Volume 3, Issue 2 June 2008

EBP briefs

A scholarly forum for guiding evidence-based practices in speech-language pathology

The Use of Phonics in the Teaching of Reading and Spelling

Greg Brooks
University of Sheffield, UK

Carole J. Torgerson
Jill Hall
University of York, UK

A publication of Pearson ISSN 1941-7756

EBP Briefs
A publication of Pearson

Cite this document as:
Brooks, G., Torgerson, C.J., & Hall, J. (2008).
The Use of Phonics in the Teaching of Reading and Spelling
EBP Briefs 3(2), 1-12

The Use of Phonics in the Teaching of Reading and Spelling

Greg Brooks
University of Sheffield, UK

Carole J. Torgerson
Jill Hall
University of York, UK

Pedagogical Context

Much ink, though fortunately little blood, has been spilt in the 'reading wars' over methods of teaching children to read and write, but much of the argument has created more heat than light. While there will undoubtedly remain irreconcilables at the 'phonics only' and 'whole language' poles, there appears now to be sufficient reliable research evidence to provide conclusions on some of the questions in the field.

Phonics can be defined as a set of *methods of teaching literacy, particularly initial reading and spelling, that focus on the relationships between letters and sounds*.

The principal positions taken on phonics can be summarised thus:

There have been long and bitter disputes over the role of phonics in teaching children to read and write – what has research to say?

- systematic phonics instruction has a crucial and central role to play in the teaching of reading and spelling, or
- phonics has either no role, or at most a peripheral and incidental role, to play in this process.

It would be fair to say that most advocates of the latter position would endorse the whole language view that providing children with a broad and rich language and literacy curriculum is essential in instilling not just the ability to read but also the enjoyment of reading; extreme proponents on this wing would say that whole language is not only essential but sufficient, and therefore that phonics is unnecessary. It would also be fair to say that many phonics advocates also stress the need for a broad and rich language and literacy curriculum, but some extreme proponents of phonics insist that the very first stage in literacy teaching should consist of 'phonics first, fast and only'.

A further important distinction in the field is between the two principal forms of phonics teaching. In **synthetic phonics**, 'sounding-out' is used. For reading, this is based on the letters in printed words and is followed by blending their sounds to produce a spoken word that the learner should recognise. The classic example is producing the sequence of sounds 'kuh – a – tuh' in response to the letters <c, a, t>, first separately, then merged to produce the whole spoken word 'cat'. For writing (spelling), sounding-out is based on a spoken word that the learner knows and is followed by writing the corresponding letter for each sound. In **analytic phonics**, by contrast, sounding-out is not used. Instead, teachers show children how to deduce the common letter and sound in a set of words that all begin (or, later, end) with the same letter and sound, e.g. *pet, park, push, pen*.

In this brief we summarise a systematic review (Torgerson *et al.*, 2006), funded by the UK's former Department of Education and Skills, and we compare it to an earlier review conducted in the United States (Ehri *et al.*, 2001).

The Questions

These can be summarised as:

Does systematic phonics teaching enable children to make better progress in reading *accuracy* (word identification) than unsystematic or no phonics teaching?

Does systematic phonics teaching enable both normally developing children and those at risk of failure to make better progress in reading *accuracy* than unsystematic or no phonics teaching?

Does systematic phonics teaching enable children to make better progress in reading *comprehension* than unsystematic or no phonics teaching?

Does systematic phonics teaching enable children to make better progress in spelling than unsystematic or no phonics teaching?

Does *systematic* synthetic phonics instruction enable children to make better progress than systematic *analytic* phonics teaching?

Searching for Evidence

Non-systematic reviews may give misleading results if it is not clear why some studies were included and others were not, and may be subject to reviewer bias. We opted for a systematic review, which allowed us to search for and include studies with the strongest research design, randomized controlled trials (RCTs) in the evidence base. Ehri *et al.* included 38 studies, of which only 13 were randomized controlled trials, and the other 25 were non-randomized controlled trials.

