Case Study 1: An Evidence-Based Practice Review Report

Theme: School (setting) based interventions for children with special educational needs (SEN)

How effective are computer or tablet-based reading interventions in raising reading standards for poor readers or those at-risk of reading difficulties?

1. Summary

Pupils with low reading attainment at the end of primary school are at higher risk of poor performance across a range of subjects in secondary school (Brännlund et al., 2017). Poor reading skills are associated with weak phonological awareness (Carroll et al., 2011). One resource to support phonological acquisition and reading development is computer or tablet-based reading interventions. They offer personalised learning dependent on the needs of the individual and provide high-levels of autonomy for the learner, resulting in increased motivation to engage in content (Cheung & Slavin, 2012). Prior reviews have focused on a broader sample of pupils with reading difficulties (Alqahtani, 2020) however, this review exclusively examined the effectiveness of computer or tablet-based interventions in improving the reading skills of those described as poor readers or at-risk of reading difficulties. Five studies met the inclusion criteria for the current review and were evaluated via Gough's (2007) Weight of Evidence framework to critically appraise their methodological rigour and applicability in answering the research question. Through this analysis, limited evidence was found for the exclusive use of computer or tablet-based interventions to improve reading outcomes for this at-risk group. These findings are discussed in more detailed alongside recommendations for further research and practice implications.

2. Introduction

Reading Skills development

In the United Kingdom (UK), one in four children leave primary school unable to read to the required standard (Department for Education [DfE], 2019). Poor reading at this age results in reduced academic engagement and motivation (Rabiner et al., 2016), and leaves pupils unable to benefit from opportunities for learning through reading. Reading below the expected standard is also negatively associated with mental wellbeing (National Literacy Trust, 2018) and poor educational attainment in secondary school (Brännlund et al., 2017). Therefore, because many reading difficulties can be prevented by early, intensive intervention, it is necessary to determine appropriate support in primary schools (Partanen & Siegel, 2013; Wanzek et al., 2018). Vital to the process of becoming a successful reader are phonological skills, which aid sight-reading accuracy (Carroll et al., 2011; Ehri et al., 2001; Department for Education and Skills, 2006). Therefore, interventions designed to improve word reading need to focus on skills underpinning this, such as grapheme-phoneme correspondence and phonological awareness (Jeffes, 2015)

Computer and tablet-based interventions

Intensive, one-to-one tutoring focusing explicitly on phonological development can have a positive effect on at-risk readers' learning outcomes (Wanzek & Vaughn, 2007). However, these interventions can be costly in terms of practitioners' time. In comparison, computer or tablet-based interventions, can be offered without high-cost implications (Cheung & Slavin, 2012) due to possible high staff to pupil ratio. The

majority of primary schools have laptops or tablets for pupils to use (DfE, 2021). Subsequently, a range of computer or tablet-based interventions have emerged to provide individual intervention to promote phonological awareness. These interventions enhance outcomes as they provide varied instruction for pupils and their content is easily accessible and user-friendly. Furthermore, they provide instant feedback, meaning each pupil has infinite opportunity to practice and refine reading skills (Cheung & Slavin, 2012).

Direct Instruction Model

The Direct Instruction Model posits that effective teaching is a result of several factors: quality and appropriateness of instruction, sufficient time for consolidation of learning and incentives for pupils to learn (Slavin, 1994). This model of effective instructional teaching underpins the content and structure of computer and tablet-based interventions (Cheung & Slavin, 2012). They are well positioned to provide well-organised and interesting lessons through their lively, structured content and deploy a direct instruction model systemically to all students thus helping to alleviate influence of individual teacher competence (Ostiz-Blanco et al., 2021). Furthermore, they cater for a range of reading rates, offer paced activities and provide greater specific explanation and scaffolding, meaning reading instruction is tailored to individual needs. Computer and tablet-based instruction also provide adequate, consistent instructional time not limited to staff capacity. Subsequently, pupils have time to fully consolidate learning dependent upon their understanding.

Motivation

Students' motivation to engage in learning tasks effects their attainment and progress (Gottfried, 1985). Effective teaching provides engaging tasks for students and increases their motivation to learn. As explored through the self-determination theory of motivation (Cook & Artino, 2016), if a child is extrinsically motivated to complete a task, they may do so because they want to earn a reward or avoid punishment. Subsequently, motivation for task completion is out of their control and driven by external factors. However, when intrinsically motivated, a child will complete a task because of the sense of personal satisfaction task-completion may bring. They feel a sense of personal motivation and will attribute task completion, and subsequent success or failure, to be in their control. Due to their design, computer or tablet-based intervention can trigger this intrinsic motivation. They intend to provide the appropriate level of challenge and target specific needs therefore, increasing pupils' feelings of autonomy over their learning and motivation for task completion (Cheung & Slavin, 2012).

Rationale

Computer and tablet-based interventions targeting phonological skills, and their impact on at-risk students' reading attainment, is less researched than their impact on other groups. Recent reviews have explored impact on general population (Ostiz-Blanco et al., 2021) or those with a range of reading difficulties, such as dyslexia (Alqahtani, 2020). As poor readers, in the early stages of primary school, are more like to develop future reading difficulties and have poorer long-term attainment, it is important to conduct a review based solely on this at-risk sample. For this group, computer or tablet-based interventions may be more motivating as they provide a level of learning autonomy not provided by traditional teaching methods. They are

also low-cost, more adaptable to individual performance, can be tailored to the needs of each child and provide opportunities for extensive practice, meaning they may more effectively target this at-risk population than alternative methods. Given this, and the possible potential of this form of reading intervention, it is important to conduct this up-to-date review to verify if they are effective in improving outcomes for poor readers or those at-risk of reading difficulties.

Link to EP practice

Educational Psychologists (EPs) must suggest appropriate intervention to promote change for children or young people. To achieve this, and make the desired evidence-based recommendations, they need to understand an intervention's empirical research (Boyle & Kelly, 2017) and whether its implementation may lead to successful outcomes in a school-context. Considering this, it is important to determine whether computer or tablet-based reading interventions, are suitable for EPs to recommend to schools for struggling readers. Through exploring the current literature base and critically appraising the methodological rigour of each study, this review will aim address the following review question:

How effective are computer or tablet-based reading interventions in raising reading standards for poor readers or those at-risk of reading difficulties?

3. Critical Review of the Evidence Base

Literature Search

A systematic review of the literature was conducted on the 6th January 2022 using three electronic databases, ERIC (EBSCOhost), PsycInfo and Web of Science. Specific search terms used are outlined in Table 1. Searches were conducted with 'written in English' and 'written after 2013' as search criteria and terms were searched under 'abstract' and 'title' for all databases.

