A systematic literature review of the organisational arrangements of primary school-based reading interventions for struggling readers

Lauren Stentiford, George Koutsouris, Brahm Norwich

Graduate School of Education, University of Exeter

Running header: Review of reading intervention delivery

Abstract

Background: This paper presents findings from a systematic literature review which explored the delivery arrangements of school-based reading interventions for struggling readers aged 5-8 years.

Methods: Databases were searched for published and unpublished studies from 1970-2017 which employed an RCT or quasi-experimental design. 64 texts that met the inclusion criteria were categorised according to targeted tier based on the Response to Intervention (RTI) model. **Results:** Findings showed that tier 2/3 interventions targeting pupils who did not respond to whole class teaching (tier 1) were almost exclusively delivered in pull-out sessions by people other than the classroom teacher. **Conclusions:** At present, certain delivery arrangements appear to be used more than others in reading interventions conducted with young pupils in schools. This could reflect practices already used in primary schools and suggests that pupils who are identified for tier 2/3 support might have less access to their teacher's time and expertise.

Highlights

What is already known about this topic

- Intervention delivery is becoming increasingly important in light of the attention that the Response to Intervention (RTI) model has been receiving

- Intervention delivery is relevant to whether schools continue using successful interventions after the end of a study

- Few papers have explored the way reading interventions have been delivered

What this paper adds

- Certain delivery arrangements appear to be used more than others in primary school-based reading intervention trials: e.g. classroom teachers delivering tier 1 sessions, and other individuals delivering tier 2/3 sessions

- This could to some extent reflect practices used in primary schools

- This suggests that pupils who are identified for tier 2/3 support might have less access to their teacher's time and expertise

Implications for theory, policy or practice

- Questions are raised about how tier 1 is designed in terms of quality and relevance of provision for the diversity of pupil needs

- And how tier 2 and 3 are then defined and designed in relation to tier 1 both with regards to their goals, methods, location of delivery and who delivers the tier 2/3 programmes

- The above reflect broader issues about the relationship between general and additional programming and teaching, that bear on the future of inclusive teaching

Introduction and rationale

There are a considerable number of research synthesis papers and meta-analyses that examine effect sizes of literacy interventions (e.g. Ciullo et al., 2016; Wanzek et al., 2013; Wanzek et al., 2016). However, we could locate only a few papers exploring the way literacy interventions have been delivered, including who was the person delivering the intervention (e.g. class teacher, teaching assistant, researcher), where interventions were delivered (e.g. in the regular classroom, in pull-out sessions, elsewhere), and when interventions were delivered in relation to core or 'Quality First' teaching (i.e. instead of or in addition to it) (Suggate, 2010; 2016; National Reading Panel, 2000).

The matter of literacy intervention delivery is becoming increasingly important in light of the attention that the wave/tier model has been receiving in the UK and the USA, as a way of effectively supporting all pupils according to their needs (Griffiths & Stuart, 2013; Fien et al. 2011). It is therefore useful to explore how literacy interventions are delivered and to what extent the arrangements for intervention delivery are relevant to schools, i.e. whether interventions are designed to align closely with existing school timetabling or whether they require changes, and whether or not the people responsible for delivering the interventions are regular school staff. The design of interventions is relevant to whether schools continue using successful interventions after the end of a study, or have been discouraged from doing so. As randomised controlled trials (RCTs) are considered to be the 'gold standard' of education evaluation (e.g. see IES, 2003), questions about the relevance of intervention delivery to schools are important, since treatment delivery methods relevant to schools might conflict with RCT designs that aim for more control over the fidelity of implementation and non-treatment group conditions.

This paper discusses these issues based on the findings of a systematic review of delivery arrangements for school-based reading interventions for struggling pupils aged 5-8 years. We are

interested in the process and variation in intervention delivery, and to what extent such decisions are relevant to the targeted wave or tier. The review in this paper addresses the following questions:

- *How are school-based reading interventions being delivered*?

    1. What are the relative proportions of tier 1, 2, 3 and multitier interventions?

    2. Who is delivering tier 1, 2, 3 and multitier reading interventions?

    3. Where are tier 1, 2, 3 and multitier reading interventions being delivered?

    4. When are tier 1, 2, 3 and multitier reading interventions being delivered?

    5. Are there any patterns in reading intervention delivery (e.g. by date, country, design of study, intervention type, intensity, duration or reported gains)?

It is important to make clear at the outset that the focus of this paper is only on mapping the delivery arrangements of reading interventions (i.e. who, what, when and where). It is, of course, important to establish whether a relationship exists between delivery arrangements and reported gains – this is critical for evaluating the efficacy of different delivery arrangements. However, whilst information on the gains obtained in the studies has been extracted and will be briefly discussed, a meta-analysis of effect sizes is beyond the scope of this paper.

Search methods

Developing the search terms

We initially compiled a list of 25 search terms (see Appendix 1). This list was produced via a process of collaborative brainstorming, scanning the titles and abstracts of known relevant articles for key terms, and synonym checking.

Electronic searching

In November 2017, the following electronic databases were searched using the 25 terms: PsycINFO, ERIC, Education Research Complete, British Education Index, Web of Science Core Collection, ASSIA, CPCI-SSH, Australian Education Index, the International Bibliography of the Social Sciences, and

ProQuest Dissertations & Theses Global. Unpublished papers, reports, conference abstracts and doctoral theses were included in the search to mitigate publication bias.

Inclusion criteria

To be included in the review, studies had to be:

- Published between 1970-2017.

- Employ RCT or quasi-experimental designs (i.e. contain a treatment and control group with random or non-random allocation of members).

- Evaluate interventions with a clear focus on reading (e.g. comprehension, fluency, phonics, etc.).

- Target children who were struggling with reading – interventions focusing on particular SEN categories or specific disability categories (e.g. dyslexia) were not included. (This decision was taken in order to provide a manageable scope for this present review).

- Conducted with participants aged 5-8 years. (We initially intended to include children aged 5-11 years, representing the entire primary phase of education in England. However, after an initial screening of texts found through the electronic databases, we identified a high volume of potentially relevant studies and therefore narrowed our inclusion criteria.)

- School-based interventions – home-based or after-school interventions were not included.

- Published in English, although reading interventions could be conducted in any OECD (Organisation for Economic Co-operation and Development) country, in the home nation's language.

- Follow-up studies were excluded from the search.

Study selection

The titles and abstracts of records retrieved through searching were screened for relevance independently by a reviewer (LS). Initially 10% of the records were screened independently by two reviewers (LS and GK), with disagreements resolved by discussion in order to pilot and agree on

screening decisions. Full text copies of potentially relevant records were then obtained. The

retrieved records were again assessed for inclusion by one reviewer (LS), with 10% of the records

initially screened independently by two reviewers (LS and GK). Discrepancies were resolved by

discussion, with involvement of a third reviewer (BN) where necessary (see Figure 1).

Additional search strategies

In order to expand the scope of the search, all included published texts obtained through database

searching were subject to backwards citation chasing. Members of the research team had also

conducted a previous small-scale literature review on the topic and the results of this search (n = 29)

were screened for inclusion (contact authors for details). In addition, the reference list of Brooks'

(2016) report *What Works for Children and Young People with Literacy Difficulties* and Torgerson et

al.'s (2006) report *A Systematic Literature Review of the Research Literature on the Use of Phonics in*

the Teaching of Reading and Spelling were scanned (see Figure 1).

FIGURE 1 TO APPEAR HERE

Data extraction

Data was extracted from the final 64 included texts by one reviewer (LS) into a bespoke data

extraction form which was piloted on several articles and refined. Extraction included details of

study design, participants, reading intervention, delivery arrangements and gains (see Tables 1 and

2).

Studies were categorised according to the Response to Intervention (RTI) model – also called the

wave model in the UK (Rose, 2006) or the *tier model* in the US (Fuchs & Fuchs, 2006) The RTI model

has recently gained traction in research and educational communities, grounded in US special

education law (i.e. the Individuals with Disabilities Education Act, *IDEA*, 2004). The RTI model is

usually seen as referring to 3 waves or tiers of educational support: wave/tier 1 concerns *'Quality*

First' teaching (UK context, Rose 2006) or *core instruction* (US context, Fien et al., 2015), which is support targeted at all pupils. Wave/tier 2 refers to individual or group support for some pupils, whereby support is time limited and has clear entry and exit criteria (Rose, 2006). Wave/tier 3 concerns highly intensive and personalised intervention for the few pupils who do not respond to the previous 2 tiers of support. The nature of teaching is expected to change at each wave/tier and become more explicit and intensive. *Multitiered* systems are designed to enable pupils to move between tiers, based on progress monitoring (Fien et al., 2015). RTI as a protocol has been particularly used in the context of reading instruction (Fuchs & Fuchs, 2006), which is the focus of this current paper.