The starting point in identifying trials for potential inclusion in this review was the 13 RCTs included in Ehri *et al.* (2001). In order to locate any further potentially relevant published or unpublished RCTs, a number of searches were undertaken in major educational databases: Education Resources and Information Center (ERIC) and PsycINFO (psychological literature) to capture published and unpublished trials (Ehri *et al.* used only published trials). Three extra databases were searched: SIGLE (System for Information on Grey Literature; N.B. This database was closed down in 2006), ASSIA (Applied Social Sciences Index and Abstracts) and BEI (British Education Index).

We also wrote to Linnea Ehri to request bibliographic details of (a) the five published studies excluded from the Ehri *et al.* (2001) review because they compared synthetic phonics teaching and analytic phonics teaching; (b) the studies she and her team identified but excluded because they were unpublished; and (c) any studies they knew of that should be included in the update.

A total of 6,114 potentially relevant studies were identified through searching the Ehri *et al.* review and the five electronic databases, and contact with authors.

Evaluating the evidence

Two of the team (CJT and JH) first screened all these studies using titles and abstracts; this reduced the database to 101 potentially relevant papers. One paper was unobtainable. The full texts of the other 100 papers were obtained, and screened (by CJT and JH) using the following pre-established inclusion and exclusion criteria.

Inclusion criteria

Trials with the following characteristics were included:

- RCTs focusing on the teaching of phonics in English, and comparing either:

  (a) the effectiveness of teaching using systematic phonics with that of teaching providing unsystematic or no phonics (but where the control condition included some alternative reading instruction); or

  (b) the effectiveness of synthetic phonics teaching compared with analytic phonics teaching.

and

- trials that measured reading as an outcome, reported statistics permitting the calculation or estimation of effect sizes, and involved interventions that might be found in schools.

Exclusion criteria

Trials were excluded if they:

- were not randomized controlled trials;
- did not evaluate either the relative effectiveness of systematic synthetic and analytic phonics teaching or of some form of systematic phonics teaching versus unsystematic or no phonics teaching (but an alternative reading instruction);
- were 'short-term laboratory studies with a limited focus' (Ehri *et al.*, 2001), e.g., a study in a psychology laboratory lasting only a few hours;
- lacked reading as an outcome;
- lacked statistics allowing calculation or estimation of effect sizes;
- primarily investigated phonemic awareness or phonological awareness instruction (such studies were also excluded in Ehri *et al.*, 2001); or
- compared two or more kinds of synthetic phonics teaching.

A total of 20 RCTs (in 19 papers) were included; full publication details of these are listed in the first part of the references. Of these 20 trials, six were excluded from further analysis because the experimental treatments were different varieties of systematic phonics instruction other than analytic and synthetic, and two more were excluded because they were cluster trials, i.e., trials in which the units were whole classes rather than individual children; these would have been difficult to compare with the studies in which the units were individual children.

Table 1 contains a summary of the quality assessments of the 12 remaining trials. This table is based on the modified CONSORT guidelines for quality assessment of

RCTs. (The Consolidated Standards for Reporting Trials are the methodological standards adopted by many medical journals for publication of randomized controlled trials; see Altman, 1996 and Altman et al., 2001.) As can be seen from Table 1, none of the 12 trials reported method of random allocation or sample size justification, and only two reported blinded assessment of outcome. Where attrition was reported it was not severe, but some authors did not report how many participants were lost. The trials were, therefore, variable in quality, but all were lacking in their reporting of some issues that are important for methodological rigour. Quality of reporting is a good but not perfect indicator of research design quality. Therefore, due to the limitations in the quality of reporting we judged the overall quality of the trials to be 'medium' to 'weak'.

12 randomized controlled trials on the impact of phonics teaching were identified for detailed statistical analysis.

