Discussion and Debate

Evidence strongly favours systematic synthetic phonics

Jennifer Buckingham provides a carefully-presented counterargument to the recent suggestion that the evidence in favour of systematic phonics instruction is weak, and that early phonics instruction should be replaced by 'Structured Word Inquiry'.

Within the community of research and practice that is informed by scientific evidence, there is a general acceptance that children need to learn the alphabetic code in order to be able to read accurately and for meaning, and that the most effective way to teach the code is through systematic and explicit phonics instruction (Castles et al., 2018).

A recent challenge to that consensus has come from Professor Jeffrey Bowers who has published a journal article that claims to show that "there is little or no evidence that systematic phonics is better than the main alternative methods used in schools, including whole language and balanced literacy" (Bowers, 2020, p. 1). Furthermore, Bowers says, "Once this is understood, my hope is that researchers and politicians will be more motivated to consider alternative methods."

Bowers is right to say that researchers should never consider that they have found the ultimate solution and stop looking for better ones. However, it is very different to propose that teachers and politicians should consider using unproven 'alternative' methods. Teaching practice and education policy should be based on the best available evidence unless and until it is superseded by new information and new evidence.

Bowers reviews major meta-analyses of studies that have looked at the effect of systematic phonics instruction on various reading outcomes. His key criticisms are that the strength of the measured effects of systematic phonics are overstated and that the studies do not directly compare systematic phonics with what he calls 'unsystematic phonics'.

However, Bowers' interpretation of the findings of these meta-analyses is not accurate. There is stronger evidence in favour of using systematic phonics in reading instruction than not using it.

What is systematic phonics?

The broad term 'systematic phonics' describes practices for the teaching of decoding and word reading. Evidence-based understandings of systematic phonics place it within a comprehensive program of instruction that includes four additional essential elements – phonemic awareness, fluency, vocabulary and comprehension. Alone, systematic phonics is not a fool proof guarantee of reading success and its effectiveness is mediated by the quality of the rest of the literacy program.

Systematic phonics does not supplant or contradict the need for instruction that develops language comprehension. Therefore, comparing the effects of systematic phonics instruction with comprehension-based programs is a false comparison. Both phonics and comprehension instruction are necessary; a finding of a positive effect of one on reading outcomes does not prove that the other is unnecessary. Measures of reading comprehension are measuring both word identification and language comprehension factors. In the early stages of reading development, word identification is the stronger predictor of reading comprehension, but once decoding is fluent, language comprehension becomes more important (Garcia and Cain, 2014).

According to Bowers (2020), "systematic phonics explicitly teaches children grapheme-phoneme correspondences prior to emphasizing the meanings of written words in text (as in whole language or balanced literacy instruction) or the meaning of written words in isolation (as in morphological instruction)." (p. 3)

This is incorrect. Systematic phonics does not preclude a focus on the meaning of words. There is no directive that learning grapheme-phoneme correspondences (GPCs) must precede all other elements of reading instruction.

The common recommendation that morphology instruction comes after a period of systematic phonics instruction (the precise optimal time for this has not yet been determined) is based on scientific evidence that the phonological pathway for decoding words is essential for beginning readers. While implicit morphological understanding is evident in young children's oral language, childrens' use of morphological

