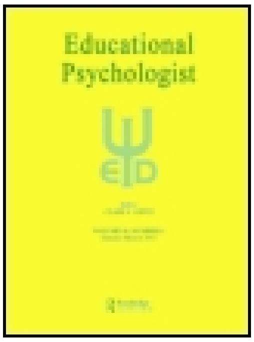
Wolf, K. M., Schroeders, U., & Kriegbaum, K. (2016). Metaanalyse zur Wirksamkeit einer Förderung der phonologischen Bewusstheit in der deutschen Sprache. *Zeitschrift für Pädagogische Psychologie, 30*(1), 9-33. doi:10.1024/1010-0652/a000165

FACHARTIKEL 1: Effectiveness of spelling interventions for learners with dyslexia: A meta-analysis and systematic review

Literaturangabe: Galuschka, K.*, Görgen, R.*, Kalmar, J., Haberstroh, S., Schmalz X., Schulte-Körne, G. (2020). Effectiveness of spelling interventions for learners with dyslexia: A meta-analysis and systematic review. *Educational Psychologist*, *55*(1), 1-20, doi: 10.1080/00461520.2019.1659794

*geteilte Erstautorenschaft

Angenommen am 05.08.2019
Online publiziert am 22.01.2020



Routledge
Taylor & Francis Group

Educational Psychologist

This is an Accepted Manuscript of an article published by Taylor & Francis in Journal of Educational Psychologist on 22 Jan 2020, available online: http://wwww.tandfonline.com/10.1080/00461520.2019.1659794

ISSN: 0046-1520 (Print) 1532-6985 (Online) Journal homepage: https://www.tandfonline.com/loi/hedp20

Effectiveness of spelling interventions for learners with dyslexia: A meta-analysis and systematic review

Katharina Galuschka, Ruth Görgen, Julia Kalmar, Stefan Haberstroh, Xenia Schmalz & Gerd Schulte-Körne

EDUCATIONAL PSYCHOLOGIST
https://doi.org/10.1080/00461520.2019.1659794

Effectiveness of spelling interventions for learners with dyslexia: A meta-analysis and systematic review

Katharina Galuschka[a], Ruth Görgen[a], Julia Kalmar[b], Stefan Haberstroh[a], Xenia Schmalz[a], and Gerd Schulte-Körne[a]

[a]Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, University Hospital, LMU Munich, Munich, Germany; [b]Center for Psychological Psychotherapy, University of Heidelberg, Heidelberg, Germany

ABSTRACT

This systematic review and meta-analysis investigated the efficacy of spelling interventions for the remediation of dyslexia and spelling deficits. Theoretically important moderators, such as the treatment approach as well as orthographic and sample characteristics, were also considered. Thirty-four controlled trials that evaluated spelling interventions in children, adolescents, and adults with dyslexia and spelling deficits were included. Results show that treatment approaches using phonics, orthographic (graphotactic or orthographic phonological spelling rules), and morphological instruction had a moderate to high impact on spelling performance. A significant influence of interventions that teach memorization strategies to improve spelling could not be confirmed. This work shows that understanding the principles of an orthography is beneficial for learners with dyslexia or spelling deficits and presents key components for effective spelling intervention.

Learning to spell is a challenging task. In most alphabetic orthographies, spelling skills are more difficult to acquire than reading skills (Bosman & Van Orden, 1997). For a substantial number of children, learning to spell remains challenging across development (Maughan et al., 2009). Difficulties with spelling (e.g., adding, omitting, or substituting letters or graphemes in written words) are a common indicator of specific learning disorders and a core characteristic of dyslexia. The term *dyslexia* is used to refer to a pattern of learning difficulties characterized by problems with word recognition, decoding, and spelling. By definition (Lyon, Shaywitz, & Shaywitz, 2003), dyslexia typically results from an underlying phonological deficit that is not due to cognitive impairment or ineffective classroom instruction. As the term *dyslexia* implies, spelling deficits often co-occur with reading problems.

Despite the frequent co-occurrence, spelling and reading difficulties can also be observed in isolation (Moll, Kunze, Neuhoff, Bruder, & Schulte-Körne, 2014). Frith (1980) proposed that this dissociation arises because children rely on a partial cue reading strategy. Mental representations of the word's orthographic form (orthographic representations), its link to phonology and semantics, are required for reading and spelling. For reading, relying on a partial cue, or weakly specified (incomplete) orthographic representations, is sufficient to quickly recognize a printed word. For spelling, high-quality orthographic representations are necessary because the readers need detailed information about all the letters within this word (Perfetti & Hart, 2002). Moll and Landerl (2009) hypothesized that problems in building up orthographic representations might encourage some learners with dyslexia to develop highly efficient decoding strategies as a compensatory mechanism. In those children, reading is in the normal range, whereas spelling remains below average.

This study extends the current knowledge of effective spelling treatment approaches for learners with dyslexia by giving a comprehensive, up-to-date overview of spelling interventions. To advance toward a theory of spelling that describes how spelling performance of learners with dyslexia and spelling deficits can be improved, we quantify the effectiveness of different spelling treatment approaches. We also consider the impact of age and severity of reading and spelling disorder on the effectiveness of spelling treatment approaches to assess whether phonics intervention should precede morphological or orthographic intervention. As a further theoretical contribution, we aim to capture the influence of orthographic consistency on spelling interventions

CONTACT Katharina Galuschka katharina.galuschka@med.uni-muenchen.de Department of Children and Adolescent Psychiatry, Psychosomatics and Psychotherapy, University Hospital Munich, Munchen 80336, Germany.
Note. Katharina Galuschka and Ruth Görgen share joint first authorship.

Supplemental data for this article can be accessed on the publisher's website.

and examine the impact of sample and intervention characteristics on the efficacy of spelling interventions.

Theoretical perspectives on the relation between spelling and text production

Spelling is very important for a person's academic and professional achievement. Severe spelling errors can make a text difficult to read and can cause readers to devalue the quality of a written composition (Graham, Harris, & Hebert, 2011; Graham & Hebert, 2011). The consequence of poor spelling skills extends beyond the readability of a text and the appearance of academic incompetence. Spelling requires considerable cognitive effort for poor spellers, so that fewer resources are left for other cognitive processes that are important for writing (e.g., content development). It is hypothesized that spelling deficits (beside handwriting skills) induce cognitive overload and influence text production fluency and quality (Berninger & Swanson, 1994; McCutchen, 1996, 2000).

Continued overload of cognitive resources during spelling may weaken the subsequent development of writing skills (Graham, Berninger, Abbott, Abbott, & Whitaker, 1997; Kellogg, 2008). Therefore, poor spelling is an obstacle for text production in the same way that difficulties in word recognition are a bottleneck for reading comprehension (Perfetti & Hogaboam, 1975). Previous research about the association between spelling and text production has shown that spelling instruction improves compositional fluency (Graham, Harris, & Adkins, 2018; Graham, Harris, & Chorzempa, 2002). However, it is unclear whether spelling ability also has an influence on composition quality (Graham & Santangelo, 2014; Graham et al., 2018). This would support the theoretical view that effective spelling instruction enhances spelling performance, which in turn frees cognitive resources and paves the way for the acquisition of advanced text production skills. The improvement of compositional quality, however, seems to depend on explicit instruction in composition.

Theoretical perspectives on learning to spell and its influence on reading

Spelling not only affects writing and compositional skills but also is closely linked to reading (Ehri, 1987; Frith, 1985). Most of the linguistic abilities that are important for reading acquisition (e.g., phonological and phoneme awareness, explicit orthographic or morphological knowledge) are also important for spelling ability. In some phases, spelling is essential for reading because it trains phoneme awareness and the alphabetic principle, whereas in other phases, reading boosts spelling because the reading exposure leads to a richer lexicon of orthographic representations.

The development of spelling and reading skills starts long before children learn to read and write in school. First, preschool children are able to differentiate between writing and drawing; for example, they learn to recall the meaning of some street signs and graphic signals. Later, children acquire phonological awareness skills and become increasingly better at detecting or manipulating syllables and phonemes in spoken words as well as at segmenting spoken words into syllables and phonemes (Treiman & Bourassa, 2000). Phonological awareness skills and the ability to segment spoken words into phonemes is an important precursor of reading and spelling acquisition (Moll, Ramus, et al., 2014). Thus, it is not surprising that *phonological awareness interventions* (including oral tasks for recognizing phonemes within words, blending phonemes into words, segmenting a word into phonemes, eliminating a phoneme from a word, or adding a phoneme to a word) are often implemented to foster spelling and reading skills. It is often assumed that phonological awareness interventions in preschool years are beneficial for early reading and spelling development. At later stages, a learner has already gained insight into the phonemic structure, and those interventions may lose their effectiveness in improving reading and spelling skills (Bus & van Ijzendoorn, 1999; Galuschka, Ise, Krick, & Schulte-Körne, 2014).

According to Frith (1985), instruction in spelling and writing trains the ability to segment spoken words into phonemes and enables the insight into the alphabetic principle: the understanding that graphemes represent phonemes. This insight is then transferred from spelling to reading and allows children to read words they have not encountered before (Frith, 1985). Share (1995) argued that knowledge of phoneme–grapheme and grapheme–phoneme correspondences make up a self-teaching mechanism, which supports the acquisition of orthographic skills and fosters sight word reading (direct retrieval of phonology and semantics associated with a given written word form from the orthographic lexicon). As reading becomes more automatized, fewer cognitive resources are required for decoding graphemes into phonemes. Thus, these resources can be allocated to reading comprehension (Ehri, 2005, 2014).

Individuals demonstrate *phoneme–grapheme knowledge* when they understand the simple mappings between phonemes and graphemes. Because of the strong influence of phoneme–grapheme and grapheme–phoneme correspondence skills on reading and spelling ability, research on reading and spelling instruction and intervention has predominantly focused on the acquisition of this fundamental knowledge. Such treatment approaches are often referred to as *phonics interventions.* A large body of research supports the efficacy of those interventions on reading and spelling performance (Ehri, Nunes, Stahl, & Willows, 2001; Galuschka et al., 2014; McArthur et al., 2012; Weiser & Mathes, 2011). At the same time, representing every phoneme with a reasonable letter is not sufficient for skilled spelling, as can be shown by incorrect spellings such as "wodr" for *water* or the German equivalent "Wasa" for *Wasser*. In many orthographies, one sound can have different spellings in different words (e.g., the phoneme /i/ in English may be spelled < ee > as in *deep* or < ea > as in *mean*; in German, /ʃ/ may be spelled < sch > as in *Schule* "school" or < s> as in *Stuhl* "chair"). According to Treiman and Kessler (2014), words that cannot be spelled correctly by simple phoneme–grapheme correspondences may follow other patterns: (a) graphotactic (e.g., consonant doubling is less common before /ɪk/ when it is spelled as <ic> [e.g., *magic*] than when it is spelled as <ick> [e.g., *gimmick*]), (b) phonological–orthographic (e.g., consonant doublets do not occur after long vowels in Dutch, English, and German, as in *late* [long *a*] vs. *latte* [short *a*]), or (c) morphological (e.g., in German vowel change [umlaut] for noun plurals are dependent on the singular *Baum* "tree" to *Bäume* "trees": An understanding of the morphological relationship will allow children to spell the plural form correctly, as opposed to a phonologically plausible but incorrect spelling such as *Beume*). To become a skilled speller, the learner has to increasingly consider those patterns or regularities (Ehri, 2000; Frith, 1985; Treiman, 2018). This is true for all learners, regardless of whether they have a diagnosis of dyslexia.

Interventions aiming to help children with dyslexia and spelling deficits deal with the deviations from one-to-one mappings between phonemes and graphemes and provide explanations for these deviations by explicitly teaching morphological or orthographic knowledge. *Morphological knowledge* refers to the awareness of the smallest meaningful language units (Carlisle, 2003; Nagy, Berninger, Abbott, Vaughan, & Vermeulen, 2003). A large body of research supports the efficacy of morphological instruction (Carlisle, 2003; Carlisle, McBride-Chang, Nagy, & Nunes, 2010; Goodwin & Ahn, 2010, 2013; Nagy et al., 2003). *Morphological interventions* for younger children often include practicing of inflections, whereas morphological interventions for older children often focus on derivations (Kirby et al., 2012). *Orthographic knowledge* refers to the understanding of the orthographic rule system that allows correct writing in terms of rules and patterns of written language. Interventions that foster orthographic skills mainly focus on graphotactic and phonological-orthographic regularities. Children have to learn the constraints and orthographic rules that govern the positions, combinations, and resulting pronunciations of letters within words. Squires and Wolter (2016) synthesized intervention studies that evaluated *orthographic interventions* and confirmed their efficacy on spelling skills.

According to the theory of Ehri (2000, 2014), reading supports spelling through better retention of written words in memory. However, the work of Ouellette, Martin-Chang, and Rossi (2017) shows that high-quality orthographic representations for spelling further support sight word reading. These results highlight, once again, how reading and spelling influence each other. The ideal process to build up complete orthographic structures is therefore of particular profound importance and has been part of deep theoretical discussions.

Theoretical discussion: Phonics first or morphology

A common view of spelling acquisition is that, at the beginning, instruction should focus on phoneme–grapheme correspondences, and orthographic and morphological knowledge should be taught after the basic correspondences are acquired (Treiman & Kessler, 2014). This view is supported by theoretical models of literacy acquisition (Castles, Rastle, & Nation, 2018; Ehri, 2000; Frith, 1985; Jackson & Coltheart, 2013; Sprenger-Charolles, Siegel, & Bonnet, 1998) and a large number of empirical findings (e.g., Castles et al., 2018; Moll, Ramus, et al., 2014; Taylor, Davis, & Rastle, 2017). These relatively simple phoneme–grapheme correspondences are sufficient to decode most of the words that beginning readers encounter in age-appropriate reading materials and provide a potent self-teaching mechanism to get at least a partial phonological representation of unfamiliar words (Rastle & Taylor, 2018). This view implies that morphological or orthographic interventions would be less effective for young children in the early grades or for more severely impaired dyslexics than for older learners in the upper grade levels or less severely impaired dyslexics.

In contrast, Bowers and Bowers (2017) postulated that phonology, morphology, and etymology should be taught jointly from the beginning of literacy instruction. They argued that the alphabetic principle does not provide a reasonable description of the words that children are confronted with first, even in age-appropriate materials (Bowers & Bowers, 2018b). The English orthography becomes more transparent when the origin and morphological structure of words are considered. For example, the spelling of the word *nation* is opaque in the sense that there are different phonologically plausible ways to spell it (e.g., "nayshun"). The spelling, however, becomes more transparent once we consider the Latin origin and the morpheme *–ion.* Thus, Bowers and Bowers (2017) proposed that children should learn the logic of the written language system in a meaningful context, and they disagreed with the assumption that children are not yet able to deal with such complex aspects at the beginning of written language aquisition. This proposal is controversal, and the question of the right starting point of phonics, orthographic, or morphological instruction resulted in an ongoing theoretical debate in reading research (Bowers & Bowers, 2018a, 2018b; Rastle, 2018; Rastle & Taylor, 2018).

To examine this theoretical debate and to approach a theory of spelling that accurately describes how learners with dyslexia and spelling deficits should be taught spelling, we aim to estimate the influence of age or grade level, as well as spelling performance on the effectiveness of phonics and morphological interventions. Irrespective of the great importance that graphemes and phonemes undoubtedly have in alphabetic orthographies, the potential of morphological interventions should not be underestimated especially for learners with reading and spelling deficits. Morphological awareness has been shown to relate to reading and spelling (McCutchen & Logan, 2011; Volkmer, Schulte-Körne, & Galuschka, 2019), and morphological interventions have been shown to foster reading and spelling skills (Carlisle et al., 2010; Goodwin & Ahn, 2010; McCutchen & Logan, 2011). There are also theoretical arguments implicating that the influence of morphological knowledge extends to writing quality (McCutchen, Stull, Herrera, Lotas, & Evans, 2014; Saddler & Graham, 2005). Thus, morphological interventions appear to be a promising way to support reading, spelling, and text production skills in learners with dyslexia and spelling deficits. This could be even more important in orthographies that adhere to the principle of morpheme consistency (i.e., where the spelling of morphemes is preserved in different word forms even if they have different pronunciations).

Theoretical perspectives on cross-linguistic differences

This contemplation leads us to another important theoretical issue on spelling interventions: the influence of orthography-level differences on the efficacy of different spelling treatment approaches. The general framework of literacy development (as just described) is valid across alphabetic orthographies: It seems that in European orthographies, alphabetic skill is a predictor of reading and spelling success (Caravolas, 2004). However, all aspects of spelling and reading are connected to the characteristics of a specific orthography. Orthographies vary in the degree to which the spellings reflect the phonology or morphology (Katz & Frost, 1992). In light of the phonics-first versus morphology-first debate, taking these orthography-level differences into account may prove pertinent.

In the previous decades, cross-linguistic studies mainly examined orthographic depth as a source of cross-linguistic differences in literacy acquisition (Serrano et al., 2011; Seymour, Aro, & Erskine, 2003). Orthographic depth is generally conceptualized as orthographic consistency: the presence of more than one pronunciation for a given letter or a cluster of letters (reading direction), or the presence of multiple spellings for one phoneme (spelling direction; Borgwaldt, Hellwig, & De Groot, 2004; Schmalz, Marinus, Coltheart, & Castles, 2015). Consistency thus measures the extent of adherence to the alphabetic principle, which varies greatly between orthographies (Caravolas, 2005).

There is a general consensus about the approximate classification of several orthographies in terms of their orthographic consistency in reading (feedforward) and spelling (feedback) direction (Seymour et al., 2003). Finnish has been shown to be on the "consistent" end of the orthographic consistency continuum, both in reading and spelling direction. English is on the opposite end of this continuum. It is a morphophonemic system in which morphemes are often preserved at the cost of phonological consistency (e.g., the same grapheme, *c*, is used to represent the phonemes /k/ and /ʃ/ in the morphologically related word pair *magic–magician*). Historical spellings are also a big source of inconsistency in the English spelling system. Historical spellings represent sounds that were pronounced differently in the past than they are pronounced today (e.g., in English, words beginning with a silent *k*, such as *knight*). Like English, German has been found to be relatively inconsistent in the feedback direction (Borgwaldt, Hellwig, & De Groot, 2005; Borgwaldt et al., 2004). This is mainly caused by morpheme consistency, but there are also some graphotactic and context-dependent phonological features to consider. Vowel-length marking is particularly challenging for German children at the beginning of literacy acquisition or children with spelling deficits. However, even in the feedback direction, German is more consistent than English.

One method for a quantitative classification that allows for formal ranking of languages on a continuum based on their orthographic consistency was presented by Borgwaldt et al. (2005)—namely, onset entropy. Onset entropy focuses on the relationship between the first letter and first phoneme of all words in a given orthography. It quantifies the extent to which words with the same first letter also have the same first phoneme as a measure of feedforward consistency, and vice versa, for feedback consistency (for further details, see the Method section). The advantage of using word onset entropy measures is that the orthographic consistency of languages with varying orthographic structures can be objectively compared without relying on intuition (Borgwaldt et al., 2005; Schmalz et al., 2015).

Alphabetic orthographies contain varying degrees of orthography-specific uncertainties induced by morphological consistency, context-dependent phonological-orthographic features, and etymology. The higher the number and complexity of rules, the number of exceptions, and sources of inconsistency of an orthography, the more difficult it is to impart this knowledge and to use it for the remediation of spelling deficits (Caravolas, 2005). For very complex orthographies such as English, it has been argued that interventions concentrating on memorization strategies of whole-word letter sequences are more efficient than interventions focusing on the teaching of complex rules and regularities (Dymock & Nicholson, 2017; Hilte & Reitsma, 2011). This implies that the interventional aspects will vary not only in their importance for a given orthographic system (Bowers & Bowers, 2017) but also in their applicability for spelling remediation. For interventional research on spelling, it is important to know if there are treatment approaches that are less effective in one orthography than in another. This would contribute to an understanding about the interaction between universal cognitive skills and language characteristics. Previous reviews and meta-analyses examined the effects of spelling interventions mainly for the English orthography (Bowers & Bowers, 2017; Graham & Santangelo, 2014; Wanzek et al., 2006; Williams, Walker, Vaughn, & Wanzek, 2017). Reviews and meta-analyses that include studies from other languages do not allow for a comparison between the efficacy of spelling interventions in different orthographies (Ehri, Nunes, Stahl, et al., 2001; Ehri, Nunes, Willows, et al., 2001; Galuschka et al., 2014; Ise, Engel, & Schulte-Körne, 2012).

Potential factors influencing treatment efficacy

Previous syntheses on the effectiveness of treatment approaches for learners with dyslexia showed that interventions should focus directly on literacy skills and not on underlying auditory or visual causes to remediate spelling and reading skills (Galuschka et al., 2014). In addition, direct, explicit instruction in spelling has been shown to be more effective for spelling improvement than implicit self-study interventions. Previous studies also report that intervention components such as immediate corrective feedback and various exercise opportunities are beneficial for spelling skills (Wanzek et al., 2006; Williams et al., 2017). Galuschka et al. (2014) and Williams et al. (2017) did not find a significant difference between interventions that were implemented individually or in small groups (two to six children) on spelling performance, so spelling deficits may be more effective when conducted in individual sessions and in small groups.

It seems likely that the implementation of interventions by professional personnel (e.g., teachers, psychologists, therapists) is of a higher quality than the implementation of interventions by paraprofessional personnel (personnel with less specific expertise like teacher assistants or students). It is therefore possible that the professionalism of the implementer influences treatment efficacy. However, previous research on this topic has shown contradictory results (Scammacca, Roberts, Vaughn, & Stuebing, 2015; Williams et al., 2017). There is also a lack of consistent research findings concerning the influence of age and grade level on treatment efficacy. Ehri, Nunes, Stahl, et al. (2001) and Goodwin and Ahn (2013) showed that early intervention has the greatest impact on reading and spelling performance, compared with interventions performed in older children. A meta-analysis by Graham and Santangelo (2014) did not confirm this finding. Other meta-analyses on the efficacy of reading and spelling treatment approaches (Galuschka et al., 2014; Wanzek et al., 2006) did not evaluate whether the efficacy of intervention is affected by age or grade level.

A further relevant topic with regard to effective treatment methods is the effect of software usage. Over the past decade, the use of software in educational and therapeutic context has increased (Jamshidifarsania, Garbayab, Limc, Blazevica, & Ritchie, 2019). This is likely to bring many advantages: First, computerized interventions appear to encourage learning through exceptional motivational aspects (e.g., a rewarding system, an attractive graphic design, or an exciting storyline) (Ecalle, Magnan, Bouchafa, & Gombert, 2009; Kast, Baschera, Gross, Jäncke, & Meyer, 2011; Thalmann, 2014). Second, algorithms can enable the implementation of specialized instructions and adaptive content in computerized interventions (Truong, 2016). To our knowledge, the effectiveness of computerized spelling interventions versus classroom settings has not yet been systematically tested using a meta-analytic approach. Wanzek et al. (2006) evaluated the influence of assistive software and demonstrated positive effects of software components on speech synthesis, word prediction, and spell-checking on measures of students' spelling accuracy and correction.

To date, it is unknown how the effectiveness of spelling interventions varies between different treatment approaches, and little is known about the impact of sample and intervention characteristics on the efficacy of spelling interventions (e.g., age and grade level of the participants, intensity or length of the intervention, setting characteristics). In addition, there is little empirical work on how the effectiveness of different spelling treatment approaches varies between orthographies.

Purpose statement

The purpose of our systematic review and meta-analysis is to extend the current knowledge about the effectiveness of spelling treatment approaches on reading and spelling performance of learners with dyslexia and spelling deficits. The specific aims of the current study are fivefold.

First, we aim to present a comprehensive up-to-date overview (in the form of a systematic review and meta-analysis) of controlled trials of all available spelling interventions evaluated on children, adolescents, and adults with dyslexia or spelling deficits. Previous syntheses indicated a moderate to high effect size of spelling interventions on spelling outcomes (Ise et al., 2012; Williams et al., 2017). We expect that, overall, spelling intervention is effective (Hypothesis 1). However, we do not have specific predictions about the size of the effect. By including a description of the quality of the literature and using study design (randomized controlled trials [RCT] vs. nonrandomized controlled trials [NRCT]), the standardization, and the blind assessment of the outcome measures as moderators, we aim to assess the validity of study findings and to identify gaps that can be used to guide the design of future studies in the field.

Second, to advance toward a theory of spelling that describes how spelling performance of learners with dyslexia and spelling deficits can be improved, we perform a meta-analysis for different treatment approaches separately (phonological awareness, phonics, morphological, orthographic, memorization, and audio-visual cue interventions, as well as assisted writing and supportive software approaches). A detailed description of the interventions is displayed in the Method section. In line with the work of Weiser and Mathes (2011), Squires and Wolter (2016), and Goodwin and Ahn (2010, 2013), we expect effects to be significant for phonics, orthographic, or morphological instructions (Hypotheses 2). Because reading and spelling influence each other in all phases of literacy development, we evaluate the effects separately for reading and spelling.

The third aim is to investigate interaction effects between treatment approaches and age, as well as severity of reading and spelling disorder. With this analysis, we aim to examine the recent theoretical discussion between Rastle and Taylor (2018) and Bowers and Bowers (2018a) and to determine whether phonics intervention should precede morphological and orthographic intervention. Thus, we are especially interested in whether the efficacy of phonics, morphological, or orthographic interventions is affected by age or severity of reading and spelling disorder. Considering the knowledge of phoneme–grapheme (and grapheme–phoneme) correspondences as a fundamental skill that affects higher order skills, such as the application of orthographic or morphological spelling rules (Castles et al., 2018; Share, 1995), we hypothesize that more severely impaired dyslexics would benefit more from a phonics approach compared to an orthographic or morphological intervention. Furthermore, the efficacy of phonics instructions should decrease with age (Hypotheses 3). As research is lacking on the evaluation of morphological and orthographic interventions in first or second grade, we do not have any a priori expectations about the efficacy of those interventions in the early years of literacy acquisition.

Fourth, we aim to contribute to an understanding how orthographic consistency influences the implementation and effectiveness of spelling interventions. Thus, we assess if treatment approaches vary between different orthographies and if the effect sizes for spelling interventions are affected by the orthographies in which they were implemented. We are especially interested in contrasting intervention effects between spelling interventions that were implemented in English and German orthography. Because a higher number of exceptions and sources of inconsistency of orthography is associated with increased difficulties to remediate spelling deficits (Caravolas, 2005), we expected that for more consistent orthographies, the effect of spelling interventions on reading and spelling would be larger than for inconsistent orthographies. Thus, we assume that intervention effects are larger if the intervention was performed in German than in English orthography (Hypotheses 4).

The fifth aim is to investigate whether effect sizes are affected by sample (age, grade level, severity of reading and spelling disorder) and intervention characteristics (amount of the intervention, setting, computer use, implementer), in an explorative way. To inform theory and practice, it is important to know whether the effects are consistent across learners and if general principles are relevant for a broad range of settings. Research on this topic is lacking for spelling interventions; therefore, we do not make any predictions about which sample and intervention characteristics should influence intervention effects.

Method

We conducted the systematic review according to the recommendations of the Cochrane group (Higgins & Altman, 2008) and following the PRISMA guidelines (Moher, Liberati, Tetzlaff, Altman, & The-PRISMA-Group, 2009).

Literature search

Galuschka, Görgen, and Kalmar searched for spelling intervention studies published from December 1989 to December 2018 in the following databases: ERIC, PsycINFO, Medline, and peDOCS. The literature search began in July 2016, and the following search terms were used in Abstract, Title, and Keywords:

(Dyslexia OR spelling disorder OR developmental spelling disorder OR specific spelling disorder OR developmental reading disorder OR reading disability OR reading disorder OR Lesestörung OR Rechtschreibstörung OR Lese-Rechtschreibstörung OR Lese-Rechtschreibschwäche OR Leseschwäche OR Rechtschreibschwäche OR Legasthenie)

AND

(Invented spelling OR spelling OR writing OR writing comprehension OR writing improvement OR writing processes OR writing strategies OR orthographic knowledge OR grammar OR orthographic processing OR word lists OR letter OR orthograph* OR written communication OR diction OR notation OR morphological OR spelling book OR spelling check OR Orthographie OR orthographisch OR rechtschreib* OR orthographic rules OR Regeltraining OR Wortlisten OR dictation OR Diktat)

AND

(Intervention OR treatment OR therapy OR therapeutics OR training OR Förderung OR Therapie OR remediation OR instruction OR exercises OR teaching)

To consider unpublished studies, we searched the website ProQuest using the aforementioned search terms. We also examined bibliographical references of systematic reviews (Graham et al., 2018; Graham & Santangelo, 2014; Ise et al., 2012; Squires & Wolter, 2016; Wanzek et al., 2006; Williams et al., 2017). In addition, we contacted experts by sending an e-mail to the members of the Society for the Scientific Studies of Reading mailing list.

Study selection criteria

To be considered for this analysis, studies had to meet the following criteria:

1. The intervention consisted at least partly of tasks and instructions that aimed to improve spelling performance. The following contents were classified as tasks and instructions that aimed to improve spelling performance: phonological and sound-based exercises, such as verbal segmentation and analysis of words and syllables; exercises in grapheme–phoneme mapping; memorization of word spellings, such as special pronunciations of orthographic markers or visualization of written words in memory; explicit instruction and application of graphotactic and orthographic spelling rules; morphological instructions in base words, suffixes and inflected, and derived words; and all tasks that were explicitly described in the original article as developed and used to improve spelling performance.
2. One or more standardized spelling tests were administered before and after treatment.
3. Pre- and posttest results of the spelling tests were reported with sufficient detail to allow for the calculation of an effect size or if this information could be obtained from the authors.
4. Study subjects were children (after starting formal schooling), adolescents, or adults with spelling deficits or dyslexia whose spelling or reading performance was below the 25th percentile or at least 1 *SD*, 1 year, or one grade below the expected age or grade level as described in the original study and with an intelligence quotient in the normal range or, if the results of intelligence tests were not reported, subjects were described as having normal intelligence by the study authors.

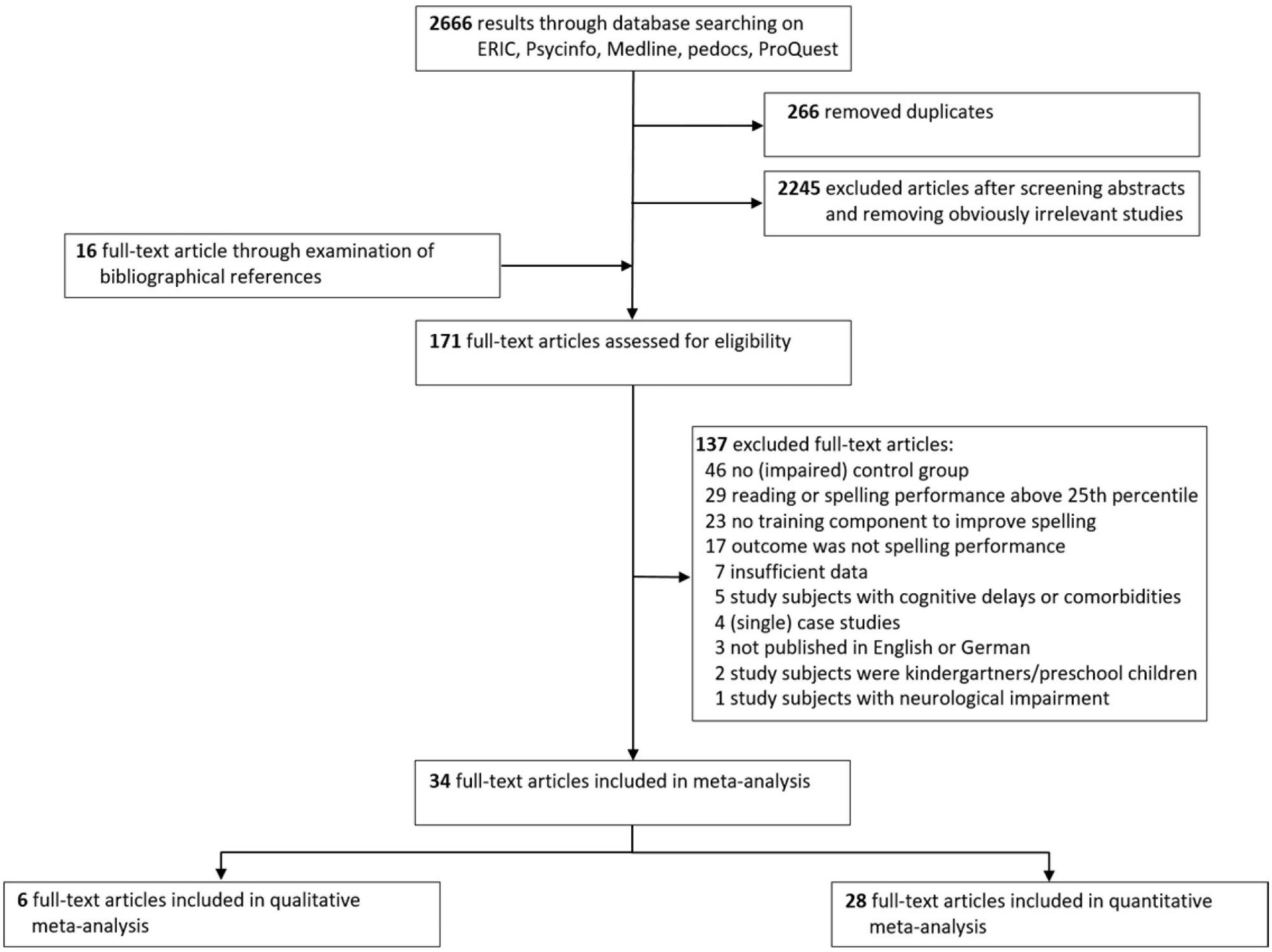


Figure 1. Flowchart of study selection process.

5. Poor spelling (and reading) performance occurred in the subjects' native language, and subjects were also taught in their native language.
6. The study included predominantly subjects without other developmental disorders (such as autism spectrum disorders), comorbid psychiatric disorders (internalizing disorders such as depression or externalizing disorders such as ADHD), neurological diseases (e.g., epilepsy), or chromosomal defects (e.g., trisomy 21).
7. The study design included a control group: an untrained control group, placebo group, regular school practice control group, or another treatment control group.
8. The manuscript was written in English or German (because of the authors' language proficiency). Studies not conducted in English or German were included as long as they were published in English or German.

Figure 1 summarizes the process of selecting studies for the meta-analysis.

Coding of the controlled trials

Following the guidelines of the Cochrane group (Higgins & Altman, 2008), we used coding sheets to record the intervention, methodological, and sample characteristics (description of the characteristics is next). To ensure reliability, study characteristics were double-coded; this was done independently by Galuschka, Görgen, and Kalmar. We determined interrater reliability for article inclusion or exclusion by calculating Cohen's kappa (Cohen, 1960) and for the coding of the different treatment approaches (the description of the treatment approaches is next) by percentage of agreement (Cohen, 1960). For the calculations, we used the independent coding results of the raters. Cohen's kappa of the interrater agreement with regard to inclusion and exclusion was $k = .68$, and the percentage of agreement for treatment approaches was 84%, on average. All discrepancies were discussed and resolved. Study details are provided in Supplement 1. Galuschka and Görgen extracted data necessary for effect size calculations (mostly means and standard deviations of pre- and posttests). The extracted intervention characteristics from the original studies contained the following:

1. The spelling treatment approach (a detailed description is next).
2. Total amount of intervention in hours: <20 hr; 20–40 hr; >40 hr.
3. Setting: class, group, or individual.
4. Computer involvement: noncomputerized, computer based, or computer assisted.

5. Implementer: paraprofessionals, professionals, or a mix of paraprofessionals and professionals.

The spelling treatment approaches were classified into categories based on the descriptions of the intervention in the publication. The category *phonological awareness intervention* included approaches with phonological and sound-based exercises without or with very limited connection to written language. The category *phonics intervention* included approaches gaining basic knowledge of the relationship between phonemes and letters/graphemes, as well as exercises combining phonological awareness with letters. The category *memorization intervention* contained approaches that fostered word-specific memorization of letter sequences and the visual appearance of written words. The category *morphological intervention* included instruction in morpheme and syllable analysis, encouraging the participants to apply knowledge of syllable and morpheme structure to spelling. The category *orthographic intervention* comprised interventions that systematically instructed the application of graphotactic or orthographic-phonological spelling rules. The category *audio-visual cue intervention* comprised multisensory activities combining auditory and visual cues with the aim to support memory strength of graphemes and phonemes. The category *assisted writing* referred to interventions containing assisted writing activities by coaches and tutors. The category *supportive software* consisted of computer-assisted writing approaches.

The extracted methodological characteristics from the original studies included the dependent variable spelling (and reading), the sample size, and the administered spelling (and reading) tests. With regard to the sample, we also extracted (a) the age, (b) the grade level of the participants, (c) the degree of severity of reading and spelling disorder, and (d) the language of instruction from the original studies and coded them as sample characteristics. If the authors did not specify the mean age of their participants, we estimated the age from other specifications like grade level or age range (Arnbak & Elbro, 2000; Faber, 2005; Gustafson, Ferreira, & Ronnberg, 2007; Rashotte, MacPhee, & Torgesen, 2001; Reuter-Liehr, 1993). The moderators grade and age level are not independent of each other. Nevertheless, we included both because grade level refers to years in formal literacy instruction and it is important to know whether the intervention effect of certain treatment approaches changes in the course of the class levels. We used age because we also included studies with adults. Using the mean age of the study population allowed us to check for interaction effects between treatment approach and age (for Hypothesis 3) with a more meaningful analysis. Finally, we considered the following spelling outcomes: phonological spelling, pseudoword spelling, and real-word spelling. Relevant reading outcomes were pseudoword reading, nonword reading, and word reading (accuracy, fluency, and speed), as well as reading comprehension.

The degree of severity of reading and spelling disorder was determined by converting the reading and spelling standard scores of the pretests to z scores. If more than one standardized reading test was administered, and it was not possible to distinguish these by their relevance, the results of the standardized tests were z-transformed and averaged. If more than one standardized spelling test was administered, severity was determined by the standardized spelling test that addresses orthographic writing. Details of our severity assessment are provided in Supplement 2.

If we could not obtain standard scores for calculating z scores, we used raw scores and grade-equivalent scores (Darch, Eaves, Crowe, Simmons, & Conniff, 2006; Gustafson et al., 2007; Hughes, Phillips, & Reed, 2013). We defined the severity categories in reading and spelling as mild (reading and spelling skills between the 25th and 16th percentile, between 0.67 *SD* and 1 *SD* or up to a 1-year curricular delay), moderate (reading and spelling skills between the 15th and 7th percentile or 1 *SD* and 1.5 *SD* or between 1 and 1.5 years curricular delay), and severe (below the 7th percentile or more than 1.5 *SD* or above a 1.5-year curricular delay).