Drawing Evidence-based Conclusions

One of the ways we can understand how research evidence is applied to the school setting is through the use of a statistic called an *effect size* (noted with the symbol *d*). This is a fairly straightforward comparison of the difference between the treated and control performance at the end of the treatment/instruction program. In general, the strength of an effect can range from small (effect size around 0.2) through medium (effect size around 0.5) to large (effect size around 0.8). However, there is one other piece of information that is often not reported that is very helpful in helping us understand the magnitude of the impact of a treatment program – the confidence interval (CI). The CI takes into account the sampling error, namely the fact that, by chance, some children or classes with certain characteristics will be sampled into one group more frequently than into another. An effect can be statistically significant at the level of $p<0.05$ (which means that, if the experiment is repeated 100 times, on average 95 times out of 100 the confidence intervals will contain the true effect), or it can be non-significant (which means that there are no statistically significant differences between the groups). The easy way of understanding the level of significance is to simply note whether or not theCI includes the value of zero (0). If it does, it means that the difference between the treated and control groups is not statistically significant. Said another way, it is possible that the treatment was not the reason for the treated groups' performance. With this in mind, the results of our review follow.

Impact of Phonics Instruction

Effect sizes were calculated for reading accuracy (word identification) (n=12) reading comprehension (n=4), and for spelling (n=3). Though some studies reported vocabulary measures and/or follow-up assessments, too few did so to make calculation of effect sizes from those data worthwhile.

Table 2 contains information about each of the 12 RCTs. The table includes information about participants, intervention and control treatments, sample sizes, and the effect sizes for word reading accuracy, reading comprehension and spelling as calculated for this review.

Using the effect sizes, four principal meta-analyses were undertaken:

- Systematic phonics teaching versus alternative reading interventions providing unsystematic or no phonics teaching, i.e. whole language/whole word ('look-and-say') methods, with reading *accuracy* as the outcome measure
- Systematic phonics teaching versus unsystematic or no phonics teaching, with reading *comprehension* as the outcome measure
- Systematic phonics teaching versus unsystematic or no phonics teaching, with *spelling* as the outcome measure
- Synthetic phonics teaching versus analytic phonics teaching, with reading *accuracy* as the outcome measure.

Phonics Instruction versus Whole Language

Reading accuracy

For measures of word reading accuracy, in nine of the 12 trials the effect size for word accuracy was positive, and ranged from extremely small (0.07) to extremely large (2.69). Only the two extremely large effect sizes were statistically significant. In three of the included studies the effect size was negative and small but in no case was it statistically significant.

The 12 studies were pooled in a meta-analysis, and estimates of effect as seen in Figure 1 show that there was a statistically significant effect of phonics instruction on reading accuracy of 0.27.

Inspection of the details of the interventions revealed that some were undertaken with children with reading difficulties or disabilities, and others with typical children. We considered that the advantage for phonics instruction might be accounted for by the student group classification. An analysis accounting for group identification showed that phonics instruction tended to produce a larger effect size on reading accuracy for typically developing children (0.45) than for children with reading disabilities and difficulties (0.21). However, the difference between the groups was not statistically significant (p=0.24). Therefore, there was no statistical evidence to support the belief that the effectiveness of phonics instruction was different for learners identified as having reading difficulties or as typical developing readers.

Comprehension

Four of the 12 RCTs used comprehension as an outcome measure at immediate post-test. The average effect size for these four trials was 0.24 but this was not statistically significant. Thus, we cannot present evidence to support or not support the use of phonics instruction to improve reading comprehension.

Spelling

In addition, three of the 12 studies included in the main analysis used spelling as an outcome measure at immediate post-test. The average effect size for these three trials was 0.09 but this was not statistically significant. Again, these data provide no evidence either way on whether phonics instruction benefits children learning to spell.

Synthetic Phonics versus Analytic Phonics instruction

Three studies directly compared systematic synthetic phonics teaching with systematic analytic phonics teaching. The estimate of effect size was 0.02, a statistically non-significant result.