Table 1

Search terms used

	Computer-based	Intervention	Reading	Primary-aged
OVID (PsycINFO)	computer-based OR tablet OR computer-assisted OR computeri?ed OR technology assisted OR technology-based	intervention OR program*	reading OR phonic* OR literacy	primary OR elementary OR school OR kindergarten
ERIC (EBSCO) and Web of Science	"computer-based" OR tablet OR "computer-assisted" OR computeri?ed OR "technology assisted" OR "technology-based"	intervention OR program*	reading OR phonic* OR literacy	primary OR elementary OR school OR kindergarten

As illustrated via the flowchart in figure 1, following this initial search, 215 studies were identified. After removal of duplicates, 157 studies were subjected to title and abstract screening using the review-specific inclusion and exclusion criteria (see Table 2). Subsequently, 18 articles were included for full-text screening using the same criteria. After full-text screening, 13studies were excluded (see Appendix A for excluded studies and rationale) and the remaining 5 studies were included in this review (see Table 3 for included studies).

Figure 1

Flowchart of the literature search process

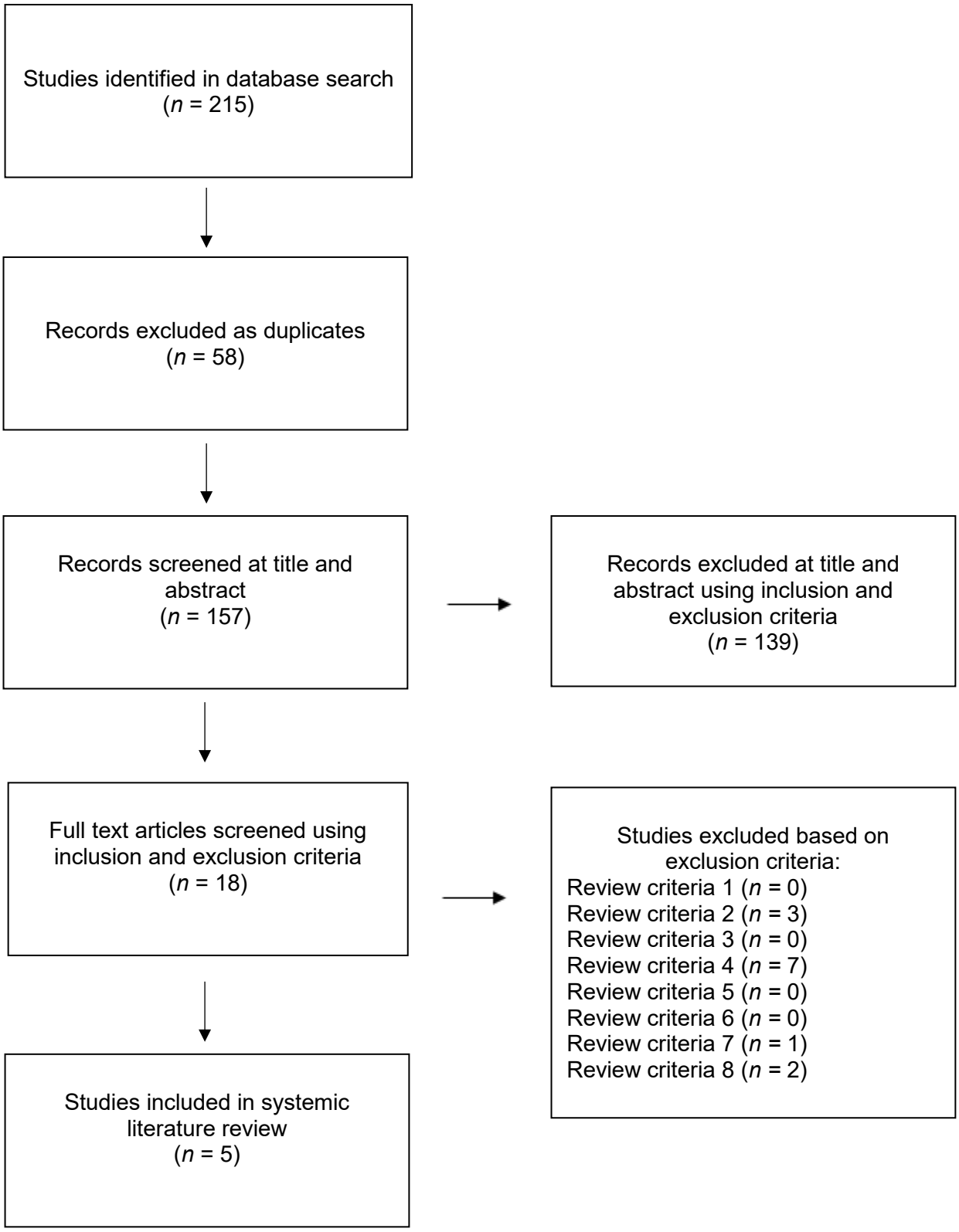

Table 2

Inclusion and exclusion criteria

Criteria	Inclusion	Exclusion	Rationale
1. Population	Primary-aged children identified as low performing or at-risk of reading difficulties	Typically developing primary-aged children	This review is looking specifically at poor readers or those at-risk of reading difficulties
2. Setting	Mainstream schools in an Organisation for Economic Co-operation and Development (OECD) country	Non-mainstream schools Schools not in an OECD country	To ensure that review findings can be generalised to UK primary school population (Gersten et al., 2005)
3. Intervention	Intervention facilitated via a computer or tablet during school hours	Interventions not facilitated via a computer or tablet during school hours	The focus of this review is on computer or technology-based intervention facilitated via a computer or tablet during school hours
4. Study design	Studies using a group experimental design	Studies not using a group experimental design	Due to the research question exploring effectiveness, group experimental designs (e.g. randomised controlled trials (RCTs), quasi-experimental designs and cohort studies) are shown to be best-suited (Petticrew, & Roberts, 2003)
5. Date	From 2014-2021	Pre-2014	Prior systematic literature review published in 2013 (Cheung & Slavin, 2013)

Criteria	Inclusion	Exclusion	Rationale
6. Language	Studies published in English	Studies not published in English	Researcher has a limited ability to accurately translate and interpret articles into English
7. Outcome measure	Studies must use a measure of reading	Any other outcome measure	The focus of this review is on reading outcomes
8. Publication	Studies have been published in a peer-reviewed journal	Studies not published in a peer review journal	As part of the reviewing process, studies published in a peer-reviewed journal, have high-levels of internal validity and originality (Kelly et al., 2014)

Table 3

Included studies

Full Study Reference
D'Agostino, J. V., Rodgers, E., Harmey, S., & Brownfield, K. (2016). Introducing an iPad app into literacy instruction for struggling readers: Teacher perceptions and student outcomes. *Journal of Early Childhood Literacy, 16*(4), 522–548. https://doi.org/10.1177/1468798415616853
Kreskey, D. D., & Truscott, S. D. (2015). Is Computer-Aided Instruction an Effective Tier-One Intervention for Kindergarten Students at Risk for Reading Failure in an Applied Setting? *Contemporary School Psychology, 20*(2), 142–151. https://doi.org/10.1007/s40688-015-0056-8
Messer, D., & Nash, G. (2017). An evaluation of the effectiveness of a computer-assisted reading intervention. *Journal of Research in Reading, 41*(1), 140–158. https://doi.org/10.1111/1467-9817.12107
Rosas, R., Escobar, J.-P., Ramírez, M.-P., Meneses, A., & Guajardo, A. (2017). Impact of a computer-based intervention in Chilean children at risk of manifesting reading difficulties / Impacto de una intervención basada en ordenador en niños chilenos con riesgo de manifestar dificultades lectoras. *Infancia Y Aprendizaje, 40*(1), 158–188. https://doi.org/10.1080/02103702.2016.1263451
Storey, C., McDowell, C., & Leslie, J. C. (2019). Headsprout Early Reading for Specific Literacy Difficulty: A Comparison Study. *Journal of Behavioral Education, 29*. https://doi.org/10.1007/s10864-019-09336-7

Weight of Evidence (WoE)

Gough's (2007) Weight of Evidence (WoE) framework was used to determine the quality and relevance of each research design and execution in answering the review question. Each study's WoE A has been evaluated using a set of methodological quality indicators for group experimental design (Gersten et al., 2005). To allow for greater specificity when evaluating each study, this indicator was adapted (see Appendix B for justification). Following this WoE A evaluation, each study was then given a WoE A rating (see Appendix B).