In this review, we adopted the language of tiers and therefore categorised reading interventions that targeted the whole classroom as tier 1, small group interventions as tier 2, and individual pupil interventions as tier 3. In reality, the boundaries between tiers 2 and 3 are not always very clear; the difference is often dependent on degrees of differentiation and intensity. However, we retained the distinction between whole class, group and one-to-one support for practical purposes, for ease of analysis. In some ambiguous cases an informed judgment was made as to which tier an intervention was targeting; for instance, when the design was different at different phases of a study. When the authors of the included studies explicitly identified an intervention at a particular tier, we respected their own categorisation. In addition, some studies were classified as multi-tier (i.e. tiers 1,2&3, tiers 1&2, or tiers 2&3) based on the definition of multitiered interventions provided by Fien et al. (2015), outlined above.

TABLE 1 TO APPEAR HERE

Findings

Findings from the 57 included journal articles and reports have been organised by theme and are discussed below. The 7 unpublished doctoral theses have been separated from this analysis and will be discussed in a section at the end of the findings.

Country

The majority of the studies that met our search criteria were US-based (32 out of 57 studies). There were also 9 studies from the UK, 4 from The Netherlands, 3 from Australia, 2 each from Canada and Ireland, and 1 each from Finland, France, Norway, South Africa and Sweden.

Year

Whilst the search criteria included studies from 1970-2017, the date of the earliest study meeting our criteria was 1994 (n = 2). In total, 5 studies were published in the 1990s, 20 published between 2000-2010, and 32 published between 2011-2017. This trend suggests greater interest in research in the field in recent decades.

Study design

The majority of the studies were randomised controlled trials (RCTs) (n = 40). 17 studies employed a quasi-experimental design.

Sample size

The sample size of the total number of children participating in each study varied considerably, ranging from 14 to 1509. Most studies were small-scale with fewer than 150 participants (33 out of 57). 16 studies were medium-scale with 150-500 participants, and 8 were large-scale with 500+ participants.

Targeted tier and way of delivery (who, where and when)

As discussed, the interventions in the studies examined were categorised according to targeted tier. We found that 4 out of 57 studies were focusing at tier 1, 19 at tier 2, and 16 at tier 3. A further 18 studies were categorised as multitiered – 1 targeting all three tiers, 8 targeting tiers 1&2, and 9 targeting tiers 2&3. More specifically:

Tier 1 studies: 4 studies were categorised as targeting tier 1. These were all US-based studies which evaluated reading interventions offered to all pupils in an attempt to respond to the diversity of student needs, and to improve core teaching in order to support struggling readers. All interventions were delivered by the regular classroom teacher. For example, Mathes et al. (1998) implemented a *peer-assisted learning strategies* (PALS) intervention amongst first graders. All children in the classrooms under study were paired (a stronger reader with a weaker reader) and completed a reading routine led by the teacher during a whole class session.

The relatively low number of tier 1 studies located might be attributed to the focus of this review, which was children who were *struggling* with reading. These children are often identified as in need of additional support at higher tiers (2 or 3).

Tier 2 studies: Many studies met our criteria of targeting tier 2 (19 out of 57 studies), in that they were designed for some pupils who did not respond to tier 1 teaching and needed extra support in small groups. Pupils were selected for tier 2 based on initial screening and, in many cases, a limit was set in terms of a standard reading score (e.g. less than *93* on the WJIII in Denton et al, 2014). 8 out of 19 were delivered outside of the regular classroom, although the exact location (e.g. empty classroom, library) tended not to be specified (e.g. Chen and Savage, 2014; Buckingham et al., 2012, 2014). Helf et al.'s (2014) intervention took place both inside and outside of the classroom – the majority of pupils received the treatment from support staff in a small tutoring room, but for practical reasons, class teachers provided supplemental instruction at the back of the general classroom in a few cases. 3 articles reported that the treatment had taken place inside the

classroom, however these papers were led by the same author (Vernon-Feagans et al. 2010, 2012, 2013) and detailed one intevention which had been piloted and then conducted on a larger scale.

It is interesting to note that the authors of 7 of the 19 studies did not state where the tier 2 intervention had taken place (i.e. inside or outside of the classroom), indicating that they did not feel this information was of importance.

The majority of tier 2 interventions (12 out of 19 studies) were delivered by specially recruited and trained instructors, i.e. non-regular school staff (e.g. Sénéchal et al., 2012; Torgesen et al., 2010). 3 studies deployed school-based staff who had been trained to implement the interventions: teaching assistants in Savage and Carless (2005); tutors in Chambers et al. (2011); and reading specialists and SEN teachers in Brinchmann et al. (2016). Class teachers were minimally involved in the delivery of Helf et al.'s (2014) intervention, as detailed above, but in Vernon-Feagans et al.'s (2010, 2012, 2013) studies, class teachers were trained to deliver the entire intervention. The authors make much of this in their papers, commenting on this 'unusual' mode of delivery (something that will be further discussed later in this paper):

> 'Unique to this intervention was the use of the classroom teacher as the primary source of delivering the intervention to children, with the support of our literacy consultant and an on-site consultant.' (2010: 185)

Most studies made clear that the evaluated tier 2 interventions were delivered in addition to core (tier 1) literacy teaching (14 out of 19 studies). For example, Vaughn et al. (2006: 461-464) wrote: 'the intervention was provided in addition to students' core reading lessons and did not conflict with their daily scheduled reading time' to emphasise that struggling pupils continued to receive tier 1 reading instruction whilst participating in the interventions. In 3 studies it was unclear whether the intervention was supplementary or instead of regular teaching. However, in Buckingham et al.'s (2012, 2014) articles – which report findings from the same study – struggling children were

'withdrawn from class during classroom literacy time' (2012: 86). The authors describe their *MiniLit* intervention as a comprehensive programme which 'prescribes both content and pedagogy' (2012: 86), which might explain the researchers' decision to remove participants from tier 1 instruction.

Tier 3 studies: 16 out of 57 studies were categorised at tier 3. Tier 3 interventions were targeting the few pupils who did not respond to either tier 1 or 2, or were designed to be more intensive and tailored to individual needs. In this review, the studies categorised as tier 3 interventions were delivered one-to-one and were, in general, implemented in a similar way to the tier 2 interventions (i.e. in addition to core teaching and outside of the regular classroom). In fact, 14 of the 16 studies explicitly stated that the reading intervention was supplementary – in the remaining 2 studies, details were unclear. 4 of the 16 studies reported that the intervention had taken place outside of the general classroom, and in 10 studies the delivery location was unspecified. Only one tier 3 intervention was conducted solely in the general classroom during usual teaching: Amendum et al. (2017) trained classroom teachers to work with struggling readers one-to-one until they made rapid progress, at which point the teacher would move on to another child. No details were provided about the usual teaching sessions or what other pupils were doing whilst teachers were working with children one-to-one.

Tier 3 interventions tended to be delivered by people other than regular classroom teachers. 10 of the 16 studies deployed individuals who were specially recruited and trained (3 of which were 1:1 computer-based interventions overseen by a trainer/s). Recruited individuals varied considerably and included: masters students (Van der Kooy-Hofland, 2012); older adult volunteers (Fives et al., 2013; Fives, 2016); and higher grade peers, student teachers and volunteers (Regtvoort et al., 2013). 4 of the 16 interventions were delivered by school staff or teaching assistants. For example, in Hatcher et al. (1994), qualified teachers from the schools under study (and one Head teacher) delivered the intervention, but these teachers were granted relief from their normal duties in order

to participate in the research. Only one tier 3 study was delivered by a regular classroom teacher – Amendum et al. (2017), which was discussed above.

Multitier 1,2&3: In the sole multitier 1,2&3 study, Houtveen et al. (2012) implemented a self-devised *Reading Acceleration Programme* (RAP) over a 3-year period in 21 schools in The Netherlands. The teachers who delivered the intervention were 'trained to improve their core instruction, to broaden their instruction for struggling learners, and to implement special measures for pupils who do not respond sufficiently to these interventions' (p. 71). The content of tier 2 and tier 3 was closely aligned with tier 1 teaching, both repeating and intensifying core content. Unfortunately, the authors provide few details of delivery arrangements; one can assume that class teachers implemented tier 1 in the regular classroom, and it is stated that tier 3 (individual) instruction was delivered by a 'specialised teacher' (p. 79). However, the location of tier 2 and 3 instruction was unspecified.

Multitier 1&2: 8 of the 57 studies were categorised as targeting tiers 1 and 2. In these studies, a tier 2 intervention tended to be used in addition to a closely designed tier 1 intervention to ensure the consistency of literacy tuition. The majority of the multitier 1&2 interventions (5 out of 8) were delivered by general classroom teachers (i.e. tier 1) alongside individuals who had been specially recruited and trained to deliver tier 2 instruction (e.g. Hay et al., 2007; Pullen et al., 2010). In one study, Baker et al. (2015) employed classroom teachers (tier 1) alongside instuctional assistants who, in most cases, already worked in the schools (tier 2). Mathes and Babyak (2001), Mathes et al. (2003) and Coyne et al. (2010) used classroom teachers to deliver both tier 1 and tier 2 interventions within the general classroom, but this appears uncommon. Coyne et al. encountered logistical problems and therefore asked classroom teachers to deliver their *extended vocabulary instruction* intervention in some schools. However, Mathes and Babyak and Mathes et al. deliberately designed reading interventions to be delivered by classroom teachers.