Effect size calculation

To evaluate the efficacy of the interventions, we calculated Hedges's g (Hedges, 1981) and 95% confidence intervals (CI). The effect size calculations were based on between-group comparisons of pre–post change scores using pooled pretest standard deviations, as recommended by Morris (2008). For studies that did not report means and standard deviations, we calculated effect sizes on the basis of F-test values (Darch et al., 2006; Oakland, Black, Stanford, Nussbaum, & Balise, 1998). If studies included more than one intervention group, we followed the recommendations of Higgins and Green (2008): We combined subgroups if they received the same intervention and if the study subjects were in same age or grade level. If we could not combine subgroups of one study, an overweighting of the effect sizes was counteracted by dividing the sample size of the control group by the number of intervention groups. Some studies used multiple reading and spelling tests to determine treatment efficacy, including standardized measures and nonstandardized measures. We preferred standardized measures for effect size calculation. However, if they were not available, nonstandardized measures were used (Arnbak & Elbro, 2000; Gustafson et al., 2007; Kast, Meyer, Vogeli, Gross, & Jancke, 2007; Oakland et al., 1998). If a study reported results for several comparable tests (e.g., several standardized tests measuring different aspects of reading or spelling), we used all tests for calculating the meta-analysis. To avoid errors with inclusion of cluster-randomized controlled trials, we followed the recommendations of Higgins (2011), and reduced the sample size to the number of clusters if statistical methods that account for clustering in the data were not presented in the original paper (Oakland et al., 1998; Reuter-Liehr, 1993; Strehlow et al., 1999; Vadasy, Sanders, & Peyton, 2005; Walter, Schliebe, & Barzen, 2007).

To estimate the average effect size and to account for the statistical dependencies because of multiple effect sizes from the same sample, we used robust variance estimation with small sample adjustments (Tipton, 2015). Robust variance estimation allows for the inclusion and synthesis of dependent effect sizes when correlations between those effect sizes are not reported (Hedges, Tipton, & Johnson, 2010). Instead, we estimated a common mean correlation of $\rho = .7$ between all pairs of effect sizes within each study, and a sensitivity analysis showed robust findings for values of ρ ranging from 0 to 1.0. We marked all mean effect sizes with an apostrophe (g').

If, in a subgroup analysis, synthesized studies reported only one outcome per study, we estimated the average effect sizes with the restricted maximum-likelihood estimator. A random effect model was chosen because we expected the effect size to vary considerably between studies.

All statistical analyses were performed using the free software R version 3.4.0 (R Core Team, 2014). We used the packages *compute.es* (Del Re, 2013), *robumeta* (Fisher & Tipton, 2015), *metafor* (Viechtbauer, 2010), *clubSandwich* (Pustejovsky, 2016) and the Shiny application *weightr* that implements the weight function model published by Vevea and Woods (2005).

Study quality and publication bias

We assessed study quality with the methodology checklist for (randomized) controlled trials published by the Scottish Intercollegiate Guideline Network (SIGN, Harbour, & Forsyth, 2008). The checklist enables the evaluation of the internal validity of the studies and the possibility of biased results. We further evaluated each study to determine whether checks for treatment fidelity were conducted. We then calculated the percentage of items, for each study, that have either not been met or specified in order to represent the risk of bias of each study. We calculated an average value from the resulting risks of bias per study, indicating the overall risk of bias. Lack of randomization, using non-standardized tests, and no blind assessment of the outcomes are likely to lead to overestimations of the true magnitude of an effect size. Thus, we included study design (RCT: yes or no), standardization of the conducted tests that evaluated the interventions' outcomes (yes or no), and blind assessment of the outcomes (yes, no, or not specified) as moderators in our analysis.

We examined publication bias using Egger, Davey Smith, Schneider, and Minder's (1997) method. In addition, we followed the recommendations of recent publications (Carter, Schönbrodt, Gervais, & Hilgard, 2018, June 4; McShane, Böckenholt, & Hansen, 2016) and performed the weight-function model by Vevea and Woods (2005).

The Egger et al. (1997) method plots a regression line between the precision of studies (independent variable) and the standardized effect (dependent variable) and evaluates if the effect size is related to the size of the study. In the weight-function model of publication bias (Vevea & Woods, 2005), the weights represent the process by which some studies are more likely to be published than others, based on statistical significance (p values). It is also possible to calculate an adjusted mean effect size estimate that offers a sense of the likely magnitude of the true effect.

Quantifying orthographic depth

There is, to date, no consensus about the best metric to quantify orthographic depth. The closest to an objective metric that can be applied to different alphabetic orthographies is onset entropy (Borgwaldt et al., 2004, 2005). Here, we considered feedback consistency because the different ways of spelling a given phonological unit should be more relevant to spelling than the different ways of pronouncing a given orthographic unit. Onset entropy focuses on the different spellings of first phonemes: For example, in English, words beginning with the phoneme /dʒ/ can be spelled either with the first letter *g* (*gist*) or *j* (*just*). This means that, for the phoneme /dʒ/, onset entropy is higher than, say, for words beginning with the phoneme /t/, which are generally spelled with the first letter *t*. The entropy for all phoneme–letter associations can be averaged to obtain a single value of orthographic inconsistency for a given orthography. Entropy (H) for each phoneme (x) is calculated with the following formula:

$$H(x) = -\sum_{i=1}^{n} \left(p_i(x = y_i) \times log_2(p_i(x = y_i))\right)$$

For each different spelling of a given phoneme, y_i, we calculated the proportion of times that this spelling occurs relative to all occurrences of this phoneme. This proportion was then multiplied by its logarithm to the second base. The resulting value was summed across all different spellings of a given phoneme. As logarithms of proportions (or, more specifically, numbers between 0 and 1) are negative numbers, the entropy value was then multiplied by -1 to give a positive number.

Here, we calculated the onset entropy in the phonology-to-orthography direction for English, German, Dutch, Danish, Swedish, and Norwegian (Bokmål). For English, German, and Dutch, the orthographic and phonological word forms were taken from the CELEX database (Baayen, Piepenbrock, & Gulikers, 1995). For Danish, Swedish, and Norwegian, the spoken and written word forms were taken from the NST Lexical Databases (available through Nasjonalbiblioteket; https://www.nb.no/sprakbanken/). For all languages, we included all monosyllabic words with greater-than-zero frequency counts. The input files and the Python script that was used to calculate entropy can be downloaded here: https://osf.io/8r6zf/ (Schmalz, 2018). Unfortunately, we could not find an appropriate database for Brazilian Portuguese; therefore we could not calculate onset entropy for Ferraz et al. (2018).

Results

Results of the literature search

The extensive literature search in several electronic databases and systematic reviews revealed 34 studies that met all inclusion criteria and were included in this review.

Four studies included more than one intervention group (Calhoon, 2005; Gustafson et al., 2007; Lange, Mulhern, & Wylie, 2009; Rashotte et al., 2001). We combined the intervention groups, in line with the Cochrane Handbook for Systematic Reviews (Higgins & Green, 2008), if they evaluated similar interventions (the same treatment approach) on participants on the same grade level. We did not combine the two intervention groups (phonological awareness intervention, orthographic intervention) of Gustafson et al. (2007) and the three subgroups (1–2 grade, 3–4 grade, 5–6 grade) of the study by Rashotte et al. (2001), to ensure that we did not lose important information. In total, we included 38 intervention and control group comparisons.

In the 34 studies, the following treatment approaches were implemented: phonics intervention ($n = 8$), morphological intervention ($n = 10$), orthographic intervention ($n = 7$), memorization intervention ($n = 5$), audiovisual cue intervention ($n = 1$), supportive software ($n = 2$), and assisted writing ($n = 1$). Seventeen studies were conducted in English, 11 studies in German, and two studies in Swedish. The other studies were each conducted in Norwegian, Danish, Dutch, and Brazilian Portuguese.

Study quality

Nineteen of the included studies reported a randomized group allocation on the individual level, and we therefore considered them RCTs. In turn, five trials assigned clusters of patients (from one classrooms, school, or clinic) to intervention or control groups (cluster controlled trials or cluster RCTs [cRCT]; Oakland et al., 1998; Reuter-Liehr, 1993; Strehlow et al., 1999; Vadasy et al., 2005; Walter et al., 2007); 10 studies did not perform a randomized group allocation, and we therefore considered them NRCTs. The moderator analysis showed that NRCTs or cRCTS found a higher mean effect size ($g' = 0.70$), 95% CI [0.26, 1.14], than RCTs ($g' = 0.53$), 95% CI [0.36, 0.70], but this difference did not reach statistical significance ($p = .510$). Five studies used nonstandardized tests to evaluate at least one of the outcome measures (Arnbak & Elbro, 2000; Ferraz et al., 2018; Gustafson et al., 2007; Kast et al., 2007; Oakland et al., 1998). Analyses showed a higher mean effect size for standardized tests ($g' = 0.62$), [0.42, 0.82], than for nonstandardized measures ($g' = 0.29$), [−0.00, 0.58], but again no significant difference ($p = .072$). Risk of detection bias is high because only six studies performed a blind assessment of treatment outcomes (Berninger, Lee, Abbott, & Breznitz, 2013; Berninger et al., 2008; Given, Wasserman, Chari, Beattie, & Eden, 2008; Gray, 2015; Gustafson et al., 2007; Hughes et al., 2013). Studies with blind assessment showed a lower mean effect size ($g' = 0.33$) than studies that did not perform a blind assessment ($g' = 0.56$), [0.34, 0.78], or did not specify if a blind assessment took place ($g' = 0.77$), [0.33, 1.22]. This difference on the descriptive level could not be statistically confirmed ($p = .278$).

Eighteen studies provided details on how the authors checked for treatment fidelity. In these 18 studies, treatment fidelity was mostly ensured through supervising sessions. However, only a few publications reported the percentage of the level of adherence to the treatment script or plan. The vast majority of our study pool did not state whether there were any differences in treatment fidelity across the control and intervention groups, or between pre- and posttest. Details of the quality assessment for each study are displayed in Supplement 3.

Effectiveness of spelling interventions

We conducted a meta-analysis that included 28 studies (31 experimental and control group comparisons). The analysis included placebo, waiting list, and untreated control groups, or compared the treatment effects against control groups that received regular school practice or other treatment approaches that did not correspond to treatment approaches of the experimental groups in other included studies. Table 1 displays the results of the meta-analysis separately for reading and spelling performance. In regards to our first hypothesis, we found, overall, a moderate effect size for spelling treatments on spelling performance and a small to moderate effect size for reading performance.

With the high I^2 values, there was evidence of substantial heterogeneity. This indicates that the analysis combined intervention effects that are likely drawn from populations with different true effect sizes. To address this issue, and to explore the general effectiveness of different spelling treatments on reading and spelling performance (Hypothesis 2), we performed subgroup analyses. We calculated separate mean effect sizes for the identified treatment approaches if they comprised a minimum of two studies. Eight studies were coded as phonics (Blachman et al., 2014; Ehri, Satlow, & Gaskins, 2009; Ferraz et al., 2018; Hughes, Phillips, & Reed, 2013; O'Shaughnessy & Swanson, 2000; Oakland, Black, Stanford, Nussbaum, & Balise, 1998; Rashotte, MacPhee, & Torgesen, 2001; Vadasy, Sanders, & Peyton, 2005), eight studies as morphological (Arnbak & Elbro, 2000; Corvacho del Toro, 2016; Gustafson, Ferreira, & Ronnberg, 2007; Kirk & Gillon, 2009; Reuter-Liehr, 1993; Tijms, 2011; Walter, Schliebe, & Barzen, 2007; Weber, Marx, & Schneider, 2002), six studies as orthographic (Calhoon, Sandow, & Hunter, 2010; Darch, Eaves, Crowe, Simmons, & Conniff, 2006; Faber, 2005; Given, Wasserman, Chari, Beattie, & Eden, 2008; Groth, Hasko, Bruder, Kunze, & Schulte-Körne, 2013; Ise & Schulte-Körne, 2010), two studies as memorization intervention (Spencer, Snart, & Das, 1989; Thaler, Landerl, & Reitsma, 2008), and two studies as supportive software (Fasting & Halaas Lyster, 2005; Lange, Mulhern, & Wylie, 2009).

As can be seen in Table 2, *phonics*, *morphological*, and *orthographic interventions* showed significant effect sizes for

Table 1. Mean effect sizes for spelling interventions on reading and spelling performance.

	g', 95% CI	SE	p	I^2	n/k
Reading/Spelling combined	0.58, [0.40, 0.76]	0.09	<.001	67%	31/107
Spelling	0.68, [0.45, 0.92]	0.12	<.001	75%	31/35
Reading	0.42, [0.27, 0.57]	0.07	<.001	54%	23/72

Note. CI = confidence interval; n = number of experimental and control group comparisons; k = number of outcomes.

Table 2. Efficacy of treatment approaches on reading and spelling outcomes.

Treatment approach	Outcome	g', 95% CI	SE	p	I^2	n/k
Phonics	Spelling	0.68, [0.15, 1.21]	0.23	.017	86%	10/14
	Reading	0.62, [0.37, 0.86]	0.10	<.001	67%	10/44
Morphological intervention	Spelling	0.80, [0.39, 1.20]	0.21	<.001	51%	8/8
	Reading	0.30, [−0.18, 0.77]	0.16	.150	27%	5/10
Orthographic intervention	Spelling	0.67, [0.05, 1.28]	0.31	.034	83%	6/6
	Reading	0.19, [0.09, 0.29]	0.02	.014	0%	4/12
Memorization intervention	Spelling	0.22, [−0.39, 0.82]	0.31	.484	0%	2/2
Supportive software	Spelling	1.25, [0.83, 1.66]	0.21	<.001	0%	2/2
	Reading	0.52, [−5.1, 6.14]	0.44	.449	77%	2/3

Note. CI = confidence interval; n = number of experimental and control group comparisons; k = number of outcomes.

reading and spelling outcomes. Effect sizes for spelling outcomes were substantially larger than for reading outcomes.

Studies that evaluated *memorization interventions* reported pre- and posttest results only for spelling tests, which made it impossible to estimate the efficacy of this treatment approach on reading. Memorization interventions did not show a significant effect on spelling. However, as there were only two studies in this category, the results have to be interpreted with caution.

A large and significant mean effect size for spelling outcomes was found for studies evaluating supportive software. The effect on spelling can be considered large and homogeneous. The effectiveness of supportive software on reading performance could not be statistically confirmed. Again, there are only two studies in this category, so results of this subgroup analysis must be interpreted with caution. The mean effect sizes of the treatment approaches phonics, morphological, orthographic, memorization interventions, and supportive software did not differ significantly from one another (see Table 2). Only one study evaluated an *audio-visual-cue intervention* (Kast et al., 2007). In this study, a nonstandardized spelling test was used. Analyses showed an effect of $g = 0.29$ on spelling of untrained words and an effect of $g = 0.52$ on trained word material. *Phonological awareness intervention* was also evaluated from only one study (Gustafson et al., 2007) and did not significantly enhance reading and spelling skills of the children (spelling: $g = -0.03$; text reading: $g = 0.03$; reading comprehension: $g = -0.12$). Jensen, Lindgren, Andersson, Ingvar, and Levander (2000) evaluated a treatment that assisted the participating Swedish adults with reading and spelling disorder in their writing and reading activities with the aim to encourage and motivate them for literacy activities. Effect sizes for reading and spelling outcomes of all individual studies and outcomes are displayed in Supplement 4.

Qualitative evaluation of studies comparing different treatment approaches

In six studies, treatment approaches that included components aiming to support spelling skills were compared directly (Abbott & Berninger, 1999; Berninger et al., 2008, 2013; Gray, 2015; Schulte-Körne, Deimel, Hülsmann, Seidler, & Remschmidt, 2001; Strehlow et al., 1999). In other words, the treatment approaches chosen for the control groups corresponded to our coded spelling intervention categories, which served as an exclusion criterion for the current meta-analysis (see the Method section). However, as they contain important information, we synthesized them narratively. In addition, we calculated effect sizes based on between-group comparisons of the pre–post change scores.

The study of Abbott and Berninger (1999) compared a phonics intervention with a combined approach of phonics and morphological intervention (coded as morphological intervention because of explicit instruction in syllable and morpheme analysis), conducted over a period of 4 months. The results indicated that spelling performance of children between 9 and 13 years of age tends to benefit more from individual instruction in syllabic and morphemic structure than from pure letter-sound relationship instruction. The combined group showed better spelling performance after the intervention than the pure phonics group: $g = 0.25$, 95% CI [−0.66, 1.14], Wide Range Achievement Test (Wilkinson, 1993); $g = 0.06$, 95% CI [−0.84, 0.96], Wechsler Individual Achievement Test (Wechsler, 1992). The results of the reading tests (decoding and reading fluency) did not reveal any differences between the two groups (g ranging between −0.02 and 0.12), but individual growth curves showed a trend toward greater individual treatment response for the group receiving combined treatment of phonics and morphological instruction.

Berninger et al. (2013) compared the effects of two treatment approaches on groups of dyslexic children between 9 and 14 years of age by stepwise introducing different spelling components and strategies over a 5-month period: First, they compared phonics instruction for spelling and reading to phonics in the reading direction only. They found no distinct advantage of one strategy over the other. Adding spelling memorization strategies (mentally photographing a word to identify letters in their correct word position and imagining a word in the mind's eye before writing) and comparing it with phonics components indicated that memorization strategies are less beneficial for improving phonological spelling ($g = -0.55$), 95% CI [−1.36, 0.3], though it may be effective for improving word decoding skills ($g = 0.45$), 95% CI [−0.39, 1.26]. In a subsequent step, they added morphological instruction and compared to memorization strategies. They observed no significant differences in reading and spelling after this step. The authors noted that the duration of morphological instruction may have not been sufficiently long to identify further improvement in reading and spelling.

Three further studies compared morphological intervention and memorization intervention (Berninger et al., 2008; Gray, 2015; Strehlow et al., 1999). The results of these three studies were largely consistent with one another. Strehlow et al. (1999) evaluated a memorization intervention that aimed to improve spelling and learning transfer to untrained words by memorizing the correct spelling of highly frequent regular and irregular German words. The control intervention was coded as a morphological intervention. Pretests revealed differences in intelligence between the two groups of third-grade children (in favor of the memorization intervention group). After 1 year of individual treatment, spelling improvement was significantly greater for children receiving memorization intervention than for children receiving morphological intervention ($g=0.93$), 95% CI [0.04, 1.76]. Considering the differences between groups on the cognitive level, effects could not solely be attributed to the interventions but could also be related to the differences in intelligence.

Gray (2015) conducted a study with adults. After 4 weeks of individual intervention, participants of the morphological group showed greater gains in reading (effect sizes of reading tests between $g=0.21$–0.40), and reading comprehension ($g=0.30$), 95% CI [−0.39, 0.98], than participants receiving memorization intervention. In reading, in contrast to the memorization group, the morphological group showed transfer effects from taught to untaught word material. No clear progress in spelling was revealed for either treatment ($g = -0.03$ to 0.10).

Berninger et al. (2008) evaluated a memorization intervention (referred to as "orthographic spelling treatment" in the original study) and a morphological intervention ("morphological spelling treatment"), conducted in a group setting with children and adolescents in Grades 4–9. The memorization intervention aimed to support precise, high-quality orthographic representations. The goal of the morphological intervention was to help children to establish representations of base words and affixes and to improve their knowledge of morphological spelling rules. After 3 weeks, the results revealed no clear differences between both groups in orthographic spelling ($g=0.16$), 95% CI [−0.48, 0.79]. However, the morphological group seemed to improve more in phonological spelling ($g=0.48$), 95% CI [−0.17, 1.12], than the memorization group. However, the memorization group showed better performance in phonological decoding compared to the morphological group ($g=0.45$), 95% CI [−0.2, 1.08].

Schulte-Körne et al. (2001) compared the effectiveness of a German orthographic intervention with a phonological awareness intervention in group setting, including children between the ages of 9 and 11, with moderate reading and severe spelling impairment. In comparison to the control group, the experimental group improved in spelling performance ($g=0.78$), 95% CI [−0.19, 1.68], and reading ($g=0.66$), 95% CI [−0.29, 1.56]. In a pre–post comparison of the phonological awareness group, spelling performance of the control group decreased. Schulte-Körne et al. concluded that an ineffective treatment, similar to no treatment, leads to degradation in performance compared to children of the same age.

In sum, morphological interventions were applied mainly to children from the upper elementary grades (Abbott & Berninger, 1999; Berninger et al., 2008, 2013) and young adults (Gray, 2015). As pointed out by Gray (2015), morphological interventions could enable transfer effects from familiar to unfamiliar word material. This appears to be especially important for reading fluency, comprehension, and spelling of unknown words (Berninger et al., 2008; Gray, 2015). Benefits of memorization interventions, in comparison to phonics and morphological interventions, could not be confirmed. The results of Schulte-Körne et al. (2001) underline the relevance of effective spelling treatment approaches and support the effectiveness of orthographic intervention to improve orthographic spelling for German children. Detailed information of the narratively synthesized studies is also provided in Supplement 1.

Table 3. Main effects and interactions of age and severity for phonics, orthographic, and morphological interventions.

	Model age				Model severity			
	B	SE	*df*	*p*	*B*	SE	*df*	*p*
Intercept	1.59	0.73	5.54	.077	1.19	0.34	3.73	.029
Morphological intervention	−1.15	1.03	4.10	.327	−1.12	0.70	5.01	.169
Orthographic intervention	−0.57	2.54	3.35	.836	−0.90	0.49	3.87	.147
Age	−0.11	0.08	4.72	.210				
Morphological Intervention × Age	0.13	0.10	4.52	.260				
Orthographic Intervention × Age	0.07	0.22	4.37	.768				
Severity					0.41	0.19	1.79	.171
Morphological Intervention × Severity					−0.90	0.50	4.41	.141
Orthographic Intervention × Severity					−0.63	0.38	3.16	.189

Influence of age and severity of reading and spelling deficit on the effectiveness of different spelling treatment approaches

To address the third aim of our systematic review, we performed meta-regressions. For meta-regressions, all outcomes (reading and spelling skills) were combined. All outcomes allowed us to use all information and to increase statistical power. Studies that did not include or did not specify a certain variable were excluded from the meta-regression in question. We did not find a significant main effect of age or severity on the magnitude of intervention effects (see Table 3).

To test Hypothesis 3, we performed a meta-regression to evaluate if the mean effect sizes of phonics, orthographic, and morphological interventions would be affected by age and severity of reading and spelling disorder. The effectiveness of phonics interventions tended to decrease with age, and the effectiveness of morphological interventions tended to increase with age, but this could be observed only on a descriptive level. Against our hypotheses, the efficacy of phonics interventions decreased with increasing severity, whereas the efficacy of orthographic and morphological interventions increased with increasing severity (see Figure 2). However, because of the small data set, main effects and interactions of the meta-regression models did not allow for clear conclusions (see Table 3).

Cross-linguistic differences

To address the fourth aim (and Hypothesis 4), we performed meta-regressions and a moderator analysis (see Tables 4 and 5) to explore if the mean effect sizes of studies vary depending on orthographic consistency. All outcomes (reading and spelling skills) were combined. Seventeen of our included studies were performed in English-speaking countries (onset entropy = 0.286) and showed a moderate mean effect size. Ten studies came from German-speaking

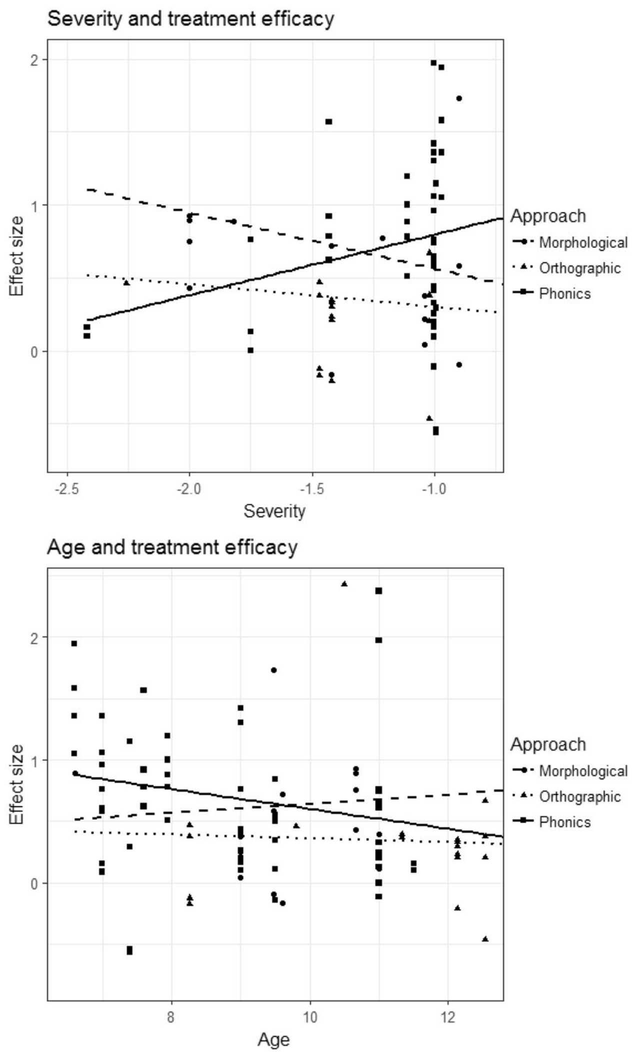


Figure 2. Scatterplot for effect sizes versus age and severity grouped by treatment approach.

Table 4. Meta-regressions for onset entropy, severity, age, and amount of the intervention for the whole study pool combined for reading and spelling outcomes.

Moderator	*B*	SE	*p*	*n/k*
Severity	−0.004	0.16	.983	27/95
Age	−0.01	0.01	.502	31/107
Onset entropy	−1.234	1.329	.363	30/101
Amount	0.00	0.00	.984	30/101

Note. n = number of experimental and control group comparisons; k = number of outcomes.

countries (onset entropy = 0.123) and showed a large mean effect size. The other studies came from Norway (onset entropy = 0.128), Denmark (onset entropy = 0.309), Sweden (onset entropy = 0.124), the Netherlands (onset entropy = 0.124), and Brazil (onset entropy could not be calculated); each orthography was represented by one study. Performing a moderator analysis, interventions in German descriptively showed larger effect sizes than interventions conducted in English, but this difference did not reach statistical significance.

We also evaluated if the effectiveness of spelling interventions varies as a function of onset entropy. Again, this could not be statistically confirmed. Phonics interventions were mainly evaluated for English-speaking children, with one exception, and orthographic and memorization interventions only in German and English. Morphological intervention components seem to be in widespread use in different orthographies like German, Danish, Dutch, and English.

Influence of intervention and sample characteristics

To investigate our fifth research aim, we performed meta-regressions and moderator analyses (see Tables 4 and 5). None of the moderator analyses indicated a significant difference. We observed a trend regarding the setting of an intervention (p = .061). Interventions that were conducted in the classroom tended to show a smaller mean effect size than interventions that were implemented in group or

Table 5 Exploratory moderator analyses combined for reading and spelling outcomes.

Moderator	Level	*g′*, 95% CI	*B*	*F*	*df*	*p*	*n/k*
Treatment approach	Phonics (intercept)	0.63, [0.33, 0.92]	0.65	0.54	2.1	.730	10/58
	Morphological intervention	0.62, [0.22, 1.02]	0.01				8/18
	Orthographic intervention	0.61, [−0.24, 1.46]	−0.05				6/18
	Memorization intervention	0.22, [−0.39, 0.82]	−0.44				2/2
	Supportive software	0.86, [−1.63, 3.36]	0.21				2/5
Language	English (intercept)	0.56, [0.36, 0.76]	0.56	0.46	14.7	.510	15/69
	German	0.81, [0.13, 1.49]	0.21				9/14
	Other	0.48, [0.13, 0.82]	−0.10				7/24
Setting	Class (intercept)	0.28, [0.02, 0.58]	0.28	3.45	16.2	.061	8/26
	Group	0.65, [0.43, 0.86]	0.37				13/50
	Individual	0.81, [0.29, 1.33]	0.52				9/25
Computer	Computer-assisted (intercept)	0.54, [−1.44, 2.52]	0.61	0.67	2.14	.60	2/3
	Computer-based	0.44, [0.15, 0.73]	−0.19				9/23
	no computer	0.68, [0.40, 0.95]	0.05				20/81
Implementer	Mixed (intercept)	0.49, [0.22, 0.78]	0.50	0.30	13.9	.743	7/38
	Paraprofessionals	0.56, [0.30, 0.82]	0.12				8/17
	Professionals	0.63, [0.27, 0.98]	0.12				15/46
Grade	Grade 1–2	0.75, [−0.17, 1.33]	0.71	0.96	8.43	.422	4/21
	Grade 2–4	0.44, [0.21, 0.66]	−0.27				12/37
	>Grade 4	0.75, [0.26, 1.24]	0.03				11/35

Note. Table comprises *F*, *df*, and *p* from a Wald-type tests of linear contrasts from a fitted linear regression model, using a sandwich estimator for the variance–covariance matrix and a small sample correction for the p value. CI = confidence interval; n = number of experimental and control group comparisons; k = number of outcomes.

individual setting. For severity, age, and amount of intervention as continuous variables, we performed meta-regressions. However, no significant link between the amount of intervention and intervention effects was revealed.

Publication bias

We evaluated publication bias in 28 studies (31 experimental and control group comparisons), which we included in the meta-analysis for reading and spelling performance. Based on Egger et al.'s (1997) regression method, we found significant publication bias (standard errors significantly predicted the effect sizes; $p = .035$). To illustrate this bias, we performed Vevea and Woods's (2005) weight function model. We applied the weight functions under consideration of the summary of them. The exact weights we used are presented in Supplement 5.

Assuming a moderate one-tailed selection, our weight function revealed a transformed mean of $g' = 0.43$. The weight function model is based on the assumption that the study's p value affects its probability of publication. It is difficult, if not impossible, to model this adequately. Furthermore, the weight function model was performed on dependent mean effect sizes, which might have an influence on the estimation of the adjusted mean effect size. Thus, the adjusted mean effect size can only be an indication of the true effect size.

Discussion

The purpose of our systematic review and meta-analysis was to extend the current knowledge about the effectiveness of spelling treatment approaches on reading and spelling performance of learners with dyslexia and spelling deficits. Testing our first hypothesis, we found that spelling interventions are effective: Learners with dyslexia or spelling deficits who take part in a spelling intervention show better reading and spelling performance compared with children who received regular school practice or no spelling instruction. This finding is consistent with previous syntheses evaluating the effectiveness of spelling interventions on children with reading or spelling disabilities or dyslexia (Squires & Wolter, 2016; Wanzek et al., 2006; Weiser & Mathes, 2011; Williams et al., 2017).

Effectiveness of spelling interventions

As we hypothesized (Hypothesis 2), mean effect sizes for spelling instructions involving phonics, orthographic, and morphological interventions are significantly greater than zero. Phonics, morphological, and orthographic interventions support children by making the spoken and written language system more transparent. This can help build and automate spoken and written language structures and, in turn, reduce cognitive load. As expected, the present meta-analysis is consistent with earlier work by Weiser and Mathes (2011), Squires and Wolter (2016), and Goodwin and Ahn (2010, 2013). However, Goodwin and Ahn (2010, 2013) showed only a small effect size for morphological trainings on spelling (Goodwin & Ahn, 2010, 2013), whereas the mean effect size of morphological interventions on spelling performance in our analyses can be considered large. This discrepancy in mean effect sizes can be explained by two major differences: First, some of the morphological interventions in our synthesis also included syllable analysis and blending exercises, which influences phonological spelling (e.g., Berninger et al., 2008). Second, instruction of morphemes was always related to their impact on spelling. To foster spelling with morphological interventions, it is especially important to explicitly display the effects of morphological regularities on spelling (Bowers & Bowers, 2017). In addition to spelling, morphological interventions seem to have a positive influence on children with different impairments. They also seem to be very promising for children with additional reading deficits (Goodwin & Ahn, 2010, 2013), specific language impairment (Berninger et al., 2013), and problems in composition or written expression (Graham & Santangelo, 2014).

Using qualitative and quantitative methods, we did not find a significant influence of memorization intervention on spelling performance. However, we found only two studies evaluating memorization interventions that could be included in the meta-analysis (Spencer, Snart, & Das, 1989; Thaler, Landerl, & Reitsma, 2008). Thus, we may have had insufficient power to detect a small but true improvement. Nevertheless, they do not seem to be effective for remediating problems with phonological spelling (Berninger et al., 2013), and they do not support the acquisition of strategies that lead to transfer effects (Berninger et al., 2008; Gray, 2015).

Phonics first or morphology

We did not find that phonics instruction is more effective than morphological interventions in the early years of formal literacy instruction or for more severe spelling deficits. According to our meta-regression analyses, it seems likely that phonics, morphological, and orthographic interventions are applicable across a wide age span. Our qualitative synthesis indicates that instruction in morphemic structure might be especially promising for children in the upper elementary and secondary grades (Abbott & Berninger, 1999). However, meta-analyses of morphological interventions demonstrated significantly larger effects for preschool and early elementary students than for older students in the upper elementary grades and high school (Goodwin & Ahn, 2010, 2013).

Phonics treatments aim to help students crack the alphabetic code and are therefore widely used in treating children with dyslexia and spelling deficits, especially during the early stages of reading and spelling acquisition (Berninger et al., 2008, 2013). The knowledge of phoneme-grapheme (and grapheme-phoneme) correspondences is a fundamental skill that affects the application of orthographic or morphological spelling rules (Castles et al., 2018; Share, 1995). However, simple grapheme–phoneme and phoneme–grapheme correspondences are not sufficient for skilled spelling. Orthographic and morphological interventions showed high to moderate effect sizes in our analysis. Thus, we argue that graphotactic and orthographic-phonological spelling rules, as well as morphological instruction, should be provided as soon as children master the basic phoneme–grapheme and grapheme–phoneme mappings and are confronted with text material for which the basic phoneme–grapheme and grapheme–phoneme mappings would lead to spelling and reading mistakes. To resolve the question if, at the beginning of literacy acquisition, phonics instruction alone leads to better literacy outcomes than a combination of phonics and morphological instruction, more primary studies are needed.

Cross-linguistic differences

In line with our fourth hypothesis, the present meta-analysis showed a higher mean effect size for spelling interventions implemented in German than for spelling interventions implemented in English. However, this difference was not statistically significant. Phonics was the most investigated treatment approach for spelling interventions in English. In German, more orthographic and morphological interventions were evaluated. The few studies evaluating orthographic or morphological interventions in English indicated moderate to large effect sizes on spelling skills (Berninger et al., 2013; Calhoon, Sandow, & Hunter, 2010; Darch et al., 2006; Kirk & Gillon, 2009). Even in English, the spelling of many words can be explained through morphological and orthographic rules, and this resource should not be ignored in spelling intervention (Bowers & Bowers, 2017). Thus, in the future, more English studies should be conducted that would assess treatment approaches dealing with orthographic and morphological regularities.

Besides the studies with German and English interventions, our meta-analysis comprised only five studies from other alphabetic orthographies (Dutch, Danish, Swedish, Norwegian, and Brazilian Portuguese). The inclusion of studies from all languages and the use of onset entropy values allowed us to assess the effect of orthography on the effectiveness of intervention along a continuum. We did not find support for our hypothesis, that with increasing orthographic consistency, the intervention effects would increase. As is typical in meta-analysis, the search procedure was limited to studies published in the languages in which the authors were sufficiently proficient to conduct the search and to extract the relevant information. Although searching for papers in a larger variety of languages would undoubtedly present a fuller picture, the current analysis offers the first step to capture the influence of orthographic consistency on spelling intervention and serves as a model of how this can be done in the future.

Nevertheless, our review showed that, across orthographies and even in inconsistent orthographies such as English and Danish, learners with dyslexia and spelling deficits benefit from interventions that are proximal contributors to spelling acquisition and focus on the logic underlying the orthographic system. In orthographies varying in consistency, morphological and orthographic interventions are effective components for spelling remediation.

Factors influencing treatment efficacy

Another aim of the present meta-analysis was to explore several factors influencing the effectiveness of interventions. Regarding the use of technology in spelling intervention, we found a large and statistically significant effect of supportive software on spelling skills. This indicates a potential for supportive software as an additional aid for patients with dyslexia and spelling deficits. Unfortunately, the limited study pool in this subgroup analyses precludes clear conclusions. Further research should tackle this topic more systematically to gain more evidence of how software can be used to help persons with dyslexia or spelling problems to cope with their spelling deficits in professional and daily life. A current meta-analysis on 5- to 16-year-old students has already shown that technological multimedia-based tools for writing can enhance the motivation and the users composing skills (Williams & Beam, 2019). Our subgroup analyses also touched on the topic of computer involvement. We found that traditional individual instruction cannot be substituted by computer-based approaches as it was also stated by Williams and Beam (2019). Nevertheless, the significant mean effect sizes on reading and spelling performance with regard to computer-based approaches are worth noting. Computer-based programs can be helpful for individuals with mild reading and spelling deficits, or they can be useful for bridging the gap until individual treatment can be provided. Individual treatment showed the largest mean effect size, which makes it very likely that computer-based approaches alone are not sufficient, especially for severely affected individuals.

Comparing different settings, interventions that were conducted in the classroom showed the smallest effects on spelling performance. Individual settings appeared to show the largest effect size, suggesting that spelling skills benefit particularly from one-to-one instructions. In classroom settings, instruction may not be specific enough and the work pace not sufficiently adapted to the individual. Our analysis revealed moderate effect sizes on spelling performance for interventions that were implemented in group settings, indicating that intervention in small groups could be an efficient alternative to individual instruction. This finding is similar to research on reading performance, which showed that one-to-one and small-group instruction tend to be particularly effective (Wanzek & Voughn, 2008; Wanzek et al., 2016).

We could not determine the effect of age or grade level on the efficacy of an intervention. Although the age of the participants in our study pool ranged from Grade 1 through adulthood, only a few studies evaluated spelling interventions on children in first grade or in adults. Early intervention seems to be important for an effective remediation (Graham & Santangelo, 2014) and dyslexia and spelling deficits last a lifetime (Maughan et al., 2009). Therefore, it is particularly important that future studies represent learners with dyslexia and spelling deficits at all ages.