WoE B is used to determine the relevance of each study's design in answering the research question. For the purpose of this review, WoE B was concluded based on the appropriateness of each study's design in answering the question pertaining to the intervention's effectiveness (Petticrew & Roberts, 2003) (see Appendix C for WoE B ratings for each study).

WoE C considers the relevance of the focus of the study in answering the review question. Studies were evaluated based on a review-specific criterion designed to determine how the focus of each study addresses the review question (Gough, 2007). (See Appendix D for criteria and each study's WoE C rating).

To determine the extent to which each study contributes to answering the review question, these three judgements average WoE A, B and C ratings were combined to provide a WoE D rating (Gough, 2007). Table 4 provides a summary of the WoE ratings for each study.

Table 4

Summary of weight of evidence (WoE) ratings

Studies	WoE A	WoE B	WoE C	WoE D
D'Agostino et al. (2016)	2 (Medium)	2 (Medium)	1.7 (Low)	1.9 (Low)
Kreskey & Truscott (2015)	1 (Low)	2 (Medium)	2 (Medium)	1.7 (Low)
Messer & Nash (2016)	3 (High)	3 (High)	2.7 (High)	2.9 (High)
Rosas et al. (2017)	1 (Low)	2 (Medium)	1.7 (Low)	1.6 (Low)
Storey et al. (2019)	2 (Medium)	2 (Medium)	2.3 (Medium)	2.1 (Medium)

Note: ≤ 2 = low, ≥ 2 to ≤ 2.5 = Medium, ≥ 2.5 = High

Participants

The five studies in this review included 337 participants from OECD countries. Two studies *(*Messer & Nash, 2016; Storey et al., 2019*)* included participants from the UK, which is reflected in their WoE A rating and increased generalisability of findings to the UK school context. In accordance with inclusion criteria, all participants were of primary age and all interventions took place in a school setting. Each study included a sample of pupils either at-risk of reading failure or those making poor progress in reading. Within their inclusion criteria, two studies (Kreskey & Truscott, 2015; Rosas et al., 2017) sampled pupils of reduced socio-economic status. Subsequently, results from these studies can only be generalised to this specific population, which contributed to their low WoE A rating.

Three studies (D'Agostino et al., 2016; Kreskey & Truscott, 2015*;* Messer & Nash, 2016*)* identified their sample via validated measures of reading attainment while one study *(*Rosas et al., 2017*)* used in-school progress and attainment data to determine an appropriate sample. One study *(*Storey et al., 2019*)* established a clear inclusion criterion for intervention. The homogeneity of this supported its higher WoE A rating (Gersten et al., 2005). Four of the studies (D'Agostino et al., 2016*;* Messer & Nash, 2016; Rosas et al., 2017; Storey et al., 2019) randomly assigned the participants to either the intervention or the control group, thus increasing the likelihood of equivalent groups in each condition. One study *(*Kreskey & Truscott, 2015*)* did not randomly assign pupils to each condition, which increased the risk of heterogeneity between groups and contributed to its low WOE A rating. Although random assignment supports equality in participants between conditions, it cannot control for it completely (Gersten et al., 2005). Two studies *(*D'Agostino et al., 2016; Kreskey & Truscott, 2015) utilised random assignment and also reported non-significant differences in the intervention and control group. This demonstrates an attempt to manage between-groups comparability and reduce the possibility that outcomes were due to pre-existing differences between groups (Gersten et al., 2005).

In reporting drop-out rates of participants, researchers are evidencing if study groups are comparable post-intervention (Gersten et al., 2005). In this review none of the studies reported attrition rates (see Appendix B for further information). In the D'Agostino et al. (2016) study, there was no participant drop out. Whereas, in the Kreskey and Truscott (2015) and Rosas et al. (2017) studies, dropout rates could not be determined due to insufficient information, contributing to their low WoE A ratings. In the Messer and Nash (2016) study, post-intervention absentees were

acknowledged and data was re-assessed to ensure this had not impacted on the mean age of both groups. Student drop out was described in the Storey et al. (2019) study and data was removed from each group so they remained of equal size. Although neither of these studies directly reported attrition, their methodological rigour, in terms of attempts to ensure comparable groups post-intervention, contributed to their higher WoE A ratings than the Kreskey and Truscott (2015) and Rosas et al. (2017) studies.

Research design

All studies conducted either a randomised, controlled trial (RCT) (Messer & Nash, 2016) or utilised a group experimental design *(*D'Agostino et al., 2016*;* Kreskey & Truscott, 2015; Rosas et al., 2017; Storey et al., 2019*)* (see Appendix E for further details). As a result, all studies received medium to high WOE B ratings (see Appendix C) as their study design was appropriate to answer the review question (Petticrew & Roberts, 2003). Four studies (D'Agostino et al., 2016, Kreskey & Truscott, 2015; Rosas et al., 2017; Storey et al., 2019) deployed a treatment as usual control group. However, one study (Messer & Nash, 2016) utilised a waitlist control group, allowing for an evaluation of difference between shorter and longer intervention periods. As this study provided further evidence of intervention fidelity, related to optimum time required for intervention effectiveness, it received the highest WOE A rating.

Intervention

Intervention type varied in most of the studies. Two studies (Kreskey & Truscott, 2015; Storey et al., 2019) implemented the 'Headsprout' (Layng et al., 2003)

internet-based reading programme to teach letter sounds, segmenting and blending. The Storey et al. (2019) study provided information about the frequency and duration of the intervention with clear guidelines for ensuring consistency of delivery, this contributed to its medium WOE A rating as this information provided a sufficient level of detail for replication in different contexts (Carroll, 2007). In contrast, Kreskey and Truscott's (2015) study provided limited specificity regarding intervention implementation and whether this adhered to guidelines. The number of recommended sessions was unclear and the average amount of completed sessions varied for each participant; this unclear intervention detail resulted in poor fidelity of implementation and contributed to its low WOE A rating.

The remaining three studies utilised different interventions. Messer and Nash (2016) used a computer-based programme 'Trainertext' (EasyRead, 2018) to support the acquisition of phonics. This programme used visual mnemonics to supply relevant clues about individual phonemes encouraging the child to read unfamiliar words. This study had high internal validity as intervention implementation was in line with the guidelines recommended by Trainertext (EasyRead, 2018) and regular monitoring visits were conducted to confirm appropriate intervention administration, contributing to its high WoE C rating.