Multi-tier 2&3: 9 of the 57 studies were categorised as targeting tiers 2 and 3. These studies utilised both small group (tier 2) and one-to-one (tier 3) delivery arrangements, and content generally aligned with whole class teaching (tier 1), but was progressively enhanced and intensified. In the 9 studies, researchers deployed the tier 2&3 system in different ways and at different times during the reading interventions. For example, Hatcher et al. (2006) alternated between small group and individual teaching sessions for struggling pupils in their *'sound-linkage'* treatment condition. In contrast, Al Otaiba et al. (2014) sought to evaluate the impact of when children were allocated to tiers 2 and 3. In the treatment condition – which the authors called *'Dynamic RTI'* – children were immediately allocated into tiers 2 and 3 on initial assessment, rather than having to participate in tier 1 and be referred on after a period of monitoring.

8 of the 9 studies explicitly stated that tier 2&3 interventions were supplementary – Clipson-Boyles (2000) was the only author leaving this detail unclear. 3 studies stated that tiers 2&3 had taken place outside of the classroom in pull-out sessions (e.g. in a quiet classroom or media centre in Al Otaiba et al., 2014), whilst 6 studies did not specify a delivery location. It is important to note that none of the studies used classroom teachers to implement tier 2&3 interventions; 3 deployed individuals who had been specially recruited and trained, 1 used school staff, and 4 deployed teaching assistants (1 study left this unspecified).

Intensity and duration

The interventions were categorised into 3 levels of intensity: *high*, *medium* and *low* (see Table 1). The shortest session lasted 10 minutes and the longest 90 minutes. The majority of the interventions (35 out of 57 studies) were of high intensity, 18 of medium intensity, and 4 of low intensity. There did not appear to be any trends linking high, medium and low intensity interventions with particular tier/s targeted.

The duration of the intervention was also recorded for each of the studies: *long*, *medium* or *short* (see Table 1). The majority of the interventions were long in duration (n=27), with 16 x medium and 14 x short. The shortest intervention lasted just 2 weeks (Pullen et al., 2010), although sessions did take place daily. The longest intervention lasted 2 school years, with 30-40 minute sessions taking place 3 times a week (Kamps et al., 2008).

Intervention type and focus

There was a high level of heterogeneity in the nature of the reading interventions evaluated in the 57 articles/reports. As can been seen in Table 1, the studies were categorised according to type and focus. Studies classified as reading interventions (RI) aimed to evaluate the efficacy of a specific reading programme (e.g. *Wizards of Words* in Fives et al. 2013 and Fives, 2016) or an approach to reading (e.g. *reading with phonology, reading alone, or phonology alone* in Hatcher et al., 1994). Some researchers devised their own reading programmes (e.g. *Reading and Language Intervention* in Duff et al., 2014) or modified existing programmes. For example, Vaughn et al. (2006a, 2006b, 2006c) modfied the *Proactive Reading* programme in order to target both native English speakers and English language learners. 11 of the RI studies were identified as computer-based (e.g. Ecalle, et al., 2009). Other studies were classified as response-to-intervention (RTI). These studies sought to evaluate the efficacy of the RTI model itself (i.e. movement of children between tiers or changes in intensity of instruction at different tiers), and the reading programme/s used were of secondary interest (e.g. Houtveen et al., 2012).

Gains

As stated in the introduction, our review focused on the delivery arrangements of reading interventions, but the studies were also categorised according to their reported gains. Only 1 of the 57 studies reported that no gains had been found that were statistically significant (Duff et al., 2014). The remaining 56 reading interventions all reported some positive gains.

Dissertations

7 unpublished doctoral dissertations were also located in this review (Table 2). These dated between 1993 and 2014. 6 of the 7 studies were conducted in the USA and employed a quasi-experimental design. 1 study was UK-based and was a RCT. All 7 studies aimed to evaluate a reading intervention programme or strategy (e.g. *Reading Recovery* in Lucas, 2004), and all interventions targeted tiers 2 and/or 3. One dissertation provided very few details of delivery arrangements (Mitchell, 2010), however of the remaining 6 studies, 3 interventions were implemented by school staff, 1 by a specially recruited individual, and 2 by the researchers themselves. 4 interventions were delivered outside the classroom, 1 inside the classroom, and 2 in unspecified locations. It is noteworthy that no interventions involved classroom teachers, and that 2 of the 6 were delivered instead of tier 1 teaching. In general, the number of participants in the studies was low, ranging from 10 to 252, and all interventions were relatively short in duration with the longest lasting 18 weeks (Burton-Archie, 2014).

One further important trend that emerged in the dissertations related to the studies' findings; 5 of the 7 dissertations reported null gains, 1 reported negative gains, and just 1 reported positive gains.

TABLE 2 TO APPEAR HERE

Discussion

The findings will now be discussed in relation to the research questions outlined in the introduction, incorporating all data sets, i.e. published studies, reports and dissertations.

Who

In our review, all tier 1 interventions – including multitier interventions which incorporated a tier 1 element – were delivered by classroom teachers in the general classroom, except for Macuruso

(2006) which was a computer-based intervention delivered in a computer lab. This is not unexpected

given that tier 1 represents core teaching or 'Quality First' instruction. However, an important

finding in this review was that tier 2 and 3 interventions were overwhelmingly conducted by people

other than the classroom teacher (52/59). Of the tier 2 and/or 3 interventions that were delivered

by other people (n=52), 18 were delivered by individuals employed in the participating schools (i.e.

teaching assistants or other school staff). Notably, the majority of these studies were Northern

European (7x UK, 2 x The Netherlands, 1 x Sweden, 1 X Norway) – the remaining 7 were US studies.

In contrast, of the 34 interventions that were delivered by individuals who had been specially

recruited and trained by the researchers, 26 studies were conducted outside Northern Europe (e.g. 3

x Australia, 1 x South Africa), with the majority originating from North America (21 x US, 2 x Canada).

This arrangement perhaps reflects particular cultural differences.

By using regular school staff such as TAs to deliver interventions – the model most commonly used

in Northern Europe (e.g. Oostdam et al., 2015) – researchers were able to offer participating schools

a long-term investment, as most programmes included training and schools could continue using the

intervention after the study had ended, if they wished. Conversely, by hiring and training other

instructional assistants to deliver interventions – the model most commonly used in North America

(e.g. Denton et al., 2013) – researchers potentially had better control over the fidelity of intervention

implementation. For instance, Torgesen et al. (2010) highlighted:

> 'Instruction in the two reading programs was provided by six certified teachers
>
> who had participated in a previous project with us and who had considerable experience
>
> working with children with reading problems…' (p.6).

This level of control might not be achievable if an intervention was delivered by regular school staff

in addition to their other responsibilities. Yet, in these cases, schools might find it difficult to use the

intervention after the study has finished, as they would not have knowledge and experience of its delivery.

Only 7 studies reported that classroom teachers had been involved in the delivery of tiers 2 and/or 3. 4 of these papers were linked to one research team and one intervention (Vernon-Feagans et al., 2010, 2012, 2013; Amendum et al., 2017), and the other 3 papers were authored by Mathes and Babyak (2001), Mathes et al. (2003) and Coyne et al. (2010). As previously noted, Vernon-Feagans et al. (2010) comment on their 'unusual' delivery approach; the researchers used teacher-interventionists to 'increase the cost effectiveness and efficiency of reading interventions in low wealth rural schools that have fewer resources to expend on ancillary personnel to help struggling readers' (p. 183). Vernon-Feagans et al. explain that the research team worked closely with teachers to help them prepare implementation, including 'problem solving sessions that allowed teachers to arrange and manage their classrooms so that they would be able to work one on one with individual children' (p.188). However, the authors still reported difficulties: 'The shortage of available classroom aids [*sic*] and the teachers' limited experience with student independent work strategies presented major barriers' (p.188).

Mathes and Babyak (2001) also encountered problems when using class teachers to deliver tier 2 interventions, experiencing some resistance from teachers:

> 'When pressed to implement mini-lessons, 2 teachers chose to be removed from the study, stating that conducting mini-lessons in addition to 1st-Grade PALS was overwhelming. Likewise, 6 of the remaining 8 teachers designated presentation for the 15- to 20- minute mini-lessons to interns from local teacher education programs. ' (p. 35)

There are some clear benefits in involving class teachers at the higher levels of the RTI model: to conserve school funds, to utilise teachers' advanced pedagogical skills, and to ensure that struggling readers have greater access to teacher time and expertise. However, based on the findings in this review, it appears that if researchers are to make greater use of this organisational arrangement in the future, reading interventions need to be carefully designed to minimise the strain placed on teachers and schools. This could perhaps be done by utilising the existing group reading organisation already present in most UK primary schools, during which the class is (often) organised into ability groups with children working on a rotating basis – some independently and others with teacher and TA support.