Judgement of Evidence

When considering the findings reported above it is important to express them in an overall judgement of the evidence based on three things: (1) the strength of the effect, (2) the statistical significance of the effect, and (3) the quality of evidence on which these are based.

The quality of evidence relates to the sample size of the individual trials, the methodological rigour of the individual trials and the number of trials included in the analysis.

With this in mind, we concluded that:

- none of the findings of the current review were based on strong evidence because there simply were not enough trials (regardless of quality or size);
- the findings on reading accuracy were based on moderate evidence (because there were a few trials of variable quality with small sample sizes);
- the findings on comprehension, spelling and analytic versus synthetic phonics were based on weak evidence (because there were very few trials with small sample sizes and variable quality); and
- on vocabulary and follow-up assessments there was insufficient evidence to support any finding.

The quality of the evidence for a finding and its effect and/or statistical significance may be independent of each other. It would be desirable to base recommendations for changes in teaching on highly statistically significant medium to large effects based on good quality of evidence (either several moderately sized, good quality trials or on one well-designed very large trial in a normal school setting). But since there are no such findings at present, it is necessary to proceed on the basis of the evidence that is available.

Summary of findings

Heeding the cautions expressed in the previous subsection, our findings can be found in Table 3.

Two of the main findings of the current review supported those of Ehri *et al.* (2001), namely, that systematic phonics instruction enables children to make better progress in reading accuracy than unsystematic or no phonics, and that this is true for both normally developing children and those at risk of failure. However, there were some important differences. Ehri *et al.* found significant findings for comprehension and spelling, where we did not. The overall effect size of 0.27 for reading accuracy was substantially lower than Ehri et al.'s estimate of 0.41 (implying

There was just enough evidence to conclude that systematic phonics teaching benefited children's reading accuracy relative to unsystematic or no phonics teaching, and that this is true for both normally developing children and those at risk of failure.

that approximately 12 extra children out of 100 rather than approximately 16 extra children out of a hundred would succeed on a relevant test. This is based on the principle of 'numbers needed to teach' or NNT. If one in ten children benefit from an intervention, ten children would need to be exposed to the intervention for one to benefit). This reduction in the effect size may have been due to the inclusion of new trials from the updated searches, and/or to some features of the Ehri *et al*. review, namely:

- the fact that they included non-randomized as well as randomized trials;
- their use of estimated rather than actual numbers in the different groups in two studies;
- their use of what was essentially an untaught control group.

Recommendations for teaching

- Since there is evidence that systematic phonics teaching within a broad and rich language and literacy curriculum benefits children's reading accuracy, it should be part of every literacy teacher's repertoire and a routine part of literacy teaching, in a judicious balance with other elements.
- Teachers who already use systematic phonics in their teaching should continue to do so; teachers who do not should add systematic phonics to their routine practices.
- There is no justification for withholding phonics from either normally developing children or those at risk of reading failure – both may benefit and it should be used with both.

Other than these recommendations, there is little evidence to justify changes to existing practice. In particular,

- There is currently no strong RCT evidence that any one form of systematic phonics is more effective than any other.
- Two other areas on which the existing research base is insufficient are whether or not phonics teaching boosts comprehension, and whether phonics should be used to teach spelling as well as reading.

Recommendations for teacher training

- The evidence that systematic phonics teaching benefits children's reading accuracy further implies that learning to use systematic phonics in a judicious balance with other elements should form part of every literacy teacher's training.

Recommendations for research

- None of the findings of this review have very strong evidence in their support. What is needed is a well-designed, large-scale RCT to shed clearer light on the key findings. We therefore recommend a large cluster-randomized controlled trial (with classes or schools assigned at random) to confirm the findings of this review and to investigate further the relative effectiveness of systematic synthetic versus analytic phonics instruction with children with different learning characteristics.