In the Rosas et al. (2017) study, GraphoGame was used to consolidate grapheme-phoneme correspondence. However, the amount of intervention received was unreported and subsequently it was unclear whether this was consistent with programme recommendations. This lack of clarity contributed to its low WoE A rating. The D'Agostino et al. (2016) study used the Reading Recovery (Clay, 1987)

intervention, whereby trained teachers focused on pupils' phonetic understanding using an iPad application. Whilst this study clearly described intervention implementation frequency and monitoring, which contributed to its medium WoE A rating, it used a technology-app to complement existing intervention. As a result, this study was deemed less applicable in answering the review question, which resulted in a low WoE C rating.

Measures

All studies measured either phonological acquisition or word-reading accuracy, either using standardised assessments (D'Agostino et al., 2016; Kreskey & Truscott, 2015, Messer & Nash, 2016; Storey et al., 2019) or curriculum-based measures (Rosas et al., 2017). Two studies determined social validity via interview (D'Agostino et al., 2016) or questionnaire (Storey et al., 2019). Due to this review's focus on reading accuracy, and the inclusion criteria pertaining to experimental design, this qualitative data was not further explored. Three studies (D'Agostino et al., 2016, Messer & Nash, 2016; Storey et al., 2019) used either blind assessors or determined inter-observer's agreement when scoring pre- and post-intervention measures. This limited any inadvertent data collection bias thus strengthening their internal validity (Gersten et al., 2005) and contributing to their medium and high WOE A ratings (see Appendix B).

Outcomes and effect sizes

A description of each study's outcome measure and effect sizes are detailed in Table 5. Cohen's d (Cohen, 1988) effect size was reported for three studies (D'Agostino et al., 2016; Messer & Nash, 2016; Storey et al., 2019) reflected in their medium to high

WOE A ratings. The remaining two studies (Kreskey & Truscott, 2015; Rosas et al., 2017), did not report effect sizes. Therefore, these were calculated using the Campbell Collaboration Calculator (Wilson, 2022) via means and standard deviations from post-intervention data (see Appendix F for further information).

Three of the studies (D'Agostino et al., 2016; Messer & Nash, 2016; Storey et al., 2019) reported large effect sizes for at least one measure supported by statistically significant differences between pre-and post-data for each intervention (see Table 5 for further detail). Therefore, suggesting that the interventions resulted in a significant change in reading accuracy for the intervention group comparative to the control group. The Storey et al. (2019) study reported a very large effect size (Sullivan & Feinn, 2012). Due to the study's small sample size and unequal intervention and control samples, these results need to be interpreted with caution (Slavin & Smith, 2009).

The remaining studies (Kreskey & Truscott, 2015; Rosas et al., 2017) reported small to medium negative effect sizes, indicating that post intervention results were lower than those pre-intervention. Three subsets of the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) were used in the Kreskey and Truscott (2015) study to determine reading accuracy. However, only two subsets were reported post-intervention, meaning intervention effectiveness could not be determined across all measures. These unclear, incoherent results contributed to its low WoE A rating. The Rosas et al. (2017) study reported effect sizes for all curriculum-based measures used. However, significant differences were reported between intervention and control groups pre-intervention. Therefore, effects cannot be solely attributed to inclusion in the intervention and need to be interpreted with caution.

Table 5

Effect sizes and WoE D quality indicators for each study

Study	Measures of reading accuracy	Number of participants	Sample	Effect size (d) (within groups, pre-post test)	Descriptor	Significance values	WOE D
D'Agostino et al. (2016)	DIBELS	50	6-7 years old	0.82	Large	p < .05	1.9 (Low)
	Observation Survey of Early Literacy Achievement (OSELA) subsets: Letter identification			0.85	Large	not reported	
	Sounding in words			0.82	Large	not reported	
Kreskey and Truscott (2015)	DIBELS subtests: Letter Naming Fluency	102	Reception	-0.37	Small	p .108	1.7 (Low)
	Nonsense Words Fluency			-0.56	Medium	not reported	

Study	Measure of reading accuracy	Number of participants	Sample	Effect size (d) (within groups, pre-post test)	Descriptor	Significance values	WOE D
Messer and Nash (2016)	Test of Word Reading Efficiency (TOWRE)	78	Mean age 7 years	0.97	Large	p < 0.001	2.9 (High)
	Phonological Assessment Battery (PHAB)			0.27	Small	p < 0.001	
Rosas et al. (2017)	Curriculum-based measures	75	6-7 years old	-0.32	Small	p 0.370	1.6 (Low)
Storey et al. (2019)	Phonics and Early Reading Assessment (PERA)	32	6-9 years old	2.65	Large	p 0.0001	2.1 (Medium)
	Sentence reading			0.96	Large	p 0.002	
	Dolch words			1.53	Large	p 0.0001	

Note: An effect size of less than 0.2 is considered small, 0.5 is medium and 0.8 or above is large (Cohen, 1988).

4. Conclusions and Recommendations

Conclusions

The purpose of this review was to determine the effectiveness of computer or tablet-based phonics-focused interventions on improving reading skills for poor readers or those at-risk of reading difficulties. The result was five studies that used a variety of computer-based programmes or tablet applications to promote phonological acquisition and improve reading skills.

From this review, it has been concluded that three out of the five studies (D'Agostino et al., 2016; Messer & Nash, 2016; Storey et al., 2019) reported a positive small to large effect sizes and statistically significant results, when using computer or tablet-based intervention to improve reading skills. Two of the five studies (Kreskey & Truscott, 2015; Rosas et al., 2017) reported small, negative effect sizes with non-significant results, indicating that the intervention had no impact on the treatment group's reading skills comparative to a control group. Although this would suggest that certain interventions are more effective than others in improving outcomes, these effect sizes need to be considered alongside reported outcomes to determine intervention effectiveness.

Results from two studies (D'Agostino et al, 2016; Messer & Nash, 2016), with a reported large effect size, were based on increases in individual word reading and phonological acquisition. Therefore, it cannot be determined if gains in sight-word reading knowledge were transferrable to sentence reading. Or, whether participants were able to use learnt phonics to read unfamiliar words. Subsequently, it is unclear if acquired skills could be generalised to sentence reading and decoding outside of

the intervention. Furthermore, neither study reported any long-term follow-up effects, therefore, it is unclear whether these interventions had a positive, long-term effect on sentence reading or phonological decoding. In contrast, the very large effect sizes reported in the Storey et al. (2019) study were based on significant increases in single word reading and sentence reading skills. This suggests that general reading skills also improved as a result of the intervention and that this intervention could be recommended to support acquisition of general reading skills. Although, yet again, follow-up data was not reported so the overall, long-term impact on reading outcomes cannot be ascertained.

Furthermore, although the remaining two studies (Kreskey & Truscott, 2015; Rosas et al., 2017) reported negative effect sizes, they also had lower, methodological quality as evidenced through their low WoE A ratings and low overall WoE D rating. Therefore, as poor methodology quality and lower internal validity may have a led to bias in interpretation of intervention outcomes (Chacón-Moscoso et al., 2016), the ineffectiveness of these interventions needs to be interpreted with caution.