When and where

Related to who is delivering an intervention is when and where the intervention is being delivered. From our literature review findings, the vast majority of tier 2 and 3 interventions were delivered outside of regular classroom sessions (e.g. tutor room, resource room). On rare occasions, studies explicitly stated that sessions were delivered in close proximity to the regular classroom. For instance, Fien et al. (2015) selected a room close to the regular classroom for small-group tier 2 sessions to underline the relation between tier 1 and tier 2 – and possibly to enable pupils to feel that they are still part of the regular classroom. Many authors of multitier studies also described an attempt to closely align content at different tiers; Pullen et al. (2010) arranged for small groups of struggling children to meet with a research assistant the day after tier 1 teaching had taken place to review word definitions taught in tier 1. This is illuminating in terms of the way tier 2 and 3 interventions are currently conceptualised and delivered; it demonstrates an effort by researchers to show the continuity and interconnectedness of instruction.

The findings of this review also give us an insight into the way that the RTI model is currently conceived of by researchers, and possibly implemented by educational professionals in practice.

Questions could be raised about whether, within the field of literacy interventions at present, 'Quality First' or core instruction programmes are sufficiently differentiated for those struggling to learn, and why additional programmes at tiers 2 and 3 are offered as supplementary to tier 1 programmes. When tier 1 provision does not necessarily differentiate enough, tier 1 attendance could be seen as time lost for those who are struggling. Al Otaiba et al.'s (2014) study could be seen as offering one way to address this issue; children were immediately allocated into tiers 2 and 3 on initial assessment, rather than having to first participate in tier 1. However, it is arguable that an RTI model needs to specify how diverse needs are provided for at tier 1, and *then* consider how higher tiers contribute relevant teaching to this. The current model of practice could be regarded as leaving children in the regular class for core teaching which they might not be able to access before offering tailored tier 2 or 3 support. The usefulness of this practice could be questioned.

Additional points

Three final points are worth discussing in light of the review findings. It is important to highlight that 6 of the 7 unpublished doctoral theses reported null or negative gains. This is potentially useful information as it raises questions about the almost uniformly positive picture of experimental research findings that are making their way into publication. It perhaps suggests the existence of a reporting bias with the authors of published studies being selective when reporting outcomes, or journals' reluctance to publish null and negative findings (see Torgerson et al., 2018).

It also seems that there is a lack of transparency in the way that reading interventions are reported in the literature. Many researchers left various aspects of delivery unspecified: the authors of 23 papers and 2 theses did not state where their intervention had taken place; the authors of 2 papers and 1 thesis did not state who had delivered their reading intervention; and the authors of 8 papers and 1 thesis did not state whether the intervention was supplemental or instead of normal teaching. Furthermore, of the 7 studies that used classroom teachers to deliver tiers 2 and/or 3, only Vernon-

Feagans et al. (2010) included details of who led whole class sessions whilst teachers were busy with small groups. Such information might not have been included because researchers did not feel that it was significant. However, it is arguable that this methodological information is of critical importance to better understand the strengths and weaknesses of particular delivery arrangements.

Finally, this review summarises the delivery arrangements utilised in research studies trialling reading interventions. Such studies should to some extent reflect the organisation of reading interventions carried out in real life by educational practitioners, but are of course not an accurate representation of actual teaching practice. The findings of this review suggest that delivery arrangements need to closely align with actual teaching practice – especially when the classroom teacher delivers the intervention – because arrangements have to fit well with the existing classroom organisation and timetable (e.g. Vernon-Feagans et al., 2010; Mathes and Babyak, 2001). This perhaps becomes less important when interventions are delivered in pull out sessions by others (e.g. TAs) because there is some space for paralell arrangements (e.g. Savage and Carless, 2005). However, even in pull out sessions, there should arguably be some fit between existing classroom arrangements and research-specific arrangements.

Conclusion

This systematic review aimed to explore the delivery arrangements of school-based reading interventions for struggling readers aged 5-8 years, and categorised studies according to the RTI model. The findings suggest that, at present, certain delivery arrangements are strongly favoured by researchers in the field. In general, teachers were asked to lead tier 1 teaching, with researchers tending to rely on school staff and teaching assistants (often in Northen European studies) or specially recruited interventionists (often in North American studies) to deliver tiers 2 and 3. This might be due to fairly entrenched views about educational 'roles' (e.g. the teacher as 'leader' and

teaching assistant as 'support'), logistical concerns (e.g. the lack of sufficiently qualified individuals to lead whole class teaching whilst teachers are working with small groups/individuals), or an anticipated resistance from teachers in relation to increased workloads. However, this also perhaps signifies researchers focusing on exercising good experimental controls, rather than investing in new and innovative delivery arrangements that schools might be able to use in the long term. It is suggested that, in the future, reading interventions should be designed in more diverse ways in order to explore the salience of delivery (i.e. who, where, when), in terms of the efficacy and real life 'usability' of interventions in schools, as well as designing rigorous experimental studies.

The findings also raise more general questions about how the RTI model is currently understood and implemented. These are questions about how tier 1 is designed in terms of quality and relevance of provision for the diversity of pupil needs, and how tier 2 and 3 are then defined and designed in relation to tier 1, in terms of their goals, methods, location of delivery and who delivers the programmes. These are broad issues about the relationship between general and additional programming and teaching, which bear on the future of inclusive teaching and require further attention.

References

Al Otaiba, S., Connor, C. M., Folsom, J. S., Wanzek, J., Greulich, L., Schatschneider, C., & Wagner, R. K. (2014). To Wait in Tier 1 or Intervene Immediately: A Randomized Experiment Examining First-Grade Response to Intervention in Reading. *Exceptional Children, 81*(1), 11-27. doi:10.1177/0014402914532234

Amendum, S. J., Bratsch-Hines, M., & Vernon-Feagans, L. (2017). Investigating the efficacy of a web-based early reading and professional development intervention for young english learners.

Reading Research Quarterly, Advance online publication

http://onlinelibrary.wiley.com/doi/10.1002/rrq.188/full

Baker, S. K., Smolkowski, K., Chaparro, E. A., Smith, J. L. M., & Fien, H. (2015). Using Regression

Discontinuity to Test the Impact of a Tier 2 Reading Intervention in First Grade. *Journal of*

Research on Educational Effectiveness, 8(2), 218-244. doi:10.1080/19345747.2014.909548

Berninger, V. W., Vermeulen, K., Abbott, R. D., McCutchen, D., Cotton, S., Cude, J., . . . Sharon, T.

(2003). Comparison of Three Approaches to Supplementary Reading Instruction for Low-

Achieving Second-Grade Readers. *Language, Speech, and Hearing Services in Schools, 34*(2),

101-116. doi:10.1044/0161-1461(2003/009)

Brinchmann, E. I., Hjetland, H. N., & Lyster, S.-A. H. (2016). Lexical quality matters: Effects of word

knowledge instruction on the language and literacy skills of third- and fourth-grade poor

readers. *Reading Research Quarterly, 51*(2), 165-180.

Brooks, G. (2016). *What Works for Children and Young People with Literacy Difficulties? The*

Effectiveness of Intervention Schemes. 5th edition. Frensham: Dyslexia-SpLD Trust. Available

online (only) at https://www.helenarkell.org.uk/documents/files/What-works-for-children-

and-young-people-with-literacy-difficulties-5th-edition.pdf

Brown, R., Pressley, M., Van Meter, P. & Schuder, T. (1995). *A Quasi-Experimental Validation of*

Transactional Strategies Instruction with Previously Low-Achieving, Second-Grade Readers.

Reading Research Report No. 33. National Reading Research Center. Retrieved from

http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=ED379636&site=ehost-

live

Buckingham, J., Wheldall, K., & Beaman, R. (2012). A randomised control trial of a Tier-2 small-group

intervention ('MiniLit') for young struggling readers. *Australian Journal of Learning*

Difficulties, 17(2), 79-99. doi: 10.1080/19404158.2012.717537

Buckingham, J., Wheldall, K., & Beaman-Wheldall, R. (2014). Evaluation of a two-phase implementation of a Tier-2 (small group) reading intervention for young low-progress readers. *Australasian Journal of Special Education, 38*(2), 169-185. doi:http://dx.doi.org/10.1017/jse.2014.13

Burroughs-Lange, S. & Douëtil, J. (2007). Literacy progress of young children from poor urban settings: a Reading Recovery comparison study. *Literacy, Teaching and Learning, 12*(1), 19-46.

Burton-Archie, S. H. (2014). *Reading intervention: The effectiveness of leveled literacy intervention.* (3611522 Ed.D.), Walden University, Ann Arbor.