References

A. The 20 RCTs

Berninger, V.W., Vermeulen, K., Abbott, R.D., McCutchen, D., Cotton, S., Cude, J., Dorn, S. and Sharon, T. (2003). Comparison of three approaches to supplementary reading instruction for low-achieving 2nd grade readers, *Language, Speech and Hearing Services in Schools*, 34(2): 101-116.

Brown, I.S. and Felton, R.H. (1990). Effects of instruction on beginning reading skills in children at risk for reading disability, *Reading and Writing: An Interdisciplinary Journal*, 2(3): 223-41.

Fayne, H.R. and Bryant, N.D. (1981). Relative effects of various word synthesis strategies on the phonics achievement of learning disabled youngsters, *Journal of Educational Psychology*, 73(5): 616-623.

Greaney, K.T., Tunmer, W.E. and Chapman, J.W. (1997). Effects of rime-based orthographic analogy training on the word recognition skills of children with reading disability, *Journal of Educational Psychology*, 89(4): 645-51.

Haskell, D.W., Foorman, B.R. and Swank, P. (1992). Effects of three orthographic/ phonological units on first-grade reading, *Remedial and Special Education*, 13(2): 40-49.

Johnston, R.S. and Watson, J.E. (2004). Accelerating the development of reading, spelling and phonemic awareness skills in initial readers, *Reading and Writing: An Interdisciplinary Journal*, 17(4): 327-57.

Leach, D.J. and Siddall, S.W. (1990). Parental involvement in the teaching of reading: A comparison of hearing reading, paired reading, pause, prompt, praise and direct instruction methods, *British Journal of Educational Psychology*, 60(3): 349-55.

Lovett, M.W., Lacerenza, L., Borden, S.L., Frijters, J.C., Steinbach, K.A. and De Palma, M. (2000). Components of effective remediation for developmental reading disabilities: Combining phonological and strategy-based instruction to improve outcomes, *Journal of Educational Psychology*, 92(2): 263-83.

Lovett, M.W., Ransby, M.R., Hardwick, N., Johns, M.S. and Donaldson, S.A. (1989). Can dyslexia be treated? Treatment-specific and generalized treatment effects in dyslexic children's response to remediation, *Brain and Language*, 37(1): 90-121.

Lovett, M.W. and Steinbach, K.A. (1997). The effectiveness of remedial programs for reading disabled children of different ages: Does the benefit decrease for older children?, *Learning Disability Quarterly*, 20(3): 189-210.

Lovett, M.W., Warren-Chaplin, P.M., Ransby, M.J. and Borden, S.L. (1990). Training the word recognition skills of reading disabled children: Treatment and transfer effects, *Journal of Educational Psychology*, 82(4): 769-80.

Martinussen, R.L. and Kirby, J.R. (1998). Instruction in successive phonological processing to improve the reading acquisition skills of at-risk kindergarten children, *Developmental Disabilities Bulletin*, 26(2): 19-39.

O'Connor, R.E. and Padeliadu, S. (2000). Blending versus whole word approaches in first grade remedial reading: Short-term and delayed effects on reading and spelling words, *Reading and Writing: An Interdisciplinary Journal*, 13(1-2): 159-82.

Skailand, D.B. (1971). *A comparison of four language units in teaching beginning reading*, Paper presented at the meeting of the American Educational Research Association, New York, USA, 4-7 February 1971.

Sullivan, H.J, Okada, M. and Niedermeyer, F.C. (1971). Learning and transfer under two methods of word-attack instruction, *American Educational Research Journal*, 8(2): 227-39.

Torgesen, J.K., Wagner, R.K., Lindamood, P., Rose, E., Conway, T. and Gravan, C. (1999). Preventing reading failure in young children with phonological processing disabilities: Group and individual responses to instruction, *Journal of Educational Psychology*, 91(4): 579-93.

Torgesen, J.K., Alexander, A.W., Wagner, R.K., Rashotte, C.A., Voeller, K.K.S. and Conway, T. (2001). Intensive remedial instruction for children with severe reading disabilities: Immediate and long-term outcomes from two instructional approaches, *Journal of Learning Disabilities*, 34(1): 33-58.