Influence of study quality and publication bias

To ensure high methodological quality of the present meta-analysis, only controlled trials were considered for analysis. Even though our analyses did not show any significant differences between study design, standardization of the conducted tests, and blind assessment of the outcomes, we descriptively saw lower mean effect sizes if a randomized group allocation and blind assessment of the outcome measures was performed. To minimize risk of bias, the magnitude of systematic errors regarding randomization as well as blinding should be considered. It is critical to ensure that an intervention was implemented as planned, and that it was delivered to each study subject in a comparable manner, to attain and demonstrate treatment fidelity. Of our included studies, 47% did not control or did not mention how they controlled for treatment fidelity. Negligence of this aspect was also criticized by previous work (Graham & Santangelo, 2014).

It is well known that publication bias and questionable research practices inflate effect size estimates of meta-analyses (Gelman & Carlin, 2014; Gelman & Loken, 2014; Schmidt, 1992). Our attempts to control for publication bias showed that the effects might be overestimated in our meta-analysis. It is likely that, if all relevant studies were included, the effect size might shift, but the key finding would probably remain unchanged. The adjusted mean effect size is an indication of the true unbiased effect size, but it should be interpreted with caution. The weight function model was performed on dependent mean effect sizes, which might have an influence on the estimation of the adjusted mean effect size. Although we are unable to estimate exactly what effect publication bias has on the results of the current study, it is an important conclusion in itself that there is publication bias in the literature. This calls for preregistered studies to assess the impact of spelling or reading treatments in the future.

Limitations

A limitation of our meta-analysis and systematic review is a lack of an extensive cross-linguistic comparison of spelling interventions. The controlled trials included in the current study mainly focused on spelling interventions in the German and English orthography because of the language proficiency of the authors. Our conclusions therefore are strongest in relation to the English and German orthographies.

We also cannot make any claims as to whether spelling instruction leads to improvements in different aspects of spelling (e.g., spelling word lists, spelling in context, phonological spelling) or other writing skills (e.g., compositional fluency or quality), because the studies we included in our meta-analysis assessed spelling mainly by word dictation and none of our included studies assessed the output of their intervention on writing measures. To resolve this question, more primary studies are needed.

The meta-analysis showed substantial heterogeneity in our study pool. It is a common criticism that the meta-analysis comprises studies that are not compatible with one another. To prevent this and to perform a meaningful analysis, we included only studies with a minimum required quality (controlled trials), applied up-to-date methodological techniques, and performed meta-regressions and subgroup analyses to identify factors that could influence the efficacy of interventions. In addition, we included a qualitative synthesis.

Nevertheless, some of the spelling interventions were part of a broad literacy treatment, whereas others focused on one aspect of spelling instruction. Our coding scheme had the main objective of accurately categorizing the interventions, but we cannot rule out that there are other factors that may have influenced the intervention effects. By focusing on learners with dyslexia or spelling deficits, the results of the current meta-analysis can be transferred only to this population. To assess the effectiveness of the intervention programs on other groups, evaluating the programs with other samples of participants beside those with dyslexia is mandatory.

Conclusion

The present meta-analysis shows that phonics, orthographic, and morphological interventions are effective components in treating spelling of learners with dyslexia and spelling deficits. Morphological interventions are researched across various orthographies, whereas pure phonics interventions were studied in English-speaking countries only. Findings from our analysis identify a gap in the literature on early implementation of spelling interventions. Given that classroom interventions do not seem sufficiently specific to an individual's needs, the implementation of small-group interventions in educational systems is essential. Of particular importance are interventions in individual settings, as shown by large effect sizes in our analysis. Computer-based and computer-assisted approaches can be a valuable alternative, or they can be used to complement individual or small-group interventions. It is striking that the support of adults with reading and spelling deficits is often neglected. Considering that dyslexia is a lifelong condition, future research should also focus on interventions addressing the question of how adults with dyslexia can be assisted. We further conclude that, to date, across orthographies an evidenced-based implementation of spelling interventions involves the application of phonics, the explicit instruction of graphotactic or orthographic-phonological spelling rules, and morphological instruction.

Acknowledgments

We thank four anonymous reviewers and the editor, Kathryn Wentzel, whose insightful comments helped us to improve our work. We also thank the authors of the studies included in this review for responding so willingly to repeated requests for information.

References

Abbott, S. P., & Berninger, V. W. (1999). It's never too late to remediate: Teaching word recognition to students with reading disabilities in grades 4–7. *Annals of Dyslexia*, *49*(1), 221–250. doi:10.1007/s11881-999-0025-x

*Arnbak, E., & Elbro, C. (2000). The effects of morphological awareness training on the reading and spelling skills of young dyslexics. *Scandinavian Journal of Educational Research*, *44*(3), 229–251. doi:10.1080/00313830050154485

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). The CELEX Lexical Database. Release 2 (CD-ROM): Linguistic Data Consortium, University of Pennsylvania.

Berninger, V. W., Lee, Y.-L., Abbott, R. D., & Breznitz, Z. (2013). Teaching children with dyslexia to spell in a reading-writers' workshop. *Annals of Dyslexia*, *63*(1), 1–24. doi:10.1007/s11881-011-0054-0

Berninger, V. W., & Swanson, H. L. (1994). Modifying Hayes and Flower's model of skilled writing to explain beginning and developing writing. *Advances in Cognition and Educational Practice*, *2*, 57–81.

Berninger, V. W., Winn, W. D., Stock, P., Abbott, R. D., Eschen, K., Lin, S.-J., … Nagy, W. (2008). Tier 3 specialized writing instruction for students with dyslexia. *Reading and Writing*, *21*(1–2), 95–129. doi:10.1007/s11145-007-9066-x

*Blachman, B. A., Schatschneider, C., Fletcher, J. M., Murray, M. S., Munger, K. A., & Vaughn, M. G. (2014). Intensive reading remediation in grade 2 or 3: Are there effects a decade later? *Journal of Educational Psychology*, *106*(1), 46. doi:10.1037/a0033663

Borgwaldt, S. R., Hellwig, F. M., & De Groot, A. (2004). Word-initial entropy in five languages: Letter to sound, and sound to letter. *Written Language and Literacy*, *7*(2), 165–184. doi:10.1075/wll.7.2.03bor

Borgwaldt, S. R., Hellwig, F. M., & De Groot, A. M. (2005). Onset entropy matters – Letter-to-phoneme mappings in seven languages. *Reading and Writing*, *18*(3), 211–229. doi:10.1007/s11145-005-3001-9

Bosman, A. M., & Van Orden, G. C. (1997). Why spelling is more difficult than reading. *Learning to Spell: Research, Theory, and Practice across Languages*, *10*, 173–194.

Bowers, J. S., & Bowers, P. N. (2017). Beyond phonics: The case for teaching children the logic of the English spelling system. *Educational Psychologist*, *52*(2), 124–141. doi:10.1080/00461520.2017.1288571

Bowers, J. S., & Bowers, P. N. (2018a). The importance of correctly characterising the English spelling system when devising and evaluating methods of reading instruction: Comment on Taylor, Davis, and Rastle (2017). *Quarterly Journal of Experimental Psychology*, *71*(7), 1497. 1747021818759477. doi:10.1177/1747021818759477

Bowers, J. S., & Bowers, P. N. (2018b). *There is no evidence to support the hypothesis that systematic phonics should precede morphological instruction: Response to Rastle and colleagues*. Retrieved from https://doi.org/10.31234/osf.io/mvdqr

Bus, A. G., & van Ijzendoorn, M. H. (1999). Phonological awareness and early reading: A meta-analysis of experimental training studies. *Journal of Educational Psychology*, *91*(3), 403–414. doi:10.1037/0022-0663.91.3.403

Calhoon, M. B. (2005). Effects of a Peer-Mediated phonological skill and reading comprehension program on reading skill acquisition for middle school students with reading disabilities. *Journal of Learning Disabilities*, *38*(5), 424–433. doi:10.1177/00222194050380050501

*Calhoon, M. B., Sandow, A., & Hunter, C. V. (2010). Reorganizing the instructional reading components: Could there be a better way to design remedial reading programs to maximize middle school students with reading disabilities' response to treatment? *Annals of Dyslexia*, *60*(1), 57–85. doi:10.1007/s11881-009-0033-x

Caravolas, M. (2004). Spelling development in alphabetic writing systems: A cross-linguistic perspective. *European Psychologist*, *9*(1), 3–14. doi:10.1027/1016-9040.9.1.3

Caravolas, M. (2005). Learning to spell in different languages: How orthographic variables might affect early literacy. In R. M. Joshi & P. G. Aaron (Eds.), *Handbook of orthography and literacy* (pp. 497–511). Mahwah, NJ: Erlbaum.

Carlisle, J. F. (2003). Morphology matters in learning to read: A commentary. *Reading Psychology*, *24*(3–4), 291–322. doi:10.1080/02702710390227369

Carlisle, J. F., McBride-Chang, C., Nagy, W., & Nunes, T. (2010). Effects of instruction in morphological awareness on literacy achievement: An integrative review. *Reading Research Quarterly*, *45*(4), 464–487. doi:10.1598/RRQ.45.4.5

Carter, E., Schönbrodt, F., Gervais, W. M., & Hilgard, J. (2018). *Correcting for bias in psychology: A comparison of meta-analytic methods*. Retrieved from https://osf.io/rf3ys/.

Castles, A., Rastle, K., & Nation, K. (2018). Ending the reading wars: Reading acquisition from novice to expert. *Psychological Science in the Public Interest*, *19*(1), 5–51. doi:10.1177/1529100618772271

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37–46. doi:10.1177/001316446002000104

*Corvacho del Toro, I. (2016). Zur qualitativen Rechtschreibfehleranalyse und einer schriftsystematischen lernförderlichen Behandlung der Rechtschreibstörung. *Zeitschrift Für Kinder- Und Jugendpsychiatrie Und Psychotherapie*, *44*(5), 397–408. doi:10.1024/1422-4917/a000457

*Darch, C., Eaves, R. C., Crowe, D. A., Simmons, K., & Conniff, A. (2006). Teaching spelling to students with learning disabilities: A comparison of rule-based strategies versus traditional instruction. *Journal of Direct Instruction*, *6*(1), 1–16. doi:10.2307/1511340

Del Re, A. (2013). compute.es: Compute effect sizes (R package version, 0.2-2).

Dymock, S., & Nicholson, T. (2017). To what extent does children's spelling improve as a result of learning words with the look, say, cover, write, check, fix strategy compared with phonological spelling strategies? *Australian Journal of Learning Difficulties*, *22*(2), 171–187. doi:10.1080/19404158.2017.1398766

Ecalle, J., Magnan, A., Bouchafa, H., & Gombert, J. E. (2009). Computer-based training with ortho-phonological units in dyslexic children: New investigations. *Dyslexia*, *15*(3), 218–238. doi:10.1002/dys.v15:310.1002/dys.373

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, *315*(7109), 629–634. doi:10.1136/bmj.315.7109.629

Ehri, L. C. (1987). Learning to read and spell words. *Journal of Reading Behavior*, *19*(1), 5–31. doi:10.1080/10862968709547585

Ehri, L. C. (2000). Learning to read and learning to spell: Two sides of a coin. *Topics in Language Disorders*, *20*(3), 19–36. doi:10.1097/00011363-200020030-00005

Ehri, L. C. (2005). Development of sight word reading: Phases and findings. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook*. Blackwell Publishing. doi:10.1002/9780470757642.ch8

Ehri, L. C. (2014). Orthographic mapping in the acquisition of sight word reading, spelling memory, and vocabulary learning. *Scientific Studies of Reading*, *18*(1), 5–21. doi:10.1080/10888438.2013.819356

Ehri, L. C., Nunes, S. R., Stahl, S. A., & Willows, D. M. (2001). Systematic phonics instruction helps students learn to read: Evidence from the National Reading Panel's meta-analysis. *Review of Educational Research*, *71*(3), 393–447. doi:10.3102/00346543071003393

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*(3), 250–287. doi:10.1598/RRQ.36.3.2

*Ehri, L. C., Satlow, E., & Gaskins, I. (2009). Grapho-phonemic enrichment strengthens keyword analogy instruction for struggling young readers. *Reading and Writing Quarterly*, *25*(2–3), 162–191. doi:10.1080/10573560802683549

*Faber, G. (2005). Systematische Rechtschreibförderung mit algorithmischen und selbstinstruktiven Lernhilfen: Die Leistungen von konzeptgemäß trainierten und untrainierten Schülern im Vergleich. *Heilpädagogische Forschung*, *42*(3), 129–144.

*Fasting, R. B., & Halaas Lyster, S.-A. (2005). The effects of computer technology in assisting the development of literacy in young struggling readers and spellers. *European Journal of Special Needs Education*, *20*(1), 21–40. doi:10.1080/0885625042000319061

*Ferraz, E., dos Santos Gonçalves, T., Freire, T., Mattar, T. D L. F., Lamônica, D. A. C., Maximino, L. P., & Crenitte, P. A. P. (2018). Effects of a phonological reading and writing remediation program in students with dyslexia: Intervention for specific learning disabilities. *Folia Phoniatrica et Logopaedica*, *70*(2), 59–73. doi:10.1159/000489091

Fisher, Z., & Tipton, E. (2015). Robumeta: Robust variance meta-regression (R package version 1.6).

Frith, U. (1980). Unexpected spelling problems. In U. Frith (Ed.), *Cognitive processes in spelling* (pp. 495–515). London, UK: Academic Press.

Frith, U. (1985). Beneath the surface of developmental dyslexia. In K. Patterson, J. Marshall, & M. Coltheart (Eds.), *Surface dyslexia: Neuropsychological and cognitive studies of phonological reading* (pp. 301–330). London, UK: Erlbaum.

Galuschka, K., Ise, E., Krick, K., & Schulte-Körne, G. (2014). Effectiveness of treatment approaches for children and adolescents with reading disabilities: A meta-analysis of randomized controlled trials. *PLoS One*, *9*(2), e89900. doi:10.1371/journal.pone.0089900

Gelman, A., & Carlin, J. (2014). Beyond power calculations assessing Type S (Sign) and Type M (Magnitude) errors. *Perspectives on Psychological Science*, *9*(6), 641–651. doi:10.1177/1745691614551642

Gelman, A., & Loken, E. (2014). The statistical crisis in science: Data-dependent analysis – A "garden of forking paths" – Explains why many statistically significant comparisons don't hold up. *American Scientist*, *102*(6), 460–465.

*Given, B. K., Wasserman, J. D., Chari, S. A., Beattie, K., & Eden, G. F. (2008). A randomized, controlled study of computer-based intervention in middle school struggling readers. *Brain and Language*, *106*(2), 83–97. doi:10.1016/j.bandl.2007.12.001

Goodwin, A., & Ahn, S. (2010). A meta-analysis of morphological interventions: Effects on literacy achievement of children with literacy difficulties. *Annals of Dyslexia*, *60*(2), 183–208. doi:10.1007/s11881-010-0041-x

Goodwin, A., & Ahn, S. (2013). A meta-analysis of morphological interventions in English: Effects on literacy outcomes for school-age children. *Scientific Studies of Reading*, *17*(4), 257–285. doi:10.1080/10888438.2012.689791

Graham, S., & Hebert, M. (2011). Writing to read: A meta-analysis of the impact of writing and writing instruction on reading. *Harvard Educational Review*, *81*(4), 710–744. doi:10.3102/0034654317746927

Graham, S., Berninger, V. W., Abbott, R. D., Abbott, S. P., & Whitaker, D. (1997). Role of mechanics in composing of elementary school students: A new methodological approach. *Journal of Educational Psychology*, *89*(1), 170. doi:10.1037/0022-0663.89.1.170

Graham, S., Harris, K. R., & Adkins, M. (2018). The impact of supplemental handwriting and spelling instruction with first grade students who do not acquire transcription skills as rapidly as peers: A randomized control trial. *Reading and Writing*, *31*(6), 1273–1294. doi:10.1007/s11145-018-9822-0

Graham, S., Harris, K. R., & Chorzempa, B. F. (2002). Contribution of spelling instruction to the spelling, writing, and reading of poor spellers. *Journal of Educational Psychology*, *94*(4), 669. doi:10.1037/0022-0663.94.4.669

Graham, S., Harris, K. R., & Hebert, M. (2011). It is more than just the message: Presentation effects in scoring writing. *Focus on Exceptional Children*, *44*(4), 1. doi:10.17161/fec.v44i4.6687

Graham, S., & Santangelo, T. (2014). Does spelling instruction make students better spellers, readers, and writers? A meta-analytic review. *Reading and Writing*, *27*(9), 1703–1743. doi:10.1007/s11145-014-9517-0

Gray, S. (2015). *The effects of morpho-phonemic and whole word instruction on the literacy skills of adult struggling readers*. New York, NY: City University of New York.

*Groth, K., Hasko, S., Bruder, J., Kunze, S., & Schulte-Körne, G. (2013). Interventionseffekte bei Lese-Rechtschreibstörung: Evaluation von zwei Förderkonzepten unter besonderer Betrachtung methodischer Aspekte. *Lernen Und Lernstörungen*, *2*(3), 161–175. doi:10.1024/2235-0977/a000038

*Gustafson, S., Ferreira, J., & Ronnberg, J. (2007). Phonological or orthographic training for children with phonological or orthographic decoding deficits. *Dyslexia*, *13*(3), 211–229. doi:10.1002/dys.339

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, *6*(2), 107–128. doi:10.2307/1164588

Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods*, *1*(1), 39–65. doi:10.1002/jrsm.5

Higgins, J. (2011). Cochrane handbook for systematic reviews of interventions. Version 5.1.0 [updated March 2011]. *The Cochrane Collaboration*. Retrieved from www.cochrane-handbook.org.

Higgins, J., & Altman, D. (2008). Assessing risk of bias in included studies. In J. P. T. Higgins & S. Green (Eds.), *Cochrane handbook for systematic reviews of interventions: Cochrane Book Series* (pp. 187–241). Hoboken, NJ: Wiley Online Library. doi:10.1002/9780470712184.ch8

Higgins, J. P. T., & Green, S. (2008). *Cochrane handbook for systematic reviews of interventions* (Vol. 5): Hoboken, NJ: Wiley Online Library.

Hilte, M., & Reitsma, P. (2011). Effects of explicit rules in learning to spell open-and closed-syllable words. *Learning and Instruction*, *21*(1), 34–45. doi:10.1016/j.learninstruc.2009.10.002

*Hughes, J. A., Phillips, G., & Reed, P. (2013). Brief exposure to a self-paced computer-based reading programme and how it impacts reading ability and behaviour problems. *PLoS One*, *8*(11), e77867. doi:10.1371/journal.pone.0077867

Ise, E., Engel, R. R., & Schulte-Körne, G. (2012). Effective treatment of dyslexia: A meta-analysis on the effectiveness of German-language intervention approaches. *Kindheit Und Entwicklung*, *21*(2), 122–136. doi:10.1026/0942-5403/a000077

*Ise, E., & Schulte-Körne, G. (2010). Spelling deficits in dyslexia: Evaluation of an orthographic spelling training. *Annals of Dyslexia*, *60*, 18–39. doi:10.1007/s11881-010-0035-8

Jackson, N. E., & Coltheart, M. (2013). *Routes to reading success and failure: Toward an integrated cognitive psychology of atypical reading*. London, UK: Psychology Press.

Jamshidifarsania, H., Garbayab, S., Limc, T., Blazevica, P., & Ritchie, J. M. (2019). Technology-based reading intervention programs for elementary grades: An analytical review. *Computers and Education*, *128*, 427–451. doi:10.1016/j.compedu.2018.10.003

*Jensen, J., Lindgren, M., Andersson, K., Ingvar, D. H., & Levander, S. (2000). Cognitive intervention in unemployed individuals with reading and writing disabilities. *Applied Neuropsychology*, *7*(4), 223–236. doi:10.1207/S15324826AN0704_4

Kast, M., Baschera, G.-M., Gross, M., Jäncke, L., & Meyer, M. (2011). Computer-based learning of spelling skills in children with and without dyslexia. *Annals of Dyslexia*, *61*(2), 177–200. doi:10.1007/s11881-011-0052-2

*Kast, M., Meyer, M., Vogeli, C., Gross, M., & Jancke, L. (2007). Computer-based multisensory learning in children with developmental dyslexia. *Restorative Neurology and Neuroscience*, *25*(3–4), 355–370.

Katz, L., & Frost, R. (1992). The reading process is different for different orthographies: The orthographic depth hypothesis. *Advances in Psychology*, *94*, 67–67. doi:10.1016/S0166-4115(08)62789-2

Kellogg, R. T. (2008). Training writing skills: A cognitive developmental perspective. *Journal of Writing Research*, *1*(1), 1–26. doi:10.17239/jowr-2008.01.01.1

Kirby, J. R., Deacon, S. H., Bowers, P. N., Izenberg, L., Wade-Woolley, L., & Parrila, R. (2012). Children's morphological awareness and reading ability. *Reading and Writing*, *25*(2), 389–410. doi:10.1007/s11145-010-9276-5

*Kirk, C., & Gillon, G. T. (2009). Integrated morphological awareness intervention as a tool for improving literacy. *Language, Speech, and*

Hearing Services in Schools, *40*(3), 341–351. doi:10.1044/0161-1461(2008/08-0009)
*Lange, A. A., Mulhern, G., & Wylie, J. (2009). Proofreading using an assistive software homophone tool. *Journal of Learning Disabilities*, *42*(4), 322–335. doi:10.1177/0022219408331035
Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2003). A definition of dyslexia. *Annals of Dyslexia*, *53*(1), 1–14. doi:10.1007/s11881-003-0001-9
Maughan, B., Messer, J., Collishaw, S., Pickles, A., Snowling, M., Yule, W., & Rutter, M. (2009). Persistence of literacy problems: Spelling in adolescence and at mid-life. *Journal of Child Psychology and Psychiatry*, *50*(8), 893–901. doi:10.1111/j.1469-7610.2009.02079.x
McArthur, G., Eve, P. M., Jones, K., Banales, E., Kohnen, S., Anandakumar, T., … Castles, A. (2012). Phonics training for English-speaking poor readers. *Cochrane Database of Systematic Reviews*, *12*. doi:10.1002/14651858.CD009115.pub2
McCutchen, D. (1996). A capacity theory of writing: Working memory in composition. *Educational Psychology Review*, *8*(3), 299–325. doi:10.1007/BF01464076
McCutchen, D. (2000). Knowledge, processing, and working memory: Implications for a theory of writing. *Educational Psychologist*, *35*(1), 13–23. doi:10.1207/S15326985EP3501_3
McCutchen, D., & Logan, B. (2011). Inside incidental word learning: Children's strategic use of morphological information to infer word meanings. *Reading Research Quarterly*, *46*(4), 334–349. doi:10.1002/RRQ.003
McCutchen, D., Stull, S., Herrera, B. L., Lotas, S., & Evans, S. (2014). Putting words to work: Effects of morphological instruction on children's writing. *Journal of Learning Disabilities*, *47*(1), 86–97. doi:10.1177/0022219413509969
McShane, B. B., Böckenholt, U., & Hansen, K. T. (2016). Adjusting for publication bias in meta-analysis: An evaluation of selection methods and some cautionary notes. *Perspectives on Psychological Science*, *11*(5), 730–749. doi:10.1177/1745691616662243
Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G, & The-PRISMA-Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, *6*(7), e1000097. doi:10.1371/journal.pmed.1000097
Moll, K., Kunze, S., Neuhoff, N., Bruder, J., & Schulte-Körne, G. (2014). Specific learning disorder: Prevalence and gender differences. *PLoS One*, *9*(7), e103537. doi:10.1371/journal.pone.0103537
Moll, K., & Landerl, K. (2009). Double dissociation between reading and spelling deficits. *Scientific Studies of Reading*, *13*(5), 359–382. doi:10.1080/10888430903162878
Moll, K., Ramus, F., Bartling, J., Bruder, J., Kunze, S., Neuhoff, N., … Landerl, K. (2014). Cognitive mechanisms underlying reading and spelling development in five European orthographies. *Learning and Instruction*, *29*, 65–77. doi:10.1016/j.learninstruc.2013.09.003
Morris, S. (2008). Estimating effect sizes from Pretest-Posttest-Control group designs. *Organizational Research Methods*, *11*(2), 364–386. doi:10.1177/1094428106291059
Nagy, W., Berninger, V. W., Abbott, R. D., Vaughan, K., & Vermeulen, K. (2003). Relationship of morphology and other language skills in at-risk second-grade readers and at-risk fourth grade writers. *Journal of Educational Psychology*, *95*(4), 730–742. doi:10.1037/0022-0663.95.4.730
*Oakland, T., Black, J. L., Stanford, G., Nussbaum, N. L., & Balise, R. R. (1998). An evaluation of the dyslexia training program: A multisensory method for promoting reading in students with reading disabilities. *Journal of Learning Disabilities*, *31*(2), 140–147. doi:10.1177/002221949803100204
*O'Shaughnessy, T. E., & Swanson, H. L. (2000). A comparison of two reading interventions for children with reading disabilities. *Journal of Learning Disabilities*, 3, 257–277. doi:10.1177/002221940003300304
Ouellette, G., Martin-Chang, S., & Rossi, M. (2017). Learning from our mistakes: Improvements in spelling lead to gains in reading speed. *Scientific Studies of Reading*, *21*(4), 350–357. doi:10.1080/10888438.2017.1306064
Perfetti, C., & Hart, L. (2002). The lexical quality hypothesis. In L. Verhoeven, C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy* (pp. 189–213). Amsterdam, The Netherlands: John Benjamin.
Perfetti, C. A., & Hogaboam, T. (1975). Relationship between single word decoding and reading comprehension skill. *Journal of Educational Psychology*, *67*(4), 461–469. doi:10.1037/h0077013
Pustejovsky, J. E. (2016). clubSandwich: Cluster-Robust (Sandwich) variance estimators with small-sample corrections (*R* Package Version 0.2. 1.9000) [Computer Software]. Retrieved from https://cran.rproject.org/package=clubSandwich
R Core Team. (2014). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.r-project.org/.
*Rashotte, C. A., MacPhee, K., & Torgesen, J. K. (2001). The effectiveness of a group reading instruction program with poor readers in multiple grades. *Learning Disability Quarterly*, *24*(2), 119–134. doi:10.2307/1511068
Rastle, K. (2018). The place of morphology in learning to read in English. *Cortex*, *116*, 45–54. doi:10.1016/j.cortex.2018.02.008
Rastle, K., & Taylor, J. (2018). Print-sound regularities are more important than print-meaning regularities in the initial stages of learning to read: Response to Bowers & Bowers (2018). *Quarterly Journal of Experimental Psychology*, *71*(7), 1501–1505. doi:10.1177/1747021818775053
*Reuter-Liehr, C. (1993). Behandlung der Lese-Rechtschreibschwäche nach der Grundschulzeit: Anwendung und Überprüfung eines Konzeptes. *Zeitschrift Für Kinder-Und Jugendpsychiatrie*, *21*(3), 135–147.
Saddler, B., & Graham, S. (2005). The effects of peer-assisted sentence-combining instruction on the writing performance of more and less skilled young writers. *Journal of Educational Psychology*, *97*(1), 43–54. doi:10.1037/0022-0663.97.1.43
Scammacca, N., Roberts, G., Vaughn, S., & Stuebing, K. (2015). A meta-analysis of interventions for struggling readers in grades 4–12: 1980–2011. *Journal of Learning Disabilities*, *48*(4), 369–390. doi:10.1177/0022219413504995
Schmalz, X. (Producer). (2018). *Python Script for calculating onset entropy*. Retrieved from https://osf.io/8r6zf/
Schmalz, X., Marinus, E., Coltheart, M., & Castles, A. (2015). Getting to the bottom of orthographic depth. *Psychonomic Bulletin and Review*, *22*(6), 1614–1629. doi:10.3758/s13423-015-0835-2
Schmidt, F. L. (1992). What do data really mean? Research findings, meta-analysis, and cumulative knowledge in psychology. *American Psychologist*, *47*(10), 1173. doi:10.1037/0003-066X.47.10.1173
Schulte-Körne, G., Deimel, W., Hülsmann, J., Seidler, T., & Remschmidt, H. (2001). Das Marburger Rechtschreib-Training–Ergebnisse einer Kurzzeit-Intervention. *Zeitschrift Für Kinder- Und Jugendpsychiatrie Und Psychotherapie*, *29*(1), 7–15. doi:10.1024//1422-4917.29.1.7
Serrano, F., Genard, N., Sucena, A., Defior, S., Alegria, J., Mousty, P., … Seymour, P. H. K. (2011). Variations in reading and spelling acquisition in Portuguese, French and Spanish: A cross-linguistic comparison. *Journal of Portuguese Linguistics*, *10*(1), 183–204. doi:10.5334/jpl.106
Seymour, P. H., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, *94*(2), 143–174. doi:10.1348/000712603321661859
Share, D. L. (1995). Phonological recoding and self-teaching: *sine qua non* of reading acquisition. *Cognition*, *55*(2), 151–218. doi:10.1016/0010-0277(94)00645-2
SIGN, Harbour, R. T., & Forsyth, L. (2008). *SIGN 50: A guideline developer's handbook*. Scottish Intercollegiate Guidelines Network.
*Spencer, F., Snart, F., & Das, J. (1989). A process-based approach to the remediation of spelling in students with reading disabilities. *Alberta Journal of Educational Research*, *35*(4), 269–282.
Sprenger-Charolles, L., Siegel, L. S., & Bonnet, P. (1998). Reading and spelling acquisition in French: The role of phonological mediation and orthographic factors. *Journal of Experimental Child Psychology*, *68*(2), 134–165. doi:10.1006/jecp.1997.2422

Squires, K. E., & Wolter, J. A. (2016). The effects of orthographic pattern intervention on spelling performance of students with reading disabilities: A best evidence synthesis. *Remedial and Special Education*, *37*(6), 357–369. doi:10.1177/0741932516631115

Strehlow, U., Haffner, J., Busch, G., Pfüller, U., Rellum, T., & Zerahn-Hartung, C. (1999). An Schwächen üben oder durch Stärken ausgleichen? Vergleich zweier Strategien in der Förderung von Kindern mit einer umschriebenen Lese-Rechtschreib-Schwäche. *Zeitschrift Für Kinder- Und Jugendpsychiatrie Und Psychotherapie*, *27*(2), 103–113. doi:10.1024//1422-4917.27.2.103

Taylor, J., Davis, M. H., & Rastle, K. (2017). Comparing and validating methods of reading instruction using behavioural and neural findings in an artificial orthography. *Journal of Experimental Psychology: General*, *146*(6), 826–858. doi:10.1037/xge0000301

*Thaler, V., Landerl, K., & Reitsma, P. (2008). An evaluation of spelling pronunciations as a means of improving spelling of orthographic markers. *European Journal of Psychology of Education*, *23*(1), 3–23. doi:10.1007/BF03173137

Thalmann, S. (2014). Adaption criteria for the personalized delivery of learning materials: A multi-stage empirical investigation. *Australasian Journal of Educational Technology*, *30*(1), 45–60. doi:10.14742/ajet.235

*Tijms, J. (2011). Effectiveness of computer-based treatment for dyslexia in a clinical care settin: Outcomes and moderators. *Educational Psychology*, *31*(7), 873–896. doi:10.1080/01443410.2011.621403

Tipton, E. (2015). Small sample adjustments for robust variance estimation with meta-regression. *Psychological Methods*, *20*(3), 375–393. doi:10.1037/met0000011

Treiman, R. (2018). Teaching and learning spelling. *Child Development Perspectives*, *12*(4), 235–239. doi:10.1111/cdep12292

Treiman, R., & Bourassa, D. C. (2000). The development of spelling skill. *Topics in Language Disorders*, *20*(3), 1–18. doi:10.1097/00011363-200020030-00004

Treiman, R., & Kessler, B. (2014). *How children learn to spell words*: New York, NY: Oxford University Press.

Truong, H. M. (2016). Integrating learning styles and adaptive e-learning system: Current developments, problems and opportunities. *Computers in Human Behavior*, *55*, 1185–1193. doi:10.1016/j.chb.2015.02.014

*Vadasy, P., Sanders, E., & Peyton, J. (2005). Relative effectiveness of reading practice or word-level instruction in supplemental tutoring: How text matters. *Journal of Learning Disabilities*, *38*(4), 364–380. doi:10.1177/00222194050380041401

Vevea, J. L., & Woods, C. M. (2005). Publication bias in research synthesis: Sensitivity analysis using a priori weight functions. *Psychological Methods*, *10*(4), 428–443. doi:10.1037/1082-989X.10.4.428

Viechtbauer, W. (2010). Metafor: Meta-analysis package for R (R Package Version, 2010, 1–0).

Volkmer, S., Schulte-Körne, G., & Galuschka, K. (2019). Die Rolle der morphologischen Bewusstheit bei Lese-und Rechtschreibleistungen. *Zeitschrift Für Kinder-Und Jugendpsychiatrie Und Psychotherapie*, *47*(4), 334–344. doi:10.1024/1422-4917/a000652

*Walter, J., Schliebe, L., & Barzen, S. (2007). Evaluation eines morphemorientiert-strategischen Rechtschreibtrainings in schulischen Fördergruppen mit Grundschülern der 3. *Heilpädagogische Forschung*, *13*(14), 3–153.

Wanzek, J., & Vaughn, S. (2008). Response to varying amounts of time in reading intervention for students with low response to intervention. *Journal of Learning Disabilities*, *41*(2), 126–142. doi:10.1177/0022219407313426

Wanzek, J., Vaughn, S., Scammacca, N., Gatlin, B., Walker, M. A., & Capin, P. (2016). Meta-analyses of the effects of tier 2 type reading interventions in grades K-3. *Educational Psychology Review*, *28*(3), 551–576. doi:10.1007/s10648-015-9321-7

Wanzek, J., Vaughn, S., Wexler, J., Swanson, E. A., Edmonds, M., & Kim, A.-H. (2006). A synthesis of spelling and reading interventions and their effects on the spelling outcomes of students with LD. *Journal of Learning Disabilities*, *39*(6), 528–543. doi:10.1177/00222194060390060501

*Weber, J.-M., Marx, P., & Schneider, W. (2002). Profitieren Legastheniker und allgemein lese-rechtschreibschwache Kinder in unterschiedlichem Ausmaß von einem Rechtschreibtraining? *Psychologie in Erziehung Und Unterricht*, *49*(1), 56–70.

Wechsler, D. (1992). *Wechsler individual achievement test: Manual*. San Antonio: TX: Psychological Corportation.

Weiser, B., & Mathes, P. (2011). Using encoding instruction to improve the reading and spelling performances of elementary students at risk for literacy difficulties: A best-evidence synthesis. *Review of Educational Research*, *81*(2), 170–200. doi:10.3102/0034654310396719

Wilkinson, G. (1993). *The wide range achievement test* (3rd ed.). Wilmington, DE: Wide Range.

Williams, C., & Beam, S. (2019). Technology and writing: Review of research. *Computers and Education*, *128*, 227–242. doi:10.1016/j.compedu.2018.09.024

Williams, K. J., Walker, M. A., Vaughn, S., & Wanzek, J. (2017). A synthesis of reading and spelling interventions and their effects on spelling outcomes for students with learning disabilities. *Journal of Learning Disabilities*, *50*(3), 286–297. doi:10.1177/0022219415619753

Supplement 1. Study details.