Cassady, J. C., Smith, L. L., & Thomas, C. L. (2017). Supporting emergent literacy for english language learners with computer-assisted instruction. *Journal of Research in Reading*. Advance online publication http://onlinelibrary.wiley.com/doi/10.1111/1467-9817.12110/abstract

Chambers, B., Slavin, R., xa, E, Madden, N., xa, . . . Gifford, R. (2011). Small-Group, Computer-Assisted Tutoring to Improve Reading Outcomes for Struggling First and Second Graders. *The Elementary School Journal, 111*(4), 625-640. doi:10.1086/659035

Chen, V., & S. Savage, R. (2014). Evidence for a simplicity principle: teaching common complex grapheme-to-phonemes improves reading and motivation in at-risk readers. *Journal of Research in Reading, 37*(2), 196-214. doi:10.1111/1467-9817.12022

Ciullo, S., Lo, Y.-L. S., Wanzek, J., & Reed, D. K. (2016). A Synthesis of Research on Informational Text Reading Interventions for Elementary Students With Learning Disabilities. *Journal of Learning Disabilities, 49*(3), 257-271. doi:10.1177/0022219414539566

Clipson-Boyles, S. (2000). The Catch Up Project: a reading intervention in Year 3 for Level 1 readers (Research Note). *Journal of Research in Reading, 23*(1), 78-84. doi:10.1111/1467-9817.00104

Coyne, M. D., McCoach, D. B., Loftus, S., Zipoli Jr, R., Ruby, M., Crevecoeur, Y. C., & Kapp, S. (2010).

Direct and Extended Vocabulary Instruction in Kindergarten: Investigating Transfer Effects.

Journal of Research on Educational Effectiveness, 3(2), 93-120.

doi:10.1080/19345741003592410

Denton, C. A., Fletcher, J. M., Taylor, W. P., Barth, A. E., & Vaughn, S. (2014). An Experimental

Evaluation of Guided Reading and Explicit Interventions for Primary-Grade Students At-Risk

for Reading Difficulties. *Journal of Research on Educational Effectiveness, 7*(3), 268-293.

doi:10.1080/19345747.2014.906010

Denton, C. A., Tolar, T. D., Fletcher, J. M., Barth, A. E., Vaughn, S., & Francis, D. J. (2013). Effects of

tier 3 intervention for students with persistent reading difficulties and characteristics of

inadequate responders. *Journal of Educational Psychology, 105*(3), 633-648.

Duff, F. J., Hayiou-Thomas, M. E., & Hulme, C. (2012). Evaluating the effectiveness of a

phonologically based reading intervention for struggling readers with varying language

profiles. *Reading and Writing, 25*(3), 621-640. doi:10.1007/s11145-010-9291-6

Duff, F. J., Hulme, C., Grainger, K., Hardwick, S. J., Miles, J. N., & Snowling, M. J. (2014). Reading and

language intervention for children at risk of dyslexia: A randomised controlled trial. *Journal

of Child Psychology and Psychiatry, 55*(11), 1234-1243.

Ecalle, J., Magnan, A., & Calmus, C. (2009). Lasting effects on literacy skills with a computer-assisted

learning using syllabic units in low-progress readers. *Computers & Education, 52*(3), 554-561.

Ehri, L. C., Dreyer, L. G., Flugman, B., & Gross, A. (2007). Reading Rescue: An Effective Tutoring

Intervention Model for Language-Minority Students Who Are Struggling Readers in First

Grade. *American Educational Research Journal, 44*(2), 414-448.

doi:10.3102/0002831207302175

Fien, H., Smith, J. L. M., Baker, S. K., Chaparro, E., Baker, D. L., & Preciado, J. A. (2011). Including

    English Learners in a Multitiered Approach to Early Reading Instruction and Intervention.

    Assessment for Effective Intervention, 36(3), 143-157. doi:10.1177/1534508410392207

Fien, H., Smith, J. L. M., Smolkowski, K., Baker, S. K., Nelson, N. J., & Chaparro, E. (2015). An

    Examination of the Efficacy of a Multitiered Intervention on Early Reading Outcomes for First

    Grade Students at Risk for Reading Difficulties. *Journal of Learning Disabilities, 48*(6), 602-

    621. doi:10.1177/0022219414521664

Fives, A. (2016). Modeling the interaction of academic self-beliefs, frequency of reading at home,

    emotional support, and reading achievement: An RCT study of at-risk early readers in first

    grade and second grade. *Reading Psychology, 37*(3), 339-370.

Fives, A., Kearns, N., Devaney, C., Canavan, J., Russell, D., Lyons, R., . . . O'Brien, A. (2013). A One-to-

    One Programme for At-Risk Readers Delivered by Older Adult Volunteers. *Review of*

    Education, 1(3), 254-280.

Fuchs, D., & Fuchs. L.S. (2006). Introduction to response to intervention: What, why, and how valid is

    it? *Reading Research Quarterly, 41*(1), 93-99. doi:doi:10.1598/RRQ.41.1.4

Gilbert, J. K., Compton, D. L., Fuchs, D., Fuchs, L. S., Bouton, B., Barquero, L. A., & Cho, E. (2013).

    Efficacy of a first-grade responsiveness-to-intervention prevention model for struggling

    readers. *Reading Research Quarterly, 48*(2), 135-154.

Griffiths, Y. & Stuart, M. (2013). Reviewing evidence-based practice for pupils with dyslexia and

    literacy difficulties. *Journal of Research in Reading, 36*(1), 96-116. doi:doi:10.1111/j.1467-

    9817.2011.01495.x

Gustafson, S., Fälth, L., Svensson, I., Tjus, T., & Heimann, M. (2011). Effects of Three Interventions on

    the Reading Skills of Children With Reading Disabilities in Grade 2. *Journal of Learning*

    Disabilities, 44(2), 123-135. doi:10.1177/0022219410391187

Hatcher, P. J., Hulme, C., & Ellis, A. W. (1994). Ameliorating Early Reading Failure by Integrating the

    Teaching of Reading and Phonological Skills: The Phonological Linkage Hypothesis. *Child*

    Development, 65(1), 41-57. doi:10.1111/j.1467-8624.1994.tb00733.x

Hatcher, P. J., Hulme, C., Jeremy, N. V. M., Carroll, J. M., Hatcher, J., Gibbs, S., . . . Snowling, M. J.

    (2006). Efficacy of small group reading intervention for beginning readers with reading-

    delay: a randomised controlled trial. *Journal of Child Psychology and Psychiatry, 47*(8), 820-

    827. doi:http://dx.doi.org/10.1111/j.1469-7610.2005.01559.x

Hay, I., Elias, G., Fielding-Barnsley, R., Homel, R., & Freiberg, K. (2007). Language delays, reading

    delays, and learning difficulties: Interactive elements requiring multidimensional

    programming. *Journal of Learning Disabilities, 40*(5), 400-409.

Helf, S., Cooke, N. L., & Konrad, M. (2014). Advantages of Providing Structured Supplemental

    Reading Instruction to Kindergarteners At Risk for Failure in Reading. *Preventing School*

    Failure, 58(4), 214-222. doi:10.1080/1045988X.2013.798773

Houtveen, A. A. M., Grift, W. J. C. M. v. d., Thoni, H., & Wim van de, G. (2012). Improving reading

    achievements of struggling learners. In (Vol. 23, pp. 71-93). Netherlands.

IES (2003). *Identifying and implementing educational practices supported by rigorous evidence: A*

    user-friendly guide. Washington, DC: US Department of Education.

Individuals with Disabilities Education Improvement Act (2004). Pub. L. 108-466.

Kamps, D., Abbott, M., Greenwood, C., Wills, H., Veerkamp, M., & Kaufman, J. (2008). Effects of

    Small-Group Reading Instruction and Curriculum Differences for Students Most at Risk in

    Kindergarten:Two-Year Results for Secondary- and Tertiary-Level Interventions. *Journal of*

    Learning Disabilities, 41(2), 101-114. doi:10.1177/0022219407313412

Kyle, F., Kujala, J., Richardson, U., Lyytinen, H., & Goswami, U. (2013). Assessing the Effectiveness of

    Two Theoretically Motivated Computer-Assisted Reading Interventions in the United

Kingdom: GG Rime and GG Phoneme. *Reading Research Quarterly, 48*(1), 61-76. doi:10.1002/rrq.038

le Roux, M., Swartz, L., & Swart, E. (2014). The Effect of an Animal-Assisted Reading Program on the Reading Rate, Accuracy and Comprehension of Grade 3 Students: A Randomized Control Study. *Child & Youth Care Forum, 43*(6), 655-673. doi:10.1007/s10566-014-9262-1

Lucas, P. J. (2004). *Responses to a school-based literacy intervention for poor readers in junior school: Evaluating academic and attitudinal impact, using a RCT methodology.* (10015675 Ph.D.), University of London, University College London (United Kingdom), Ann Arbor.