Umbach, B., Darch, C. and Halpin, G. (1989). Teaching reading to low performing first graders in rural schools: A comparison of two instructional approaches, *Journal of Instructional Psychology*, 16(3): 112-21.

Walton, P.D., Walton, L.M. and Felton, K. (2001). Teaching rime analogy or letter recoding reading strategies to pre-readers: Effects on pre-reading skill and word reading, *Journal of Educational Psychology*, 93(1): 160-80.

B. Other references

Altman, D.G. (1996). Better reporting of randomised controlled trials: The CONSORT statement, *British Medical Journal*, 313: 570-1.

Altman, D.G., Schulz, K.F., Moher, D., Egger, M., Davidoff, F., Elbourne, D., Gotzsche, P.C. and Lang, T. (2001). The revised CONSORT statement for reporting randomized trials: Explanation and elaboration, *Annals of Internal Medicine*, 134(8): 663-94.

Ehri, L.C., Nunes, S.R., Stahl, S.A. and 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.

Torgerson, C.J., 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. ISBN 1 84478 659 5

http://www.dfes.gov.uk/research/data/uploadfiles/RR711_.pdf

Table 1. Quality assessment of the 12 RCTs used for analysis

Author, date	Reporting of method of allocation	Sample size justification	Blinded assessment of outcome	Comments
Greaney *et al.* (1997)	N/S	N/S	Y	No attrition
Haskell *et al.* (1992)	N/S	N/S	N/S	
Johnston and Watson (2004), Exp. 2	N/S	N/S	N/S	Attrition n = 7. Random allocation only confirmed through contact with author
Leach and Siddall (1990)	N/S	N/S	N/S	
Lovett *et al.* (1989)	N/S	N/S	N/S	Numbers in each of the treatment groups requested and received from authors. Numbers only available for first battery of tests
Lovett *et al.* (1990)	N/S	N/S	N/S	Numbers in each of the treatment groups requested and received from authors
Martinussen and Kirby (1998)	N/S	N/S	N/S	Attrition n = 2 in phonics group. Results at floor for word attack test (meaning group)
O'Connor and Padeliadu (2000)	N/S	N/S	N/S	
Skailand (1971)	N/S	N/S	N/S	
Torgesen *et al.* (1999)	N/S	N/S	Y	
Torgesen *et al.* (2001)	N/S	N/S	N/S	Attrition n = 10 for two-year follow-up test
Umbach et al. (1989)	N/S	N/S	N/S	

N/S = not stated

Table 2. Characteristics of the 12 RCTs included in the first meta-analysis

Author, date	Participants	Intervention/control treatments	Sample size	Effect sizes and confidence intervals as calculated by the reviewers
Greaney *et al.* (1997)	'Disabled readers' * G3 – G6 (ages 8-11)	Rime analogy training or item-specific training (onset-rime versus look-and-say)	36 (18 in each group)	Accuracy: 0.29 (-0.37 to 0.95)
Haskell *et al.* (1992)	Normally attaining first grade (age 6) pupils	Phoneme level training group versus whole-word level training group	24 (12 in each group)	Accuracy: 0.07 (-0.73 to 0.87)
Johnston and Watson (2004), Exp. 2	Normally attaining Primary 1 (age 5) children	Synthetic phonics group versus no-letter training group (look-and-say)	92	Accuracy: 0.96 (0.42 to 1.50) Synthetic versus analytic: 1.32 (0.77 to 1.82)
Leach and Siddall (1990)	Normally attaining first grade (age 6) pupils	Direct instruction versus paired reading (look-and-say)	20 (10 in each group)	Accuracy: 0.80 (-0.11 to 1.71) Comprehension: 0.56 (-0.33 to 1.45)
Lovett *et al.* (1989)	'Disabled readers' **, mean age 10.8 years	Decoding skills programme group (DS) versus oral and written language stimulation group (OWLS, whole language)	121 (DS n = 60, OWLS n = 61)	Accuracy: 0.22 (-0.14 to 0.57) Comprehension: 0.08 (-0.28 to 0.44) Spelling: 0.07 (-0.29 to 0.42)
Lovett *et al.* (1990)	'Disabled readers' **, mean age 8.4 years	REG π EXC versus REG = EXC (look-and-say)	36 (18 in each group)	Accuracy: -0.19 (-0.85 to 0.46)
Martinussen and Kirby (1998)	Kindergarten (age 5) pupils assessed as low performers on phonological processing measures	Successive phonological group versus meaning group (whole language)	28 (13 in phonics group; 15 in meaning group). Attrition n = 2 from phonics group	Accuracy: 0.44 (-0.31 to 1.19) Spelling: 0.30 (-0.44 to 1.05)