Limitations

A limitation of this review is the variability in how the computer or tablet-based technology was implemented and utilised across studies. Two of the studies (Rosas et al., 2016; Storey et al., 2019) used trained researchers to conduct the intervention. School studies conducted using researchers instead of staff typically have weaker external validity due to their inability to replicate features of everyday practice. Subsequently, positive outcomes from these studies can be hard to replicate in real-life contexts (Fredrickson, 2002). This suggests that if the interventions were

conducted by school staff, they may not garner the same results. Therefore, reported effect sizes need to be interpreted with caution due to limited generalisability of findings and implementation to school contexts. The D'Agostino et al. (2016) study, randomised teachers to the experimental and control conditions to try to control differences between groups. However, teacher implementation proficiency was not assessed or evaluated at any time. Subsequently, the impact of individual teacher ability to deliver the programme acted as an uncontrolled variable. This means that results cannot be wholly attributed to the intervention as teacher effectiveness may have influenced outcomes.

A suggested benefit of computer or tablet-based interventions is their cost-effectiveness, as they require limited staff-pupil interaction. A further limitation of this review was each study's inability to define whether the intervention was conducted with minimal staff input or if it was complemented by additional one-to-one intervention. Studies (D'Agostino et al., 2016; Messer & Nash, 2017; Storey et al., 2019), which reported large effect sizes, also included more interaction between students and intervention facilitator than those reporting smaller, negative effect sizes (Kreskey & Truscott, 2015; Rosas et al., 2017). Within the Messer and Nash (2017) study, staff provide an unspecified level of individual attention to pupils who struggled. Whereas the Story et al. (2019) study included an element of independent reading post-session, without reported clarity on the specific skills targeted in these sessions. This means that improvements cannot be entirely attributed to the computer-based intervention and that the interventions required higher staff input, equating to higher running costs.

Recommendations

In schools, highly valued amongst school staff is an EPs ability to recommend of strategies to support pupils (Ashton & Roberts, 2006). However, EPs need to ensure that the practice shared is based on evidence of its effectiveness (Fox, 2003). While several studies within this review, with high WoE D ratings, found that computer or tablet-based interventions can be used to improve reading outcomes for poor readers or those at-risk of reading difficulties, due to the limiting factors discussed, there is insufficient evidence for EPs to recommend their use for this at-risk group.

As a rationale for their use, all studies suggested computer or tablet-based interventions to be more motivating for students than traditional methods. Using anecdotal evidence from those implementing the intervention, D'Agostino et al. (2016) considered the inherent motivational qualities of tech-based interventions as a key mediating factor in letter-knowledge acquisition for the intervention group. Although the specific contributing, motivational features, were not specified, Messer and Nash (2016) suggested the trialled intervention's ability to adapt to individual pupil knowledge as a key factor in promoting pupil motivation. It was stated this led to increased keenness to engage and greater letter knowledge acquisition. However, data regarding student motivation and engagement, and any specific, observed changes, were not explored. Therefore, limited conclusions can be reached regarding the specific motivational qualities of these interventions. Further research could explore the role of motivation in computer or tablet-based interventions and which specific features increased pupil's motivation to learn and engage.

Within this review, the largest effect sizes were reported for interventions that supplemented the computer or tablet-based instruction with non-computer based, adult-led instruction (D'Agostino et al., 2016; Messer & Nash, 2017; Storey et al., 2019). This suggests that these programmes cannot be conducted with an adult simply overseeing implementation but require elements of one-to-one instruction between adult and pupil to improve overall reading outcomes. The role of staff, alongside the time devoted to computer-based and noncomputer interactions, is integral to their success. Further research is needed to explore the effectiveness of computer or tablet-based interventions, which also utilise a staff-led instructional element. This intervention design may show further promise as it would maintain the immediate feedback and instructionally effective strategies inherent to computer-based games whilst providing the additional support necessary as suggested in this review.

5. References

References

Alqahtani, S. S. (2020). Technology-based interventions for children with reading difficulties: a literature review from 2010 to 2020. *Educational Technology Research and Development*, *68*(6), 3495–3525. https://doi.org/10.1007/s11423-020-09859-1

Ashton, R., & Roberts, E. (2006). What is Valuable and Unique about the Educational Psychologist? *Educational Psychology in Practice*, *22*(2), 111–123. https://doi.org/10.1080/02667360600668204

Brännlund, A., Strandh, M., & Nilsson, K. (2017). Mental-health and educational achievement: the link between poor mental-health and upper secondary school completion and grades. *Journal of Mental Health*, *26*(4), 318–325. https://doi.org/10.1080/09638237.2017.1294739

Boyle, J & Kelly, B. (2017). The Role of Evidence in Educational Psychology. In B Kelly, L Woolfson & J Boyle (Eds.), *Frameworks for Practice in Educational Psychology: A Textbook for Trainees and Practitioners* (pp. 29-43). Jessica Kingsley.

Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., & Balain, S. (2007). A conceptual framework for implementation fidelity. *Implementation Science*, *2*(1).

https://doi.org/10.1186/1748-5908-2-40

Chacón-Moscoso, S., Sanduvete-Chaves, S., & Sánchez-Martín, M. (2016). The Development of a Checklist to Enhance Methodological Quality in Intervention Programs. *Frontiers in Psychology, 7.* https://doi.org/10.3389/fpsyg.2016.01811

Cheung, A. C. K., & Slavin, R. E. (2012). How features of educational technology applications affect student reading outcomes: A meta-analysis. *Educational Research Review*, *7*(3), 198–215. https://doi.org/10.1016/j.edurev.2012.05.002

Cheung, A. C. K., & Slavin, R. E. (2013). Effects of Educational Technology Applications on Reading Outcomes for Struggling Readers: A Best-Evidence Synthesis. *Reading Research Quarterly*, *48*(3), 277–299. https://doi.org/10.1002/rrq.50

Clay, M. (1987). Implementing Reading Recovery: Systemic adaptations to an educational innovation. *New Zealand Journal of Educational Studies, 22*(1), 35–58.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). L. Erlbaum Associates.

Cook, D. A., & Artino, A. R. (2016). Motivation to learn: an overview of contemporary theories. *Medical Education*, *50*(10), 997–1014. https://doi.org/10.1111/medu.13074

D'Agostino, J. V., Rodgers, E., Harmey, S., & Brownfield, K. (2016). Introducing an iPad app into literacy instruction for struggling readers: Teacher perceptions and student outcomes. *Journal of Early Childhood Literacy*, *16*(4), 522–548. https://doi.org/10.1177/1468798415616853

Department for Education. (2019). *National curriculum assessments at key stage 2 in England, 2019 (revised).* https://www.gov.uk/government/statistics/national-curriculum-assessments-key-stage-2-2019-revised

Department for Education. (2021). *Education Technology (EdTech) Survey 2020-21 Research report.* https://www.gov.uk/government/publications/education-technology-edtech-survey-2020-to-2021

Department for Education and Skills. (2006). *Independent review of the teaching of early reading. Final Report.* Rose, J. https://dera.ioe.ac.uk/5551/

Dumville, J. C., Torgerson, D. J., & Hewitt, C. E. (2006a). Reporting attrition in randomised controlled trials. *BMJ, 332*(7547), 969–971. https://doi.org/10.1136/bmj.332.7547.969

EasyRead. (2018, March 6). *Easyread*. David Morgan Education. https://dm-ed.com/info/easyread/

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. https://doi.org/10.3102/00346543071003393

Fox, M. (2003). Opening Pandora's Box: Evidence-based practice for educational psychologists. *Educational Psychology in Practice*, *19*(2), 91–102. https://doi.org/10.1080/02667360303233

Frederickson, N. (2002). Evidence-based practice and educational psychology. *Educational and Child Psychology, 19*(3), 96–111.