Macaruso, P., Hook, P. E., & McCabe, R. (2006). The efficacy of computer-based supplementary phonics programs for advancing reading skills in at-risk elementary students. *Journal of Research in Reading, 29*(2), 162-172. doi:10.1111/j.1467-9817.2006.00282.x

Mathes, P. G., & Babyak, A. E. (2001). The Effects of Peer-Assisted Literacy Strategies for First-Grade Readers With and Without Additional Mini-Skills Lessons. *Learning Disabilities Research & Practice, 16*(1), 28-44. doi:10.1111/0938-8982.00004

Mathes, P. G., Howard, J. K., Allen, S. H., & Fuchs, D. (1998). Peer-Assisted Learning Strategies for First-Grade Readers: Responding to the Needs of Diverse Learners. *Reading Research Quarterly, 33*(1), 62-94. doi:10.1598/RRQ.33.1.4

Mathes, P. G., Torgesen, J. K., Clancy-Menchetti, J., Santi, K., Nicholas, K., Robinson, C., & Grek, M. (2003). A Comparison of Teacher-Directed versus Peer-Assisted Instruction to Struggling First-Grade Readers. *The Elementary School Journal, 103*(5), 459-479.

May, H., Gray, A., Sirinides, P., Goldsworthy, H., Armijo, M., Sam, C., . . . Tognatta, N. (2015). Year one results from the multisite randomized evaluation of the i3 Scale-Up of Reading Recovery. *American Educational Research Journal, 52*(3), 547-581.

Messer, D., & Nash, G. (2017). An evaluation of the effectiveness of a computer-assisted reading

    intervention. *Journal of Research in Reading*. Advance online publication

    http://onlinelibrary.wiley.com/doi/10.1111/1467-9817.12107/full

Ming, K. (2007). *The effects of a fluency intervention on the oral reading fluency of first grade*

    students at risk for reading failure. (3268792 Ed.D.), Florida Atlantic University, Ann Arbor.

Mitchell, V. (2010). *A quasi-experimental study of the use of "Dr. Cupp's Readers" in comparison to*

    traditional instruction of at -risk second grade students' test scores. (3397991 Ed.D.), Walden

    University, Ann Arbor.

National Reading Panel (2000). *Teaching Children to Read: An Evidence-Based Assessment of the*

    Scientific Research Literature on Reading and its Implications for Reading Instruction, NIH.

    Available at https://www1.nichd.nih.gov/publications/pubs/nrp/Documents/report.pdf

Oostdam, R., Blok, H., & Boendermaker, C. (2015). Effects of individualised and small-group guided

    oral reading interventions on reading skills and reading attitude of poor readers in grades 2–

    4. *Research Papers in Education, 30*(4), 427-450. doi:10.1080/02671522.2014.953195

Pinnell, G. S., Lyons, C. A., DeFord, D. E., Bryk, A. S., & Seltzer, M. (1994). Comparing Instructional

    Models for the Literacy Education of High-Risk First Graders. *Reading Research Quarterly,*

    29(1), 9-39. doi:10.2307/747736

Pullen, P. C., Tuckwiller, E. D., Konold, T. R., Maynard, K. L., & Coyne, M. D. (2010). A tiered

    intervention model for early vocabulary instruction: The effects of tiered instruction for

    young students at risk for reading disability. *Learning Disabilities Research & Practice, 25*(3),

    110-123.

Regtvoort, A., Zijlstra, H., & Leij, A. (2013). The Effectiveness of a 2-year Supplementary Tutor-

    assisted Computerized Intervention on the Reading Development of Beginning Readers at

Risk for Reading Difficulties: A Randomized Controlled Trial. *Dyslexia (10769242), 19*(4), 256-280. doi:10.1002/dys.1465

Rehmann, R. (2005). *The effect of Earobics™ Step 1, software on student acquisition of phonological awareness skills.* (3181124 D.Ed.), University of Oregon, Ann Arbor.

Rimm-Kaufman, S. E., Kagan, J., & Byers, H. (1998). The effectiveness of adult volunteer tutoring on reading among "at risk" first grade children. *Reading Research and Instruction, 38*(2), 143-152. doi:10.1080/19388079909558284

Rose, J. (2006). *Independent review of the teaching of early reading: Interim report*. London: DFES.

Saine, N. L., Lerkkanen, M.-K., Ahonen, T., Tolvanen, A., & Lyytinen, H. (2011). Computer-Assisted Remedial Reading Intervention for School Beginners at Risk for Reading Disability. *Child Development, 82*(3), 1013-1028. doi:10.1111/j.1467-8624.2011.01580.x

Samuelson, T. C. (2010). *An Examination of an Early Intervention Reading Program Focusing on the Progress Monitoring of Literacy Skills and the Reading Self-Concepts of Struggling Readers.* (10618030 Ed.D.), Western Connecticut State University, Ann Arbor.

Savage, R., & Carless, S. (2005). Learning support assistants can deliver effective reading interventions for 'at-risk' children. *Educational Research, 47*(1), 45-61. doi:10.1080/0013188042000337550

Schwartz, R. M. (2005). Literacy Learning of At-Risk First-Grade Students in the Reading Recovery Early Intervention. *Journal of Educational Psychology, 97*(2), 257-267.

Sénéchal, M., Ouellette, G., Pagan, S., & Lever, R. (2012). The role of invented spelling on learning to read in low-phoneme awareness kindergartners: A randomized-control-trial study. *Reading and Writing, 25*(4), 917-934.

Smith, J. L. M., Nelson, N. J., Smolkowski, K., Baker, S. K., Fien, H., & Kosty, D. (2016). Examining the efficacy of a multitiered intervention for at-risk readers in grade1. *The Elementary School Journal, 116*(4), 549-573.

Suggate, S. P. (2010). Why what we teach depends on when: Grade and reading intervention modality moderate effect size. *Developmental Psychology, 46*(6), 1556-1579. doi:10.1037/a0020612

Suggate, S. P. (2016). A Meta-Analysis of the Long-Term Effects of Phonemic Awareness, Phonics, Fluency, and Reading Comprehension Interventions. *Journal of Learning Disabilities, 49*(1), 77-96. doi:10.1177/0022219414528540

Torgerson, C., Brooks, G. and Hall, J. (2006). A Systematic Review of the Research Literature on the Use of Phonics in the Teaching of Reading and Spelling. London: Department for Education and Skills Research Report 711. Available online at http://dera.ioe.ac.uk/14791/1/RR711.pdf

Torgerson, C., Brooks, G., Gascoine, L., & Higgins, S. (2018). Phonics: reading policy and the evidence of effectiveness from a systematic 'tertiary' review. *Research Papers in Education*, DOI: 10.1080/02671522.2017.1420816

Torgesen, J. K., Wagner, R. K., Rashotte, C. A., Herron, J., & Lindamood, P. (2010). Computer-assisted instruction to prevent early reading difficulties in students at risk for dyslexia: Outcomes from two instructional approaches. *Annals of Dyslexia, 60*(1), 40-56. doi:10.1007/s11881-009-0032-y

Van der Kooy-Hofland, V. A. C., Bus, A. G., & Roskos, K. (2012). Effects of a Brief but Intensive Remedial Computer Intervention in a Sub-Sample of Kindergartners with Early Literacy Delays. *Reading and Writing: An Interdisciplinary Journal, 25*(7), 1479-1497.

Vaughn, S., Cirino, P. T., Linan-Thompson, S., Mathes, P. G., Carlson, C. D., Hagan, E. C., . . . Francis, D. J. (2006a). Effectiveness of a Spanish Intervention and an English Intervention for English-

Language Learners at Risk for Reading Problems. *American Educational Research Journal, 43*(3), 449-487. doi:10.3102/00028312043003449

Vaughn, S., Linan-Thompson, S., Mathes, P. G., Cirino, P. T., Carlson, C. D., Pollard-Durodola, S. D., . . . Francis, D. J. (2006b). Effectiveness of Spanish Intervention for First-Grade English Language Learners at Risk for Reading Difficulties. *Journal of Learning Disabilities, 39*(1), 56-73. doi:10.1177/00222194060390010601

Vaughn, S., Mathes, P., Linan, x, Thompson, S., Cirino, P., . . . Francis, D. (2006c). Effectiveness of an English Intervention for First Grade English Language Learners at Risk for Reading Problems. *The Elementary School Journal, 107*(2), 153-180. doi:10.1086/510653

Vernon-Feagans, L., Gallagher, K., Ginsberg, M. C., Amendum, S., Kainz, K., Rose, J., & Burchinal, M. (2010). A Diagnostic Teaching Intervention for Classroom Teachers: Helping Struggling Readers in Early Elementary School. *Learning Disabilities Research & Practice, 25*(4), 183-193. doi:10.1111/j.1540-5826.2010.00316.x

Vernon-Feagans, L., Kainz, K., Amendum, S., Ginsberg, M., Wood, T., & Bock, A. (2012). Targeted Reading Intervention:A Coaching Model to Help Classroom Teachers With Struggling Readers. *Learning Disability Quarterly, 35*(2), 102-114. doi:10.1177/0731948711434048

Vernon-Feagans, L., Kainz, K., Hedrick, A., Ginsberg, M., & Amendum, S. (2013). Live webcam coaching to help early elementary classroom teachers provide effective literacy instruction for struggling readers: The Targeted Reading Intervention. *Journal of Educational Psychology, 105*(4), 1175-1187. doi:10.1037/a0032143

Wanzek, J., Vaughn, S., Scammacca, N. K., 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. https://doi.org/10.1007/s10648-015-9321-7

Wanzek, J., Vaughn, S., Scammacca, N. K., Metz, K., Murray, C. S., Roberts, G., & Danielson, L. (2013). Extensive Reading Interventions for Students With Reading Difficulties After Grade 3. *Review of Educational Research, 83*(2), 163-195. doi:10.3102/0034654313477212

Wille, G. A. (1993). *Project Read as an early intervention program.* (1353782 M.S.), California State University, Fullerton, Ann Arbor.