Table 2. Characteristics of the 12 RCTs included in the first meta-analysis (continued)

Author, date	Participants	Intervention/control treatments	Sample size	Effect sizes and confidence intervals as calculated by the reviewers
O'Connor and Padeliadu (2000)	G1 (age 6) children nominated as 'very poor readers'	Blending versus whole word conditions	12 (6 in each group)	Accuracy: 0.53 (-0.62 to 1.68) Spelling: -0.15 (-1.28 to 0.99)
Skailand (1971)	Normally attaining kindergarten (age 5) children	Grapheme/ phoneme group versus whole word (look-and-say) group	42	Accuracy: -0.17 (-0.78 to 0.44) Synthetic versus analytic: -1.03 (-1.64 to –0.41)
Torgesen *et al.* (1999)	Kindergarten (age 5) children with weak phonological skills	PASP versus RCS	90 (45 in each group)	Accuracy: 0.07 (-0.34 to 0.48)
Torgesen *et al.* (2001)	Children between the ages of 8 and 10 identified as 'learning disabled' (= having learning difficulties)	Embedded phonics versus Auditory Discrimination in Depth Program	50	Accuracy: -0.31 (-0.87 to 0.45) Comprehension: 0.05 (-0.50 to 0.60) Synthetic versus analytic: -0.25 (-0.66 to 0.17)
Umbach *et al.* (1989)	First grade (age 6) students having difficulty with reading	Reading mastery (direct instruction) versus Houghton-Mifflin (look-and-say)	31 (15 in direct instruction, 16 in basal programme)	Accuracy: 2.69 (1.72 to 3.67) Comprehension: 1.08 (0.33 to 1.84)

* In the New Zealand context, these are 'children who fall within the bottom 1% to 2% of beginning readers' (Greaney et al., 1997, p.646).
** Children referred to the Learning Disabilities Research Program at the Hospital for Sick Children in Toronto, Canada who scored at least 1.5 years below expectation on EITHER word recognition accuracy OR reading speed (see Lovett et al., 1989, pp.97-98).