Gersten, R., Fuchs, L. S., Compton, D., Coyne, M., Greenwood, C., & Innocenti, M. S. (2005). Quality Indicators for Group Experimental and Quasi-Experimental Research in Special Education. *Exceptional Children*, *71*(2), 149–164. https://doi.org/10.1177/001440290507100202

Gottfried, A. E. (1985). Academic intrinsic motivation in elementary and junior high school students. *Journal of Educational Psychology*, *77*(6), 631–645. https://doi.org/10.1037/0022-0663.77.6.631

Gough, D. (2007). Weight of Evidence: a framework for the appraisal of the quality and relevance of evidence. *Research Papers in Education*, *22*(2), 213–228. https://doi.org/10.1080/02671520701296189

Jeffes, B. (2015). Raising the reading skills of secondary-age students with severe and persistent reading difficulties: evaluation of the efficacy and implementation of a

phonics-based intervention programme. *Educational Psychology in Practice*, *32*(1), 73–84. https://doi.org/10.1080/02667363.2015.1111198

Kelly, J., Sadeghieh, T., & Adeli, K. (2014). Peer Review in Scientific Publications: Benefits, Critiques, & A Survival Guide. *EJIFCC*, *25*(3), 227–243.

Kreskey, D. D., & Truscott, S. D. (2015). Is Computer-Aided Instruction an Effective Tier-One Intervention for Kindergarten Students at Risk for Reading Failure in an Applied Setting? *Contemporary School Psychology*, *20*(2), 142–151. https://doi.org/10.1007/s40688-015-0056-8

Layng, J., Twyman, J., & Stikeleather, G. (2003). Headsprout early reading: Reliably teaching children to read. *Behavioral Technology Today, 3,* 7–20.

Messer, D., & Nash, G. (2017). An evaluation of the effectiveness of a computer-assisted reading intervention. *Journal of Research in Reading*, *41*(1), 140–158. https://doi.org/10.1111/1467-9817.12107

Ostiz-Blanco, M., Bernacer, J., Garcia-Arbizu, I., Diaz-Sanchez, P., Rello, L., Lallier, M., & Arrondo, G. (2021). Improving Reading Through Videogames and Digital Apps: A Systematic Review. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.652948

Partanen, M., & Siegel, L. S. (2013). Long-term outcome of the early identification and intervention of reading disabilities. *Reading and Writing*, *27*(4), 665–684.

https://doi.org/10.1007/s11145-013-9472-1

Petticrew, M., & Roberts, H. (2003). Evidence, hierarchies, and typologies: horses for courses. *Journal of Epidemiology & Community Health*, *57*(7), 527–529. https://doi.org/10.1136/jech.57.7.527

Rabiner, D. L., Godwin, J., & Dodge, K. A. (2016). Predicting Academic Achievement and Attainment: The Contribution of Early Academic Skills, Attention Difficulties, and Social Competence. *School Psychology Review*, *45*(2), 250–267. https://doi.org/10.17105/spr45-2.250-267

Rosas, R., Escobar, J.-P., Ramírez, M.-P., Meneses, A., & Guajardo, A. (2017). Impact of a computer-based intervention in Chilean children at risk of manifesting reading difficulties / Impacto de una intervención basada en ordenador en niños chilenos con riesgo de manifestar dificultades lectoras. *Infancia Y Aprendizaje*, *40*(1), 158–188. https://doi.org/10.1080/02103702.2016.1263451

Slavin, R. E. (1994). Quality, appropriateness, incentive, and time: A model of instructional effectiveness. *International Journal of Educational Research*, *21*(2), 141–157. https://doi.org/10.1016/0883-0355(94)90029-9

Slavin, R., & Smith, D. (2009). The Relationship between Sample Sizes and Effect Sizes in Systematic Reviews in Education. *Educational Evaluation and Policy Analysis*, *31*(4), 500–506. https://doi.org/10.3102/0162373709352369

Solheim, O. J., Frijters, J. C., Lundetræ, K., & Uppstad, P. H. (2018). Effectiveness of an early reading intervention in a semi-transparent orthography: A group randomised controlled trial. *Learning and Instruction*, *58*, 65–79. https://doi.org/10.1016/j.learninstruc.2018.05.004

Storey, C., McDowell, C., & Leslie, J. C. (2019). Headsprout Early Reading for Specific Literacy Difficulty: A Comparison Study. *Journal of Behavioral Education*, *29*. https://doi.org/10.1007/s10864-019-09336-7

Sullivan, G. M., & Feinn, R. (2012). Using effect size—or why the P value is not enough. *Journal of Graduate Medical Education*, *4*(3), 279–282. https://doi.org/10.4300/jgme-d-12-00156.1

Wanzek, J., & Vaughn, S. (2007). Research-Based Implications From Extensive Early Reading Interventions. *School Psychology Review*, *36*(4), 541–561. https://doi.org/10.1080/02796015.2007.12087917

Wanzek, J., Stevens, E. A., Williams, K. J., Scammacca, N., Vaughn, S., & Sargent, K. (2018). Current Evidence on the Effects of Intensive Early Reading Interventions. *Journal of Learning Disabilities*, *51*(6), 612–624. https://doi.org/10.1177/0022219418775110

Wilson, D.B. (2022, May 7). *This is a web-based effect size calculator.* Campbell Collaborati