Appendix 1 - Search terms

1. **Design**

RCT

"Random* control*"

"Quasi-experiment*"

"Control* trial*"

2. **Reading**

Read*

3. **Population descriptor**

Struggl*

"At-risk"

Delay*

Remedial*

Poor

Retard*

"Learn* disab*"

"Read* disab*"

"Read* difficult*"

"Learn* difficult*"

"Low-attain*"

"Low-achiev*"

"Low ability"

4. **Age**

Primary

Elementary

Pupil*

Student*

Child*

Grade*

Year*

Figure 1 - Flow diagram depicting records identified, included and excluded

```
┌─────────────────────────────────┐
│ Records identified through      │
│ database searching              │
│ (n = 1108)                      │
└─────────────────────────────────┘
                 │
                 ▼
┌─────────────────────────────────┐
│ Records after duplicates removed│
│ (n = 949)                       │
└─────────────────────────────────┘
                 │
                 ▼
┌─────────────────────────────────┐        ┌──────────────────────┐
│ Records screened - titles and   │───────▶│ Records excluded     │
│ abstracts                       │        │ (n = 818)            │
│ (n = 949)                       │        └──────────────────────┘
└─────────────────────────────────┘
                 │
                 ▼
┌─────────────────────────────────┐   ┌──────────────────┐   ┌──────────────────────────────┐
│ Full-text records assessed for  │──▶│ Unable to obtain │──▶│ Full-text records excluded,  │
│ eligibility                     │   │ (n =1)           │   │ with reasons                 │
│ (n = 131)                       │   └──────────────────┘   │ (n = 87)                     │
└─────────────────────────────────┘                          │                              │
                 │                                            │ -not relevant design (n = 22)│
                 ▼                                            │                              │
┌─────────────────────────────────┐                          │ -participant ages not        │
│ Additional records identified   │                          │ relevant (n = 38)            │
│ through other sources (n = 113) │                          │                              │
│                                 │                          │ -not 'struggling' readers    │
│ -a previous review conducted by │                          │ (n = 5)                      │
│ the research team (n = 29)      │                          │                              │
│                                 │                          │ -not OECD country (n = 13)   │
│ -backwards citation chasing     │                          │                              │
│ (n = 57)                        │                          │ -not school-based (n = 9)    │
│                                 │                          └──────────────────────────────┘
│ -other literature searching     │
│ (n = 27)                        │
└─────────────────────────────────┘
                 │
                 ▼
┌─────────────────────────────────┐   ┌──────────────────┐   ┌──────────────────────────────┐
│ Additional full-text records    │──▶│ Unable to obtain │──▶│ Full-text records excluded,  │
│ assessed for eligibility        │   │ (n =9)           │   │ with reasons                 │
│ (n = 113)                       │   └──────────────────┘   │ (n = 83)                     │
└─────────────────────────────────┘                          │                              │
                 │                                            │ -not relevant design (n = 44)│
                 ▼                                            │                              │
┌─────────────────────────────────┐                          │ -participant ages not        │
│ **Total number of studies       │                          │ relevant (n = 14)            │
│ included in data synthesis**    │                          │                              │
│ **(n = 64)**                    │                          │ -not 'struggling' readers    │
│                                 │                          │ (n = 21)                     │
│ **-published studies (n = 56)** │                          │                              │
│                                 │                          │ -duplicates (n = 4)          │
│ **-reports (n = 1)**            │                          │                              │
│                                 │                          │ -not OECD country (n = 1)    │
│ **-doctoral dissertations       │                          └──────────────────────────────┘
│ (n = 7)**                       │
└─────────────────────────────────┘
```

35

First Author	Year	Country	Design[1]	Sample size	Age in years	Intervention type/focus[2]	Targeted tier	Intensity[3]	Duration[4]	Delivery location	Delivered by[5]	Supplementary/ instead of[6]	Gains[7]
Al Otaiba	2014	USA	RCT	522	6-7	RTI - Typical RTI and Dynamic RTI	Multi-tier - 2,3	Medium	Long	Out of the classroom	Specially recruited and trained	Supplementary	Positive
Amendum	2017	USA	RCT	108	5-7	RI - Targeted Reading Intervention	Tier 3	Medium	Short	In the classroom	Class teachers	Supplementary	Positive
Baker	2015	USA	RCT	1509	6-7	RTI - Enhanced Core Reading Instruction	Multi-tier - 1,2	High	Long	In classroom and unclear	Class teachers and school staff	Supplementary	Positive
Berninger	2003	USA	RCT	96	7-8	RI – Explicit word recognition, explicit reading comprehension, both	Tier 2	Medium	Medium	Unspecified	Specially recruited and trained	Supplementary	Positive
Brinchmann	2016	Norway	Quasi	118	7-8	RI - Word Knowledge Intervention	Tier 2	Medium	Short	Out of the classroom	School staff	Unclear	Positive
Brown	1995	USA	Quasi	60	7-8	RI - Students Achieving Independent Learning	Tier 1	High	Long	In the classroom	Class teachers	Instead of	Positive
Buckingham	2014	Australia	RCT	14	5-6	RI -MinLit	Tier 2	High	Long	Out of the classroom	Specially recruited and trained	Instead of	Positive
Buckingham	2012	Australia	RCT	22	5-7	RI -MinLit	Tier 2	High	Long	Out of the classroom	Specially recruited and trained	Instead of	Positive
Burroughs-Lange	2007	UK	Quasi	234	5-6	RI - Reading Recovery	Tier 3	High	Medium	Unspecified	Specially recruited and trained	Unclear	Positive

[1] RCT = randomised control trial; Quasi = quasi-experimental [2] RI = evaluation of a reading intervention program/s; RTI = evaluation of Response to Intervention model. * signifies a computer-based intervention.
[3] High intensity = 4-5 sessions a week; medium intensity = 2-3 sessions a week; low intensity = 1 or fewer times a week, n.b. if studies were on the cusp of an intensity bracket (e.g. 3-4 sessions a week) or employed different intensities at different tiers, the lower number was used for categorisation. [4] Long duration = 21+ weeks; medium duration = 20 or fewer weeks; short duration = 10 or fewer weeks.[5] Class teachers = regular classroom teacher; school staff = other staff employed in the schools (e.g. reading specialists, SEN teachers); specially recruited and trained = individuals not employed in the schools who had been trained to deliver the intervention (e.g. research assistants, graduate students, paraprofessionals, older volunteers, peers); teaching assistants = learning support staff employed in the schools; doctoral researcher = the doctoral candidate. [6] Supplementary = intervention delivered in additional to regular teaching; instead of = intervention delivered instead of regular teaching. [7] Positive gains = some statistically significant gains; null gains = no significant gains; negative gains = control group scored higher than treatment group.

Table 1 – Published journal articles and reports

Cassady	2017	USA	RCT	1490	5-7	RI - Imagine Learning*	Tier 1	High	Long	In classroom and computer labs	Class teachers	Unclear	Positive
Chambers	2011	USA	RCT	646	6-8	RI - Team Alpine*	Tier 2	High	Long	Unspecified	School staff	Supplementary	Positive
Chen	2014	Canada	RCT	38	6-8	RI - grapheme-to-phoneme correspondences	Tier 2	Medium	Short	Out of the classroom	Specially recruited and trained	Unclear	Positive
Clipson-Boyles	2000	UK	Quasi	74	7-8	RI - Catch Up	Multi-tier - 2,3	Medium	Short	Unspecified	Unclear	Unclear	Positive
Coyne	2010	USA	Quasi	124	5-6	RI - extended vocabulary instruction	Multi-tier - 1,2	Medium	Medium	In and out of the classroom	Class teachers and specially recruited and trained	Unclear	Positive
Denton	2014	USA	RCT	380	6-7	RI - Guided Reading	Tier 2	High	Long	Out of the classroom	Specially recruited and trained	Supplementary	Positive
Denton	2013	USA	RCT	72	7-8	RI - adaptation of Responsive Reading Instruction and Read Naturally	Tier 3	High	Long	Out of the classroom	Specially recruited and trained	Supplementary	Positive
Duff	2014	UK	RCT	145	5-7	RI - Reading and Language Intervention	Multi-tier - 2,3	High	Short	Unspecified	Teaching assistants	Supplementary	Null
Duff	2012	UK	Quasi	59	6	RI - Reading Intervention	Multi-tier - 2,3	High	Short	Unspecified	Teaching assistants	Supplementary	Positive
Ecalle	2009	France	RCT	28	6	RI - Computer-Assisted Learning (syllabic units highlighted)*	Tier 3	High	Short	Unspecified	Unclear	Unclear	Positive
Ehri	2007	USA	Quasi	186	6-7	RI - Reading Rescue	Tier 3	High	Medium	Unspecified	School staff	Supplementary	Positive
Fien	2015	USA	RCT	267	6-7	RI - Enhanced Core Reading Instruction	Multi-tier - 1,2	High	Long	In and out of the classroom	Class teachers and specially recruited and trained	Supplementary	Positive