Figure 1: Meta-analysis of the 12 randomized trials

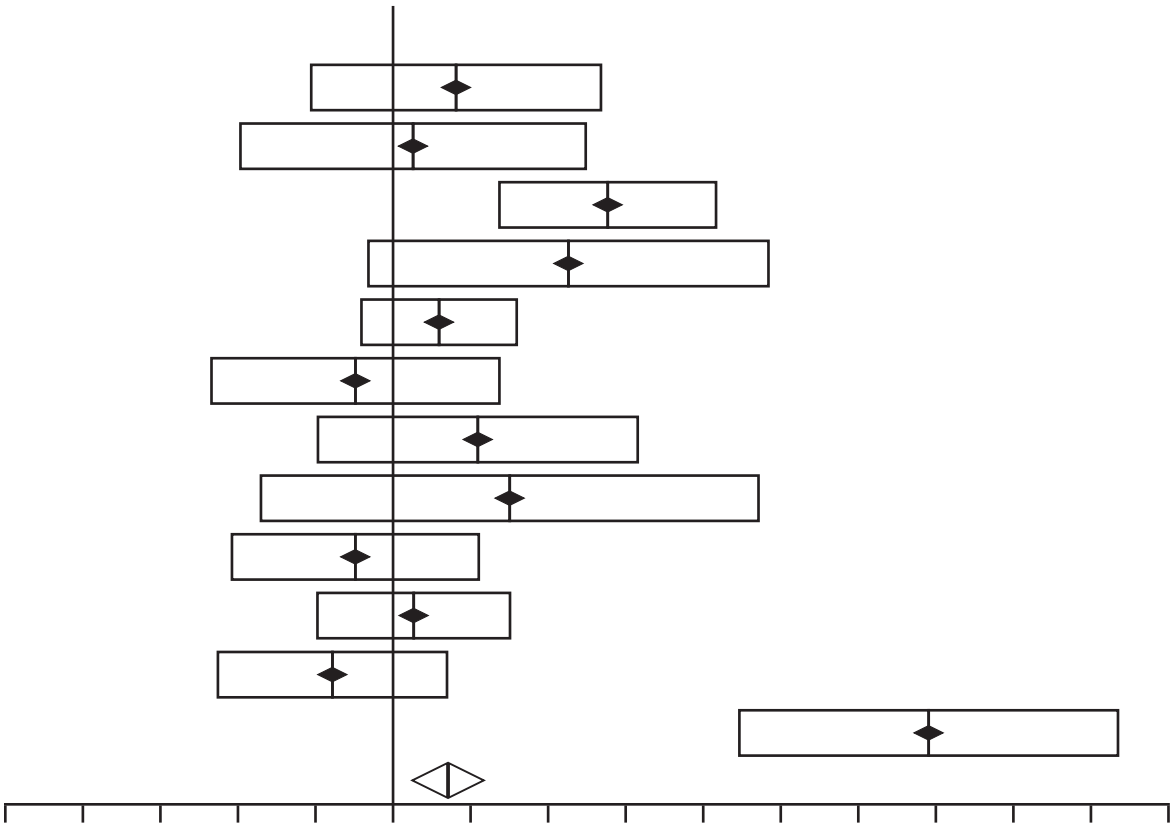

Table 3. Summary of findings, by review question, answer, quality of evidence, strength of effect, statistical significance, and implications

Review question	Answer	Quality of evdience	Strength of effect	Statistical signficance	Implications for teaching
Does systematic phonics teaching enable children to make better progress in reading accuracy than unsystematic or no phonics teaching?	Yes *	Moderate	Small (effect size = 0.27)	Highly statistically significant (p=0.002)	No warrant for NOT using phonics – it should be a routine part of literacy teaching
Does systematic phonics teaching enable both normally developing children and those at risk of failure to make better progress in reading accuracy than unsystematic or no phonics teaching?	Yes *	Moderate	Medium and small (effect sizes = 0.45 and 0.21 respectively)	Not statistically significant (p>0.05) N.B. The non-significant value implies no difference between groups.	No warrant for NOT using phonics with either group – both normally-developing children and those at risk of failure can benefit
Does systematic phonics teaching enable children to make better progress in reading comprehension than unsystematic or no phonics teaching?	Not clear	Weak	Small (effect size = 0.24)	Not statistically significant (p=0.08)	No clear finding from research on whether or not phonics boosts progress in comprehension
Does systematic phonics teaching enable children to make better progress in spelling than unsystematic or no phonics teaching?	Not clear	Weak	Very small (effect size = 0.09)	Not statistically significant (p=0.56)	No warrant from research for either using or not using phonics to teach spelling
Does systematic synthetic phonics teaching enable children to make better progress in reading accuracy than systematic analytic phonics teaching?	Not clear	Weak	Very small (effect size = 0.02)	Not statistically significant (p=0.87)	No warrant from research for choosing between these varieties of systematic phonics

* Finding supports one reported by Ehri et al. (2001), but with a reduced effect size