Study	EG approach	EG (original study)	CG	Short description of the intervention	Content of the intervention and amount of intervention components (if specified)	Focus reading/ spelling	Total amount	Setting	Conductor	Inclusion criterion	Age/ class	NT
Abbott & Berninger, 1999	MoI	Structural analysis training	PhI	Both, EG and CG, received phonics intervention for PGC. The EG also received structural analysis instruction focusing on syllables and morphemes.	Each lesson (60 minutes) included phonological exercises (6 minutes) as well as letter-sound tasks (5 minutes) on subword level and phonological decoding exercises (10 minutes) on word level. The lessons mainly focused on Henrys word program (15 minutes) which contained instruction in syllable types and morpheme patterns including checking for affixes and roots and dividing words into syllables. During spelling activities children were explicitly encouraged to make use of syllable and morpheme patterns instead of relying on PGC. Each lesson further included reading tasks (24 minutes) for oral reading (error correction) and reading comprehension (comprehension monitoring and rereading).	Reading and spelling	16 hours	Individual setting	Psychologists, students, teachers	> 1 SD discrepancy between WRMT-R subtests and WISC-III verbal IQ	4th - 7th grade	USA
Arnbak & Elbro, 2000	MoI	Morphological awareness training program	OT	The intervention focused on morphological components, especially on semantics of morphemes, affixes, inflections and root morphemes.	The training comprised three modules with different focuses: 1. Part one: Semantically transparent type of morpheme, root morphemes 2. Part two: Affixes 3. Part three: Inflections	Reading and spelling	9 hours	Group setting	Teachers	Reading at least 2 years below expected reading level	4th - 5th grade	Denmark
Berninger et al., 2013	MI	Orthographic strategy training	PhI	Both groups received phonics intervention. Different instructional	The treatment included the following different steps:	Reading and spelling	40 hours	Group setting	Teacher assistants	Reading, spelling or below the mean and	4th - 9th grade	USA

				components were added sequentially.	1. Grapheme and phoneme correspondences in reading direction (6 lessons) 2. Phoneme-Grapheme correspondences for spelling and oral reading (8 lessons) 3. Orthographic strategies → The children were taught to create a "silent orthography" in their minds					at least 1 SD below the child's verbal IQ		
Berninger et al., 2008	MI	Mark Twain Writers'	MoI	The intervention focused on PGC for oral reading and spelling as well as on orthographic spelling strategies for written spelling.	The treatment included the following lessons: 1. Alphabetic principle strategies: Associating phonemes and graphemes as well as using a visual matrix with pictured key words containing each phoneme and the associated grapheme 2. Orthographic strategies: Visualizing words in their mind's eye as well as spelling words forwards and backwards and other orthographic activities (e.g., anagrams and visual letter search) 3. Morphological strategies: Memorizing wordparts by writing them down on worksheets and other morphological activities (e.g., word-building and word-dissecting) 4. Probes: writing taught words from dictation) as well as composition instruction (strategies for planning,	Reading and Spelling	28 hours	Group setting	Teacher	At least 1 SD below verbal IQ and below average performance in reading or spelling	4th – 9th grade	USA

					drafting, reviewing and revising).							
Blachman et al., 2014	PhI	Reading intervention	RSP	The treatment focused on phonologic and orthographic connections in words as well as accurate decoding, word recognition, fluency, spelling and text-based reading.	Each lesson (50 minutes) followed a five step plan 1. Sound-Symbol associations 2. Phoneme analysis and blending 3. Reading fluency by quickly reading words 4. Oral reading practice (texts, books) 5. Word dictation	Reading and spelling	133.3 hours	Individual setting	Teachers	Below the 25th percentile on one WRMT-R subtest	2nd - 3rd grade	USA
Calhoon et al., 2010	OI	RAMP-UP	OT	The intervention combined components of linguistic skills, spelling, fluency and comprehension.	Four different intervention components: 1. Linguistic skill component (phonetics, phonology, English orthography, e.g., historical layers, orthographic conventions) → 37 hours, 38% 2. Reading Fluency component (repeated reading technique) → 8 hours, 8% 3. Spelling component (explicitly teaching linguistic/orthographic skills, e.g., double vowels, silent e etc.) → 13 hours, 14% 4. Reading comprehension component (echo reading, paragraph summarizing, prediction relay) → 39 hours, 40%	Reading and spelling	97 hours	Classroom setting	Teachers and assistants	Achievement 3.5 years below grade level in reading	6th - 8th grade	USA
Corvacho del Toro, 2016	MoI	Experimental Group	UCG	The intervention concentrated mainly on the acquisition of morphological skills by utilizing an individualized proceeding.	Explicit instruction in morphological and orthographic knowledge based on a qualitative analysis of spelling mistakes. Morphological knowledge plays an important role throughout the intervention (e.g., for	Spelling	15 hours	Individual setting	Trained students	Pre-diagnosed isolated spelling disorder	6th- 8th grade	Germany

					long and short vowels, consonant doubling, grammar etc.)							
Darch et al., 2006	OI	Spelling Mastery, Level D	OT	The program employed systematic rule-based strategies for teaching spelling words.	Explicit instruction of spelling rules and orthographic knowledge according to word types. Different word types were introduced per week: 1. Week 1: Two letter combinations within 24 phonetically regular words 2. Week 2: Morphographs (adding prefixes or suffixes to base words) 3. Week 3: Formed words by combining word parts and applying different spelling rules (e.g., doubling consonants) 4. Week 4: Irregular words including common letters or letter combinations not represented by the most common sounds	Spelling	10 hours	Group setting	Graduate students	At least 1 SD discrepancy between IQ and achievement in spelling	2^{nd} - 4^{th} grade	USA
Ehri et al., 2009	PhI	KEY-PLUS	OT	The intervention focused on word decoding ability by learning keywords. Alphabetic spelling regularities in words were taught and utilizing spelling activities were discussed.	The KEY PLUS program includes 3 major components: 1. Analyzing grapheme-phoneme constituents of keywords 2. Discovering of spelling regularities in words and discussing these 3. Spelling activities to reinforce phonemic segmentation skills	Reading and spelling	Sessions throughout 4 years	Classroom setting	Teachers	Pre-diagnosed struggling readers	1^{st} - 3^{rd} grade	USA
Faber, 2005	OI	Remedial spelling training	UCG	Children learned metacognitive strategies to apply spelling rules and control their own spellings.	Explicit instruction of spelling rules and orthographic knowledge, for example consonant doubling. Capitalization principles and phonemic writing	Spelling	34 hours	Individual setting	Author	Below 1.5 SD below age and class norm in spelling	3^{rd} – 8^{th} grade	Germany

					were addressed only in the case of individual needs.							
Fasting & Lyster, 2005	SS	MultiFunk	RSP	The computer-reading software aimed to support reading and writing by making it an easier task based on individual adaption of text appearance.	To the software program, texts can be imported and subsequently adapt to the individual profile of the child, e.g., with different modes of word highlighting and reading aloud difficult words. A text editor can be used to summarize a text or to write new texts.	Reading and spelling	16.3 hours	Classroom setting, on PC	Teachers	Identified struggling readers and spellers, reading skills $z = -1.05$ (SD=0.722) and spelling skills $z=-1.5$ (SD=0.74)	5^{th} – 7^{th} grade	Norway
Ferraz et al., 2018	PhI	Phonological Reading and Writing Remediation Program	WCG	The program focused on auditory and visual discrimination of sounds, letters, syllables and words. Additionally, silent and oral reading activities as well as dictation of sentences and texts were included.	The curriculum included: 1. Session 1-8: Phonological training: Auditory discrimination and manipulation of phonemes and syllables and identification of corresponding letters 2. Session 9-16: Phonological and reading training: Auditory discrimination and manipulation of phonemes, syllables and words, identification of corresponding letters or pictures, rapid letter naming, oral and silent text reading 3. Session 17-24: Phonological reading and writing training: Auditory discrimination and manipulation of phonemes,	Reading and spelling	18 hours	Probably group setting (not specified)	Not specified	Reading and writing disability, DSM-IV criteria	8 - 14 years	Brazil

					syllables and words, grapheme discrimination, oral and silent text reading, dictation of syllables, real- and non-words, sentences, writing stories						
Given et al., 2008	OI	Success Maker	UCG	The program focused on vocabulary skills, contextual reading, spelling, and writing skill development.	The curriculum of SuccessMaker included: 1. Phonological awareness: Blending and segmenting phonemes 2. Phonics: Basic principles of phonics, connecting sounds and letters 3. Fluency: Speed drills of letters, words and phrases, retelling component, timed reading of familiar text 4. Vocabulary: Exposure to and mastery of vocabulary words 5. Comprehension: Reading strategies and assessing text understanding 6. Spelling: Reinforce students understanding of spelling rules and conventions, e.g., vowel combinations etc. 7. Grammar: Sentence structure and punctuation	Reading and spelling	88 hours	Classroom setting, on PC	Teachers	Identified struggling readers and spellers, spelling skills were: *z* = - 1.125, PR=13 (WRAT3, spelling)	Middle school	USA
Gray, 2016	MoI	Morphophonemic treatment	MI	The treatment focused on analysing word origins, parsed morpheme and syllable structures and extracted base words in morphologically related words. It included word reading, definitions, synonyms, sentence context, word	During this intervention each word was learned according to the following teaching elements: 1. Define the word 2. Synonym: Reading aloud and write synonyms 3. Sentence Context: Reading the word in a sentence	Reading and spelling	8 hours	Individual setting	Tutors (trained research assistants)	Prediagnosed, recent involvement in a GED program.	18 – 31 years	USA

				origin, word sum, suffix role, morphological relatives, syllabification and passage reading.	4. Word Origin: Reading aloud and write the word origin and base 5. Word Sum: Reading aloud and write the word sum as well as circle its base 6. Suffix Role: Observing changes in word class, spelling or pronunciation 7. Morphological Relatives: Finding, reading and writing morphologically related words with the same root or base, circle its base 8. Syllabification: Clapping and writing words in syllables, underlining syllables with primary stress 9. Passage Reading: Reading aloud passage with word embedded							
Groth et al., 2013	OI	Marburg Spelling Training	WCG	This intervention focused on the acquisition and application of orthographic regularities.	Explicit instruction of spelling rules and orthographic knowledge, for example differentiation of long and short vowels or capitalization.	Spelling	28.5 hours	Individual setting	Trained scientific assistants (pedagogy and linguistic sciences)	At least 1 SD below IQ, below average performance in reading and spelling	3rd grade	Germany
Gustafson et al., 2007	PA	COMPHOT (Phonological training program)	RSP	This treatment focused on phonology including rhyme, addition, position and segmentation. The focused units were phonemes, word segments and words.	The phonological training program included the following 4 different sections: 1. Rhyme (4 exercises) 2. Addition (5 exercises) 3. Position (8 exercises) 4. Segmentation (3 exercises)	Reading	21.6 hours	Classroom setting but individually on	Teachers (special instruction)	Reading disabled children	2nd – 3rd grade	Sweden

	MoI	DOT (Orthographic training program)		The intervention focused on orthography including word reading, text reading, word parts and building words. The focused units were words, texts, morphemes and letters/ words.	The orthographic training program included 4 sections: 1. Word reading (4 exercises) 2. Text reading (2 exercises) 3. Word parts (2 exercises) 4. Building words (3 exercises) Language units were written letters, morphemes and words.	Reading and spelling	17.3 hours	computer				
Hughes et al. 2013	PhI	"Self-learn read and spell" computer program	RSP	The intervention mainly focused on the improvement of phonetic discrimination. It included exercises repeating the alphabet and teaching basic as well as more complex phonic words.	23 different levels of difficulty: 1. Level 1-13: Basic sounds and sounds with special markers (e.g., double sounds) 2. Level 14-18: Analyzing and adding word endings 3. Level 19-23 Silent letters, homophones, capital letters and plurals	Reading and spelling	12 hours	Group setting	Teachers and learning support assistant	Identified as struggling readers	11 – 12 years	UK
Ise & Schulte-Körne, 2010	OI	Marburg Spelling Training	WCG	This intervention focused on the acquisition and application of orthographic regularities.	Explicit instruction of spelling rules and orthographic knowledge, for example differentiation of long and short vowels or capitalization.	Spelling	12 hours	Individual setting	Author, students in pedagogy, psychology, psycholinguistics, German literature	At least 1 SD below class norm in spelling	5th – 6th grade	Germany
Jensen et al., 2000	AW	Educational program	UCG	The first phase of the intervention aimed to encourage motivation and cooperation, enhancing the participants' self-confidence. Writing activities included process writing with increase of complexity of texts and additionally focusing on	The training included the following activities: 1. Writing: Process writing which aimed to stimulate and increase writing by increasing the complexity of texts; process writing was followed by individual help with punctuation marks, sentence structure, grammar, spelling;	Reading and spelling	20 weeks, (different amounts of time – not further described – for each	Group setting	Teachers (trained to act as coaches)	Reading and writing disability, DSM-IV criteria	Mean age 34.6 years	Sweden

				grammar, and spelling. Reading activities aimed to improve reading habits.	daily diaries or logs were written 2. Computer usage as an assisting writing tool for spell check and synonym features as well as computerized exercises 3. Reading: Regular reading of books and newspapers In addition, the training included activities to enhance motivation as well as mathematics. Participants further were provided with information about their dysfunction, history of language and economics.		part of intervention)					
Kast et al., 2007	AVI	Dybuster	WCG	A multimodal learning software was used to improve spelling by addressing different sensory channels.	Isolated letters as well as words were learned with the help of colours, shapes and melody. The program consists of 3 different games: 1. Colour game: Learning associations between letters and colours 2. Graph game: Graphically segmenting a word into syllables and letters 3. Learning game: Combining graphs, colours, shapes and sounds before entering the word with a keyboard	Spelling	14 hours	Group setting	Not specified	At least 2 SD below expected age norm in reading	9 - 11 years	Germany
Kirk & Gillon, 2009	MoI	Integrated morphological awareness intervention	WCG	This intervention focused on the identification of vowel length and further provided an insight into morphology.	Each session included the following two types of activities: 1. Word sorting in terms of morphological, phonological, orthographic, syntactic or semantic patterns: Analysing the words due to shared patterns/rules etc.	Reading and spelling	19.5 hours	Individual and group setting (4 children)	Author or trained students	At least 1 SD below expected age norm in spelling	8 - 11 years	New Zealand

					2. Prompted spelling: Receiving a series of prompts before spelling a word						
Lange at al., 2009	SS	Homophone proofreading	OT	The intervention aimed to improve automatic word recognition, fluent reading and reading comprehension using an assistive homophone tool. Assistive homophone tool included identifying correctly or incorrectly spelled homophones.	With the software, homophones in a text were highlighted and presented with options as well as definitions. The tool included identifying correctly or incorrectly spelled homophones. It mainly aimed to improve fluent word recognition and reading comprehension.	Reading	2.3 – 5.8 hours	Not further specified)	Authors	Low reading ability, one year below expected age norm in reading	13 – 15 years	UK
Oakland et al., 1998	PhI	Dyslexia Treatment Program	RSP	This treatment approach emphasized phonetic analysis of written language. It was presented in a structured, multisensory sequence of alphabet, reading, spelling, cursive handwriting, listening, verbal expression and review activities.	The one-hour sessions started with basic exercises (e.g., letter recognition) and extended sequentially to sophisticated levels of linguistic knowledge (e.g., phonological decoding and syllabicating of polysyllabic words). Students are taught vocabulary, code marks indicating speech sounds, symbols plus abbreviations related to word decoding and formulas for syllable division and spelling. When decoding skills reached a minimal level of accuracy and automaticity, reading comprehension activities started.	Reading and spelling	2 years, 10 months a year, 5 times a week, (not further specified)	Group setting	Teachers, dyslexia therapists	Diagnosis of dyslexia (1.5 SD discrepancy between IQ and reading)	11 years	USA
O'Shaughnessy & Swanson, 2000	PhI	Phonological Awareness Training (PAT)	PI	The goal of the intervention was to help children acquiring awareness of speech sounds and of PGC. Exercises were sound blending, sound segmenting and PGC.	PAT is divided into four sets of activities: 1. Rhyming 2. Sound blending 3. Sound segmenting 4. Reading and spelling (transferring sound knowledge to	Reading	9 hours	Group setting	Teachers trained and observed by authors	Scores below the 25th percentile on reading tests	2nd grade	USA

					letters → phoneme-grapheme knowledge)							
Rashotte et al., 2001	PhI	Spell Read P.A.T. Program	RSP	The intervention aimed to improve of phonological and auditory skills. It integrated elements of phonemic awareness, phonics, reading and writing. Phonemic activities, for example involved blending, segmentation, dissecting. Reading and writing activities included shared reading followed by free writing activities.	Each session consisted of: 1. Phonemic activities: Analyzing, blending and manipulating phonemes and syllables (30 minutes). 2. Share reading: Reading out loud and discussing what was read (15 minutes). 3. Free writing: Writing about what was read before (5-6 minutes).	Reading and spelling	35 hours.	Group setting	Teachers and one supervisor with experience in the Spell Read program	Selection for the program was based on below-average reading skills as measured by subtest of the WRMT	1st - 6th grade	Canada
Reuter-Liehr, 1993	MoI	Phonemic reading and spelling training	UCG	This approach was based on practising phonemic writing, syllable analyses and blending, followed by morpheme segmentation and knowledge about regular deviations.	Not further specified.	Spelling	90 hours	Group setting	Author and assistants	Below the 25th percentile in spelling	5th grade	Germany
Schulte-Körne et al., 2001	OI	Marburg Spelling Training	PA	The intervention focused on the acquisition and application of orthographic regularities.	Explicit instruction of spelling rules and orthographic knowledge, for example differentiation of long and short vowels or capitalization.	Spelling	25 hours	Group setting	Medical students and author (supervisor)	Children with spelling disorder (criteria of regression, at least 1 SD)	2nd - 4th grade	Germany
Spencer et al., 1989	MI	Experimental Group	RSP	During the spelling intervention, the children had to memorize and reproduce words that	Each session students received 5 minutes of training on CAP (Coding, Attention, Planning training) task and then 10 minutes on the associated	Spelling	20 hours	Classroom	Teacher and assistant	Academic achievement in reading and / or	8 – 12 years	Canada

				had been learned and shown before.	spelling task. 5 spelling tasks were created modelled after each of the CAP tasks (e.g., copying words, correcting misspelled words or choose correct letters etc.).					spelling at least two years below the expected level on standardized measures		
Strehlow, 1999	MI	Simultaneous processing method	MoI	Learning the correct spelling in high frequent regular and irregular German words by heart.	Words were practiced in context with the help of stories; training words were highlighted in the text. Children were first made familiar with the stories and then tasks were carried out analyzing the training words: 1. Comparing/Rhyming 2. Remembering the spelling 3. Remembering the word meaning 4. Recognizing the words	Spelling	22.5 hours	Individual setting	Trainer (not specified)	Dyslexic children (criteria of discrepancy, at least 1 SD)	3rd grade	Germany
Thaler et al., 2008	MI	Spelling pronunciation intervention	OT	The focus of the intervention was orthographic marking of long vowel phonemes using special pronunciation of orthographic marker aiming to remind correct spellings.	Spelling of 45 words were learned divided in 5 subsets. Each word was presented visually and auditorily, each letter of a word was pronounced, and critical graphemes highlighted to remember the spellings of the word.	Spelling	15 days, daily, different amounts of sessions and duration (differences in subsets)	Probably individual setting (not specified)	Not specified	Below the 15th percentile)	3rd grade	Austria
Tijms, 2011	MoI	LEXY	WCG	The intervention focused on knowledge about Dutch phonemes. The intervention included letter-sound mappings and the usage of these mappings and	1. 5 stages with different topics: 2. Phonemic knowledge/PGC (12 sessions) 3. Rules for monosyllabic words: Mapping the phonetic structure of a word onto the	Reading and spelling	43 hours (average)	Individual setting	Speech therapists or psychologists	At least 1.5 SD below expected age norm in reading or 1 SD	9 - 12 years, 3rd - 6th grade	Netherlands

				morphological knowledge in reading and spelling.	corresponding orthographic word (6 sessions) 4. Syllabification/Rules for multisyllabic words: Dividing words into syllables and identifying syllables, introduction to rules and using these in multisyllabic words (10 sessions). 5. Morphological knowledge/Verb structure: Morphological knowledge, which is essential for reading and spelling and can overrule phonological structures. Furthermore, knowledge of written form of verbs was taught (10 sessions) 6. Loan words: Bound morphemes in loan words from Greek, Latin, English and French (4 sessions)					below expected age norm in reading and 1.5 SD below expected age norm in spelling		
Vadasy et al., 2005	PhI	Sound Partners	RSP	The intervention focused on the improvement of phonological and phonics skills. Most tasks included teaching PGC, phonological decoding, phoneme segmentation.	Treatment consists of the following components: 1. Word study: Letter sounds, decoding, segmenting, spelling, sight words (15-20 minutes per session) 2. Added word study: Letter-sounds, words with silent-e, letter consonant and vowel pairs, repeated word list reading	Reading	64 hours	Individual setting	Paraprofessional tutors, trained by the research staff	Students who scored at or below the 25th percentile on the Reading subtest of the WRAT-R	1st grade	USA
Walter et al., 2007	MoI	REMO-2	OT	The intervention aimed to improve spelling using orthographic-morphological instruction. The focus of the inter-	Explicit instruction in morphological and orthographic knowledge → Sessions included recognizing morphemes and breaking down words into their morphemes (affixes, suffixes	Spelling	12 hours.	Group setting, work with the	Students just before examination	Spelling performance below the 25th percentile	3rd grade	Germany

				vention was the instruction in morpheme analysis encouraging applying knowledge of morpheme structure.	and roots) as well as finding analogies between words.			computer				
Weber et al., 2002	MoI	Phonemic reading and spelling training	WCG	This approach aimed to improve phonemic and morphemic awareness skills.	The approach was based on practising phonemic writing, syllable analyses and blending, followed by morpheme segmentation and knowledge about regular deviations.	Spelling	22.5 hours	Group setting	Psychologists, research assistants	Spelling performance below the 15th percentile respective below the 25th percentile if the reading performance was also below the 15th percentile	3rd grade	Germany

Note. AVI = Audio-visual cue intervention; AW = Assisted writing; CG = control group; DSM-IV = Diagnostic Statistical Manual of Mental Disorders, 4th edition; EG = experimental group; GED = General Educational Degree; IQ = intelligence quotient; MI = Memorization intervention; MoI= Morphological intervention; OI = Orthographic intervention; OT = Other treatment; PA = Phonological awareness intervention; PGC = phoneme-grapheme-correspondences; PhI = Phonics intervention; PR = percentile rank; PI = Placebo intervention; RSP = Regular school practice; SD = standard deviation; SS = Supportive software; UK = United Kingdom; USA = United States of America; UCG = Untreated control group; WISC-III = Wechsler Intelligence Scale for Children = 3rd edition; WCG = Waiting list control group; WRAT3 = Wide Range Achievement Test, 3rd edition; WRMT-R = Woodcock Reading Mastery Tests-Revised

Supplement 2. Severity assessment of reading and spelling disorder.

Study	Reading performance	Based on the results of…	Degree of severity in reading	Spelling performance	Based on the results of…	Degree of severity in spelling
Abbott & Berninger, 1999	Reading on average 1 SD below the age norm	- WRMT-R, subtest Word Identification (z = -1.25, PR = 10.56) - WRMT-R, subtest Word Attack (z = -1.073, PR = 14.16)	Moderate	Spelling on average 1 SD below the age norm	- WIAT, subtest Spelling (z = -1.18, PR = 11.9) - WRAT-3, subtest Spelling (z = -1.08, PR = 14.01)	Moderate
Arnbak & Elbro, 2000	Reading at least 2 years below expected reading level	/	Severe	/	- WRAT, subtest Spelling (z = --0.967, PR = 16.69) - QIWK average spelling level = 0.3 (72% scored on Spelling Level 0 of 8), shows severe spelling deficits	Mild
Berninger et al., 2013	Accuracy or rate of single word reading or oral text reading below the population mean and at least 1 SD below the child's verbal IQ	- KTEA-II, subtest Decoding Fluency (z = -1.083, PR = 13.94)	Moderate	Accuracy or rate of spelling below the population mean and at least 1 SD below the child's verbal IQ	- WJ III, subtest Spell Sounds (z = -.683, PR = 24.73) - WIAT-II, subtest Spelling (z = -1.172, PR = 12.06)	Moderate
Berninger et al., 2008	Accuracy and/or rate of oral reading of real and pseudowords on a list or connected text and/or spelling below the population mean and discrepant from verbal IQ by at least 1 SD	- TOWRE, subtest Phonemic Efficency (z = -1.215, PR = 11.22)	Moderate	Accuracy and/or rate of oral reading of real and pseudowords on a list or connected text and/or spelling below the population mean and discrepant from verbal IQ by at least 1 SD	- WJ III, subtest Spelling Sounds (z = -.691, PR = 24.48,) - WRAT-3, subtest Spelling (z = -.978, PR = 16.4)	Mild
Blachman et al., 2014	Standard score below the 25th percentile on either the Word Identification or the Word Attack subtest of the WRMT-R	- WRMT-R, subtest Word Identification (z = -1.009, PR = 15.65) - WRMT-R, subtest Word Attack (z = -1.071, PR = 14.21)	Moderate	/	- WRAT-3, subtest Spelling (z = -1.182, PR = 11.86)	Moderate
Calhoon et al., 2010	Achievement 3.5 grades below their own grade	- WJ III, subtest Letter Word Identification (z = -1.292, PR = 9.82)	Moderate	/	- WJ III, subtest Spelling (z = -1.306, PR = 9.58)	Moderate

	on a combined average of three reading tests	- WJ III, subtest Word Attack (z = -1.106, PR = 13.44) - GSRT (z = -2.167, PR = 1.51)				
Corvacho del Toro, 2016	Not specified	/	/	Pre-diagnosed spelling disorder	- HSP 5-9 (z = 1.21, PR=11.31)	Moderate
Darch et al., 2006	At least 1 SD discrepancy between IQ and achievement	/	/	At least 1 SD discrepancy between IQ and achievement	- WRAT 3, mean spelling grade level across all 2nd to 4th graders: 1.6 (z = -2.26, based on grade equivalent score)	Moderate
Ehri et al., 2009	Private school of struggling readers with history of difficulty in learning to read after at least 1 year of instruction, despite average or above-general intelligence	- WRAT-R and the WRAT-3, subtest Reading (Key-Plus group z = -.707, PR = 23.98; Key group z = -1.353, PR = 8.8)	Mild	/	WRAT-3 subtest Spelling Words: z = -.93, PR 17.67 (Key-Plus group z = -.64, PR = 26.11 , Key group z = -1.327, PR = 9.23)	Mild
Faber, 2005	Not specified	/	/	On average 1.5 SD below age / class norm	- WRT 3+, DRT 4 and DRT 5, subtest A2 of the RST 6-7, HSP 5-9 (control group z = -1.87, PR = 3.07; experimental group z = -1.54, PR = 6.18)	Severe
Fasting & Lyster, 2005	Pupils' reading skills were 1.05 SD below the population mean	- OS400 (silent, single word-decoding test) - LS60 (sentence reading test)	Moderate	Spelling skills on average 1.5 SD below the population mean	- Sivertsen's 20-word spelling test with addition of 12 selected words (z = -1.50)	Severe
Ferraz et al. 2018	Dyslexia classified with DSM-IV	/	Moderate	Dyslexia classified with DSM-IV	/	Moderate
Given et al., 2008	Reading skills were at least below the 25th percentile	- WJ-R, subtest Letter-Word-Identification (z = -.901, PR = 18.38) - WJ-R, subtest Word Attack (z = -1.12, PR = 13.14) - WJ-R, subtest Passage Comprehension (z = -.695, PR = 24.35)	Mild	Spelling skills were on average 1.125 SD below the population mean	- WRAT-3 (z = -1.125, PR = 13.03)	Moderate
Gray, 2016	Struggling readers	- WJ-R, subtest Letter Word Identification (z = -.988, PR = 16.16)	Mild	/	- WJ-R, subtest Spelling (z = -.651, PR = 25.75)	Mild

		- WJ-R, subtest Word Attack (z = -.996, PR = 15.96) - WJ-R, subtest Passage Comprehension (z = -.847, PR = 19.85) - WJ-R, subtest Reading Vocabulary (z = -1.2, PR = 11.51)			- WJ-R, subtest Spell Sounds (z = -.941, PR = 17.34)	
Groth et al., 2013	T-value less the same 40 in reading, regression criterion one standard deviation	- SLRT II, subtest Word Reading (z = -1.71, PR = 4.36) - SLRT II, subtest Pseudo Word Reading (z = -1.36, PR = 8.69) - ELFE 1-6 (z = -1.36, PR = 8.69)	Moderate	T-value less the same 40 in reading, regression criterion one standard deviation	WRT 2+ (z = -1.46, PR = 7.21)	Moderate
Gustafson et al., 2007	Reading disabled children	- Text reading (based on raw score z = -1.26) - General word decoding (based on raw score z = -1.14)	/	/	No standardized spelling test conducted (based on raw score z = -0.87, PR 19.22)	Mild
Hughes et al., 2009	Reading problems (2-3 years below expected performance)	- Mean reading age on BAS: 8,11 (z-score based on grade equivalent score = 2.42)	Severe	/	Mean spelling age on BAS: 8,11 (z-score based on grade equivalent score = 2.42)	Severe
Ise & Schulte-Körne, 2010	Children with specific spelling disorder or dyslexia	Study 1: ELFE 1-6; treatment group (z = -.02, PR = 49.2), control group (z = -.65, PR = 25.78)	Study 1 & 2: average reading performance	Children with specific spelling disorder or dyslexia	Study 1: RST 4-7; treatment group (z = -1.92, PR = 2.74), control group (z = -1.55, PR = 6.06)	Study 1: moderate
		Study 2: ELFE 1-6; treatment group (z = -.27, PR = 39.36), control group (z = -.57, PR = 28.43)			Study 2: RST 4-7 (z = -2.01, PR = 2.22 both groups)	Study 2: severe
Jensen et al., 2000	Dyslexia classified with DSM-IV	/	Moderate	Dyslexia classified with DSM-IV	No standardized spelling test conducted	Moderate
Kast et al., 2007	Performance below 2 SD in one of the two reading subtests (SLRT, ZLT)	- ZLT, subtest Wordlist (z = -4.1, PR < .1) - ZLT, subtest Text (z = -2.7, PR = .35)	Severe	/	- SLRT and DRT 5 (z = -1.3, PR = 9.68)	Moderate
Kirk & Gillon, 2009	Reading achievement of all participants greater than 1SD below the mean (except 1 child)	- WRMT-R, basic skill cluster (merged Word Identification and Word Attack subtest) (z = -.32, PR = 37.5)	Average reading performance	≥ 1 SD below expected age norm in spelling	- TWS-4 (z = -1.477, PR = 6.99)	Moderate
Lange, A., 2009	Low reading ability, 1 year below expected age norm in reading	- WORD, subtest Basic Reading Ability (z= - 1,56, PR = 5.9)	Severe	/	- WORD, subtest Spelling (z= -1,2, PR = 11.51)	Moderate

Oakland et al., 1998	Below the 25th percentile in reading achievement	- WRAT-R, subtest Word Recognition (control group z = -1.569, PR = 5.83, experimental group z = -1.872, PR = 3.06)	Severe	/	- WRAT-R, subtest Spelling (control group z = -1.703, PR = 4.43; experimental group z = -1.784, PR = 3.72)	Severe
O'Shaughnessy & Swanson, 2000	Scores below the 25th percentile in subtests of the WRMT-R and at least 1year below grade level on CBM of Oral Reading Fluency	- WRMT-R, subtest Word Identification (z = -1.427, PR = 7.68) - WRMT-R, subtest Word Attack (z = -1.267, PR = 10.26) - WRMT-R, subtest Passage Comprehension (z = -1.522, PR = 6.4)	Moderate	/	- PIAT-R, subtest Spelling (z = -1.443, PR = 7.45)	Moderate
Rashotte et al., 2001	Below-average phonetic decoding and word-level reading skills	- WJ III, subtest Word Attack (z = -1.126, PR = 13) - WJ III, subtest Word Identification (z = -.878, PR = 19)	Mild		Average performance in Schonell Spelling Inventory (calculated from raw scores) - Grade 1/2: 5;11 y - Grade 3/4: 7;8 y - Grade 5/6: 9;0 y	Grade 1/2:Mild Grade 3/4: Mild Grade 5/6: Moderate
Reuter-Liehr,1993	Not specified	/	/	Spelling disabled children	- WRT 4/5 (z = -1.616, PR = 5.3)	Severe
Schulte-Körne et al., 2001	Discrepancy criterion 1 SD	- SLRT II, subtest Reading (z = -1.43, PR = 7.64)	Moderate	Discrepancy criterion 1 SD	- WRT (z = -1.62, PR = 5.26)	Severe
Spencer et al., 1989	2 or more years behind their age peers in reading	/	/	/	Schonell Spelling Inventory and TWS-2 show mild to moderate spelling deficits	Mild
Strehlow, 1999	Only spelling performance inclusion criterion	- DLF1-2, subtest Reading Accuracy (z = -.598, PR = 27.49) - Reading speed (z = -.035, PR = 48.6)	Mild	Spelling performance under 1 SD	- DRT/WRT (z = -1.463, PR = 7.17)	Moderate
Thaler et al., 2008	Not specified	/	/	Below the 15th percentile	- SLRT Spelling (z = -1.65, PR = 4.95)	Severe
Tijms, 2011	Word reading rate at least 1 SD below the mean	- Word reading skills measured by the 1 min test (z = -1.691, PR = 4.54)	Severe	/	- Standard Dictation of Harel & Schaap (z = -2.313, PR = 1.04)	Severe
Vadasy et al., 2005	Scoring on the lowest quartile for reading skills	WRAT-R and the WRAT-3 - subtest Reading (z = -1.296, PR = 9.75)	Mild	/	WRAT-R and the WRAT-3 - Subtest Spelling (z = -.92, PR = 17.88)	Mild

		- subtest Word Attack (z = -.8, PR = 21.19) - subtest Word Identification (z = -.964, PR = 16.75)				
Walter et al., 2007	Not specified	/	/	Quantitative overall performance of minimum penalty points (PR < 25)	- DRT (PR = 15)	Severe
Weber et al., 2002	Reading performance PR < 15	- WLLP (z = -1.36, PR = 8.69)	Moderate	Spelling performance <15 or <25 if reading performance <15	- DRT (z = -1.48, PR = 6.94) for experimental group and control group	Moderate

Notes. CBM = Curriculum-Based Measurement; DLF 1-2 = Der Diagnostische Lesetest zur Frühdiagnose [Diagnostic Reading Test for Reading Disorders]; DRT 4 = Diagnostischer Rechtschreibtest für 4. Klassen [DRT 4: diagnostic spelling test for 4th graders]; DRT 5 = Diagnostischer Rechtschreibtest für 5. Klassen [DRT 5: diagnostic spelling test for 5th graders]; DSM-IV = Diagnostic Statistical Manual of Mental Disorders, 4th edition; ELFE 1-6 = Ein Leseverständnistest für Erst- bis Sechstklässler [ELFE 1–6: A reading comprehension test for first to sixth graders]; GSRT = Gray Silent Reading Tests; HSP 5–9 = Hamburger Schreibprobe für die Klassen 5 bis 9 [Hamburger Writing Sample Test for class 5 through 9]; IQ = intelligence quotient; KTEA-II = Kaufman Test of Educational Achievement = 2nd Edition; LS60 = Setningsleseprøve [Test of silent sentence reading]; MAT6 = Metropolitan Achievement Test 6; OS-400 = Ordstillelæsingsprøve [Test of silent word reading]; PIAT-R = Peabody Individual Achievement Test-Revised; PR = percentile rank; QIWK = Qualitative Inventory of Word Knowledge; RST 4-7 = Rechtschreibtest für 4-7. Klassen [RST 4-7. Spelling test for grades 4 to 7]; RST 6-7 = Rechtschreibtest für 6. und 7. Klassen [RST 4-7. Spelling test for grades 6 to 7]; SD = standard deviation; SLRT = Salzburger Lese- und Rechtschreibtests [Salzburg reading and spelling test]; SLRT-II = Weiterentwicklung des Salzburger Lese- und Rechtschreibtests (SLRT) [Reading and Spelling Test, SLRT-II]; TOWRE = Test of Word Reading Efficiency; TWS-2: Test of Written Spelling – 2; WIAT = Wechsler Individual Achievement Test; WIAT-II = Wechsler Individual Achievement Test, 2nd edition; WJ III = Woodcock–Johnson Tests of Cognitive Abilities, 3rd edition; WJ-R = Woodcock-Johnson Psycho-Educational Battery-Revised; WLLP = Würzburger Leise Leseprobe [Würzburger Silent Reading Test]; WORD = Wechsler Objective Reading Dimensions; WRAT-3 = Wide Range Achievement Test, 3rd Edition; WRAT-R = Wide Range Achievement Test-Revised; WRMT-R = Woodcock Reading Mastery Test-Revised; WRT 2+ = Weingartener Grundwortschatz Rechtschreibtest für zweite und dritte Klassen [Weingarten Basic Vocabulary Spelling Test for Second and Third Grades]; WRT 3+ = Weingartener Grundwortschatz Rechtschreibtest für dritte und vierte Klassen [Weingarten Basic Vocabulary Spelling Test for Third and Fourth Grades (WRT 3+)]; WRT 4/5 = Westermann-Rechtschreibtest 4/5 [Westermann spelling test for fourth and fifth graders]; ZLT = Zürcher Lesetest [Zurich Reading Test].

Supplement 3. Quality assessment.

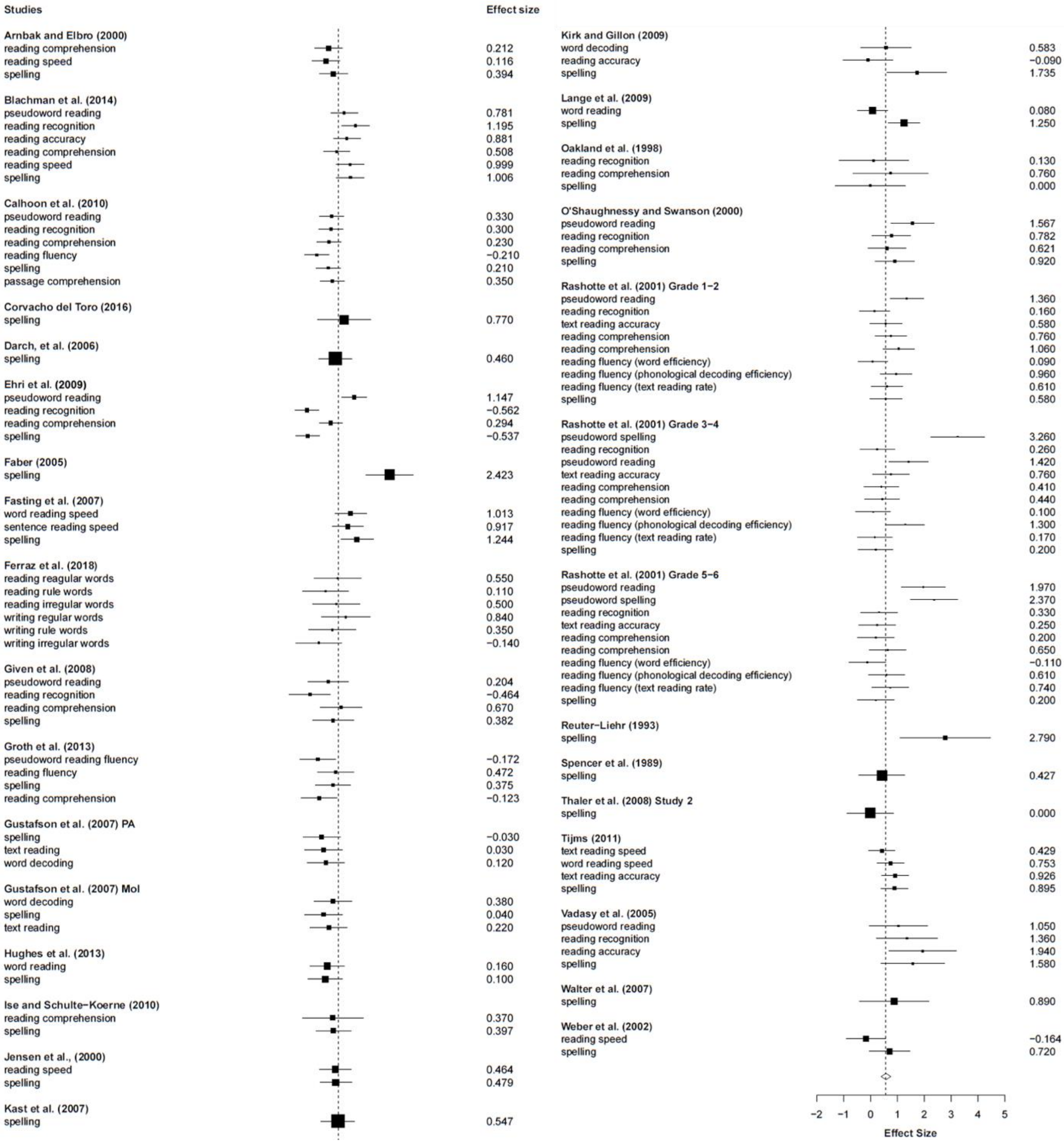
Study	Study type	Randomisation (selection bias)	Adequate concealment method (selection bias)	Similarity of groups at trial start (selection bias)	Blinding (detection bias)	Group differences in provided care (performance bias)	Standard. measures	All outcomes reported (reporting bias)	Drop-out (attrition bias)	Analysis of subjects in their group	Treatment fidelity [a]	Overall risk of bias
Abbott et al., 1999	RCT	✓	✓	✓	N/S	✓	✓	✓	0%	✓	✓	10 %
Arnbak et al., 2000	CT	×	×	✓	×	×	✓ reading × spelling	✓	0%	✓	N/S	60 %
Berninger et al., 2013	RCT	✓	✓	✓	N/S	✓	✓	✓	N/S	✓	✓	20 %
Berninger et al., 2008 (study 1)	RCT	✓	✓	✓	✓	✓	✓	✓	0%	✓	✓	0 %
Blachman et al., 2014	RCT	✓	✓	✓	✓	✓	✓	✓	0%	×	✓ 96%[a]	10 %
Calhoon et al., 2010	CT	×	×	✓	×	✓	✓	✓	N/S	✓	✓ 86-95% [a]	40 %
Corvacho del Toro, 2016	RCT	✓	✓	✓	N/S	✓	✓	✓	0%	✓	✓	10 %
Darch et al., 2006	RCT	✓	✓	✓	×	✓	✓	✓	5%	✓	✓ 98%[a]	10 %
Ehri et al., 2010	CT	×	×	differ in age and reading	×	×	✓	✓	N/S	✓	✓	60 %
Farber 2005	CT	×	×	differ in spelling	N/S	×	✓	✓	N/S	✓	N/S	70 %
Fasting et al., 2005	RCT	✓	✓	differ in IQ	×	✓	✓	✓	0%	✓	N/S	30 %
Ferraz et al., 2018	RCT	✓	✓	✓	N/S	✓	×	✓	0%	✓	N/S	30 %
Given et al., 2008	RCT	✓	✓	✓	✓	✓	✓	✓	14%	✓	N/S	10 %
Gray et al., 2015	RCT	✓	✓	✓	✓	✓	✓	✓	26%	✓	✓ 92% [a]	10 %
Groth et al., 2013	RCT	✓	✓	✓	×	✓	✓	✓	7%	✓	✓	10 %
Gustafson et al., 2007	RCT	✓	✓	✓	✓	✓	✓ reading × spelling	✓	N/S	✓	×	20 %
Hughes et al., 2013	RCT	✓	✓	differ in externalising behavior	✓	✓	✓	✓	0%	✓	N/S	20 %
Ise et al., 2010	CT	×	×	differ in age	N/S	✓	✓	✓	0%	✓	✓	40 %

Jensen et al., 2000	CT	×	×	✓	×	✓	✓	✓	30%	✓	N/S	50 %
Kast et al., 2007	RCT	✓	✓	✓	×	✓	×	N/S	5%	✓	N/S	40 %
Kirk & Gillon, 2009	RCT	✓	✓	✓	N/S	✓	✓	✓	0%	✓	✓	10 %
Lange et al., 2009	RCT	✓	✓	✓	N/S	✓	✓	✓	20%	✓	N/S	20 %
Oakland et al., 1998	cRCT	×	×	differ in SES, oral language	×	×	×	✓	N/S	✓	N/S	80 %
O'Shaughnessy, 1998	RCT	✓	✓	✓	×	✓	✓	✓	0%	✓	✓ 94-97% [a]	10 %
Rashotte et al., 2001	RCT	✓	✓	✓	×	✓	✓	✓	0.8%	✓	✓	10 %
Reuther-Liehr, 1993	cCT	×	×	✓	×	✓	✓	✓	N/S	✓	N/S	50 %
Schulte-Körne et al., 2001	CT	N/S	N/S	✓	×	✓	✓	✓	0%	✓	✓	30 %
Spencer et al., 1989	CT	×	×	✓	×	✓	✓	✓	0%	✓	N/S	40 %
Strehlow et al., 1999	cRCT	×	×	differ in IQ	×	N/S	✓	✓	0%	✓	N/S	60 %
Thaler et al., 2008	RCT	✓	✓	✓	N/S	✓	N/S	✓	0%	✓	N/S	20 %
Tijms et al., 2011	CT	×	×	✓	N/S	✓	✓	✓	2.6%	✓	N/S	40 %
Vadasy et al., 2005	cRCT	×	×	✓	N/S	N/S	✓	✓	21%	✓	✓ 95%[a]	50 %
Walter et al., 2007	cRCT	×	×	✓	N/S	✓	✓	✓	0%	✓	N/S	40 %
Weber et al., 2002	CT	N/S	N/S	✓	N/S	✓	✓	✓	17%	✓	✓	40 %
Overall risk of bias		44 %	44 %	21 %	82 %	18 %	15 %	3 %	12%	3 %	47 %	

Notes. CT = Controlled trial; cRCT = cluster randomized controlled trial; IQ = intelligence quotient; N/A = not applicable; N/S = not specified; RCT = randomized controlled trial; SES = social-economic status. To estimate overall risk of bias the percentage of drop-outs were counted. A loss of > 20% poses serious threats to validity (Sacket, 2000) and was rate as inadequate.