Appendices contents list

Appendix A

List of studies excluded at full-text screening

Excluded studies reference	Reason for exclusion (criteria number)
Bennett, J. G., Gardner, R., Cartledge, G., Ramnath, R., & Council, M. R. (2017). Second-Grade Urban Learners: Preliminary Findings for a Computer-Assisted, Culturally Relevant, Repeated Reading Intervention. *Education and Treatment of Children, 40*(2), 145–185. https://doi.org/10.1353/etc.2017.0008	4. Concurrent multiple probe experimental design
Council, M. R., Gardner, R., Cartledge, G., & Telesman, A. O. (2019). Improving reading within an urban elementary school: computerized intervention and paraprofessional factors. *Preventing School Failure: Alternative Education for Children and Youth, 63*(2), 162–174. https://doi.org/10.1080/1045988x.2018.1540392	4. Multiple probe design
Council, M. R., Cartledge, G., Green, D., Barber, M., & Gardner, R. (2016). Reducing Risk through a Supplementary Reading Intervention: A Case Study of First- and Second-Grade Urban Students. *Behavioral Disorders, 41*(4), 241–257. https://doi.org/10.17988/bedi-41-04-241-257.1	4. Descriptive study
Mize, M., Bryant, D. P., & Bryant, B. R. (2019). Teaching reading to students with learning disabilities: Effects of combined iPad-assisted and peer-assisted instruction on oral reading fluency performance. *Assistive Technology, 13*(2), 1–8. https://doi.org/10.1080/10400435.2018.1559896	8. Not peer-reviewed
Gibbon, J. M., Duffield, S., Hoffman, J., & Wageman, J. J. (2017) Effects of educational games on sight word reading achievement and student motivation. *Journal of Language and Literacy Education*, *13*(2), 1-27.	2. Non-school-based setting
Keyes, S. E., Cartledge, G., Gibson, L., & Robinson-Ervin, P. (2016). Programming for Generalization of Oral Reading Fluency Using Computer-Assisted Instruction and Changing Fluency Criteria. *Education and Treatment of Children, 39*(2), 141–172. https://doi.org/10.1353/etc.2016.0011	4. Multiple baseline design across participants
Pindiprolu, S. S., & Forbush, D. E. (2021). Comparative Effects of Computer-Based Reading Programs on the Early Literacy Skills of At-Risk Students. *Journal of Educational Technology Systems, 50*(2), 004723952110400. https://doi.org/10.1177/00472395211040048	4. Pre–post-test design

Excluded studies reference	Reason for exclusion (criteria number)
Reutzel, D. R., Petscher, Y., & Spichtig, A. N. (2012). Exploring the Value Added of a Guided, Silent Reading Intervention: Effects on Struggling Third-Grade Readers' Achievement. *The Journal of Educational Research, 105*(6), 404–415. https://doi.org/10.1080/00220671.2011.629693	7. Outcome measure of reading comprehension
Tijms, J., Pavlidou, E. V., & Hoette, H. A. I. (2020). Improvements in reading and spelling skills after a phonological and morphological knowledge intervention in Greek children with spelling difficulties: a pilot study. *European Journal of Special Needs Education, 35*(5), 1–11. https://doi.org/10.1080/08856257.2019.1709702	4. Single group repeated measurement design
Vanden Bempt, F., Economou, M., Van Herck, S., Vanderauwera, J., Glatz, T., Vandermosten, M., Wouters, J., & Ghesquière, P. (2021). Digital Game-Based Phonics Instruction Promotes Print Knowledge in Pre-Readers at Cognitive Risk for Dyslexia. *Frontiers in Psychology, 12*(1). https://doi.org/10.3389/fpsyg.2021.720548	2. Intervention took place at home not in-school
Vincent, K. (2020). Closing the gap: supporting literacy through a computer-assisted-reading-intervention. *Support for Learning, 35*(1), 68–82. https://doi.org/10.1111/1467-9604.12286	8. Not peer-reviewed

Appendix B

Weight of Evidence (WoE) A

Gersten et al.'s (2005) non-review specific coding protocol was used to determine the methodological quality of each study, generating a WoE A. As this framework, and resulting sub-categories are non-review specific, Gertsen et al. (2005) suggests that they can be adapted dependent on research project. Therefore, one of the desirable quality indicators was adapted accordingly (see Table 1 for further detail). Subsequently, each study was coded using the adapted coding protocol (see Table 2 for evidence of each study's coding). Coding was completed, using the Gersten et al. (2005) defined criteria (see Table 3) and each study was given a final WOE A rating (see Table 4).

Table 1

Adaptions made to WoE A coding protocol

Desirable Quality Indicator	Rationale
Did the study provide not only internal consistency reliability but also test–retest reliability and interrater reliability (when appropriate) for outcome measures? Were data collectors and/or scorers blind to study conditions and equally (un)familiar to examinees across study conditions?	Separated into two distinct questions to allow for two specific aspects of study fidelity to be considered

Table 2

Coding protocol (Gersten et al. (2005)

Author	D'Agostino et al. (2016)	Kreskey & Truscott (2015)	Messer & Nash (2017)	Rosas et al. (2017)	Storey et al. (2019)
Essential Quality Indicators					
Quality Indicators for Describing Participants					
1. Was sufficient information provided to determine/confirm whether the participants demonstrated the disability(ies) or difficulties presented?	Yes	Yes	Yes	Yes	Yes
2. Were appropriate procedures used to increase the likelihood that relevant characteristics of participants in the sample were comparable across conditions?	Yes	Yes	Yes	Yes	Yes
3. Was sufficient information given characterizing the interventionists or teachers provided? Did it indicate whether they were comparable across conditions?	Yes	No	Yes	No	Yes
Quality Indicators for Implementation of the Intervention and Description of Comparison Conditions					
1. Was the intervention clearly described and specified?	Yes	Yes	Yes	Yes	Yes

Author	D'Agostino et al. (2016)	Kreskey & Truscott (2015)	Messer & Nash (2017)	Rosas et al. (2017)	Storey et al. (2019)
2. Was the fidelity of implementation described and assessed?	Yes	No	Yes	Yes	Yes
3. Was the nature of services provided in comparison conditions described?	Yes	Yes	Yes	Yes	Yes
Quality Indicators for Outcome Measures					
1. Were multiple measures used to provide an appropriate balance between measures closely aligned with the intervention and measures of generalized performance?	Yes	Yes	Yes	Yes	Yes
2. Were outcomes for capturing the intervention's effect measured at the appropriate times?	Yes	Yes	Yes	Yes	Yes

Author	D'Agostino et al. (2016)	Kreskey & Truscott (2015)	Messer & Nash (2017)	Rosas et al. (2017)	Storey et al. (2019)
Quality Indicators for Data Analysis					
1. Were the data analysis techniques appropriately linked to key research questions and hypotheses? Were they appropriately linked to the unit of analysis in the study?	Yes	No	Yes	Yes	Yes
2. Did the research report include not only inferential statistics but also effect size calculations?	Yes	No	Yes	No	Yes
Number of essential indictors met out of 10	10	6	10	8	10
Desirable Quality Indicators					
1. Was data available on attrition rates among intervention samples? Was severe overall attrition documented? If so, is attrition comparable across samples? Is overall attrition less than 30%?	No	No	No	No	No

Author	D'Agostino et al. (2016)	Kreskey & Truscott (2015)	Messer & Nash (2017)	Rosas et al. (2017)	Storey et al. (2019)
2. Did the study provide not only internal consistency reliability but also test–retest reliability and interrater reliability (when appropriate) for outcome measures?	No	No	Yes	No	No
Were data collectors and/or scorers blind to study conditions and equally (un)familiar to examinees across study conditions?	Yes	No	Yes	No	Yes
3. Were outcomes for capturing the intervention's effect measured beyond an immediate posttest?	No	No	Yes	No	Yes
4. Was evidence of the criterion-related validity and construct validity of the measures provided?	No	No	No	No	No
5. Did the research team assess not only surface features of fidelity implementation (e.g., number of minutes allocated to the intervention or teacher/interventionist following procedures specified), but also examine quality of implementation?	No	No	Yes	No	No