Fives	2016	Ireland	RCT	229	6-8	RI - Wizards of Words	Tier 3	Low	Long	Out of the classroom	Specially recruited and trained	Supplementary	Positive
Fives	2013	Ireland	RCT	229	6-8	RI - Wizards of Words	Tier 3	Low	Long	Out of the classroom	Specially recruited and trained	Supplementary	Positive
Gilbert	2013	USA	RCT	212	6-7	RTI - packaged program	Multi-tier - 2,3	Medium	Medium	Out of the classroom	Specially recruited and trained	Supplementary	Positive
Gustafson	2011	Sweden	RCT	130	6-8	RI - COMPHOT and OmegaIS*	Tier 3	High	Short	Unspecified	School staff	Supplementary	Positive
Hatcher	2006	UK	RCT	77	5-6	RI - Modified RI programme (Jolly Phonics and Sound Linkage)	Multi-tier - 2,3	High	Medium	Unspecified	Teaching assistants	Supplementary	Positive
Hatcher	1994	UK	Quasi	125	7	RI - Reading with phonology, reading alone, phonology alone	Tier 3	Medium	Medium	Unspecified	School staff	Supplementary	Positive
Hay	2007	Australia	Quasi	116	6-7	RI - Blank's theory of language enhancement	Multi-tier - 1,2	High	Long	In classroom and unclear	Class teachers and specially recruited and trained	Supplementary	Positive
Helf	2014	USA	Quasi	303	5-6	RI - Early Reading Tutor	Tier 2	High	Long	In and out of the classroom	Class teachers and school staff	Supplementary	Positive
Houtveen	2012	Netherlands	Quasi	1021	7-8	RTI - Reading Acceleration Programme	Multi-tier - 1,2,3	High	Long	In classroom and unclear	Class teachers and school staff	Instead of	Positive
Kamps	2008	USA	Quasi	83	5-8	RI - Reading Mastery, Early Interventions in Reading, and Read Well	Multi-tier - 2,3	Medium	Long	Unspecified	School staff	Supplementary	Positive
Kyle	2013	UK	Quasi	31	6-7	RI - English GraphoGame Rime and English GraphoGame Phoneme*	Tier 3	High	Medium	Out of the classroom	Specially recruited and trained	Supplementary	Positive

le Roux	2014	South Africa	RCT	102	8	RI - Animal-Assisted Reading Program	Tier 2	Low	Short	Out of the classroom	Specially recruited and trained	Unclear	Positive
Macaruso	2006	USA	Quasi	167	6-7	RI - Lexia Learning Systems*	Tier 1	Medium	Long	Out of the classroom	Class teachers and school staff	Supplementary	Positive
Mathes	2003	USA	RCT	89	6-7	RI - PALS and Teacher Directed Instruction	Multi-tier - 1,2	Medium	Medium	In the classroom	Class teachers	Instead of	Positive
Mathes	1998	USA	Quasi	96	6-7	RI - PALS	Tier 1	Medium	Medium	In the classroom	Class teachers	Instead of	Positive
Mathes	2001	USA	RCT	130	6-7	RI - PALS and Mini-Skills Lessons	Multi-tier - 1,2	Medium	Medium	In the classroom	Class teachers	Instead of	Positive
May	2015	USA	RCT	866	6-7	RI - Reading Recovery	Tier 3	High	Medium	Unspecified	Specially recruited and trained	Supplementary	Positive
Messer	2017	UK	RCT	78	7.7	RI - Trainertext*	Tier 3	Medium	Long	Unspecified	Teaching assistants	Supplementary	Positive
Oostdam	2015	Netherlands	RCT	265	5-7	RI - Continuous Reading or Repeated Reading	Multi-tier - 2,3	High	Medium	Out of the classroom	Teaching assistants	Supplementary	Positive
Pinnell	1994	USA	RCT	324	6-7	RI - Reading Recovery	Multi-tier - 2,3	High	Medium	Unspecified	Specially recruited and trained	Supplementary	Positive
Pullen	2010	USA	Quasi	224	6-7	RTI - Storybooks and Target Vocabulary	Multi-tier - 1,2	Medium	Short	In classroom and unclear	Class teachers and specially recruited and trained	Supplementary	Positive
Regtvoort	2013	Netherlands	RCT	137	6	RI - Build!*	Tier 3	High	Long	Unspecified	Specially recruited and trained	Supplementary	Positive
Rimm-Kaufmann	1998	USA	RCT	42	6-7	RI - Adult Volunteer Tutoring	Tier 3	Medium	Long	Unspecified	Specially recruited and trained	Supplementary	Positive
Saine	2011	Finland	RCT	166	7	RI - GraphoGame*	Tier 2	High	Long	Out of the classroom	Specially recruited and trained	Supplementary	Positive
Savage	2005	UK	Quasi	108	6	RI - Soundworks, Rime, Mixed	Tier 2	High	Short	Unspecified	Teaching assistants	Supplementary	Positive
Schwartz	2005	USA	RCT	148	6-7	RI - Reading Recovery	Tier 3	High	Medium	Unspecified	Specially recruited and trained	Supplementary	Positive

Sénéchal	2012	Canada	RCT	59	5-6	RI - Invented Spelling	Tier 2	Medium	Short	Unspecified	Specially recruited and trained	Supplementary	Positive
Smith	2016	USA	RCT	811	6-7	RTI - Enhanced Core Reading Instruction	Multi-tier - 1,2	High	Long	In classroom and unclear	Class teachers and specially recruited and trained	Instead of	Positive
Torgesen	2010	USA	RCT	112	6-7	RI - Read Write and Type*	Tier 2	High	Long	Out of the classroom	Specially recruited and trained	Supplementary and instead of	Positive
Van der Kooy-Hofland	2012	Netherlands	RCT	110	5	RI - Living Letters*	Tier 3	Low	Medium	In and out of the classroom	Specially recruited and trained	Supplementary	Positive
Vaughn	2006a	USA	RCT	171	6-7	RI - modified Proactive Reading	Tier 2	High	Long	Unspecified	Specially recruited and trained	Supplementary	Positive
Vaughn	2006b	USA	RCT	64	6	RI - modified Proactive Reading	Tier 2	High	Long	Unspecified	Specially recruited and trained	Supplementary	Positive
Vaughn	2006c	USA	RCT	41	6-7	RI - modified Proactive Reading	Tier 2	High	Long	Unspecified	Specially recruited and trained	Supplementary	Positive
Vernon-Feagans	2010	USA	RCT	200	5-7	RI - Targeted Reading Intervention	Tier 2	High	Short	In the classroom	Class teachers	Supplementary	Positive
Vernon-Feagans	2012	USA	RCT	276	5-7	RI - Targeted Reading Intervention	Tier 2	High	Short	In the classroom	Class teachers	Supplementary	Positive
Vernon-Feagans	2013	USA	RCT	631	5-7	RI - Targeted Reading Intervention	Tier 2	High	Long	In the classroom	Class teachers	Supplementary	Positive

Studies	Year	Country	Design	Sample size	Age in years	Intervention type/focus	Targeted tier	Intensity	Duration	Delivery location	Delivered by	Supplementary/ instead of	Gains
Burton-Archie	2014	USA	Quasi	252	7-8	RI - Leveled Literacy Intervention	Tier 3	High	Medium	Out of the classroom	Specially recruited and trained	Supplementary	Negative
Lucas	2004	UK	RCT	126	7-8	RI - modified Reading Recovery	Multi-tier - 2,3	High	Medium	Out of the classroom	School staff	Supplementary	Null
Ming	2007	USA	Quasi	34	6-7	RI - combining sight word and spelling with timed repeated readings	Tier 2	High	Short	In the classroom	Doctoral researcher	Instead of	Positive
Mitchell	2010	USA	Quasi	20	7-8	RI - Dr Cupp's Readers	Tier 2	Unspecified	Short	Unspecified	Unspecified	Unspecified	Null
Rehmann	2005	USA	Quasi	66	5-7	RI - Earobics Step 1*	Tier 3	Medium	Short	Out of the classroom	School staff	Instead of	Null
Samuelson	2010	USA	Quasi	40	6-8	RI - progress monitoring	Multi-tier - 2,3	High	Medium	Unspecified	School staff	Supplementary	Null
Wille	1993	USA	Quasi	10	6-7	RI - Project Read	Tier 2	High	Short	Out of the classroom	Doctoral researcher	Supplementary	Null

Table 2 - Doctoral dissertations