[a] Percentages of adherence to the treatment script or plan.

Supplement 4. Forest-Plot.



Notes. MoI = Morphological Intervention; PA = Phonological awareness intervention.

Supplement 5. Weight function for publication bias.

Based on the assumption of a moderate one-tailed selection (Vevea & Hedges, 2005), we applied a weight function assuming that studies with *p*-values of less than .05 are always observed, studies with *p*-values between .050 to .100 are observed in 90% of the cases, .100 to .250 are observed in 80% of the cases, .250 to .500 are observed in 65% of the cases, .500 to .750 are observed 55% of the cases, and .750 to 1.000 are observed with a probability of 50% of the cases. This resulted in a transformed mean of *g'* = 0.43.

	Probability of observing effect
p interval	Moderate one-tailed selection
.000 - .005	1.0
.005 – .050	1.0
.050 - .100	.90
.100 - .250	.80
.250 - .500	.65
.500 - .750	.55
.750 - .900	.50
.900 - 1.000	.50
Adjusted ES	.44

FACHARTIKEL 2: Predictors of reading and spelling skills in German: the role of morphological awareness

Literaturangabe:
Görgen, R., De Simone, E., Schulte-Körne, G., Moll, K. (2021). Predictors of reading and spelling skills in German: the role of morphological awareness. *Journal of Research in Reading 44*(1), 210-227. doi: 10.1111/1467-9817.12343

Angenommen am 01.12.2020
Online publiziert am 11.01.2021

Journal of Research in Reading, ISSN 0141-0423 *DOI:10.1111/1467-9817.12343*
Volume 44, Issue 1, 2021, pp 210–227

Predictors of reading and spelling skills in German: the role of morphological awareness

Ruth Görgen

Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, University Hospital, LMU Munich, Munich, Germany and Department Language and Education System, Mercator Institute for Literacy and Language Education, University of Cologne, Cologne, Germany

Elisabetta De Simone

Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, University Hospital, LMU Munich, Munich, Germany

Gerd Schulte-Körne

Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, University Hospital, LMU Munich, Munich, Germany

Kristina Moll

Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, University Hospital, LMU Munich, Munich, Germany

Background: The role of morphological awareness for literacy development is non-controversial, but it is likely to depend on the characteristics of a specific orthography. Previous studies analysing the role of morphological awareness are mainly based on English samples; thus, it is unclear how generalisable these results are. In the current study, we evaluated the predictive pattern of morphological awareness on reading and spelling in German, which is characterised by high consistency between morphology and orthography. While many words cannot be spelled correctly by relying on phoneme–grapheme mappings, morphological awareness can be reliably used to infer the correct spelling for many words. In contrast, morphological awareness seems less important for reading in German given the high consistency of grapheme–phoneme mappings. Therefore, we hypothesised that the predictive pattern of morphological awareness for spelling is higher than for reading due to the structure of German orthography. In addition, we examined whether the association between morphological awareness and spelling reflects specific awareness about morphemic units or rather general knowledge about frequent and permissible letter sequences in words.
Method: We developed web-based tasks that allowed us to assess morphological awareness using pseudowords. Based on the data of 3,122 third and fourth graders, we analysed the predictive pattern of morphological awareness on reading and spelling after controlling for non-verbal cognitive abilities, age/grade, phoneme awareness and sublexical orthographic sensitivity.

Results: We found that morphological awareness accounted for significant amounts of unique variance over and above sublexical orthographic sensitivity in both literacy measures but was a better predictor for spelling than for reading.
Conclusions: The role of morphological awareness depends on the structure of a given orthography. In German, this is reflected by differences in the predictive pattern between reading and spelling skills. Furthermore, results support the specific role of morphological awareness for spelling in the German orthography.

Keywords: morphological awareness, sublexical orthographic sensitivity, reading, spelling, German orthography

Highlights

What is already known about this topic

- Words are linguistic signs made up of smaller units, such as letters, syllables and morphemes.
- Morphological awareness is recognised as an important predictor of reading, spelling and reading comprehension.

What this paper adds

- Novel, web-based measures can be used to assess literacy as well as related cognitive skills, such as phonological and morphological awareness, and sublexical orthographic sensitivity.
- In German, morphological awareness is a better predictor for spelling than for reading (fluency). The relative importance of morphological awareness might differ according to the orthographic structure of the language.
- Morphological awareness predicts literacy and does not simply reflect general sublexical skills.

Implications for theory, policy or practice

- Morphological awareness is a specific predictor of literacy skills in a consistent orthography like German. Inclusion in diagnostic batteries might therefore be considered.
- Training of morphological awareness and highlighting the morphemic units in words might be especially useful for spelling intervention programmes.

In recent years, researchers have become increasingly interested in understanding the facilitating role of morphological awareness in literacy acquisition (e.g., Deacon et al. 2008; Reed 2008). Containing phonological, semantic and syntactical information, morphemes represent the smallest meaningful language units. *Morphological awareness* refers to an individual's awareness of the meaning and the role of morphemes in a certain word. This includes the capability to manipulate these morphemes in a meaningful way, for example, by creating new words using derivational (e.g., worker = work + er), inflectional (e.g., working = work + ing) or compound morphology (e.g., worksheet = work + sheet) (Carlisle 1995, 2010; Nagy et al. 2013).

Development of morphological awareness begins early during spoken language acquisition (Carlisle 2010) and is closely linked to vocabulary growth (Bertram et al. 2000; Sparks & Deacon 2015). During school years, the awareness of the morphological structure of words increases and so does its contribution to literacy skills during secondary education (e.g., Rispens et al. 2008; Berninger et al. 2010).

Morphological awareness has been shown to affect the following literacy skills: (1) reading comprehension, (2) word recognition and (3) spelling (for reviews refer to, e.g., Bowers et al. 2010; Goodwin & Ahn 2010). Here, we focus on word recognition (word reading accuracy and fluency) and spelling.

How can the role of morphological awareness for word recognition be explained? In less consistent orthographies like English, where grapheme–phoneme correspondences are inconsistent, readers need to rely on larger units, such as whole words and morphemes, in order to read a word correctly, because decoding a word using smaller sublexical units such as single graphemes does not result in correct word reading (Nagy et al. 2003). In accordance, English-based studies consistently report that morphological awareness plays an important role for word recognition (i.e., reading accuracy). However, the role of morphological awareness for word recognition has been reported not only in English (e.g., Kirby et al. 2012; Casalis et al. 2015) but also in other orthographies along the transparency continuum, including French (Quémart et al. 2012; Casalis et al. 2015), Portuguese (de Freitas et al. 2017) and Italian (Burani et al. 2002; Angelelli et al. 2014).

Compared with word reading accuracy, less is known about the role of morphological awareness for *fluent* reading, although it was shown that, with regard to the English orthography, the use of morphemes contributes to reading fluency as well (Nunes et al. 2012). For reading fluency, orthographies that are characterised by a high consistency between graphemes and phonemes (e.g., Finnish, Italian and German) are of special interest, given that reading accuracy and fluency are not as intertwined as in less consistent orthographies, such as English. This is because in consistent orthographies reading accuracy is close to ceiling after 1 year of reading instruction (Seymour et al. 2003), so that reading fluency can be assessed irrespective of accurate reading.

How can the role of morphological awareness for fluent reading be explained? Although in consistent orthographies words can be read accurately using smaller sublexical units (e.g., single graphemes), there is evidence that morphological awareness might still play a role in *fluent* reading (Italian: e.g., Marcolini et al. 2011; German: Volkmer et al. 2019). This seems reasonable given that the use of larger written units might speed up the recognition of unknown words and is likely to be faster than decoding a word by breaking it down into single graphemes (Carlisle & Stone 2005). The use of morphemes may be especially important when reading long compound words, which frequently occur in orthographies like Finnish or German (e.g., German: *Staßenbahnhaltestelle* – 'street-train-stop-place' = tram stop).

With respect to *spelling*, the majority of alphabetic orthographies are less consistent in the spelling direction than in the reading direction (Bosman & van Orden 1997; Ziegler et al. 1997; Moll & Landerl 2009), suggesting that morphological awareness might play a particularly important role in spelling. Indeed, robust relationships between morphological awareness and spelling were reported across orthographies (English: Deacon et al. 2009; French: Pacton et al. 2013; German: Volkmer et al. 2019). One exception is the Finnish orthography, which is highly consistent in both reading and spelling direction. In line with the idea that morphological awareness is especially important when phoneme–grapheme consistencies are inconsistent, Lethonen and Bryant (2005) did not find a significant contribution of morphological awareness to spelling in Finnish.

Together, these findings suggest that although the mechanisms underlying the association between morphological awareness and literacy skills apply to all alphabetic orthographies, the relative importance of morphological awareness for the different literacy skills differs between orthographies (Kuo & Anderson 2006; Bowers et al. 2010). Further evidence for differences between orthographies comes from the few existing cross-linguistic studies directly comparing the role of morphological awareness for literacy skills in orthographies of varying orthographic consistency (Casalis et al. 2015; Desrochers et al. 2018; Manolitsis et al. 2019). For example, Desrochers et al. (2018) compared the role of morphological awareness in English, French and Greek, with English representing the least consistent, and Greek representing the most consistent orthography of the three orthographies. Results indicated that morphological awareness predicted reading accuracy only in English, reading fluency in English and French, and reading comprehension and spelling in all three languages. Thus, the role of morphological awareness depends on the characteristics of a specific orthography (Seymour et al. 2003), namely, the consistency between graphemes and phonemes as well as the consistency between morphology and orthography.

Against this background, the first aim of the current study is to examine the role of morphological awareness in German orthography. German represents an interesting case, as the asymmetry of consistency between reading and spelling direction is especially striking. In German, grapheme–phoneme correspondences (reading direction) are highly consistent, thus allowing to assess reading fluency independent of reading accuracy. In contrast, phoneme–grapheme correspondences (spelling direction) are rather inconsistent. For example, the long vowel /a:/ can be represented by three different graphemes: <a>, <aa> and <ah> as in 'Tal – valley', 'Saal – hall' and 'Wahl – election'. In order to produce the correct spelling, larger units (e.g., morphemes or whole words) need to be applied. The use of morphological strategies in spelling is especially promising in orthographies like German, which are characterised by high consistency between morphology and orthography so that morphological awareness can be reliably used to infer the correct spelling of a word (Landerl & Reitsma 2005). For instance, German has a phonologically inconsistent, but morphologically consistent spelling pattern of voicing alteration of the plural. This means that words that have a voiced syllable-final consonant are devoiced in their singular form (e.g., 'Hund' [/hunt/ – dog] is pronounced with a final /t/). In the plural form, the voicing value and the grapheme match, as 'Hunde' is pronounced with a /d/. Thus, children will spell the singular correctly if they take the plural spelling into account (Kargl et al. 2017)

Based on the characteristics of the German orthography and previous research in different orthographies, we expect morphological awareness to impact on both reading fluency and spelling, but will play a larger role for spelling, given the inconsistency of phoneme–grapheme correspondence in the spelling direction.

The second aim of the current study is to assess the association between morphological awareness and other cognitive predictors of literacy skills. Previous findings suggest that morphological awareness affects literacy skills over and above other well-known predictors, including phonological skills, vocabulary knowledge and lexical orthographic processing (e.g., Nagy et al. 2003; Deacon & Kirby 2004; Roman et al. 2009). This suggests that morphological awareness impacts on literacy skills after controlling for awareness of smaller units (phonemes) as well as after controlling for whole word knowledge. However, only few studies have investigated whether morphological awareness still accounts for individual differences in literacy skills once sublexical orthographic sensitivity has been taken into account (Pacton et al. 2005). Sublexical orthographic sensitivity refers to the sensitivity to sublexical orthographic conventions in a given orthography, such as permissible letter sequences and letter positions in a word (Perfetti 1984; Shahar-Yames & Share 2008). Compared with lexical orthographic tasks, which measure word-specific knowledge (similar to spelling tests) (Burt 2006), sublexical orthographic tasks provide information whether a child is sensitive to orthographic regularities irrespective of word specific knowledge. This is for example the case when a child selects a pseudoword with a frequent double consonant (e.g., <ll> in German) more often as a plausible German spelling than a pseudoword with a rare double consonant (e.g., <dd> in German), (Rothe et al. 2013). This implicit sublexical orthographic sensitivity is likely to help children's spelling, and it seems that even when children can rely on morphological information, they are still influenced by their sensitivity to orthographic regularities (Cassar & Treiman 1997; Pacton et al. 2005; Rothe et al. 2013). As children use both knowledge about morphemes and orthographic regularities, controlling for sublexical orthographic sensitivity enables to specify the role of morphological awareness for literacy skills (Treiman et al. 1994; Kemp & Bryant 2003; Deacon & Leung 2013). For example, in German, there is no difference when pronouncing the letters 'f' and 'v' (e.g., *Fisch* [*fɪʃ*] – fish vs *Verbot* = [*fɛɐ̯*'bo:t] – ban), which often causes confusion in spelling. The letter pattern 'fer', however, is rarely used initially (only in roots, such as '*Ferse* – heel'), while in final positions, it occurs more frequently (in roots as in '*Ufer* – shore' or resulting from suffixation with -er as in '*tiefer* – deeper'). Thus, being aware of 'ver' as a morpheme, as well as being sensitive to the position of the sublexical unit 'fer', can help to spell words correctly. When analysing the unique contribution of morphological awareness to literacy skills, it is therefore important to control for sublexical orthographic sensitivity (Deacon & Leung 2013).

Aims of the study

The present study aims to explore the unique variance of morphological awareness on reading fluency and spelling skills in the asymmetric German orthography by examining third and fourth grade children (age 9 and 10 years) using novel web-based measures. This age range is especially interesting for the present study, as children more and more acquire orthographic knowledge and regularities (Ehri 2005). Thus, the use of morphological awareness is likely to be important and might increase from third to fourth grade.

First, we take advantage of the asymmetry of German orthography (i.e., consistent in the reading direction and inconsistent in the spelling direction) in order to investigate how orthographic consistency affects the role of morphological awareness within one language in a within-subject design. Based on the few existing studies that have investigated the impact of morphological awareness on reading fluency (Marcolini et al. 2011; Nunes et al. 2012; de Freitas et al. 2017; Desrochers et al. 2018; Vernice & Pagliarini 2018; Volkmer et al. 2019), we assume that morphological awareness will explain unique variance in reading fluency. However, given the asymmetry of German orthography, we predict that morphological awareness plays a more important role for spelling than for reading.

Second, we extend previous findings about the specific role of morphological awareness for reading and spelling skills by controlling for sublexical orthographic sensitivity irrespective of lexical knowledge.

Method

Participants

Within a large project using a web-based application to assess children's academic skills and psychopathological profile, 52,734 families with third and/or fourth grade children from the two German federal states Bavaria (n = 27,734) and Hesse (n = 25,000) were invited for study participation. Of the 4,621 families who initially downloaded the application, 4,542 (8.6% of those invited) families started study participation by giving consent to study procedures and providing demographic information about their child. We excluded children who did not complete all tests relevant for the present study (N = 1,341; 29.5%) and children whose data did not survive plausibility checks (N = 79; 1.7%), which were applied to the standardised tests. Thus, in total, 1,420 children (31.3%) were excluded, resulting in a final sample of 3,122 children.

According to parental statements via the application, 72.3% of participating children were monolingual German, and 23.3% were bilingual. For the remaining 4.4%, this information was missing. The socioeconomic status of families participating in the study was relatively high with 1,743 mothers (55.8%) that have completed the highest German school degree.

Within the sample, 6.7% scored at least 1.5 *SD* below the expected grade level (American Psychiatric Association 2013; Galuschka & Schulte-Körne 2016) on reading (*Wuerzburger silent reading test – revised* [WLLP-R]), and 6.9% scored below the grade level on spelling (*Weingarten spelling test for basic vocabulary for years 3 and 4/4 and 5* [WRT 3+/4+]). The contacted families were randomly chosen by local registration offices of Bavaria and the Hessian Ministry of Culture. The study was approved by the ethics committees of the University Hospital of the Ludwig Maximilian University Munich and the DIPF|Leibniz Institute for Research and Information in Education, Frankfurt am Main, and was performed in accordance with the latest version of the Declaration of Helsinki and in compliance with the national legislation.

Procedure

Standardised paper–pencil tests and questionnaires were transformed to the web-based application by the software company Meister Cody GmbH. In addition, non-standardised tasks were developed explicitly for web-based usage on tablets and smart phones. The families were asked to download and use the web-based application independently within 8 weeks on their own tablets or smart phones. The application consisted of five sessions for children, each session lasting 30 to 45 min. The five sessions had to be completed on five different days. The children were guided through the web-based application by magician Meister Cody and received detailed instructions per task. All tasks started with two to four practice items (depending on the task) that the children had to solve correctly before the actual task was started. To prevent children from going off-task, there was an automatic reminder to continue with the task if there was no reaction for a certain amount of time (15 or 25 s, depending on the task).

Measures

Morphological awareness. Morphological awareness was assessed using pseudowords in three different domains with in total 22 items. *Plural formation* (seven items) measured the ability to choose the correct plural form out of three options. Single fantasy creatures were shown and introduced (e.g., *Das ist ein Roto* – That is a Roto). Subsequently, the children were asked to choose the correct name for two equivalent creatures while seeing a matching picture (e.g., *Das sind zwei …* – These are two … 'Roto' – 'Roton' – 'Rotos'). Similar tasks are used in prior work on plural formation (Berko 1958) and in a common standardised German language test (SET 5-10; Petermann 2010). For the second domain, *verb inflection* (eight items), the children were supposed to identify the correct inflection of a pseudoverb. First, a verbs' infinitive (e.g., dannen) was presented orally and in written form. Second, a sentence frame was given and the children had to select the appropriate inflected verb out of a selection of three (e.g., Ben … 'dannt' – 'dannst' – 'dannen', refer to Diamanti et al. 2017 for a similar procedure). The third domain, *derivation* (seven items), measured the ability to recognise correct derivational changes of words, comparable with procedures used by Diamanti et al. (2017) and Fink et al. (2012). The children heard and read a sentence containing a pseudoword (e.g., *Wir kresen* – We kresen). Slightly adapted sentence frames were presented, subsequently, and children had to select the appropriate derived form of the previous used pseudoword fitting to the adapted sentence frame (e.g., *Das war eine tolle …* – This was a great … 'Kreserei' – 'Krest' – 'Kresel'). All used pseudowords followed the structure of existing German words. The response alternatives were constructed close to the target response but were either illegal in the specific sentence frame, or irregular, or infrequent, or did not represent permissible word forms according to German word formation conventions. Cronbach's alpha was .61, after deleting the two items that showed the lowest correlation with the total scale. Deleting more items would have considerably reduced the number of items without a clear increase in reliability.

Phoneme awareness. For the present study, a phoneme deletion task (e.g., Mann & Wimmer 2002; Casalis et al. 2011), was designed. Twenty-one pseudowords were presented orally, and for each pseudoword, children were asked to identify the correct word after deleting a specified sound. Children had to choose the picture that corresponds to the target response out of three orally and visually presented real word response alternatives, which were closely related in terms of pronunciation (e.g.*, what word do you hear when you delete the [/f/] in [vɛlfə]?* [*vɛlpə*] – pup, [*vɛlə*] – wave, or [*bɛlə*] – balls). Cronbach's alpha was .79.

Sublexical orthographic sensitivity. This task measured the ability to differentiate between permissible and non-permissible letter combinations according to German written language rules using pseudowords. Therefore, 60 pseudowords were presented on the screen, and the child was asked to decide whether the pseudowords' spelling corresponds to German real word spellings by pressing a yes or no response button (e.g., [*solbst*] with permissible letter pattern or [*vuulɐ*], which has a non-permissible double 'u' within one morpheme). For tasks setup in a similar way, refer to, for example, Pacton et al. (2001) and Rothe et al. (2013). Cronbach's alpha was .86.

Reading. The standardised reading test WLLP-R (Schneider et al. 2011; parallel-test reliability r = .93 for grade 3 and r = .82 for grade 4) was administered, which measures word reading speed. Children were asked to silently read up to 180 simple and high frequency words, choosing the corresponding image out of four options for each word. One written word was displayed simultaneously with the four images before the next written word with the associated images was presented. The task was terminated after 5 min, and the number of correct responses was calculated.

Spelling. The long versions of the standardised 'Weingarten spelling test for basic vocabulary' (WRT 3+ for grade 3 children; Birkel 2007a; parallel-test reliability $r > .91$, and WRT4+ for grade 4 children; Birkel 2007b; parallel-test reliability $r > .90$) were administered. Children had to fill in missing words (55 words in third grade/60 words in fourth grade) by using the virtual keyboard presented on the screen of their device. Children were allowed to correct their typing and had to confirm each typed word by ticking a box. The sentence frames and target words were read aloud to the children.

Non-verbal intelligence. Non-verbal cognitive abilities were estimated with the subtests *Series*, *Classification*, and *Matrices* of the German version of the *Culture fair intelligence test* (CFT 20-R; Weiß 2006; retest reliability: $r$tt = .92–.96).

For the standardised tests (WLLP-R, WRT 3+/4+ and CFT 20-R), reliability information as reported in the manual was used. For the experimental measures, we calculated Cronbach's alpha based on the study sample.

In order to validate the results from online testing, a subsample (N = 237) was retested with the corresponding traditional paper–pencil tests for reading and spelling after the online testing. This was especially important for the spelling test in order to ensure that writing and typing words does not result in different responses. The results of the online and paper–pencil tests were highly correlated (WLLP-R: r^2 = .79; WRT3+: r^2 = .74; WRT 4+: r^2 = .87), and the internal consistency of the online versions was also high (WLLP-R: α = .96; WRT3+: α = 93; WRT4+: α = .95).

Data preparation

Because we were unable to monitor the child's behaviour while performing the tests, we applied plausibility checks to the standardised tests to filter out unreliable data. For WLLP-R (reading speed test), data were considered as implausible when a child repeatedly (i.e., more than 10 times) selected the same answer alternative and/or when total testing time deviated from the maximum time limit of 5 min due to technical problems. For WRT3+/4+ (spelling accuracy test), data were considered as implausible, when a child just typed a key or random letters in more than 15% of the dictated items. Random typing was defined by a Jaro–Winkler distance (Winkler 1990) of more than 0.41 between the answer and the target word.

Results

Descriptives

Descriptive statistics for literacy outcomes, age and predictors are displayed separately for grades 3 and 4 in Table 1. Due to negative skew across literacy and predictor variables, logarithmic data transformation was performed (e.g., Tabachnick & Fidell 2013). Data transformation did not result in normally distributed data, but improved skewness and/or kurtosis. For this reason, correlational as well as regression analyses are reported with transformed data but were also confirmed using raw data (refer to section Regression analyses).

Correlation analyses

Table 2 presents the correlations between literacy measures, non-verbal cognitive abilities, sublexical orthographic sensitivity, phoneme awareness and morphological awareness separated by grades. Correlations were comparable between grades but differed significantly between non-verbal cognitive abilities and each of the specific predictor and between spelling and morphological awareness with higher correlations in grade 4. Furthermore, correlations of literacy measures were higher for phoneme awareness and morphological awareness than for sublexical orthographic sensitivity and non-verbal cognitive abilities.

TABLE 1. Descriptive statistics for the two grade levels.

Variable	Grade 3 (*N* = 1,477)		Grade 4 (*N* = 1,645)	
	M (*SD*)	Range	M (*SD*)	Range
Age (months)	110.73 (4.4)	97.0–130.0	121.98 (4.6)	98.0–140.0
Reading (WLLP-R)[1]	92.66 (16.5)	23–156	103.72 (15.6)	7–166
Spelling (WRT 3+/4+)[1]	36.71 (11.4)	1–55	40.82 (13.03)	0–60
Non-verbal cognitive abilities[1]	25.33 (5.2)	12–42	26.90 (5.5)	9–43
Sublexical orthographic sensitivity[1]	45.73 (7.9)	13–60	47.32 (7.6)	24–60
Phonological awareness[1]	15.52 (3.9)	3–21	16.32 (3.8)	3–21
Morphological awareness[1]	18.63 (2.7)	8–22	19.44 (2.2)	9–22
	Frequencies			
Male/female	748/729		833/812	
Monolingual/bilingual	1,051/362		1,207/366	
Reading performance below average[2] (yes/no)	102/1,375		108/1,537	
Spelling performance below average[3] (yes/no)	99/1,378		115/1,530	

WLLP-R, Wuerzburger Silent Reading Test – Revised; WRT 3+, Weingarten spelling test for basic vocabulary for years 3 and 4; WRT 4+, Weingarten spelling test for basic vocabulary for years 4 and 5.
[1]Raw scores.
[2]Reading performance below average = at least 1.5 *SD* below the expected grade level measured with WLLP-R.
[3]Spelling performance below average = at least 1.5 *SD* below the expected grad level measured with WRT 3+/4+.

TABLE 2. Correlations (Pearson) between literacy measures and predictors for the whole sample ($N = 3{,}122$) separated for third (upper row) and fourth (lower row) grade.

	1	2	3	4	5	6
1. Reading		.476**	.164**	.096**	.286**	.310**
		.512**	.185**	.100**	.287**	.327**
2. Spelling			.285**	.224**	.441**	.432**
			.323**	.197**	.479**	.486**
3. Non-verbal cognitive abilities				.178**	.237**	.191**
				.273**	.318**	.279**
4. Sublexical orthographic sensitivity					.243**	.213**
					.248**	.235**
5. Phonological awareness						.422**
						.423**
6. Morphological awareness						

Note: Analysis was performed with grade-specific *z*-scores (based on raw scores) after logarithmic data transformation.
** $p \leq 0.01$.

Regression analyses

The predictive pattern for reading and spelling was examined by converting raw scores into grade-specific *z*-scores in a series of stepwise regression analyses. Separate regression analyses were performed for reading and spelling, each controlling for differences in non-verbal cognitive abilities and age in steps 1 and 2, respectively. The specific predictors, sublexical orthographic sensitivity, phoneme awareness and morphological awareness were added simultaneously in step 3, in order to calculate the unique variance for each predictor variable as well as the shared variance between predictors.

TABLE 3. Three step regression analyses for non-verbal cognitive abilities, age and specific predictors with reading as dependent variable.

		Reading			
		R^2 change (%)	*p*	*B*	*SE B*
Step 1	Non-verbal cognitive abilities	3.0	<.001	.174	.018
Step 2	Age	0.0	.731	.006	.000
Step 3	Unique sublexical orthographic sensitivity	0.0	.548	−.011	.018
	Unique phoneme awareness	2.3	<.001	.173	.019
	Unique morphological awareness	4.0	<.001	.233	.019
	Shared variance step 3	3.9			
	Total variance step 3	10.2			
	Total variance steps 1–3	13.2			

Note: Combined analysis for third and fourth grade was performed with grade-specific *z*-scores (based on raw scores) after logarithmic data transformation.

As shown in Tables 3 and 4, the total contribution of all predictors was higher for spelling (31.8%) than for reading (13.2%). The unique variance of non-verbal cognitive abilities differed substantially between reading (3.0%) and spelling (9.2%). The variance explained by age was negligible. The variance explained by sublexical orthographic sensitivity alone was small (0–0.2%) and only significant for spelling. Phoneme awareness accounted for more unique variance in spelling (6.3%) than in reading (2.3%). Morphological awareness was the most important unique predictor for both reading and spelling, with higher amounts of variance accounted for in spelling (6.8%) than in reading (4.0%). The shared variance between all three predictors was 3.9% for reading and 9.2% for spelling. The pattern was the same for raw data, resulting in a total contribution of 31.7% for spelling and 13.5% for reading (Tables S4 and S5).

As phoneme awareness and morphological awareness accounted for the highest amounts of unique variance, we also calculated the shared variance of these two predictors, after controlling for non-verbal cognitive abilities, age and sublexical orthographic sensitivity. The shared variance of phoneme and morphological awareness was 3.5% for reading and 7.3% for spelling.

In order to examine whether the impact of predictors on literacy skills differs between grades, further regression analyses were conducted, analysing interaction effects between grade and the specific predictor variables (Pedhazur & Schmelkin 1991). Therefore, we calculated deviation scores by subtracting the mean from each of the independent variables so that each variable had a mean of 0 and a standard deviation of 1. We then multiplied each of the specific predictors by the dummy-coded variable grade in order to compute the interaction term. The interaction for each specific predictor with grade was calculated in a separate regression analysis after controlling for non-verbal cognitive abilities in step 1 and simultaneously entering predictors and the interaction term in step 2. The explained variance by all interaction terms was small (0–0.2%) and were neither significant for sublexical orthographic sensitivity × grade (Reading: p = .984, Spelling: p = .237) nor for phonological awareness × grade (Reading: p = .635, Spelling: p = .073). The interaction

TABLE 4. Three-step regression analyses for non-verbal cognitive abilities, age and the specific predictors with spelling as dependent variable.

		Spelling			
		R^2 change (%)	*p*	*B*	*SE B*
Step 1	Non-verbal cognitive abilities	9.2	<.001	.303	.017
Step 2	Age	0.1	.030	-.037	.000
Step 3	Unique sublexical orthographic sensitivity	0.2	.007	.042	.016
	Unique phoneme awareness	6.3	<.001	.286	.017
	Unique morphological awareness	6.8	<.001	.293	.017
	Shared variance step 3	9.2			
	Total variance step 3	22.5			
	Total variance steps 1–3	31.8			

Note: Combined analysis for third and fourth grade was performed with grade-specific *z*-scores (based on raw scores) after logarithmic data transformation.

term for morphological awareness × grade was not significant for reading (p = .097), but for spelling ($p < .01$). Examining the significant interaction of morphological awareness × grade for spelling by performing separate regression analyses for the two grades revealed that morphological awareness accounted for more unique variance in grade 4 (8.1%) than in grade 3 (6.0%).

Discussion

The current study examined the predictive pattern of morphological awareness on reading fluency and spelling skills while controlling for non-verbal cognitive abilities, age/grade, phoneme awareness and sublexical orthographic sensitivity. Our findings extend previous research by showing that morphological awareness contributes to word recognition and especially to spelling in a consistent but asymmetric orthography like German, and by providing evidence for the role of morphological awareness for reading fluency.

First, we asked to what extent morphological awareness influences reading and spelling in the asymmetric German orthography. In line with our prediction, morphological awareness accounted for more variance in spelling than in reading fluency. This goes in line with recent results reported by Volkmer et al. (2019), who also found significant contribution of morphological awareness on spelling for the German orthography and by Desrochers et al. (2018), who emphasised the influence of morphological awareness on spelling for the consistent Greek orthography. In German, the contribution of morphological awareness on spelling is likely to result from the strong asymmetry in German orthography, which is characterised by clearly lower phoneme–grapheme consistency (spelling direction) than grapheme–phoneme consistency (reading direction). Consequently, spelling in German depends more strongly on word specific representations and larger sublexical units, such as morphemes, than reading.

Correlation and regression analyses further indicated that the impact of morphological awareness on spelling increased from third to fourth grade. This supports the idea that morphological awareness gets more important with literacy experience and increases once phoneme–grapheme correspondences are mastered and the correct spelling of a word relies on the use of larger written language units (Kirby et al. 2012; Treiman 2018).

Although we found that the amount of variance explained in reading was clearly smaller than in spelling, morphological awareness still accounted for unique variance (4.0%) in reading fluency. This suggests that even in a consistent orthography where readers do not have to rely on larger units in order to read a word correctly, the use of larger written units still seems to facilitate word recognition speed (as measured by reading fluency). In line with this idea, De Freitas et al. (2017) indicated that in more advanced reading stages, when reading gets more automatised (i.e., consolidated phase of Ehri's model of word recognition, 2005), morphological awareness becomes more important for word recognition. Furthermore, it should be noted that our standardised reading measure contains many short and high frequent words. As outlined in the introduction, morphological awareness might be especially important when reading long compound words (Gilbert et al. 2013). Thus, future studies may include experimental reading tests with different word lists that allow comparing the role of morphological awareness for reading simple and reading compound words.

Our second aim was to assess the contribution of morphological awareness on reading and spelling while taking phoneme awareness and sublexical orthographic sensitivity into account in order to identify the specific role of morphemic units during reading and spelling. The results showed that morphological awareness accounted for the highest amount of unique variance in both reading and spelling models, confirming the relevance of morphological awareness for literacy acquisition shown in previous research (e.g., Bowers et al. 2010; Carlisle 2010; Goodwin & Ahn 2010; Nagy et al. 2013).

In accordance with earlier findings (e.g., Caravolas et al. 2005; Georgiou et al. 2008), phoneme awareness also turned out to be an important predictor for literacy attainment and accounted for unique variance in both literacy measures. As was the case for morphological awareness, phoneme awareness was more important for spelling than for reading, a finding that has been reported in German before (Moll et al. 2009; Moll et al. 2014).

In our analyses, the amount of variance explained by sublexical orthographic sensitivity was small and only significant for spelling. For spelling, the data indicate that it is indeed the morphemic unit that is important and that accounts for individual differences in spelling, rather than the knowledge about permissible and non-permissible letter combinations.

Furthermore, the current study provides evidence that literacy skills of third and fourth grade children can reliably be assessed using a web-based application. This is particularly important given the use of advanced technologies in educational and therapeutical contexts has increased over the past decades (Chauhan 2016) and has become indispensable as a result of temporarily school closures all over the world during COVID-19 pandemic (UNESCO 2020). We encourage to further develop innovative tools to assess literacy skills as we were able to reach a large number of children by having them use the application independently at home. Although we were not able to control for parental support or surrounding factors, such as background noises or interruptions, we have taken several actions to ensure high quality of the data, such as plausibility checks and validation with standardised paper–pencil measures.