Author	D'Agostino et al. (2016)	Kreskey & Truscott (2015)	Messer & Nash (2017)	Rosas et al. (2017)	Storey et al. (2019)
6. Was any documentation of the nature of instruction or series provided in comparison conditions?	Yes	No	No	No	No
7. Did the research report include actual audio or videotape excerpts that capture the nature of the intervention?	No	No	No	No	No
8. Were results presented in a clear, coherent fashion?	Yes	No	Yes	Yes	Yes
Number of desirable indictors met out of 9	3	0	5	1	3

Table 3

Weight of Evidence (WoE) A criteria for Group Experimental and Quasi-Experimental Research in Special Education

WoE A Rating	Criteria
3 (High)	1. Study meets a9 or more of the essential quality indicators 2. Study meets 4 or more of the desirable quality indicators
2 (Medium)	1. Study meets at 9 or more of the essential quality indicators 2. Study meets between 1 and 3 4 of the desirable quality indicators
1 (Low)	1. Study meets at 9 or less of the essential quality indicators

Table 4

Summary of WoE A for all studies

Study	Essential criteria met	Desirable criteria met	WoE A rating
D'Agostino et al. 2016	**10**	**3**	2 (Medium)
Kreskey & Truscott, 2015	**6**	**0**	1 (Low)
Messer & Nash, 2016	**10**	**5**	3 (High)
Rosas et al. 2017	**8**	**1**	1 (Low)
Storey et al. 2019	**10**	**3**	2 (Medium)

Appendix C

Weight of Evidence (WoE) B

WoE B is a review-specific judgement to determine the appropriateness of a study's design in answering the review question. When determining this, typologies are preferable to hierarchies as they consider the type of study required to answer a specific review question. Table 1 illustrates the preferred study type to determine an intervention's effectiveness derived from a pre-existing 'typology of evidence' (Petticrew & Roberts, 2003). Table 2 provides the WOE B rating for each study.

Table 1

Rationale and criteria for WoE B

Study Design	WoE B rating	Rationale
Randomised controlled trial (RCT)	3 (High)	To determine the effectiveness of a specific-type of intervention, RCTs are considered the most robust study design.
Quasi-experimental and cohort studies	2 (Medium)	
Qualitative research, survey, case-control studies and non-experimental evaluations	1 (Low)	

Table 2

WoE B rating for each study

Study	WoE B rating
D'Agostino et al. 2016	2 (Medium)
Kreskey & Truscott, 2015	2 (Medium)
Messer & Nash, 2016	3 (High)
Rosas et al. 2017	2 (Medium)
Storey et al. 2019	2 (Medium)

Appendix D

Weight of Evidence (WoE) C

WoE C is a review-specific judgement regarding how relevant each study and its findings are in answering a review question. Table 1 outlines the specific criteria used for this review with a rationale for each. Table 2 provides the WoE C rating for each study.

Table 1

WoE C criteria

Criteria	WoE Rating	Descriptor	Rationale
Location of intervention	3	UK schools	Findings from UK schools can be better generalised to inform UK intervention, practice and EP recommendations for appropriate intervention
	2	Schools in OECD countries	
	1	Schools not in OECD countries	
Intervention	3	Computer or tablet-based intervention used in isolation	Focus of the review is to determine the effectiveness of computer or tablet-based interventions. Additional intervention act as confounding variables and limit ability to determine effectiveness.
	2	Computer or tablet-based intervention used alongside additional intervention	
	1	Computer or tablet-based intervention used to complement existing intervention	

Criteria	WoE Rating	Descriptor	Rationale
Outcome measures, reliability and validity	3	High reliability and validity reported for outcome measures	
	2	An element of reliability or validity reported for most outcome measures	
	1	Reliability and validity of outcome measures not reported	

Table 2

WoE C rating for each study

Study	Location of Intervention	Intervention	Outcome measures, reliability and validity	Overall WoE C Rating and Qualitative Descriptor
D'Agostino et al. (2016)	2	1	2	1.7 (Low)
Kreskey & Truscott (2015)	2	2	2	2 (Medium)
Messer & Nash (2016)	3	2	3	2.7 (High)
Rosas et al. (2017)	2	2	1	1.7 (Low)
Storey et al. (2019)	3	2	2	2.3 (Medium)

Note: ≤ 2 = low, ≥ 2 to ≤ 2.5 = Medium, ≥ 2.5 = High

Appendix E

Mapping the field

Author	Study design	Geographical distribution	Participants	Intervention investigated	Context of intervention	Outcome variables measured (relevant to the review question)
D'Agostino et al.(2016)	Experimental study with a double random assignment	USA	**Sample:** *n* = 50 (treatment = 25, control = 25) **Age:** First-grade students (6-7yrs)	Reading recovery using the Letter-works app	**Facilitators:** Reading recovery teachers **Time:** During school day **Pre-post:** 20weeks	Dynamic Indicators of Basic Literacy Skills, Sixth Edition (DIBELS) Observation Survey of Early Literacy Achievement (OSELA)

Author	Study design	Geographical distribution	Participants	Intervention investigated	Context of intervention	Outcome variables measured (relevant to the review question)
Kreskey and Truscott (2015)	Quasi-experimental design	USA	**Sample:** *n* = 102 (treatment = 51 control = 51) **Age:** Kindergarten students	Headsprout Early Reading – computer aided instruction	**Facilitators:** Teachers **Time:** Unclear - some in-school, some before school **Pre-post:** Autumn – Spring	Dynamic Indicators of Basic Literacy Skills, Sixth Edition (DIBELS)
Messer and Nash (2017)	RCT	UK – London	**Sample:** *n* = 78 (treatment = 45 control = 33) **Age**: Mean age 7 years	Trainertext and interactive, multimedia materials.	**Facilitators:** Teaching assistants **Time:** During school day **Pre-post:** 10months	Test of Word Reading Efficency (TOWRE) Phonological Assessment Battery (PHAB)

Author	Study design	Geographical distribution	Participants	Intervention investigated	Context of intervention	Outcome variables measured (relevant to the review question)
Rosas et al. (2016)	Quasi-experimental design	Chile	**Sample:** *n* = 75 (treatment = 41 control = 34) **Age:** 6-7 years old	GraphoGame - computer game supporting phonics aqusition	**Facilitators:** Researcher **Time:** End of school day **Pre-post:** 3months	Curriculum based measures
Storey et al. (2019)	Between-subjects experimental control design	Northern Ireland	**Sample:** *n* = 32 (treatment = 17 control = 15) **Age:** Primary-aged children (6-9yrs)	Headsprout Early Reading – computer aided instruction	**Facilitators:** Researcher **Time:** During school day **Pre-post:** 7months	Phonics and Early Reading Assessment (PERA) Dolch Sight Words

Appendix E

Campbell Collaboration Calculator

A web-based effect-size calculator designed to facilitate the computation of effect-sizes so that they can be compared as part of a systematic literature review. This calculator was used to determine effect sizes for two studies (Kreskey & Truscott, 2015; Rosas et al., 2016) using means and standard deviations from post-intervention data.