Finally, we would like to discuss some limitations and, by doing so, provide suggestions for future research. First, we reduced the confound of vocabulary knowledge on morphological awareness by using pseudowords. This enabled us to assess the child's ability to generalise morphological rules to unfamiliar word material (Berko 1958; Kargl et al. 2017), but we cannot completely rule out that semantics were still partly activated by analogies and by the real word sentence frames. Second, although previous research has consistently shown that rapid automatised naming (RAN) is a strong predictor for reading fluency across orthographies (Moll et al. 2009; Kirby et al. 2010) and fosters fluent reading independent of the units (i.e., whole words, morphemes, or single graphemes), we did not include RAN measures in our study. Supported by the findings of Roman et al. (2009), we assume that the impact of morphological awareness on reading and spelling survives the inclusion of RAN measure as a control. Future studies will have to clarify to what extend morphological awareness indeed accounts for unique variance in reading fluency and spelling after controlling for RAN and vocabulary knowledge. In our morphological awareness task, we have included different dimensions of morphological awareness (i.e., verb inflection, plural formation and derivation) that have been assessed in the literature (Kirby et al. 2012). Due to time limitations, it was not possible to include a large number of items for each dimension, so that we could not analyse the predictive pattern separately for each dimension. Furthermore, even though German orthography is morphologically regular, there are exceptions such as the plural formation with various irregular forms. In our morphological awareness task, irregular but still possible forms have been avoided to a large extent but were not completely excluded. Moreover, as in prior work on morpheme recognition (e.g., Derwing 1976; Diamanti et al. 2017), we used judgement instead of production tasks because production tasks were not suitable for the current web-based format. Future research may combine judgement and production tasks in order to prevent guessing. It should also be noted that we do not have detailed information whether the children used tablets or smart phones. Although children were mainly asked to select an answer by simply tapping, we cannot rule out that the type of device might have had an impact on test performance, for example, when typing words in the spelling test.

Overall, the current findings add to the existing literature by showing that morphological awareness predicts reading and spelling in a consistent orthography. While the total amount of variance explained in reading is rather small, the amount of variance explained in spelling is comparable with prior research in consistent orthographies (Desrochers et al. 2018; Volkmer et al. 2019), although smaller than reported in less consistent orthographies (Desrochers et al. 2018). The inclusion of further relevant predictors such as RAN instead of sublexical orthographic sensitivity is expected to increase the total amount of variance accounted for in the models, especially for reading fluency. Our results, however, show that morphological awareness does impact not only on recognising words accurately but also on reading words fluently. Results further indicate that the role of morphological awareness seems to depend on the structure of the orthography. In the asymmetric German orthography, morphological awareness plays a more important role for spelling than for reading with an important implication for practice: in orthographies with high consistency between morphology and orthography, focusing on the morphemic structure of words will be very promising in teaching spelling.

Acknowledgements

We thank the children and parents for their participation in this study. We also thank Marcus Hasselhorn, Julia Kalmar, Josefine Rothe, Linda Visser, Katharina Grunwald and Janosch Linkersdörfer for their contribution in the preparation and execution of data collection. Furthermore, we thank Katharina Galuschka and Sini Huemer for their contribution in the development of the experimental tasks. This project was supported by a grant from the Federal Ministry of Education and Research (BMBF, grant no. 01GJ1601A).

References

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders (DSM) 5*. Washington DC: American Psychiatric Association.

Angelelli, P., Marinelli, C.V. & Burani, C. (2014). The effect of morphology on spelling and reading accuracy: A study on Italian children. *Frontiers in Psychology*, 5(1373), 1–10. https://doi.org/10.3389/fpsyg.2014.01373

Berko, J. (1958). The child's learning of English morphology. *Word*, 14, 150–177. https://doi.org/10.1080/00437956.1958.11659661

Berninger, V.W., Abbott, R.D., Nagy, W.E. & Carlisle, J.F. (2010). Growth in phonological, orthographic, and morphological awareness in grades 1 to 6. *Journal of Psycholinguistic Research*, 39, 141–163. https://doi.org/10.1007/s10936-009-9130-6

Bertram, R., Laine, M. & Virkkala, M.M. (2000). The role of derivational morphology in vocabulary acquisition: Get by with a little help from my morpheme friends. *Scandinavian Journal of Psychology*, 41, 287–296. https://doi.org/10.1007/s10936-009-9130-6

Birkel, P. (2007a). *Weingartener grundwortschatz rechtschreib-test für dritte und vierte klassen*. Göttingen: Hogrefe.

Birkel, P. (2007b). *Weingartener grundwortschatz rechtschreib-test für vierte und fünfte klassen*. Göttingen: Hogrefe.

Bosman, A.M.T. & van Orden, G.C. (1997). Why spelling is more difficult than reading. In C.A. Perfetti, L. Rieben & M. Fayol (Eds.), *Learning to spell*, (pp. 173–194). Mahwah, NJ: Erlbaum.

Bowers, P.N., Kirby, J.R. & Deacon, S.H. (2010). The effects of morphological instruction on literacy skills: A systematic review of the literature. *Review of Educational Research*, 80(2), 144–179. https://doi.org/10.3102/0034654309359353

Burani, C., Marcolini, S. & Stella, G. (2002). How early does morpholexical reading develop in readers of a shallow orthography? *Brain and Language*, 81(1-3), 568–586. https://doi.org/10.1016/S0093-934X(02)00050-0

Burt, J.S. (2006). What is orthographic processing skill and how does it relate to word identification in reading? *Journal of Research in Reading*, 29(4), 400–417. https://doi.org/10.1111/j.1467-9817.2006.00315.x

Caravolas, M., Volín, J. & Hulme, C. (2005). Phoneme awareness is a key component of alphabetic literacy skills in consistent and inconsistent orthographies: Evidence from Czech and English children. *Journal of Experimental Child Psychology*, 92(2), 107–139. https://doi.org/10.1016/j.jecp.2005.04.003

Carlisle, J.F. (1995). Morphological awareness and early reading achievement. In L.B. Feldman (Ed.), *Morphological aspects of language processing*, (pp. 189–209). Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.

Carlisle, J.F. (2010). Effects of instruction in morphological awareness on literacy achievement: An integrative review. *Reading Research Quarterly*, 45(4), 464–487. https://doi.org/10.1598/RRQ.45.4.5

Carlisle, J.F. & Stone, C.A. (2005). Exploring the role of morphemes in word reading. *Reading Research Quarterly*, 40(4), 428–449. https://doi.org/10.1598/RRQ.40.4.3

Casalis, S., Deacon, S.H. & Pacton, S. (2011). How specific is the connection between morphological awareness and spelling? A study of French children. *Applied PsychoLinguistics*, 32(3), 499–511. https://doi.org/10.1017/S014271641100018X

Casalis, S., Quémart, P. & Duncan, L.G. (2015). How language affects children´s use of derivational morphology in visual word and pseudoword processing: Evidence from a cross-language study. *Frontiers in Psychology*, 6, 1–10. https://doi.org/10.3389/fpsyg.2015.00452

Cassar, M. & Treiman, R. (1997). The beginnings of orthographic knowledge: Children's knowledge of double letters in words. *Journal of Educational Psychology*, 89(4), 631–644. https://doi.org/10.1037/0022-0663.89.4.631

Chauhan, S. (2016). A meta-analysis of the impact of technology on learning effectiveness of elementary students. *Computers & Education*, 105, 14–30. https://doi.org/10.1016/j.compedu.2016.11.005

de Freitas, P.V., da Mota, M.M.P.E. & Deacon, S.H. (2017). Morphological awareness, word reading, and reading comprehension. *Applied PsychoLinguistics*, 39(3), 507–525. https://doi.org/10.1017/S0142716417000479

Deacon, S.H. & Kirby, J.R. (2004). Morphological awareness: Just "more phonological"? The roles of morphological and phonological awareness in reading development. *Applied PsychoLinguistics*, 25(2), 223–238 10.1017/S0124716404001117.

Deacon, S.H., Kirby, J.R. & Casselman-Bell, M. (2009). How robust is the contribution of morphological awareness to general spelling outcomes? *Reading Psychology*, 30(4), 301–318. https://doi.org/10.1080/02702710802412057

Deacon, S.H. & Leung, D. (2013). Testing statistical learning of spelling patterns by manipulating semantic and orthographic frequency. *Applied PsychoLinguistics*, 34, 1093–1108. https://doi.org/10.1017/S0142716412000173

Deacon, S.H., Parrila, R. & Kirby, J.R. (2008). A review of the evidence on morphological processing in dyslexics and poor readers: A strength or weakness? In G. Reid, A.J. Fawcett, F. Manis & L.S. Siegel (Eds.), *The Sage handbook of dyslexia*, (pp. 212–237). London: SAGE Publications.

Derwing, B.L. (1976). Morpheme recognition and the learning of rules for derivational morphology. *Canadian Journal of Linguistics*, 21(1), 38–66. https://doi.org/10.1017/S0008413100008045

Desrochers, A., Manolitsis, G., Gaudreau, P. & Georgiou, G. (2018). Early contribution of morphological awareness to literacy skills across languages varying in orthographic consistency. *Reading and Writing*, 31(8), 1695–1719. https://doi.org/10.1007/s11145-017-9772-y

Diamanti, V., Mouzaki, A., Ralli, A., Antoniou, F., Papaioannou, S. & Protopapas, A. (2017). Preschool phonological and morphological awareness as longitudinal predictors of early reading and spelling development in Greek. *Frontiers in Psychology*, 8(2039), 1–12. https://doi.org/10.3389/fpsyg.2017.02039

Ehri, L.C. (2005). Development of sight word reading: Phases and findings. In M.J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook*. Oxford: Blackwell.

Fink, A., Pucher, S., Reicher, A., Purgstaller, C. & Kargl, R. (2012). Development of a test for the assessment of morphological awareness. *Empirische Pädagogik*, 26(4), 423–451.

Galuschka, K. & Schulte-Körne, G. (2016). The diagnosis and treatment of reading and/or spelling disorders in children and adolescents. *Deutsches Aerzteblatt International*, 113, 279–286. https://doi.org/10.3238/arztebl.2016.0279

Georgiou, G.K., Parrila, R. & Papadopoulos, T.C. (2008). Predictors of word decoding and reading fluency across languages varying in orthographic consistency. *Journal of Educational Psychology*, 100(3), 566–580. https://doi.org/10.1037/0022-0663.100.3.566

Gilbert, J.K., Goodwin, A.P., Compton, D.L. & Kearns, D.M. (2013). Multisyllabic word reading as a moderator of morphological awareness and reading comprehension. *Journal of Learning Disabilities*, 47, 34–43. https://doi.org/10.1177/0022219413509966

Goodwin, A.P. & Ahn, S. (2010). A meta-analysis of morphological interventions: Effects on literacy achievement of children with literacy difficulties. *Annals of Dyslexia*, 60(2), 183–208. https://doi.org/10.1007/s11881-010-0041-x

Kargl, R., Wendtner, A., Purgstaller, C. & Fink, A. (2017). Der einfluss der morphematischen bewusstheit auf die rechtschreibleistung. *Lernen und Lernstörungen*, 7(1), 45–54. https://doi.org/10.1024/2235-0977/a000202

Kemp, N. & Bryant, P. (2003). Do beez buzz? Rule-based and frequency-based knowledge in learning to spell plural -s. *Child Development*, 74(1), 63–74. https://doi.org/10.1111/1467-8624.00521

Kirby, J.R., Deacon, S.H., Bowers, P.N., Izenberg, L., Wade-Woolley, L. & Parrila, R. (2012). Children's morphological awareness and reading ability. *Reading and Writing*, 25(2), 389–410. https://doi.org/10.1007/s11145-010-9276-5

Kirby, J.R., Georgiou, G.K., Martinussen, R. & Parrila, R. (2010). Naming speed and reading: From prediction to instruction. *Reading Research Quarterly*, 45(3), 341–362. https://doi.org/10.1598/RRQ.45.3.4

Kuo, L.-j. & Anderson, R.C. (2006). Morphological awareness and learning to read: A cross-language perspective. *Educational Psychologist*, 41(3), 161–180. https://doi.org/10.1207/s15326985ep4103_3

Landerl, K. & Reitsma, P. (2005). Phonological and morphological consistency in the acquisition of vowel duration spelling in Dutch and German. *Journal of Experimental Child Psychology*, 92(4), 322–344. https://doi.org/10.1016/j.jecp.2005.04.005

Lethonen, A. & Bryant, P. (2005). Active players or just passive bystanders? The role of morphemes in spelling development in a transparent orthography. *Applied PsychoLinguistics*, 26, 137–155. https://doi.org/10.1017/s0142716405050113

Mann, V.A. & Wimmer, H. (2002). Phoneme awareness and pathways into literacy: A comparison of German and American children. *Reading and Writing*, 15(7), 653–682. https://doi.org/10.1023/A:1020984704781

Manolitsis, G., Georgiou, G.K., Inoue, T. & Parrila, R. (2019). Are morphological awareness and literacy skills reciprocally related? Evidence from a cross-linguistic study. *Journal of Educational Psychology*, 111(8), 1362–1381. https://doi.org/10.1037/edu0000354

Marcolini, S., Traficante, D., Zoccolotti, P. & Burani, C. (2011). Word frequency modulates morpheme-based reading in poor and skilled Italian readers. *Applied PsychoLinguistics*, 32(3), 513–532. https://doi.org/10.1017/S0142716411000191

Moll, K., Fussenegger, B., Willburger, E. & Landerl, K. (2009). RAN is not a measure of orthographic processing. Evidence from the asymmetric German orthography. *Scientific Studies of Reading*, 13(1), 1–25. https://doi.org/10.1080/10888430802631684

Moll, K. & Landerl, K. (2009). Double dissociation between reading and spelling deficits. *Scientific Studies of Reading*, 13(5), 359–382. https://doi.org/10.1080/10888430903162878

Moll, K., Ramus, F., Bartling, J., Bruder, J., Kunze, S., Neuhoff, N. et al. (2014). Cognitive mechanisms underlying reading and spelling development in five European orthographies. *Learning and Instruction*, 29. https://doi.org/10.1016/j.learninstruc.2013.09.003

Nagy, W.E., Berninger, V.W., Abbott, R.D., Vaughan, K. & Vermeulen, K. (2003). Relationship of morphology and other language skills to literacy skills in at-risk second-grade readers and at-risk fourth-grade writers. *Journal of Educational Psychology*, 95(4), 730–742. https://doi.org/10.1037/0022-0663.95.4.730

Nagy, W.E., Carlisle, J.F. & Goodwin, A.P. (2013). Morphological knowledge and literacy acquisition. *Journal of Learning Disabilities*, 47(1), 3–12. https://doi.org/10.1177/0022219413509967

Nunes, T., Bryant, P. & Barros, R. (2012). The development of word recognition and its significance for comprehension and fluency. *Journal of Educational Psychology*, 104(4), 959–973. https://doi.org/10.1037/a0027412

Pacton, S., Foulin, J.N., Casalis, S. & Treiman, R. (2013). Children benefit from morphological relatedness when they learn to spell new words. *Frontiers in Psychology*, 4(696). https://doi.org/10.3389/fpsyg.2013.00696

Pacton, S., Michel, F. & Perruchet, P. (2005). Children's implicit learning of graphotactic and morphological regularities. *Child Development*, 76, 324–339. https://doi.org/10.1111/j.1467-8624.2005.00848_a.x

Pacton, S., Perruchet, P., Fayol, M. & Cleeremans, A. (2001). Implicit learning out of the lab: The case of orthographic regularities. *Journal of Experimental Psychology: General*, 130(3), 401–426. https://doi.org/10.1037/0096-3445.130.3.401

Pedhazur, E.J. & Schmelkin, L.P. (1991). *Measurement, design, and analysis: An integrated approach*. (1st edn). Hillsdale, NJ: Lawrence Erlbaum.

Perfetti, C.A. (1984). Reading acquisition and beyond: Decoding includes cognition. *American Journal of Education*, 93(1), 40–60. https://doi.org/10.1086/443785

Petermann, F. (2010). *Sprachstanderhebungstest für fünf-bis zehnjährige (SET 5-10)*. Göttingen: Hogrefe.

Quémart, P., Casalis, S. & Duncan, L.G. (2012). Exploring the role of bases and suffixes when reading familiar and unfamiliar words: Evidence from French young readers. *Scientific Studies of Reading*, 15(5), 424–442. https://doi.org/10.1080/10888438.2011.584333

Reed, D.K. (2008). A synthesis of morphology interventions and effects on Reading outcomes for students in grades K–12. *Learning Disabilities Research & Practice*, 23(1), 36–49. https://doi.org/10.1111/j.1540-5826.2007.00261.x

Rispens, J.E., McBride-Chang, C. & Reitsma, P. (2008). Morphological awareness and early and advanced word recognition and spelling in Dutch. *Reading and Writing*, 21(6), 587–607. https://doi.org/10.1007/s11145-007-9077-7

Roman, A.A., Kirby, J.R., Parrila, R., Wade-Woolley, L. & Deacon, S.H. (2009). Toward a comprehensive view of the skills involved in word reading in grades 4, 6, and 8. *Journal of Experimental Child Psychology*, 102, 96–113. https://doi.org/10.1016/j.jecp.2008.01.004

Rothe, J., Schulte-Körne, G. & Ise, E. (2013). Does sensitivity to orthographic regularities influence reading and spelling acquisition? A 1-year prospective study. *Reading and Writing*, 27, 1141–1161. https://doi.org/10.1007/s11145-013-9479-7

Schneider, W., Blanke, I., Faust, V. & Küspert, P. (2011). *Würzburger leise leseprobe – revision (WLLP-R)*. Göttingen: Hogrefe.

Seymour, P.H., Aro, M. & Erskine, J.M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, 94(2), 143–174. https://doi.org/10.1348/000712603321661859

Shahar-Yames, D. & Share, D. (2008). Spelling as a self-teaching mechanism in orthographic learning. *Journal of Research in Reading*, 31(1), 22–39. https://doi.org/10.1111/j.1467-9817.2007.00359.x

Sparks, E. & Deacon, S.H. (2015). Morphological awareness and vocabulary acquisition: A longitudinal examination of their relationship in English-speaking children. *Applied PsychoLinguistics*, 36(2), 299–321. https://doi.org/10.1017/S0142716413000246

Tabachnick, B.G. & Fidell, L.S. (2013). *Using multivariate statistics*. (6th edn). New Jersey: Pearson.

Treiman, R. (2018). Teaching and learning spelling. *Child Development Perspectives*, 12(4), 1–5. https://doi.org/10.1111/cdep.12292

Treiman, R., Cassar, M. & Zukowski, A. (1994). What types of linguistic information do children use in spelling? The case of flaps. *Child Development*, 65(5), 1318–1337. https://doi.org/10.1111/j.1467-8624.1994.tb00819.x

UNESCO. (2020). COVID-19 impact on education: Global monitoring of school closures by COVID-19 [cited 2020 Oct 04]. Retrieved from https://en.unesco.org/covid19/educationresponse/#

Vernice, M. & Pagliarini, E. (2018). Is morphological awareness a relevant predictor of reading fluency and comprehension? New evidence from Italian monolingual and Arabic–Italian bilingual children. *Frontiers in communication*, 3(11), 1–15. https://doi.org/10.3389/fcomm.2018.00011

Volkmer, S., Schulte-Körne, G. & Galuschka, K. (2019). Die rolle der morphologischen bewusstheit bei lese-und rechtschreibleistungen. *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie*, 47(4), 334–344. https://doi.org/10.1024/1422-4917/a000652

Weiß, R.H. (2006). *Grundintelligenztest skala 2-revision CFT 20-R [culture fair intelligence test scale 2-revision]*. Göttingen: Hogrefe.

Winkler, W.E. (1990). *String comparator metrics and enhanced decision rules in the Fellegi–Sunter model of record linkage*, (pp. 354–359). American Statistical Association: Proceedings of the Section on Survey Research Methods.

Ziegler, J.C., Stone, G.O. & Jacobs, A.M. (1997). What is the pronunciation for –ough and the spelling for /u/? A database for computing feedforward and feedback consistency in English. *Behavior Research Methods, Instruments, & Computers*, 29(4), 600–618. https://doi.org/10.3758/BF03210615

Ruth Görgen has an MA in speech and language therapy and works as a research assistant at the Mercator Institute for Literacy and Language Education of the University of Cologne and is a PhD

student at the Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy of the University Hospital of Munich. She is interested in typical and atypical oral and written language acquisition including diagnostics and therapy as well as game-based learning.

Elisabetta De Simone is a research assistant with an MA in linguistics. She belongs to the Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy of the University Hospital of Munich. She is interested in dyslexia, orthographic depth and morphological processing.

Gerd Schulte-Körne is the director of the Clinic of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy of the University Hospital of Munich and holds the chair in the Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy of the University Hospital of Munich. His research interests are genetic and neurobiological processes underlying learning disorders, depression and eating disorders as well as the development and evaluation of intervention programmes.

Kristina Moll has a PhD in psychology and works as a researcher at the Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy of the University Hospital of Munich. Her research interests lie in the field of typical and atypical development of reading, spelling and arithmetic, comorbidity between developmental disorders and cross-linguistic research in literacy development.

Received 4 December 2019; revised version received 24 October 2020.

Address for correspondence: Ruth Görgen, Department Language and Education System, Mercator Institute for Literacy and Language Education, University of Cologne, Albertus-Magnus-Platz, 50923 Cologne, Germany. E-mail: ruth.goergen@mercator.uni-koeln.de

Supplementary material

Table 1

Phoneme awareness task items.

Item no	Response real word alternatives	Pseudoword
1	ˈ*kɛlɐ* - ˈtɛlɐ - ˈkɛlbɐ	ˈkɛl(z)ɐ
2	ˈvɛlpə - ˈ*vɛlə* - ˈbɛlə	ˈvɛl(f)ə
3	ˈmʊstɐ - ˈbʊtɐ - ˈ*mʊtɐ*	ˈmʊ(k)tɐ
4	ˈ*vanə* -ˈkanə -ˈvant͡sə	ˈvan(t͡)ə
5	blaʊ̯ - zaʊ̯ - *baʊ̯*	b(ʁ)aʊ̯
6	ˈvapn̩ - ˈ*lapn̩* - ˈlampn̩	ˈla(f)pn̩
7	ˈbalkn̩ - ˈt͡sakn̩ - ˈ*bakn̩*	ˈba(m)kn̩
8	ʃlaːf - gʁaːf - *ʃaːf*	ʃ(n̩)aːf
9	ˈkʁalə - ˈ*falə* - ˈfaltə	ˈfal(b)ə
10	gluːt - huːt - *guːt*	g(ʁ)uːt
11	ˈ*vɔlə* - ˈvɔlkə - ˈʁɔlə	ˈvɔl(z)ə
12	ˈ*kaʁən* - ˈkaʁtən - ˈbaʁən	ˈkaʁ(f)ən
13	ˈp͡fʏt͡sə - ˈ*mʏt͡sə* - ˈmʏnt͡sə	ˈmʏ(ʁ)t͡sə
14	aɪ̯ns - maɪ̯s - *aɪ̯s*	aɪ̯(p)s
15	ˈtastə - ˈ*tasə* - ˈklasə	ˈtas(k)ə
16	bʁoːt - ʁoːt - *boːt*	b(l)oːt
17	vant – *bant* - bʁant	b(l)ant
18	ˈbɪsn̩ - ˈ*kɪsn̩* - ˈkɪstən	ˈkɪs(p)n
19	ˈ*kʏsə* - ˈkʏstə - ˈnʏsə	ˈkʏs(p)ə
20	ˈzʊpə - ˈpʊmpə - ˈ*pʊpə*	ˈpʊ(s)pə
21	ˈtanə - ˈ*kanə* - ˈkantə	ˈkan(f)ə

Note: The correct answer (on which scores are based) is in italics.
Section in brackets is the sound that the child was asked to delete.

Table 2

Morphological awareness task items.

Item no	Plural formation task		
	Pseudoword singular Das ist ein … (This is a ...)	Response alternatives plural Das sind zwei … (These are two …)	Conventions for solving the tasks
1	Zante	*Zanten* – Zants – Zante	Noun ending on “schwa-e” → regular plural -n; both -s/zero allomorph no regular option
2	Tofa	Tofen – *Tofas* – Tofer	Noun ending in a full vowel → regular plural -s; -en → competing pluralform; both -en/-er less likely
3	Miegel	Miegelen – *Miegel* – Miegeler	Noun ending with -el → regular plural form zero allomorph (or -n); both -en/er no regular option
4	Roto	Roto – Roton – *Rotos*	Noun ending in a full vowel → regular plural -s; both -n/zero allomorph no regular option

5	Lerd		*Lerde* – Lerden – Lerd	Noun ending with a consonant (except for -ung, -en, -el, -er) → regular plural -e; both -en/zero allomorph no regular option
6	Tipör		Tipör – *Tipöre* – Tipöres	Noun ending with a consonant (except for -ung, -en, -el, -er) → regular plural -e; both -en/zero allomorph no regular option
7	Nengel		Nengeler – Nengelen – *Nengel*	Noun ending with -el → regular plural form zero allomorph (or -n); both -en/er no regular option
	⇨ 2*plural -s, 1*plural –(e)n, 2*plural -e, 2*zero allomorph			

	Verb inflection task			
	Verb infinitive	Sentence frame	Response alternatives	Conventions for solving the tasks
8	dannen	Ben …	*dannt* – dannen – dannst	Weak verbs in simple present 3. pers. sg. → legal ending -t; both st-/en no legal endings for 3. pers. sg. (context information)
9	felfen	Gestern … er. (Yesterday, he … .)	felfest – *falf* – falfen	Strong verbs in simple past 3. pers. sg. → vowel change in stem + deletion of the ending -en; both st-/en no legal endings for 3. pers. sg. (context information)
10	wennen	Gestern … du. (Yesterday, you … .)	wannten – wennte – *wanntest*	Irregular verbs in simple past 2. pers. sg. → vowel change in stem + ending -test; both -ten/-te no legal endings for 2. pers. sg. (context information)
11	tallen	Du … (You …)	*tällst* – tallt – tällen	Strong verbs in simple present 2. pers. sg. → vowel change in stem + ending -st; both -t/-en no legal endings for 2. pers. sg. (context information)
12	glafen	Lisa …	gläfen – glafst – *gläft*	Strong verbs in simple present 3. pers. sg. → vowel change in stem + ending -t; both -st/-en no legal endings for 3. pers. sg. (context information)
13	rielen	Gestern … du. (Yesterday, you … .)	rielten – rielte – *rieltest*	Weak verbs in simple past 2. pers. sg. → legal ending -test; both -ten/-te no legal endings for 2. pers. sg. (context information)
14	negen	Gestern … er. (Yesterday, he … .)	negtest – *negte* – negten	Weak verbs in simple past 3. pers. sg. → legal ending -te; both -test/-ten no legal endings for 3. pers. sg. (context information)
15	tochen	Du … (You …)	*tochst* – tocht – tochen	Weak verbs in simple present 2. pers. sg. → legal ending -st ;

				both -t/-en no legal endings for 2. pers. sg. (context information)
	⇨ 2*weak simple present (2./3. pers. sg.), 2*weak simple past (2./3. pers. sg.), 2*strong simple present (2./3. pers. sg.), 1*strong simple past (3. pers.), 1*irregular simple past (2.pers. sg.)			
	Derivation			
	Initial sentence	Adapted sentence frame	Response alternatives	Conventions for solving the tasks
16	Leo kann gut pongen. (Leo is good at pongen.)	Leo ist ein guter … (Leo is a good … .)	Pong – *Ponger* – Pongung.	Ending -er most probable ending in agent nouns (context information)
17	Das sind Beuse. (These are Beuse.)	Sie sind sehr … (They are very …)	beuser – beusel – *beuslich*	Noun to adjective → legal option -lich; both -er/-el no legal options
18	Ich bin ein Gauler. (I am a Gauler.)	Heute habe ich … (Today I …)	*gegault* – gaulen – gegaulst	Verbs in present participle → no stem change, adding prefix ge- and ending -t; both infinitive -en/ prefix ge- + ending -st no legal options
19	Ich bin kein guter Geker. (I am not a good Geker).	Ich kann nicht gut … (I am not good at …)	gegekst – *geken* – gegekt	Verbs in infinitive → ending -en; adding prefix ge- or endings -st/-t
20	Wir kresen. (We kresen.)	Das war eine tolle … (That was a great …)	*Kreserei* – Kresel – Krest	Ending -erei most probable → el/-t less likely → ending -el usually for indicating diminutive form/ ending -t unproductive suffix only used in fossilized word-forms
21	Papa kann schnell nulfen. (Dad is fast at nulfen.)	Papa ist ein schneller … (Dad is a fast …)	*Nulfer* – Nulfung – Nulf	Ending -er most probable ending in agent nouns (context information)
22	Dort gibt es viele Lahne. (There are many Lahne there.)	Es ist sehr … (It is very …)	lahnel – *lahnig* – lahner	Noun to adjective → legal option -ig/-lich; both -er/-el no legal options
	⇨ 2*V->N with suffix -er, 1*N->V in present participle, 1*N->V in infinitive, 1*N->A with suffix -lich, 1*N->A with suffix -ig			

Note: The correct answer (on which scores are based) is in italics.
Section in brackets is the English translation.

Table 3

Sublexical orthographic sensitivity task items.

Item no	Item	Item no	Item	Item no	Item
1	uule	2	*solbst*	3	*Abst*
4	Fääne	5	lezz	6	*sucken*
7	Muzz	8	*kecke*	9	Wuuler

10	göölen	11	*rels*	12	*Teck*
13	Flezze	14	Züüme	15	*gielen*
16	*Tiele*	17	süüren	18	*käck*
19	*Krals*	20	*Sprack*	21	*derst*
22	trizz	23	*Ricker*	24	Kööne
25	kruuze	26	fäämen	27	*Sien*
28	Prezz	29	fruul	30	*Fonste*
31	*krutze*	32	Tekk	33	Tiile
34	*letz*	35	Fonßte	36	*uhle*
37	*sühren*	38	derßt	39	giilen
40	*Köhne*	41	kekke	42	Sprakk
43	*göhlen*	44	*fähmen*	45	*Fähne*
46	relß	47	*Mutz*	48	*Tritz*
49	Kralß	50	Abßt	51	*Fletze*
52	*Wuhler*	53	*Pretz*	54	sukken
55	käkk	56	solbßt	57	Siin
58	*Zühme*	59	Rikker	60	*fruhl*

Note: Pseudowords with permissible letter patterns are in italics.

Table 4

Three step regression analyses for non-verbal cognitive abilities, age and specific predictors with reading as dependent variable.

		Reading			
		R2-change %	*p*	*B*	*SE B*
Step 1	Non-verbal cognitive abilities	3.0	<.001	.174	.018
Step 2	Age	0.0	.731	.006	.000
Step 3	Unique sublexical orthographic sensitivity	0.0	.456	-.013	.018
	Unique phoneme awareness	2.3	<.001	.171	.019
	Unique morphological awareness	4.3	<.001	.233	.019
	Shared variance step 3	3.9			
	Total variance step 3	10.5			
	Total variance step 1-3	13.5			

Note: Combined analysis for 3rd and 4th grade was performed with grade-specific z-scores (based on raw scores) after logarithmic data transformation

Table 5

Three step regression analyses for non-verbal cognitive abilities, age and specific predictors with spelling as dependent variable.

		Spelling			
		R2-change %	*p*	*B*	*SE B*
Step 1	Non-verbal cognitive abilities	9.2	<.001	.303	.017
Step 2	Age	0.1	.030	-.037	.000
Step 3	Unique sublexical orthographic sensitivity	0.1	.010	.040	.016
	Unique phoneme awareness	6.4	<.001	.287	.017
	Unique morphological awareness	6.7	<.001	.290	.017
	Shared variance step 3	9.2			
	Total variance step 3	22.4			
	Total variance step 1-3	31.7			

Note: Combined analysis for 3rd and 4th grade was performed with grade-specific z-scores (based on raw scores).

FACHARTIKEL 3: Evaluation of a digital game-based reading training for German children with reading disorder

Literaturangabe: Görgen, R., Huemer, S., Schulte-Körne, G., Moll, K. (2020). Evaluation of a digital game-based reading training for German children with reading disorder. *Computers & Education, 150,*1-14. doi: 10.1016/j.compedu.2020.103834

Angenommen am 03.02.2020
Online publiziert am 04.02.2020

Computers & Education 150 (2020) 103834

Contents lists available at ScienceDirect

Computers & Education

journal homepage: http://www.elsevier.com/locate/compedu

Evaluation of a digital game-based reading training for German children with reading disorder

Ruth Görgen *, Sini Huemer, Gerd Schulte-Körne, Kristina Moll

Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, University Hospital, LMU Munich, Nußbaumstraße 5a, 80336, Munich, Germany

A R T I C L E I N F O

Keywords:
Elementary education
Evaluation methodologies
Interactive learning environments

A B S T R A C T

Technological tools have the potential to efficiently support learning performance in children and are therefore recognized as being beneficial for children with special needs, such as reading disorders. The present study investigated the effectivity of a novel digital game-based reading training in second- and third-grade children with reading disorder ($N = 50$) carried out in home environment over a short training period of eight to eleven weeks. Children were randomly assigned to either a training group ($N = 25$) receiving the digital game-based reading training or a control group performing digitalized mind games ($N = 25$). Results showed significant improvement of reading performance for trained word material and a trend for transfer effects to untrained words. The digital game-based reading training was also found to be highly motivating and feasible for use in home environment. This study extends existing knowledge about digital game-based reading support and underlines that using a multi-component reading approach combined with flow principles is beneficial for children with reading disorder. Furthermore, the evaluated program seems to be a promising addition for children who do not have access to individual reading support.

1. Introduction

Acquiring reading skills is essential for participation in modern society. However, around 4–8% of children do not master the challenge of learning to read adequately and are diagnosed with a reading disorder (Fischbach et al., 2013; Landerl & Moll, 2010), typically characterized by deficits in reading accuracy, fluency and/or comprehension (American Psychiatric Association, 2013). These deficits are not due to inadequate schooling, neurological, visual or auditory impairment (American Psychiatric Association, 2013). Without adequate diagnosis and treatment, reading deficits have a negative impact on academic careers (Wyschkon et al., 2017) as well as on personal well-being (Bäcker & Neuhäuser, 2003; Goldston et al., 2007). It is therefore important to develop efficient treatment approaches for children with reading disorder (Jamshidifarsani, Garbaya, Lim, Blazevic, & Ritchie, 2019). As attention to innovative learning support through technological tools is growing (Chauhan, 2016), it has also been recognized that technological tools have enormous potential for the treatment of children with special needs, such as reading difficulties (Jamshidifarsani et al., 2019). Whereas much is known about the effective treatment components in standard classroom or individual reading instruction, more research is needed in order to evaluate the impact of digital reading interventions. This study extends current

* Corresponding author.
E-mail addresses: ruth.goergen@med.uni-muenchen.de (R. Görgen), sini.huemer@med.uni-muenchen.de (S. Huemer), gerd.schulte-koerne@med.uni-muenchen.de (G. Schulte-Körne), kristina.moll@med.uni-muenchen.de (K. Moll).

https://doi.org/10.1016/j.compedu.2020.103834
Received 6 September 2019; Received in revised form 13 January 2020; Accepted 3 February 2020
Available online 4 February 2020
0360-1315/

knowledge about technology-based treatment of reading disorders by evaluating a novel, innovative reading training program combining evidence-based components of standard reading instruction with recent digital game-based learning findings.

1.1. Components of effective reading instruction

Skilled reading requires the development of efficient (accurate and fast) word identification, which represents the basic process underlying text comprehension (Vellutino, Fletcher, Snowling, & Scanlon, 2004). Therefore, the developed reading training used in the current study focuses on enhancing accurate and fast word reading skills.

Meta-analyses (Galuschka et al., 2020; Galuschka, Ise, Krick, & Schulte-Körne, 2014; Ise, Engel, & Schulte-Körne, 2012) underline that effective treatment approaches for supporting children with reading disorder are symptom-oriented and directly targeted to reading skills and proximal cognitive skills directly associated with reading. Cognitive deficits associated with problems in word identification are difficulties in rapid automatized naming as well as in phonological and morphological processing. Although rapid automatized naming (i.e. naming a matrix of repeated letters, digits, objects or colors) is a reliable predictor of reading skills (Araújo & Faísca, 2019), it seems hard to improve and does not generalize to word reading skills (de Jong & Vrielink, 2004). In contrast, training phonological awareness (i.e. the ability to recognize and manipulate smaller units of spoken words such as syllables, onset and rimes as well as phonemes) has been shown to be highly effective for improving reading skills as it contributes to the child's understanding of the relationship between spoken and written language units (Landerl et al., 2013; National Institute of Child Health and Human Development, 2000; Snowling, 2014; Vellutino et al., 2004). However, the effectiveness of phonological awareness training on reading also depends on the language and is lower in transparent orthographies (such as German), characterized by high grapheme-phoneme consistency, compared to less transparent orthographies characterized by low grapheme-phoneme consistency such as English (Georgiou, Parrila, & Papadopoulos, 2008; Mann & Wimmer, 2002). In transparent orthographies phonological awareness trainings alone should only be provided to children before or at the very beginning of literacy acquisition (Galuschka et al., 2014; Wolf, Schroeders, & Kriegbaum, 2016). In school-aged children phonological training is most effective when combined with symptom-oriented reading training, known as *systematic phonics instruction* (Galuschka et al., 2014; National Institute of Child Health and Human Development, 2000). Systematic phonics instruction encourages children's insight in the alphabetic principle by combining phonological awareness training with the teaching of grapheme-phoneme correspondence, which is the basis for reading larger written language units such as syllables or words (National Institute of Child Health and Human Development, 2000).

Although word decoding (i.e. mapping graphemes to their corresponding phonemes) gets more automatized by repetition (Share, 1995), decoding strategies are not sufficient to become an efficient and fluent reader as it mainly contributes to reading accuracy (Thaler, Ebner, Wimmer, & Landerl, 2004). In order to be able to read more complex words, word-specific orthographic representations of whole words and of larger sublexical units such as morphemes (Thaler et al., 2004) need to be build-up and stored in long-term memory (Ehri, 1995; Mann & Singson, 2003). Morphemes represent the smallest meaningful language units and carry phonologic, semantic and syntactical information. A large body of research (Bowers, Kirby, & Deacon, 2010; De Freitas, Da Mota & Deacon, 2017; Fink, Pucher, Reicher, Purgstaller, & Kargl, 2012; Goodwin & Ahn, 2010; Kirby et al., 2012; Kirk & Gillon, 2009; Mann & Singson, 2003; Reed, 2008) has shown positive effects of morphological instruction (e.g. splitting long words such as *excitement* into the consisting morphemes: *excite – ment*) (Mann & Singson, 2003)). It is assumed that building mental representations of base words and affixes, speeds the identification of words and facilitates the recognition of new words (Carlisle & Stone, 2005).

Thus, reading practice by repeated exposure to words and morphemes has been the most commonly used and evaluated method to enhance reading speed (Berends & Reitsma, 2006; Lemoine, Levy, & Hutchinson, 1993; Levy, Bourassa, & Horn, 1999; Thaler et al., 2004; Van Gorp, Segers, & Verhoeven, 2017) and is also recommended by the National Institute of Child Health and Human Development (2000). This approach seems especially relevant in transparent orthographies such as German, where a deficit in reading speed is the main diagnostic feature for reading disorder because reading accuracy is close to ceiling after one year of reading instruction (Aro & Wimmer, 2003; Landerl & Wimmer, 2008; Seymour, Aro, & Erskine, 2003).

In sum, reading approaches should foster the build-up of word-specific orthographic representations by taking advantage of *phonological* and *morphological instruction* components combined with *repeated word reading activities*.

1.2. The potential of technological tools to improve learning

Technological tools can impart knowledge or specific skills and are at the same time highly motivating to children because they are learning in a playful and engaging way (Chauhan, 2016; Susi, Johannesson, & Backlund, 2007; Wouters, van Nimwegen, van Oostendorp, & van der Spek, 2013; Zheng & Spires, 2014). The use of advanced technological tools in therapeutic or educational contexts is also known as *edutainment*, *digital game - based learning*, or *serious games* (Susi et al., 2007) meaning the integration of motivational play features with an educational aim. Meta-analyses (Chauhan, 2016; Chung, Hwang, & Lai, 2018; Jamshidifarsani et al., 2019; Sung, Chang, & Liu, 2015; C. Williams & Beam, 2019) indicate that the use of technology can be a powerful tool for effective learning. It was found that a digital game-based learning approach is more effective in gaining knowledge (Papastergiou, 2009) and more motivational as children show greater willingness to learn compared to children participating in a non-digital approach (Wrzesien & Alcañiz Raya, 2010). An important aspect which contributes to learning success is to ensure that the game-based content does not distract the learner from the learning objectives (Zheng & Spires, 2014). Features such as a reminder to focus attention on the task can help self-regulation and prevent learners from going off-task (Zheng & Spires, 2014) and in this way facilitate learning processes (Chauhan, 2016). This is in line with earlier findings indicating that high levels of engagement and motivation result in better learning performance (Kuh, 2009; Lepper, Henderlong Corpus, & Iyengar, 2005), while reduced engagement results in reduced level of active processing (Ke & Abras, 2013). Following the *Flow theory* of Csikszentmihalyi (1975), flow describes an ideal experience of deep engagement when fully immersed in an activity (Zheng & Spires, 2014). In their proposed *GameFlow Model*, Sweetser and Wyeth (2005) suggest the following principles that optimize engagement in digital game-based learning: *Concentration, challenge, player skills, control, clear goals, feedback, immersion* and *social interaction.* Among those principles, feedback is one of the most powerful factors for effective learning because it enables the learner to monitor his/her own learning progress (Hattie & Timperley, 2007). Feedback can be given immediately or with the help of records that can be included in a rewarding system (Wouters et al., 2013), which is found to enhance flow experiences by helping the learner to feel pleased and joyful (Kiili, de Freitag, Arnab, & Lainema, 2012). Constant rewarding integrated in a fantasy game-world seems to be most appealing to children with special learning needs (Ke & Abras, 2013). Another relevant principle is the role of challenge-skill balances (Chang, Liang, Chou, & Lin, 2017; Kiili et al., 2012; Sweetser & Wyeth, 2005; Wouters et al., 2013) indicating that tasks should neither be too easy nor too difficult so that the learner doesn't get bored or frustrated (Kiili et al., 2012). Ideally, games should adapt the difficulty level to the learner's performance (Sweetser & Wyeth, 2005). According to Maertens, Vandewaeterea, Cornillie, and Desmet (2014), adaptive item sequencing is given if a) the tasks get easier when the learner repeatedly fails and b) the tasks get more difficult after the learner successfully completed a task. Taking these elements into account, high flow experiences and successful learning can be achieved (Admiraal, Huizenga, Akkerman, & ten Dam, 2011).

1.3. Research findings on digital game-based reading approaches

Previous studies on the effectivity of technology-based literacy interventions emphasize that reading approaches should be individualized and adaptive (Blok, Oostdam, Otter, & Overmaat, 2002; Jamshidifarsani et al., 2019; MacArthur, Ferretti, Okolo, & Cavalier, 2001; Perelmutter, McGregor, & Gordon, 2017; Santoro & Bishop, 2010). This can be enabled through advanced technologies such as smartphones and tablet computers, which has been largely neglected so far (Jamshidifarsani et al., 2019). Jamshidifarsani et al. (2019) noted that less than one-fourth of the reviewed studies used innovative game-based reading programs. Instead, computer-assisted reading training with less innovative features (e.g. adaptivity) is still predominantly used. The potential to support children with reading disorder who don't have access to special reading support is therefore not used sufficiently and more studies evaluating the effectivity of innovative digital game-based reading approaches are needed (Jamshidifarsani et al., 2019). Of the reviewed studies by Jamshidifarsani et al. (2019) most of the digitalized reading approaches target phonological awareness skills or phonics only (see section 1.1). Jamshidifarsani et al. (2019) recommend developing and evaluating more digital game-based reading approaches that address the enhancement of reading fluency and draw attention to the implementation of more holistic approaches.

Given that children with reading difficulties often avoid reading activities (Morgan, Fuchs, Compton, Cordray, & Fuchs, 2008), motivational components are likely to be highly relevant for acceptance of the reading intervention (Ronimus, Eklund, Pesu, & Lyytinen, 2019). However, evaluating motivational components in reading treatment approaches is mostly overlooked (Jamshidifarsani et al., 2019; Ronimus et al., 2019). Consistent with the proposed GameFlow Model by Sweetser and Wyeth (2005), appealing reading approaches for children with reading disorder need to comprise *challenges*, *rewards* and *visible cues of progress*, which need to be integrated rather than just added to the learning content (Jamshidifarsani et al., 2019; Ke & Abras, 2013; Ronimus, Kujala, Tolvanen, & Lyytinen, 2014).

1.4. Aims of the current study

Taking the evidence of standard reading instruction and technology-based learning into account, a multi-component digital game-based training program focusing on phonological awareness, grapheme-phoneme mapping, and morphological instruction combined with repeated word reading activities was developed. Importantly, the specific reading content was integrated in a game-based learning environment that incorporates relevant design principles, namely immediate feedback and rewarding to induce high flow experiences, and challenge-skill balance through adaptivity.

In the current study the effectivity of this novel digital game-based reading training program was assessed. Due to its focus on word reading skills (accuracy and speed), we expected that children improve word reading performance for trained words as a result of the structured training sessions and of the high repetition rate of word material. As transfer effects of reading approaches have not yet been sufficiently investigated (K. J. Williams, Walker, Vaughn, & Wanzek, 2017), we also analyzed transfer effects on untrained words. Besides word reading accuracy and speed, the training also targets phonological abilities and grapheme-phoneme mapping. Therefore, we examined the effects of the training on these skills. In addition to the word reading outcome, we assessed the effect of the training on related literacy skills, including reading comprehension and spelling skills.

Little is known about the implementation of digital game-based reading trainings for children with reading disorder as previous research mainly focused on standard training methods in schools or therapeutic practices mostly guided by a professional. Therefore, we assessed the feasibility of implementing an adaptive digital game-based reading training that is conducted independently in a home environment. Furthermore, we evaluated whether the children's attitude towards reading is influenced by the training as a consequence of the generated flow experiences.

1.5. Research questions

To specify, the following research questions were addressed:

RQ1: Does participation in a digital game-based reading training result in significant improvement of word reading skills for (a) trained words and (b) untrained words (transfer effect)?
RQ2: Do precursor skills (phonological awareness skills and grapheme-phoneme knowledge) benefit from the training?
RQ3: Does the reading training has a positive impact on untrained skills (reading comprehension and spelling) which are related to word reading skills?
RQ4: Is the digital game-based reading training feasible in home environment and does it influence the child's attitude towards reading?

2. Method

2.1. Description of the digital game-based reading training

Meister Cody-Namagi is an adaptive multi-component digital game-based reading training developed by the software company Meister Cody GmbH and the dyslexia research group at the Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, University Hospital LMU Munich, Germany, targeting phonological awareness, phoneme-grapheme correspondence and word reading skills in primary school children.

Modules and tasks: The program consists of three different modules (phonological awareness, phoneme-grapheme mapping and word reading) with in total 22 tasks. The first module, phonological awareness, focuses on exercises to improve basic phonological skills (e.g. segmenting and counting syllables, phoneme identification, blending and segmenting phonemes and differentiation between long and short vowels). The second module is targeting phoneme-grapheme and grapheme-phoneme mappings. The third module consists of exercises for reading accuracy (e.g. lexical decision and word building tasks) as well as reading speed (e.g. semantic mapping of words and word-picture mapping within limited time). Examples of tasks of each module are presented in Fig. 1. Supplement 1 provides an overview of all tasks.

Training schedule: The training comprised 30 training sessions, with each session consisting of three pre-defined exercises. In the first half of the training schedule, the three exercises comprised one exercise from each module (phonological awareness, phoneme-grapheme correspondence, and word reading accuracy). The second half of the training schedule focused stronger on word reading activities and therefore each session consisted of one exercise from either module one or two and two exercises from the word reading module, including both reading accuracy and reading speed exercises. In total, the training schedule comprised 17 phonological awareness and 28 phoneme-grapheme correspondence tasks as well as 30 word reading accuracy and 15 word reading speed tasks. The training schedule is illustrated in Fig. 2.

Adaptivity: A special feature of the program is its adaptivity: All tasks have different levels of difficulty, which automatically adapt to the child's abilities. This keeps the tasks challenging enough while maintaining sufficient success rate as suggested by the flow

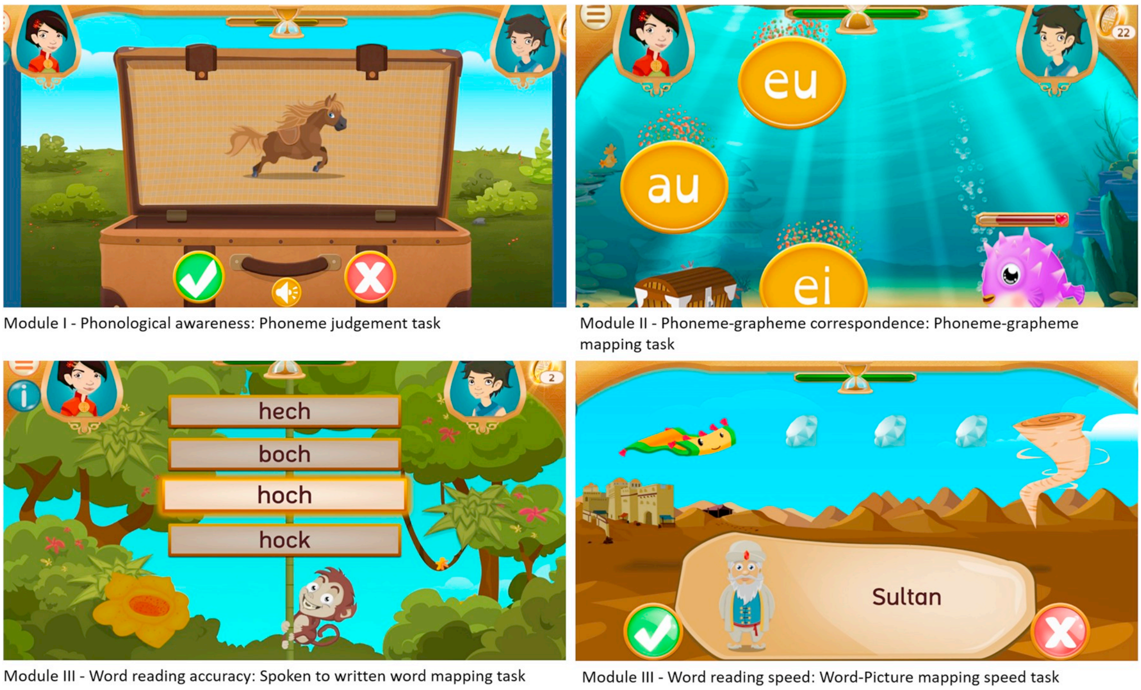


Fig. 1. Examples of tasks of the three modules phonological awareness, phoneme-grapheme correspondence and word reading (accuracy and speed).

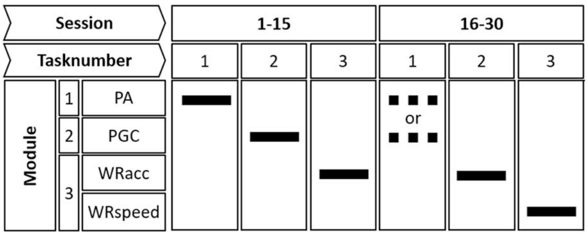


Fig. 2. Scheme of the training schedule (PA = Phonological awareness, PGC = Phoneme-grapheme correspondence, WRacc = Word reading accuracy, WRspeed = Word reading speed).

theory (Kiili et al., 2012; Sweetser & Wyeth, 2005). In order to move on to the next difficulty level, children have to score on average more than 80% correct (range 50–90% depending on task demands) in one item block. When scoring correct less than 55% on average (range 32–70% depending on task demands) in one item block, the difficulty level decreases. Otherwise the difficulty level remains the same. The number of difficulty levels per task is listed in Supplement 1. The number of items in one block ranges between four and 20 items, varying between the different tasks. The complexity of the items was determined by grapheme-phoneme structure and word length following the structuring method for wordmaterial applied in two evaluated German treatment approaches *("Lautgetreue Lese-Rechtschreibförderung"* (Reuter-Liehr, 2000, 2006)*; "Kieler Leseaufbau"* (Dummer-Smoch & Hackethal, 2007). The word material was selected using the corpus childLex (Schroeder, Würzner, Heister, Geyken, & Kliegl, 2015) which provides a database representing the core vocabulary of German school children based on German children's books. The minimum normalized word frequency (i.e. absolute frequency/(sum (absolute frequency)/1000000) for selected words was >10 in modules one and two, and >1 in module three.

Training narrative: The training is embedded in a fictional story around the magician *Meister Cody*, Princess *Namea* and Prince *Fandales* traveling through the land of *Talasia*. On their travels, the children meet a variety of other characters. By performing magical exercises, the children improve their magic language *Namagi* and solve the mystery of the dragon *Sordan*. As the magician Meister Cody explains the exercises using his magical crystal ball, the training does not require an external instructor, so that children can do the training on their own.

Feedback and Reminder: The children receive immediate feedback regarding their choice after each item in order to let the children monitor their own learning process. Immediate feedback is given by notifying the children whether the choice is correct or not. In some tasks, the children receive corrective feedback or the correct answer is presented in case of an incorrect choice. When repeatedly failing on a task, the explanation of the task is repeated together with an example. When there is no reaction for a certain amount of time an automatic reminder alerts the children to continue with the exercise to avoid children going off-task.

Reward system: The children are given regular positive reinforcement during all exercises by collecting magical power and gold coins. The children are further rewarded for each successfully completed task. In an additional game, the gold coins can be used to create a magical garden.

2.2. Participants

Participants from grade two and three were referred by therapeutic practices (psychologists and psychiatrists, speech- and language therapists, occupational therapists and learning therapists) as well as afternoon care centres, who agreed to inform families who were concerned about the literacy skills of their child. Interested families who contacted the study team were informed in detail about the procedures and the aims of the study and gave their written consent for participation. Children received vouchers as a reward for taking part in the assessments. The study was approved by the institutional review board of the local ethics committee and was performed in accordance with the latest version of the Declaration of Helsinki and in compliance with national legislation.

Participants were selected based on their reading fluency performance, which had to be below average (≤ 16th percentile rank in at least one of the two subtests and ≤ 30th percentile rank in the other subtest) measured with a standardized word reading test (SLRT-II; Moll & Landerl, 2014). Based on these results, children's reading disorder was categorized into mild/moderate (reading fluency performance ≥ 8th percentile rank) or severe (reading fluency performance < 8th percentile rank) to control for severity level during group allocation. Further selection criteria were: Nonverbal IQ ≥ 85 (CFT-1R/CFT-20R; Weiβ, 2006), not yet or less than 4 months reading and spelling therapy received, German as the first language, normal or corrected to-normal vision, absence of neurological deficits. Children who did not fulfill the criteria were excluded but were still given the opportunity to complete the training without participating in the study. Fifty-six of the 76 children examined met all inclusion criteria and initially started participation in the study. Attention-Deficit-(Hyperactivity)-Disorder (AD(H)D), which frequently co-occurs with reading disorder was not an exclusion criterion. However, symptoms of attention problems, hyperactivity and impulsivity were assessed based on a standardized parental questionnaire (DISYPS-II; Döpfner et al., 2008; see section 2.4). Children fulfilling inclusion criteria were then allocated alternately to the training group (TG) or the control group (CG) in a pseudorandom manner: Severity of reading disorder (mild/moderate or severe, as described above) and grade level were considered for the group assignment. Of the participating 56 children, two children dropped out (due to lack of motivation) and four children did not meet the training criterion of at least 25 scheduled sessions within eleven weeks (see section 2.5.1). Thus, the final sample consisted of 50 children with reading disorder, with 25 children each in the TG and CG.

Detailed characteristics of the groups are presented in Table 1.

2.3. Procedure

The children were tested individually in a silent room at the Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, University Hospital LMU Munich, Germany. The assessments were administered in a fixed order and lasted between 60 and 90 min. They were carried out by six advanced students in psychology or pedagogy as well as three experienced research assistants under the supervision of the first and last author. Importantly, testers were blind to the treatment assignments of the children.

The study used a pre-test-training-post-test design with two test time points for the TG and three test time points for the CG. The design is presented in Fig. 3.

Between mid-February and mid-May 2018, participants started the training in seven subgroups, each consisting of an equal number of children in the TG and CG. Pre- and post-tests were carried out within two weeks before and after the training period.

2.4. Measures

The following standardized and experimental tests and questionnaires were administered:

Word reading. Word reading skills were assessed using three different reading measures. First, an individually administered standardized 1-min reading speed test was given (SLRT-II; Moll & Landerl, 2014; paralleltest-reliability: rtt = 0.90-0.98). It contains a word and a pseudoword reading list with increasing item length and complexity. The task is to read both lists aloud as fast as possible without making errors. The number of items read correctly within 1 min was scored for each subtest. Second, an experimental reading fluency test, which contained two lists of trained word material was administered, Children were asked to read aloud the words as quickly and accurately as possible. Words from word reading list 1 were simple in word structure (e.g. *Maus* – mouse, *Auto* – car, *Sonne* – sun) while words from word reading list 2 were longer and/or more complex in structure (e.g. *Flasche* – bottle, *Papagei* – parrot, *Prinzessin* – princess). Each list contained 25 words to be read. Time and reading errors were recorded and the number of words read correctly per minute for each list was calculated. Performance on the lists was highly correlated ($r = 0.86$). Third, a computer-based word-picture mapping task including the trained words increasing in length and difficulty was conducted (retest-reliability: rtt = 0.88). Children were asked to read silently and to select the target word that matches the picture from four alternative words. The number of correctly matched items within 3 min was calculated.

Reading comprehension. The standardized reading comprehension test (VSL; Walter, 2013; paralleltest-reliability: rtt = 0.78-0.80) measures the ability to quietly read up to two texts within 4 min while selecting missing keywords at regular intervals from a selection of three words. The final score is the number of correctly selected keywords.

Spelling. In the standardized spelling test children had to write 24 (2nd graders) or 48 (3rd graders) words dictated in sentence frames (SLRT-II; Moll & Landerl, 2014; paralleltest-reliability: rtt = 0.67-0.86). The total number of misspelled words was calculated.

Phonological awareness and grapheme-phoneme mapping. These skills were assessed using four non-standardized computerized subtests (split-half-reliability: r = 0.64). The first subtest, *syllable counting*, measures if the children can divide spoken words into syllables by asking the children to choose the correct number of syllables out of a selection from one to five. In the second subtest, *phoneme identification*, the children had to decide if they hear a certain sound in a spoken word. In the subtest *vowel length* children have to decide whether they hear a long or short vowel in a spoken word. In the fourth subtest, *grapheme-phoneme mapping*, the children were

Table 1
Descriptive statistics for the two groups.

Variables	TG (N = 25)		CG (N = 25)		F	p
	M (SD)	Range	M (SD)	Range		
Age (years)	8.50 (0.67)	7.58–9.75	8.55 (0.66)	7.42–9.58	0.16	.779
Reading (SLRT-II)[a]	8.8 (5.79)	1.0–20.5	8.5 (7.60)	0.5–24.3	2.66	.859
Non-verbal IQ[b]	110.2 (11.0)	91.0–126.0	108.7 (12.8)	87–138	0.36	.654
Digit Span forward[c]	6.72 (1.67)	5.0–12.0	7.12 (1.59)	4.0–11.0	0.17	.390
Digit Span backward[c]	5.52 (1.33)	3.0–8.0	5.52 (0.96)	4.0–8.0	3.15	1.00
RAN digits (item/s)	1.56 (0.35)	1.04–2.39	1.48 (0.34)	0.62–2.06	0.001	.403
RAN objects (item/s)	0.89 (0.23)	0.43–1.38	0.83 (0.19)	0.37–1.19	2.05	.351
	Frequencies				χ^2	**p**
Male/female	15/10		13/12		0.33	.569
Grade 2/grade 3	12/13		11/14		0.81	.777
Reading disorder mild to moderate/severe	13/12		11/14		0.32	.571
Isolated reading disorder/combined reading and spelling disorder	5/20		6/19		0.12	.733
Monolingual/bilingual	19/6		18/7		0.10	.747
ADHD/no ADHD (DISYPS-II)	5/20		5/20		0.00	1.00
Special reading support/no special reading support	8/17		9/16		0.37	.544

[a] Reading: Percentile rank.
[b] Non-verbal IQ scores: M = 100, SD = 15.
[c] Digit Span: Raw scores (maximum score = 16).

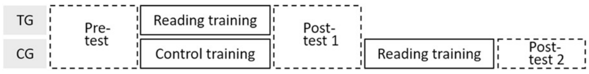


Fig. 3. Study design.

asked to map single sounds or groups of sounds to the corresponding letters out of a selection of three. All four subtests have different levels of difficulty, which are determined by the item structure, length or frequency. In each task the number of correctly given answers was summed up.

Nonverbal IQ. Non-verbal cognitive abilities were measured at pretest with the age-appropriate German version of the Culture Fair Intelligence Test (CFT 1-R for children < 8.5 years, Weiβ, 2006; and CFT 20-R for children ≥ 8.5 years, Weiβ & Osterland, 2012; retest-reliability: rtt = 0.88-0.96).

Verbal memory. The Subtest Digit Span from the German version of the Wechsler Intelligence Scale for Children (WISC-IV; (Petermann & Petermann, 2011) retest-reliability: rtt = 0.83) was administered at pretest to assess verbal short-term and working memory. The children were required to repeat a series of spoken digits of increasing length forwards or backwards. The number of correctly repeated digits forwards and backwards was summed up.

Rapid automatized naming (RAN). A standard RAN paradigm with two subtests was administered (Denckla & Rudel, 1976). Children had to name a matrix of 40 monosyllabic digits and 40 monosyllabic objects as fast and accurately as possible. Items named correctly per second were calculated. The two subtests (RAN digits and RAN objects) were highly correlated ($r = 0.70$).

Training experiences. A self-report and a parental questionnaire were designed to evaluate the children's training experience. The self-report questionnaire consists of eight questions to be answered on a 4-point rating scale, represented by circles of different sizes and colors (smallest red circle = strongly disagree, second smallest orange circle = somewhat disagree, second largest yellow circle = somewhat agree, largest green circle = strongly agree). After reading each item aloud to the children, they were asked to point to one of the four circles presented on a sheet of paper. The parents were asked to fill in a similar questionnaire, which consisted of 14 questions and the same 4-point rating scale from 0 (strongly disagree) to 3 (strongly agree). The questions were grouped into five categories and subsequently the mean was calculated for each category: *Enjoyment* (2 questions children, 3 questions parents), *difficulty-level* (2 questions each), *user-friendliness* (3 questions each), *perceived reading impact* (1 question children, 2 questions parents), *feasibility* (3 questions parents). Additional qualitative feedback could be given by parents and children based on two open questions about positive and/or negative aspects of the training.

Attention rating. Attentional (sub-)clinical problems were assessed with a standardized ADHD questionnaire of the Diagnostic System of Mental Disorders for Children and Adolescents – II (DISYPS-II; Döpfner et al., 2008; Cronbach's alpha = .87-.94 for parental ratings of ADHD symptoms). The questionnaire consists of 20 questions assigned to the three main dimensions of ADHD symptoms (attention deficit, hyperactivity and impulsivity). Questions were answered on a 4-point rating scale by one of the participant's caregivers.

2.5. Training conditions

The study compared two groups: (a) children who followed the reading training with Meister Cody-Namagi (TG) and (b) children who received an unspecific training using mind games (CG).

Due to ethical reasons, children of both groups continued to receive their regular reading instruction in school or individualized special reading instruction (with a maximum of once a week). Trainings of both TG and CG were conducted independently at home on a tablet computer (with Android or Apple operating system). Families were neither required to have a tablet computer nor to be experienced in using any kind of technical device; if necessary, tablet computers as well as technical support were provided to each family for the duration of the study.

2.5.1. Reading training

Children of the TG received the digital game-based reading training. Children were expected to train two to four times a week and thus complete the 30 training sessions ideally over a period of eight weeks. Criteria for inclusion in data analyses was to complete a minimum of 25 scheduled training sessions within a maximum of eleven weeks. Each session lasted about 25–30 min depending on the different tasks. In order to track their training progress, children got a training schedule and were asked to mark every training session. The study team received notifications of each child's session including user-id, date, time and duration of the task, task name and number of session, the difficulty level, ratio of success, as well as information about the operating system of the used tablet computer. Parents were informed in advance and had to give their consent that these data will be collected in order to track the training progress. During the training period, families were contacted regularly via e-mail or phone to avoid any difficulties with the app and to clarify questions that might arise. Due to the close monitoring of each child's training progress and the regular contact to the families, treatment fidelity can be considered as being high for an in-home training (Jamshidifarsani et al., 2019).

2.5.2. Control training

For the CG, an unspecific digital game-based training was prepared focusing on logic and attention skills. For this purpose, different commercially available apps *("Hama Universe"*,(Malte Haaning Plastic A/S, 2018); *"Petterssons Erfindungen 2"*, (Filimundus, 2015); *"Petterssons Erfindungen 3"*, (Filimundus, 2017); *"Richtig konzentrieren"*, (Spielend Lernen Verlag, 2018)) were selected. It was ensured that these apps were free of reading activities. The training schedule including the frequency of the training and the total training period was comparable to the TG. Again, families were contacted at regular intervals via e-mail or phone and the regularity of the training was assessed based on the training schedules.

2.6. Statistical analyses

To check for group differences in age, non-verbal cognitive abilities, verbal memory, RAN and literacy skills, independent samples t-tests were used. A series of univariate repeated measures analyses of variance (ANOVA) were conducted to evaluate significant change in reading and reading related skills from pre- to posttest, with group (TG vs. CG) as between-subjects factor and test time point (pretest vs. posttest) as within-subjects factor. Main outcome variables were the word reading measures for trained words (Word reading list 1 & 2, word-picture mapping) and untrained words (SLRT-II, word reading and pseudoword reading). In addition, the same analyses were run for reading comprehension (VSL) and spelling (SLRT-II) as dependent variables, although these skills were not directly trained. Raw scores were transformed into grade-specific z-scores in order to account for the two grade levels in our sample. Z-scores were calculated based on the norm sample for standardized measures and based on the current sample for experimental measures. Non-parametric tests for non-normally distributed data showed the same results as the conducted ANOVA's. Analyses were performed with IBM SPSS Statistics 25.

3. Results

3.1. Participant descriptives

As shown in Table 1, there were no significant differences between groups in any of the cognitive, behavioral and environmental pretest measures. Most importantly, both groups were similarly impaired in reading performance (8.8 vs. 8.5 percentiles).

3.2. Training exposure times

Exposure times to the digital game-based reading training were regularly monitored and analyzed after the training period. Twenty-five children of the TG fulfilled the training criteria and performed the reading training with a minimum of 25 scheduled sessions ($M = 29.28$, $SD = 1.31$) over a period of eight to eleven weeks ($M = 8.8$, $SD = 0.76$) with a frequency of two to four times per week ($M = 3.35$, $SD = 0.38$).

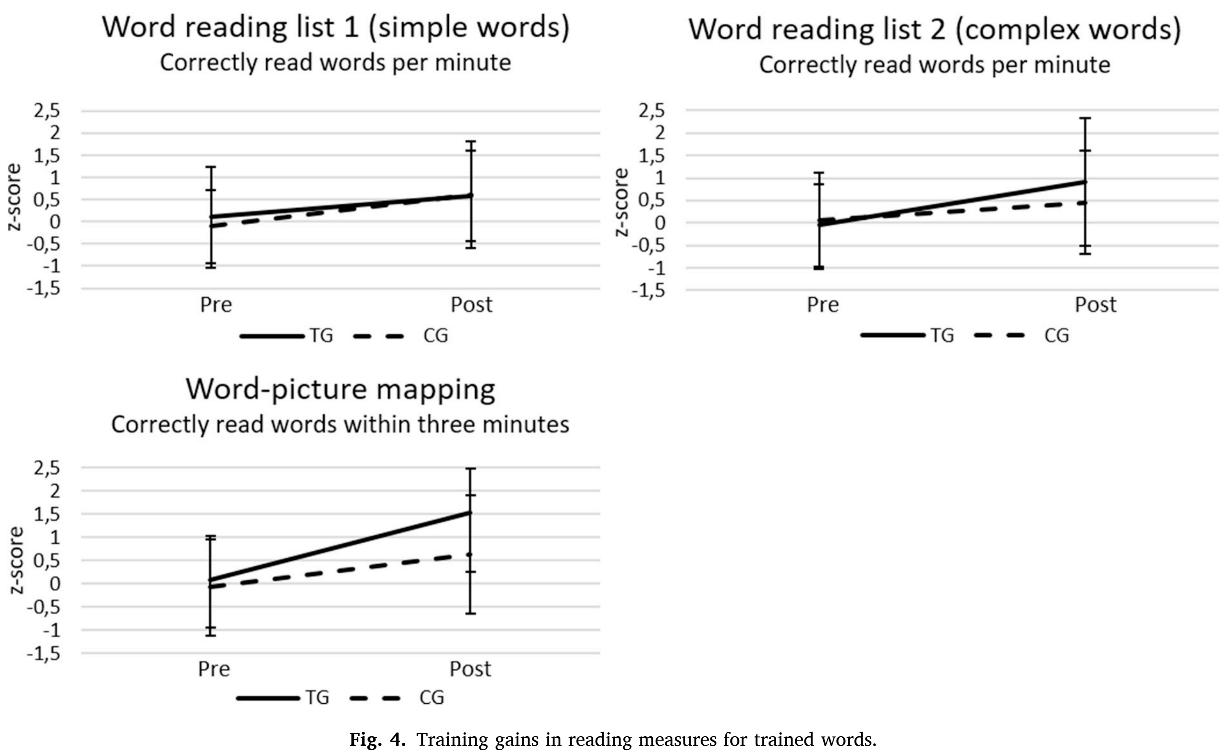


Fig. 4. Training gains in reading measures for trained words.

3.3. Training effects

3.3.1. Effects on word reading

The conducted ANOVA's for the trained words of the word reading list 1 (simple and short words) revealed a statistically significant main effect of time, F (1, 48) = 52.34, $p < .001$, $\eta^2_p = .52$, but neither a significant main effect of group, F (1, 48) = 0.10, $p = .759$, $\eta^2_p = .00$, nor a significant time × group interaction, F (1, 48) = 2.21, $p = .144$, $\eta^2_p = .044$. The results indicate that reading skills for simple and short words improved significantly irrespective of the training. For the word reading list 2 (complex and long words) a statistically significant main effect was found for time, F (1, 48) = 11.58, $p < .001$, $\eta^2_p = .53$, but not for group, F (1, 48) = 0.32, $p = .575$, $\eta^2_p = .01$. Importantly, a significant time × group interaction for the word-reading list 2, F (1, 48) = 9.50, $p < .01$, $\eta^2_p = .17$, was found, suggesting that reading of complex and longer words improved more in the TG than in the CG. Likewise, in the computerized word-picture mapping test, results showed a significant main effect of time, F (1, 48) = 118.94, $p < .001$, $\eta^2_p = .71$, no significant main effect of group, F (1, 48) = 3.49, $p = .068$, $\eta^2_p = .07$, but a significant time × group interaction, F (1, 48) = 13.39, $p < .01$, $\eta^2_p = .22$, indicating that word-picture mapping improved more in the TG than in the CG.

Training gains in the reading measures for the trained words are illustrated in Fig. 4.

Analyses of untrained words, measured with the standardized reading test SLRT-II, showed significant main effects of time on word reading and pseudoword reading, F (1, 48) = 70.66, $p < 0.001$, $\eta^2_p = .60$ and F (1, 48) = 17.26, $p < .001$, $\eta^2_p = .26$, but no significant main effect of a group on word reading and pseudoword reading was found, F (1,48) = 0.07, $p = .793$, $\eta^2_p = .00$ and F (1, 48) = 0.91, $p = .345$, $\eta^2_p = .02$. The main effects of time indicate that word and pseudoword reading skills improved from pretest to posttest in both groups. A marginal significant time × group interaction was found for word reading, F (1, 48) = 3.99, $p = .051$, $\eta^2_p = .08$, but not for pseudoword reading, F (1, 48) = 0.63, $p = .433$, $\eta^2_p = .01$, reflecting a trend of transfer effects from trained to untrained words for the TG on a standardized word reading measure.

3.3.2. Effects on phonological awareness and grapheme-phoneme mapping

Regarding phonological awareness skills and grapheme-phoneme mapping, repeated measure ANOVA's showed significant main effects of time for phoneme identification and grapheme-phoneme mapping, respectively, F (1, 48) = 16.39, $p < .001$, $\eta^2_p = .36$ and F (1, 48) = $p < .001$, $\eta^2_p = .25$, whereas the main effects of group were not significant, F (1, 48) = 0.02, $p = .88$, $\eta^2_p = .00$ and F (1,48) = 0.16, $p = .694$, $\eta^2_p = .00$. Analyses further revealed a significant time × group interaction for phoneme identification, F (1, 48) = 7.40, $p < .01$, $\eta^2_p = .13$, but not for grapheme-phoneme mapping, F (1, 48) = 0.70, $p = .409$, $\eta^2_p = .01$, indicating that children from the TG improved more on phoneme identification task than children from the CG. For the tasks syllable counting and vowel length there was neither a significant main effect of time, F (1, 48) = 3.22, $p = .08$, $\eta^2_p = .06$ and F (1, 48) = 1.75, $p = .19$, $\eta^2_p = .04$, nor a significant

Table 2
Raw scores (mean and standard deviations) for the two groups by grade and test time point (pre and post).

Variables	TG (N = 25)				CG (N = 25)			
	Grade 2 (N = 12)		Grade 3 (N = 13)		Grade 2 (N = 11)		Grade 3 (N = 14)	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post
	Raw score M (SD)				Raw score M (SD)			
Word reading[a] (SLRT-II)	19.75 (7.93)	25.67 (9.44)	29.85 (8.05)	36.46 (10.88)	21.73 (9.93)	25.55 (10.36)	31.86 (7.88)	35.79 (10.04)
Pseudoword reading[a] (SLRT-II)	19.25 (3.70)	21.50 (5.23)	22.85 (7.11)	24.62 (5.24)	17.09 (17.09)	20.36 (4.99)	21.36 (4.97)	24.07 (7.13)
Word reading list 1[a]	36.02 (19.42)	44.88 (18.37)	50.20 (20.32)	57.69 (16.81)	26.60 (13.07)	37.98 (20.18)	51.59 (15.01)	64.78 (21.59)
Word reading list 2[a]	16.25 (9.80)	25.58 (15.35)	26.74 (12.82)	41.12 (19.09)	14.89 (9.98)	20.13 (12.03)	31.49 (16.74)	35.71 (16.52)
Word-picture mapping (max. 66)	26.83 (3.41)	35.67 (5.18)	34.31 (9.25)	46.15 (8.49)	23.55 (9.13)	27.73 (9.96)	35.07 (5.68)	41.07 (8.32)
Reading comprehension[b] (VSL)	4.58 (2.61)	5.67 (2.93)	8.92 (4.61)	11.62 (6.05)	3.91 (1.58)	3.73 (2.90)	7.79 (1.72)	11.07 (5.57)
Spelling[c] (SLRT-II)	13.83 (2.79)	11.58 (2.84)	25.08 (8.03)	23.23 (7.92)	15.00 (5.73)	13.18 (6.21)	25.29 (8.32)	23.71 (6.60)
Syllable counting (max. 18)	13.17 (3.97)	15.67 (1.37)	15.85 (1.73)	16.23 (1.83)	14.64 (3.36)	14.73 (3.04)	14.07 (2.81)	14.36 (3.61)
Phoneme identification (max. 18)	16.42 (0.90)	17.50 (0.68)	15.85 (1.07)	17.31 (0.86)	16.73 (0.91)	16.73 (1.01)	16.36 (1.22)	17.14 (0.66)
Vowel length (max. 19)	12.58 (2.84)	13.67 (2.31)	14.85 (2.73)	14.54 (3.36)	13.00 (2.49)	14.09 (2.39)	12.57 (2.24)	12.64 (2.98)
Grapheme-phoneme mapping (max. 26)	22.92 (2.35)	23.58 (1.78)	23.77 (1.48)	24.62 (1.12)	23.09 (2.02)	24.73 (1.01)	23.57 (1.34)	24.50 (1.23)

[a] Number of correctly read words per minute.
[b] Number of correctly selected words within 4 min.
[c] Number of spelling errors, maximum of 24 words for 2nd Graders and 48 words for 3rd Graders.

main effect of a group, F (1, 48) = 2.34, p = .133, η^2_p = .05 and F (1,48) = 1.80, p = .186, η^2_p = .04, and no significant time × group interaction, F (1, 48) = 1.52, p = .224, η^2_p = .03 and F (1, 48) = 0.05, p = .818, η^2_p = .00. The results indicate that both groups did not improve in syllable counting and vowel length.

3.3.3. Effects on reading comprehension and spelling

Repeated measure ANOVA's were also conducted to assess the influence of the training on untrained literacy skills such as reading comprehension and spelling. Results revealed that both skills improved over time, F (1, 48) = 9.95, $p < .01$, η^2_p = .17 and F (1, 48) = 16.38, $p < .001$, η^2_p = .25, but there was neither a significant main effect of group, F (1, 48) = 1.58, p = .215, η^2_p = .03 and F (1, 48) = 0.39, p = 535, η^2_p = .01, nor a significant time × group interaction, F (1, 48) = 0.08, p = .777, η^2_p = .00 and F (1, 48) = 0.19, p = .662, η^2_p = .00, suggesting that both groups developed similarly in measures of reading comprehension and spelling irrespective of the training.

Raw scores for reading, spelling and phonological awareness tests by group and test timepoint are presented in Table 2.

3.4. Training experiences

Training experiences were measured with a self-report for children and a parent questionnaire using a four-point rating scale ranging from 0 (strong disagreement) to 3 (strong agreement). The category *enjoyment* on average was rated with 2.44 (*SD* = 0.79) by children and 2.60 (*SD* = 0.47) by parents indicating that the children enjoyed the training and were willing to continue. The self-reported mean rating for *difficulty level* was 2.26 (*SD* = 0.68) and the adult-observed mean rating was 2.42 (*SD* = 0.43), showing that tasks were appropriate in terms of difficulty level and duration. Average scores of the category *user-friendliness* were 2.68 (*SD* = 0.37) for children and 2.88 (*SD* = 0.29) for parents. This result indicates that instructions in the program were clear and children were able to use it unassisted. Children's mean rating in the category *perceived reading impact* was 1.92 (*SD* = 0.95) and parent's mean rating was 2.17 (*SD* = 0.60). This finding reflects a perceived positive influence on children's reading ability and parents seem to consider the training content suitable with regard to their child's reading disorder. The category *feasibility* was rated by parents with an average of 2.60 (*SD* = 0.42), showing that it was possible to integrate the training into everyday life and that the duration and frequency of the training were appropriate. The open questions revealed that the *story frame* and *reward system* were particularly important since these aspects of the training were mentioned by 64% of the children and 56% of the parents.

4. Discussion

The primary goal of the current study was to assess whether a digital game-based and independently performed reading training can improve reading skills of second- and third-graders with reading disorder. In line with Ronimus et al. (2019), our data confirm that children with reading impairment benefit from digital game-based reading training. The results further show that the training can be carried out independently in home environment. This is an important finding since reading skills of children with reading disorder are often hard to remediate (Galuschka et al., 2014; Ise et al., 2012; Ronimus et al., 2019). Results of the current study extend existing knowledge about digital game-based reading approaches by using a holistic reading approach (Jamshidifarsani et al., 2019), which combines different components identified as being effective to improve reading skills with digital game-based learning. Importantly, the training resulted in significant training gains on the computerized word-picture mapping test clearly underlining that word reading fluency for the digital presented word material increased. Transfer effects to non-digital word material can be partially confirmed. While reading skills of the simple and short words did not significantly differ between the training and control group after the training, reading skills of complex and long words improved significantly more for children using the digital game-based reading program compared to children in the control group. The difference between the two lists can be explained by the fact that word reading fluency for high-frequent short words develops more quickly than reading fluency for low-frequent words which are encountered less often (Berends & Reitsma, 2006). As a consequence, exposure to such simple and short words has likely taken place in both groups regardless of the training (see also Heikkilä, Aro, Närhi, Westerholm, & Ahonen, 2013). Thus, we assume, that reading of simple words was already relatively fluent prior to the training, while for complex and long words there was more room for improvement (Heikkilä et al., 2013). The significant improvement of complex words is especially important given that complex written language units are processed less efficiently by children with reading disorder (Spinelli et al., 2005; Ziegler, Perry, Ma-Wyatt, Ladner, & Schulte-Körne, 2003) and are more difficult to learn to read (Manis, Custodio, & Szeszulski, 1993). The finding that complex words can be trained is especially relevant for languages like German which are characterized by long compound words with complex syllable structure and many consonant clusters.

In addition to the training effect on trained word material there was also a trend of transfer from trained to untrained word material, suggesting that the training has the potential not only to improve word specific reading but also general word reading abilities. Inducing significant gains in general word reading is hard to accomplish and is influenced by several conditions, such as training intensity and setting as well as the level of severity of the reading deficit (Berends & Reitsma, 2006). Generating transfer effects to untrained word material is crucial for skilled reading but often neglected when evaluating reading trainings. For this reason, future studies evaluating the current reading training should tackle this issue by slightly extending the program and adapting the training schedule as follows: First, given the clear trend of transfer to untrained word material after only a short intervention period of eight to eleven weeks, it is likely that children will make greater progress over a longer training period as a high amount of repetition is needed to improve word reading fluency (Thaler et al., 2004). Thus, the training period should be extended while at the same time some tasks could be complemented by additional levels of difficulty to keep the tasks challenging enough for less severe impaired children.

Second, bearing in mind that reading speed is especially hard to improve (Thaler et al., 2004), the number of speeded reading tasks should be increased as the training schedule of the present study contained twice as many reading accuracy as speed tasks (see supplement 1). Third, compared to phonological and word reading tasks, the training currently contains relatively few morphological tasks (see supplement 1). It is therefore promising to complement the training by additional morphological tasks as morphological instruction seems to speed the recognition of new words (Carlisle & Stone, 2005) and contributes to reading fluency especially in later stages of literacy development (Mann & Singson, 2003).

No significant training effects between groups were found on the standardized pseudoword reading measure. This is not unexpected since the focus of the training was on fluent real word recognition, while explicit reading of pseudowords was targeted in one task only.

In the current age group (second and third grade children), performance on measures of phonological awareness was already high prior to the training. Therefore, children did not show significant gains on several of these measures, including syllable counting and vowel length discrimination. Although we found significant gains in grapheme-phoneme mapping and phoneme identification as well as a stronger improvement of the training group compared to the control group in the phoneme identification task, it should be noted that overall performance was relatively high even before the training. This supports the view that in transparent orthographies, children mainly benefit from phonological awareness training in earlier stages of literacy development (Galuschka et al., 2014; Georgiou et al., 2008; Mann & Wimmer, 2002; Wolf et al., 2016). Thus, phonological awareness tasks of the evaluated training program might be more beneficial for children in first or beginning of second grade or for the more severely affected children.

The training did not have significant impact on untrained literacy skills, such as reading comprehension and spelling. Research indicates that due to different deficit profiles associated with each literacy skill, different methods targeting the specific symptoms are needed (Moll & Landerl, 2009; Moll, Wallner, & Landerl, 2012; Vellutino et al., 2004). For this reason, the program should be complemented by modules specifically targeting reading comprehension and spelling skills, thus extending the current focus on word reading only.

With respect to motivational aspects of the training, children's self-reports and parent reports showed that children enjoyed the training and were motivated to continue the training after the scheduled training time. Additional qualitative feedback revealed that the story frame and reward system particularly contributed to the child's interest in the program which is consistent with previous findings in digital game-based learning (Jamshidifarsani et al., 2019; Ke & Abras, 2013; Ronimus, Kujala, Tolvanen, & Lyytinen, 2014, 2019). Since motivation to read is often low in children with reading difficulties (Morgan et al., 2008; Polychroni, Koukoura, & Anagnostou, 2006), the high motivation to use the program reported by children and parents is very encouraging and should be given special consideration when designing digital game-based reading approaches. To further increase motivation, children may be involved more actively by taking on a more active role as it was suggested by Ronimus, Kujala, Tolvanen, & Lyytinen, 2014, for example by adding more features to the magical garden. Mean ratings of self- and parent-reports show that the difficulty level and user-friendliness (e.g. explanation of tasks) were perceived as adequate. High values in the mean ratings of the perceived adequate difficulty level imply that the adaptivity system of the program worked well and the challenge-skill balance was appropriate. The relevance of these principles has been highlighted before (Jamshidifarsani et al., 2019; Maertens et al., 2014; Sweetser & Wyeth, 2005). Ratings of user-friendliness and parental ratings of feasibility reveal that the children can perform the training at home without an external instructor which might simplify access to individual reading support for many children. However, it should be noted that although our study team received notifications of each child's training progress, we were unable to directly observe the children while using the program. Therefore, we cannot rule out that parental support may have differed between children, although all parents were carefully instructed that children are supposed to perform the training on their own. The perceived positive influence on reading ability indicates that the training enhanced a positive attitude towards reading. However, it should be noted that reading attitudes were assessed by only one question in self-reports and two questions in parental reports. Thus, future studies might develop a more fine-grained questionnaire to capture the children's attitudes towards reading in more detail (see also Ronimus, Kujala, Tolvanen, & Lyytinen, 2014). Finally, according to the parents' ratings on feasibility, it seems that the training can be easily integrated into everyday life which is a crucial factor for families to use the program.

One limitation of the training program to be considered is the lack of having children reading aloud during the training. Although automatic speech recognition systems would allow to assess reading aloud (Kim, 2006), this technology is still challenging (Li, Deng, Haeb-Umbach, & Gong, 2016). Due to the limited reliability of automatic speech recognition systems (Forsberg, 2003; Li et al., 2016; Petkar, 2016) and the lack of evidence showing an advantage of reading aloud over silent reading (Berends & Reitsma, 2007; Huemer, Landerl, Aro, & Lyytinen, 2008; Thaler et al., 2004), no such system was integrated into our reading training. Share (2004), however, noted that availability of an oral word form supports orthographic learning. Consistent with Huemer et al. (2008), we argue that in our training the phonological word form is likely activated through silent reading as well as through additional auditory presentation of the written language units. However, incorporating speech recognition technology would make it possible to monitor precisely what children read and would in turn allow for even more individualized analyses of reading performance. Thus, when technology allows more accurate speech recognition, taking into account the variety of human speech and disruptive factors (e.g. background noises) (Forsberg, 2003; Petkar, 2016), such a system should be integrated in the evaluated reading program.

5. Conclusions

The present study shows that a multi-component digital game-based reading training carried out independently at home can improve reading skills of children with reading disorder. A holistic approach integrating evidence-based findings from standard reading instruction with flow principles seems to facilitate enjoyable reading experiences as well as word specific reading performance.

By incorporating innovative features such as adaptivity, immediate feedback and rewarding, the training doesn't need an external instructor but can be performed independently in home environment. Thus, it can be particularly useful for children who don't have access to special reading support or for bridging the gap until individual support can be provided. It can also be used as an additional tool to individual therapy to increase frequency of support. The program seems promising to induce transfer effects, thus future studies should especially address the issue of transfer effects by evaluating the training using an extended training period. General word reading skills might further benefit from a higher number of difficulty levels and morphological tasks as well as a stronger focus on reading speed tasks. In order to improve other literacy skills such as reading comprehension and spelling, the development of specific modules is mandatory.

CRediT authorship contribution statement

Ruth Görgen: Conceptualization, Methodology, Investigation, Formal analysis, Writing - original draft, Writing - review & editing. **Sini Huemer:** Conceptualization, Writing - original draft. **Gerd Schulte-Körne:** Conceptualization, Writing - original draft. **Kristina Moll:** Conceptualization, Formal analysis, Writing - original draft, Writing - review & editing.

Acknowledgements

We wish to thank the children for their participation in this study. We also thank their parents as well as therapeutic practices and afternoon care centres for the good cooperation. Special thanks go to the following research assistants for their help with data collection: Jessica Feichtmayr, Sarah Ghezih, Elisabeth Hecht, Veronika Jäger, Linda Schuster, Carolina Silberbauer, Petra Wagenbüchler, Dana Winogradow, Björn Witzel. This project was supported by grants from the Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung (BMBF), grant no. 80167009).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compedu.2020.103834.

References

Admiraal, W., Huizenga, J., Akkerman, S., & ten Dam, G. (2011). The concept of flow in collaborative game-based learning. *Computers in Human Behavior, 27*(3), 1185–1194. https://doi.org/10.1016/j.chb.2010.12.013.

Denckla, M. B., & Rudel, R. G. (1976). Rapid "automatized" naming (R.A.N.): Dyslexia differentiated from other learning disabilities. *Neuropsychologia, 14*(4), 471–479. https://doi.org/10.1016/0028-3932(76)90075-0.

American Psychiatric Association, A. E.. (2013). *Diagnostic and statistical manual of mental disorder, DSM-5* (5th ed.). Washington, DC: American Psychiatric Pub.

Araújo, S., & Faísca, L. (2019). A meta-analytic review of naming-speed deficits in developmental dyslexia. *Scientific Studies of Reading, 23*(5), 1–20. https://doi.org/10.1080/10888438.2019.1572758.

Aro, M., & Wimmer, H. (2003). Learning to read: English in comparison to six more regular orthographies. *Applied PsychoLinguistics, 24*, 621–635. doi:10.1017. S0142716403000316.

Bäcker, A., & Neuhäuser, G. (2003). Internalisierende & externalisierende Syndrome bei Lese- und Rechtschreibstörungen. *Praxis der Kinderpsychologie und Kinderpsychiatrie, 52*(5), 329–337.

Berends, I., & Reitsma, P. (2006). Remediation of fluency: Word specific or generalised training effects? *Reading and Writing: An Interdisciplinary Journal, 19*, 221–234. https://doi.org/10.1007/s11145-005-5259-3.

Berends, I., & Reitsma, P. (2007). Orthographic analysis of words during fluency training promotes reading of new similar words. *Journal of Research in Reading, 30*(2), 129–139. https://doi.org/10.1111/j.1467-9817.2006.00324.x.

Blok, H., Oostdam, R., Otter, M. E., & Overmaat, M. (2002). Computer-assisted instruction in support of beginning reading instruction: A review. *Review of Educational Research, 72*(1), 101–130. https://doi.org/10.3102/00346543072001101.

Bowers, P. N., Kirby, J. R., & Deacon, H. S. (2010). The effects of morphological instruction on literacy skills: A systematic review of the literature. *Review of Educational Research, 80*(2), 144–179. https://doi.org/10.3102/0034654309359353.

Carlisle, J., & Stone, A. C. (2005). Exploring the role of morphemes in word reading. *Reading Research Quarterly, 40*(4), 428–449. https://doi.org/10.1598/RRQ.40.4.3.

Chang, C.-C., Liang, C., Chou, P.-N., & Lin, G.-Y. (2017). Is game-based learning better in flow experience and various types of cognitive load than non-game-based learning? Perspective from multimedia and media richness. *Computers in Human Behavior, 71*, 218–227. https://doi.org/10.1016/j.chb.2017.01.031.

Chauhan, S. (2016). A meta-analysis of the impact of technology on learning effectiveness of elementary students. *Computers & Education, 105*, 14–30. https://doi.org/10.1016/j.compedu.2016.11.005.

Chung, C.-J., Hwang, G.-J., & Lai, C.-L. (2018). A review of experimental mobile learning research in 2010-2016 based on the activity theory framework. *Computers & Education, 129*, 1–13. https://doi.org/10.1016/j.compedu.2018.10.010.

Csikszentmihalyi, M. (1975). *Beyond boredom and anxiety - Experiencing flow in work and play*. San Francisco: Jossey-Bass.

De Freitas, P. V., Da Mota, M. M. P. E., & Deacon, H. S. (2017). Morphological awareness, word reading, and reading comprehension. *Applied PsychoLinguistics, 39*(3), 1–19. https://doi.org/10.1017/S0142716417000479.

Döpfner, M., & Görtz-Dorten, A. (2008). *DISYPS-III Diagnostik System für psychische Störungen nach ICD-10 und DSM-IV für Kinder und Jugendliche – III*. Göttingen: Hogrefe.

Dummer-Smoch, L., & Hackethal, R. (2007). *Kieler Leseaufbau*. Kiel: Veris.

Ehri, L. C. (1995). Phases of development in learning to read words by sight. *Journal of Research in Reading, 18*(2), 116–125.

Filimundus, A. B. (2015). *Petterssons Erfindungen 2 (version 1.11)*. Stockholm: Filimundus AB. Retrieved from https://play.google.com/store/apps/details?id=se.filimundus.pi2&hl=de.

Filimundus, A. B. (2017). *Petterssons Erfindungen 3 (version 1.1)*. Stockholm. Retrieved from https://play.google.com/store/apps/details?id=se.filimundus.pi3&hl=de.

Fink, A., Pucher, S., Reicher, A., Purgstaller, C., & Kargl, R. (2012). Development of a test for the assessment of morphological awareness. *Empirische Pädagogik, 26*(4), 423–451.

Fischbach, A., Schuchardt, K., Brandenburg, J., Klesczweski, J., Balke-Melcher, C., Schmidt, C., et al. (2013). Prävalenz von Lernschwächen und Lernstörungen: Zur Bedeutung der Diagnosekriterien. *Lernen und Lernstörungen, 2*(2), 65–76.

Forsberg, M. (2003). *Why is speech recognition difficult*. Retrieved from https://www.researchgate.net/publication/228763868.

Galuschka, K., Görgen, R., Kalmar, J., Haberstroh, S., Schmalz, X., & Schulte-Körne, G. (2020). *Effectiveness of spelling treatment approaches for learners with dyslexia and spelling deficits – a meta-analysis and systematic review* (1st, (1st,. *55* pp. 1–20). Educational Psychologist.

Galuschka, K., Ise, E., Krick, K., & Schulte-Körne, G. (2014). Effectiveness of treatment approaches for children and adolescents with reading disabilities: A meta-analysis of randomized controlled trials. *PloS One, 9*(2), 1–12. https://doi.org/10.1371/journal.pone.0089900.

Georgiou, G. K., Parrila, R., & Papadopoulos, T. C. (2008). Predictors of word decoding and reading fluency across languages varying in orthographic consistency. *Journal of Educational Psychology, 100*(3), 566–580. https://doi.org/10.1037/0022-0663.100.3.566.

Goldston, D. B., Walsh, A., Mayfield, A. E., Reboussin, B., Sergent, D. S., Erkanli, A., … Wood, F. (2007). Reading problems, psychiatric disorders, and functional impairment from mid- to late adolescence. *Journal of the American Academy of Child & Adolescent Psychiatry, 46*(1), 25–32. https://doi.org/10.1097/01.chi.0000242241.77302.f4.

Goodwin, A. P., & Ahn, S. (2010). A meta-analysis of morphological interventions: Effects on literacy achievement of children with literacy difficulties. *Annals of Dyslexia, 60*(2), 183–208. https://doi.org/10.1007/s11881-010-0041-x.

Hattie, J., & Timperley, H. (2007). The power of Feedback. *Review of Educational Research, 77*(1), 81–112. https://doi.org/10.3102/003465430298487.

Heikkilä, R., Aro, M., Närhi, V., Westerholm, J., & Ahonen, T. (2013). Does training in syllable recognition improve reading speed? *A Computer-Based Trial With Poor Readers From Second and Third Grade, 17*(6), 1–37. https://doi.org/10.1080/10888438.2012.753452.

Huemer, S., Landerl, K., Aro, M., & Lyytinen, H. (2008). Training reading fluency: Is it important to practice reading aloud and is generalization possible? *Annals of Dyslexia, 58*(1), 59–79. doi:10.1007/s11881-008-0012-7. doi: 10.1026/0942-5403/a000077.

Ise, E., Engel, R. R., & Schulte-Körne, G. (2012). Was hilft bei der Lese- und Rechtschreibstörung? Ergebnisse einer Metaanalyse zur Wirksamkeit deutschsprachiger Förderansätze. *Kindheit und Entwicklung, 21*, 122–136.

Jamshidifarsani, H., Garbaya, S., Lim, T., Blazevic, P., & Ritchie, J. M. (2019). Technology-based reading intervention programs for elementary grades: An analytic review. *Computers & Education, 128*, 427–451. https://doi.org/10.1016/j.compedu.2018.10.003.

de Jong, P. F., & Vrielink, L. O. (2004). Rapid automatic naming: Easy to measure, hard to improve (quickly). *Annals of Dyslexia, 54*(1), 65–88. https://doi.org/10.1007/s11881-004-0004-1.

Ke, F., & Abras, T. (2013). Games for engaged learning of middle school children with special learning needs. *British Journal of Educational Technology, 44*(2), 225–242. https://doi.org/10.1111/j.1467-8535.2012.01326.x.

Kiili, K., de Freitag, S., Arnab, S., & Lainema, T. (2012). The design principles for flow experience in educational games. *Procedia Computer Science, 15*, 78–91. https://doi.org/10.1016/j.procs.2012.10.060.

Kim, I.-S. (2006). Automatic speech recognition: Reliability and pedagogical implications for teaching pronunciation. *Educational Technology & Society, 9*(1), 322–334.

Kirby, J. R., Deacon, H. S., Bowers, P. N., Izenberg, L., Wade-Woolley, L., & Parrila, R. (2012). Children's morphological awareness and reading ability. *Reading and Writing, 25*(2), 389–410. https://doi.org/10.1007/s11145-010-9276-5.

Kirk, C., & Gillon, G. T. (2009). Integrated morphological awareness intervention as a tool for improving literacy. *Language, Speech, and Hearing Services in Schools, 40*, 341–351. https://doi.org/10.1044/0161-1461(2008/08-0009.

Kuh, G. D. (2009). The national survey of student engagement: Conceptual and empirical foundations. *New Directions for Institutional Research, 141*, 5–20. https://doi.org/10.1002/ir.283.

Landerl, K., & Moll, K. (2010). Comorbidity of learning disorders: Prevalence and familial transmission. *Journal of Child Psychology and Psychiatry, 51*(3), 287–294. https://doi.org/10.1111/j.1469-7610.2009.02164.x.

Landerl, K., Ramus, F., Moll, K., Lyytinen, H., Leppanen, P. H., Lohvansuu, K., … Schulte-Körne, G. (2013). Predictors of developmental dyslexia in European orthographies with varying complexity. *Journal of Child Psychology and Psychiatry, 54*(6), 686–694. https://doi.org/10.1111/jcpp.12029.

Landerl, K., & Wimmer, H. (2008). Development of word reading fluency and spelling in a consistent orthography: An 8-year follow-up. *Journal of Educational Psychology, 100*(1), 150–161. https://doi.org/10.1037/0022-0663.100.1.150.

Lemoine, H. E., Levy, B. A., & Hutchinson, A. (1993). Increasing the naming speed of poor readers: Representations formed across repetitions. *Journal of Experimental Child Psychology, 55*(3), 297–328. https://doi.org/10.1006/jecp.1993.1018.

Lepper, M. R., Henderlong Corpus, J., & Iyengar, S. S. (2005). Intrinsic and extrinsic motivational orientations in the classroom. *Journal of Educational Psychology, 97*(2), 184–196. https://doi.org/10.1037/0022-0663.97.2.184.

Levy, B. A., Bourassa, D. C., & Horn, C. (1999). Fast and slow namers: Benefits of segmentation and whole word training. *Journal of Experimental Child Psychology, 73*, 115–138. https://doi.org/10.1006/jecp.1999.2497.

Li, J., Deng, L., Haeb-Umbach, R., & Gong, Y. (2016). Learn more about automatic speech recognition. In J. Li, L. Deng, R. Haeb-Umbach, & Y. Gong (Eds.), *Robust automatic speech recognition* (pp. 1–7). Oxford: Academic Press.

MacArthur, C. A., Ferretti, R. P., Okolo, C. M., & Cavalier, A. R. (2001). Technology applications for the students with literacy problems: A critical review. *The Elementary School Journal, 101*(3), 273–301. https://doi.org/10.1086/499669.

Maertens, M., Vandewaeterea, M., Cornillie, F., & Desmet, P. (2014). From pen-and-paper content to educational math game content for children: A transfer with added difficulty. *International Journal of Child-Computer Interaction, 2*, 85–92. https://doi.org/10.1016/j.ijcci.2014.04.001.

Malte Haaning Plastic A/S. (2018). *Hama Universe (version 1.9.6). Nykøbing Mors: Malte Haaning plastic A/S*. Retrieved from https://play.google.com/store/apps/details?id=com.hama.universe&hl=de.

Manis, F., Custodio, R., & Szeszulski, P. (1993). Development of phonological and orthographic skill: A 2-year longitudinal study of dyslexic children. *Journal of Experimental Child Psychology, 56*(1), 64–86. https://doi.org/10.1006/jecp.1993.1026.

Mann, V., & Singson, M. (2003). Linking morphological knowledge to English decoding ability: Large effects of little suffixes. In E. M. H. Assink, & D. Sandra (Eds.), *Reading complex words* (Vol. 22). Boston: Springer.

Mann, V., & Wimmer, H. (2002). Phoneme awareness and pathways into literacy: A comparison of German and American children. *Reading and Writing, 15*(7), 653–682. https://doi.org/10.1023/A:1020984704781.

Moll, K., & Landerl, K. (2009). Double dissociation between reading and spelling deficits. *Scientific Studies of Reading, 13*(5), 359–382. https://doi.org/10.1080/10888430903162878.

Moll, K., & Landerl, K. (2014). *SLRT-II: Lese- und Rechtschreibtest*. Bern: Huber.

Moll, K., Wallner, R., & Landerl, K. (2012). Kognitive Korrelate der Lese-, Leserechtschreib- und der Rechtschreibstörung. *Lernen und Lernstörungen, 1*, 7–19. https://doi.org/10.1024/2235-0977/a000002.

Morgan, P. L., Fuchs, D., Compton, D. L., Cordray, D. S., & Fuchs, L. S. (2008). Does early reading failure decrease children's reading motivation? *Journal of Learning Disabilities, 41*(5), 387–404. https://doi.org/10.1177/0022219408321112.

National Institute of Child Health and Human Development. (2000). Report of the national reading panel. In *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction (NIH Publication No. 00–4769)*. Washington, DC: US Government Printing Office. Retrieved from.

Papastergiou, M. (2009). Digital Game-Based Learning in high school Computer Science education: Impact on educational effectiveness and student motivation. *Computers & Education, 52*, 1–2. https://doi.org/10.1016/j.compedu.2008.06.004.

Perelmutter, B., McGregor, K. K., & Gordon, K. R. (2017). Assistive technology interventions for adolescents and adults with learning disabilities: An evidence-based systematic review and meta-analysis. *Computers & Education, 114*, 139–163. https://doi.org/10.1016/j.compedu.2017.06.005.

Petermann, F., & Petermann, U. (2011). *Wechsler intelligence scale for children* (Fourth). Frankfurt am Main: Pearson Assessment & Information GmbH.

Petkar, H. (2016). A review of challenges in automatic speech recognition. *International Journal of Computer Application, 151*(3), 23–26.
Polychroni, F., Koukoura, K., & Anagnostou, I. (2006). Academic self-concept, reading attitudes and approaches to learning of children with dyslexia: Do they differ from their peers? *European Journal of Special Needs Education, 21*(4), 415–430. https://doi.org/10.1080/08856250600956311.
Reed, D. K. (2008). A synthesis of morphology interventions and effects on reading outcomes for students in grades K-12. *Learning Disabilities Research & Practice, 23* (1), 36–49. https://doi.org/10.1111/j.1540-5826.2007.00261.x.
Reuter-Liehr, C. (2000). *Lautgetreuer Lese-Rechtschreibförderung*. Bochum: Dr. Winkler Verlag.
Reuter-Liehr, C. (2006). *Lautgetreuer Lese-Rechtschreibförderung, band 5. Bochum*. Dr. Winkler Verlag.
Ronimus, M., Eklund, K., Pesu, L., & Lyytinen, H. (2019). Supporting struggling readers with digital game-based learning. *Educational Technology Research & Development, 67*(3), 639–663. https://doi.org/10.1007/s11423-019-09658-3.
Ronimus, M., Kujala, J., Tolvanen, A., & Lyytinen, H. (2014). Children's engagement during digital game-based learning of reading: The effects of time, rewards, and challenge. *Computers & Education, 71*, 237–246. https://doi.org/10.1016/j.compedu.2013.10.008.
Santoro, L. E., & Bishop, M. J. (2010). Selecting software with caution: An empirical evaluation of popular beginning reading software for children with early literacy difficulties. *Computers in the Schools, 27*(2), 99–120. https://doi.org/10.1080/07380561003801582.
Schroeder, S., Würzner, K. M., Heister, J., Geyken, A., & Kliegl, R. (2015). childLex: A database of German read by children. *Behavior Research Methods, 47*, 1085–1094. https://doi.org/10.3758/s13428-014-0528-1.
Seymour, P. H., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology, 94*(2), 143–174. https://doi.org/10.1348/000712603321661859.
Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition, 55*, 151–218.
Share, D. L. (2004). Orthographic learning at a glance: On the time course and developmental onset of self-teaching. *Journal of Experimental Child Psychology, 87*, 267–298. https://doi.org/10.1016/j.jecp.2004.01.001.
Snowling, M. J. (2014). Dyslexia: A language learning impairment. *Journal of the British Academy, 2*, 43–58. https://doi.org/10.5871/jba/002.043.
Spielend Lernen Verlag. (2018). *Richtig konzentrieren (version 1.1)*. Wörthsee: Spielend Lernen Verlag. Retrieved from https://play.google.com/store/apps/details?id=de.fragenbaer.konzentrieren.HD&hl=de.
Spinelli, D., De Luca, M., Di Filippo, G., Mancini, M., Martelli, M., & Zoccolotti, P. (2005). Length effect in word naming in reading: Role of reading experience and reading deficit in Italian readers. *Developmental Neuropsychology, 27*(2), 217–235. https://doi.org/10.1207/s15326942dn2702_2.
Sung, Y.-T., Chang, K.-E., & Liu, T.-C. (2015). The effects of integrating mobile devices with teaching and learning on students' learning performance: A meta-analysis and research synthesis. *Computers & Education, 94*, 252–275. https://doi.org/10.1016/j.compedu.2015.11.008.
Susi, T., Johannesson, M., & Backlund, P. (2007). *Serious games - an overview (HS- IKI -TR-07-001)*. Retrieved from https://www.diva-portal.org/smash/get/diva2:2416/FULLTEXT01.pdf.
Sweetser, P., & Wyeth, P. (2005). GameFlow: A model for evaluating player enjoyment in games. *Computer Entertainment, 3*(3). https://doi.org/10.1145/1077246.1077253.
Thaler, V., Ebner, E. M., Wimmer, H., & Landerl, K. (2004). Training reading fluency in dysfluent readers with high reading accuracy: Word specific effects but low transfer to untrained words. *Annals of Dyslexia, 54*(1), 89–113. https://doi.org/10.1007/s11881-004-0005-0.
Van Gorp, K., Segers, E., & Verhoeven, L. (2017). The role of feedback and differences between good and poor decoders in a repeated word reading paradigm in first grade. *Annals of Dyslexia, 67*(2017), 1–25. https://doi.org/10.1007/s11881-016-0129-z.
Vellutino, F. R., Fletcher, J. M., Snowling, M. J., & Scanlon, D. M. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychology and Psychiatry, 45*(1), 2–40. https://doi.org/10.1046/j.0021-9630.2003.00305.x.
Walter, J. (2013). *Verlaufsdiagnostik sinnerfassenden Lesens*. Göttingen: Hogrefe.
Weiß, R. H. (2006). *Grundintelligenztest skala 2-Revision CFT 20-R*. Göttingen: Hogrefe.
Weiß, R. H., & Osterland, J. (2012). Grundintelligenztest skala 1-Revision CFT 1-R [culture fair intelligence test scale 1-revision]. *Göttingen: Hogrefe*.
Williams, C., & Beam, S. (2019). Technology and writing: Review of research. *Computers & Education, 128*, 227–242. https://doi.org/10.1016/j.compedu.2018.09.024.
Williams, K., Walker, M. A., Vaughn, S., & Wanzek, J. (2017). A synthesis of reading and spelling interventions and their effects on spelling outcomes for students with learning disabilities. *Journal of Learning Disabilities, 50*(3), 286–297. https://doi.org/10.1177/0022219415619753.
Wolf, K. M., Schroeders, U., & Kriegbaum, K. (2016). Metaanalyse zur Wirksamkeit einer Förderung der phonologischen Bewusstheit in der deutschen Sprache. *30*(1), 9–33. https://doi.org/10.1024/1010-0652/a000165.
Wouters, P., van Nimwegen, C., van Oostendorp, H., & van der Spek, E. D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology, 105*(2), 249–265. https://doi.org/10.1037/a0031311.
Wrzesien, M., & Alcañiz Raya, M. (2010). Learning in serious virtual worlds: Evaluation of learning effectiveness and appeal to students in the E-Junior project. *Computers & Education, 55*, 178–187. https://doi.org/10.1016/j.compedu.2010.01.003.
Wyschkon, A., Schulz, F., Sunnyi Gallit, F., Poltz, N., Kohn, J., Moraske, S., … Esser, G. (2017). 5-Jahres-Verlauf der LRS: Stabilität, Geschlechtseffekte, Schriftsprachniveau und Schulerfolg. *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie*, 1–16. https://doi.org/10.1024/1422-4917/a000535.
Zheng, M., & Spires, H. A. (2014). Fifth Graders' flow experience in a digital game-based science learning environment. *International Journal of Virtual and Personal Learning Environments, 5*(2), 69–86. https://doi.org/10.4018/ijvple.2014040106.
Ziegler, J. C., Perry, C., Ma-Wyatt, A., Ladner, D., & Schulte-Körne, G. (2003). Developmental dyslexia in different languages: Language-specific or universal? *Journal of Experimental Child Psychology, 86*(3), 169–193. https://doi.org/10.1016/S0022-0965(03)00139-5.

Supplementary material

	Task name	Description	Skills	Module	Levels	Task frequency (n=90)
1	Ice mirror	The children help to solve the ice mirrors riddle by segmenting words into syllables.	Syllable segmentation	PA	3	1
2	Crystal ball	Children are asked to decrypt secret messages from the crystal ball by counting syllables.	Syllable counting	PA	3	3
3	Port	By loading boxes on boats, the children are asked to identify the correct position of sounds within words.	Identification of phoneme positions	PA	3	2
4	Suitcase	The children pack a suitcase by sorting heard words based on a specified sound.	Phoneme identification	PA	3	2
5	Bedouin camp	At a market, the children are supposed to help the Bedouins to shorten their secret language by deciding whether a word is pronounced correctly after omitting a certain sound.	Phoneme deletion	PA	3	2
6	Cliff keeper 1	Children learn the magical language Namagi by choosing the correct word that corresponds to the series of orally presented sounds.	Phoneme blending	PA	2	3
7	Cliff keeper 2	Children learn to understand the magical language Namagi by deciding whether an orally presented word is correctly broke down into its component sounds.	Phoneme segmentation	PA	2	2
8	Frog love	Children are asked to help the frog crossing the lotus pond by choosing lotus leaves representing sound pairs of long and short vowels.	Long vowel identification	PA	2	1
9	Morning exercises	Children are supposed to decide whether the vowel in an orally presented word is short or long. By doing this the children help the frog to carry out morning exercises.	Long and short vowel differentiation	PA	2	1
10	Maze	To find their way out of a maze, the children have to choose one of three orally presented sounds that	Phoneme-grapheme mapping	PGC	8	4

		correspond to a certain letter.				
11	Fire ghosts	By the rapid allocation of heard sounds to letters, fire spirits are frozen on a volcano.	Automatization of phoneme-grapheme mapping	PGC	24	6
12	Fire seeds	The children help a fire seed to cross lava by assigning single sounds or sound combinations to the corresponding letters.	Grapheme-phoneme mapping	PGC	9	4
13	Fish friends	The children feed fish by quickly tapping on the food package with the correct letter or letter combination corresponding to the heard sound or sound combination.	Automatization of grapheme-phoneme mapping	PGC	24	6
14	Ice quiz	In this task a letter needs to be allocated to one of two orally and visually presented words.	Grapheme-phoneme mapping and phoneme identification in words	PGC	3	4
15	Magical stone circle	A magical stone circle presents pictures of words that are presented orally. Different letters are displayed on the screen and the children are supposed to select the letters that occur in the presented word.	Phoneme-grapheme mapping and phoneme identification in words	PGC	4	4
16	Magical cauldron	The children brew a magic potion by selecting syllables to build a real word and by doing this collecting the ingredients for the magic potion.	Syllable reading and sorting	WR accuracy	8	7
17	Temple gold	Similar to a memory game, the children are supposed to find matching pairs of words and pictures. Thereby they help the chameleon to solve the old mysteries of the temple.	Word-picture mapping	WR accuracy	16	6
18	Kobold apes	A word is orally presented, and the correct written form of the word must be chosen from alternatives to help the apes climbing on trees.	Spoken to written word mapping	WR accuracy	13	9
19	Magical feather	To conjure up magical creatures, the children must identify the magic words (pseudowords) presented within a list of real words.	Identification of pseudowords	WR accuracy	14	8

20	Pearl diver	The children collect pearls by first recognizing prefixes and suffixes and then adding base words to prefixes or suffixes.	Identification of prefixes, suffixes and creation of new words	WR speed	8	5
21	Enchanted stories	The children are supposed to fix up enchanted stories by sorting words to a certain semantic field as fast as possible.	Semantic mapping of words	WR speed	9	5
22	Flying carpet	The children have to evade sandstorms with the flying carpet by quickly deciding whether a displayed picture and a written word match.	Word reading speed	WR speed	7	5

PA=Phonological awareness, PGC=Phoneme-grapheme correspondence, WR=Word reading

DANKSAGUNG

An dieser Stelle möchte ich mich bei all denjenigen bedanken, die mich bei meinem Promotionsvorhaben unterstützten.
Mein besonderer Dank gilt Herrn Prof. Dr. Schulte-Körne, der mir die Erstellung dieser Dissertation ermöglichte und sich stets mit seiner Expertise einbrachte. Ganz herzlich möchte ich mich auch bei meiner Betreuerin Frau PD Dr. Kristina Moll bedanken, die mich während der gesamten Promotionszeit hervorragend und engagiert betreute und somit sehr zum Gelingen dieser Arbeit beitrug.

Auch danke ich den Ko-Autorinnen und Ko-Autoren sowie allen weiteren Kolleginnen und Kollegen für die intensive und bereichernde Zusammenarbeit. Ebenso möchte ich mich bei allen beteiligten Studienassistentinnen, Hilfskräften sowie den Praktikantinnen und Praktikanten für die Mithilfe und das Engagement bedanken. Eine besondere Anerkennung gilt außerdem den Kindern und Eltern, die an den Studien teilnahmen und somit einen wertvollen Beitrag zu dieser Dissertation leisteten.

Von ganzem Herzen möchte ich mich bei meiner Familie, meinem Freund, meiner Freundin Katharina und meinen anderen lieben Freundinnen bedanken, die mir mit Geduld, Zuversicht und Hilfsbereitschaft zur Seite standen. Danke, dass ihr allzeit ein offenes Ohr für mich habt und immer an mich glaubt.

LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Promotionsbüro Medizinische Fakultät

MMRS

EIDESSTATTLICHE VERSICHERUNG

Ich, Ruth Maria Görgen, erkläre hiermit an Eides statt, dass ich die vorliegende Dissertation mit dem Titel

„Zur Entwicklung digitaler, evidenzbasierter Instrumente für die Erfassung und Förderung von Lese- Rechtschreibfähigkeiten bei Grundschulkindern“

selbständig verfasst, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

Ich erkläre des Weiteren, dass die hier vorgelegte Dissertation nicht in gleicher oder in ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades eingereicht wurde.

Köln, 17.12.2021

Ruth Maria Görgen

Ruth Maria Görgen