Study	Effect size (Cohen's d or Hedge's g)
National Reading Panel (2000)/ Ehri (2001)	Overall d = 0.67 (decoding regular words) d = 0.60 (decoding pseudowords) d = 0.40 (irregular words) d = 0.51 (reading comprehension) Type of phonics Synthetic phonics d = 0.45 (average for all measures) Analytic phonics d = 0.35 (average for all measures) Grade level Kindergarten d = 0.56 (average for all measures) First grade d = 0.54 (average for all measures) Grade 2-6 d = 0.27 (average for all measures)
Camilli, Vargas & Yurecko (2003)	d = 0.24 (average for all measures)
Camilli, Wolfe & Smith (2006)	d = 0.123 * phonics only instruction
Torgerson, Brooks & Hall (2006)	d = 0.27 / 0.38 (fixed effects/random effects; word reading accuracy) d = 0.24 / 0.35 (fixed effects/random effects; reading comprehension)
Suggate (2010)	d = 0.5 (average for all measures) d = 0.59 (pre-reading) d = 0.42 (reading) d = 0.41 (comprehension) d = 0.32 (average for all measures; follow up)
Adesope, Lavin, Thompson, & Ungerleider (2011)	g = 0.40
Galuschka, Ise, Krick & Schulte-Korne (2014)	g = 0.322 (average for all measures)
Suggate (2016)	Post-test d = 0.44 (average for all measures) d = 0.48 (pre-reading) d = 0.45 (reading skills) d = 0.48 (comprehension) Follow up d = 0.25 (average for all measures) d = 0.26 (pre-reading) d = 0.30 (reading skills) d = -0.03 (comprehension) * included unpublished and published studies ** 8 out of 22 phonics interventions were computer-based phonics training
McArthur et al (2018)	d = 0.51 (mixed/regular word reading accuracy) d = 0.67 (nonword reading accuracy) d = 0.84 (irregular word reading accuracy) d = 0.45 (mixed/regular word reading fluency) d = 0.39 (non-word reading fluency) d = 0.28 (reading comprehension)

Table 1: Effect sizes for reading outcomes associated with systematic phonics instruction

knowledge in word reading is demonstrated later (Rastle, 2019).

Systematic phonics can include synthetic and analytic approaches, which differ in the unit of sub-word analysis. Synthetic phonics begins with phonemes – the smallest sub-word level. Children learn the associations between speech sounds (phonemes) and the letters or letter clusters that represent them in writing (graphemes), and that this is a reversible process. They learn to synthesise the phonemes and graphemes to read and spell words. Synthetic phonics instruction has a defined sequence for teaching grapheme-phoneme correspondences.

Analytic phonics uses larger sub-word units such as onset-rime for word analysis. For example, rather than learning to read the word rat as a composition of three letters and sounds, r-a-t, children would learn that the word rat is in a 'word family' with the rime -at, such as r-at, s-at, c-at, and so on.

There are far fewer GPCs than there are 'word families' and learning phonics at the phoneme level is more systematic and efficient than onset-rime families (Vousden et al., 2011). The vast majority of rimes can be read using their component GPCs (Brooks, 2015). Knowledge of phonemes is a stronger predictor of early reading acquisition than knowledge of rimes (Nation & Hulme, 1997).

Is Bowers' interpretation of the meta-analyses fair?

In his review of evidence on systematic phonics, Bowers looks in detail at meta-analyses conducted over the past twenty years, starting with the National Reading Panel (2000) later published as Ehri et al. (2001).

Bowers argues that the effect sizes in these studies are not large and do not justify the authors' conclusions that systematic phonics has the strongest evidence in its favour. However, the effect sizes in these studies are certainly stronger than the evidence found for any other method, including whole language. Subsequent studies have added to the evidence in favour of including systematic phonics in reading instruction (for example, Hjetland et al. 2019). Detailed descriptions of these meta-analyses and an explanation of the flaws in Bowers' interpretation of them is provided in Buckingham (2020).

Table 1 presents the main findings of meta-analyses and systematic reviews of phonics instruction and intervention. These are the same studies reviewed by Bowers, with the exception of the McArthur et al (2018) study which supersedes the earlier study included in Bowers (2020). The effect sizes relate to the difference in reading outcomes associated with systematic phonics instruction/intervention as opposed to non-systematic or no phonics instruction.

The most common interpretation of effect sizes is that proposed by Cohen (1969): 0.2 is small; 0.5 is medium or moderate; and 0.8 is large. Some of the studies in the above table are reported as Hedges' 'g'. According to Torgerson et al. (2018), the difference between these types of effect size estimates is minimal. By Cohen's interpretation, the effect sizes in the table fall mostly in the moderate range. The outlier is Camilli et al. (2006) who achieved a small effect size by making multiple coding manipulations to the studies that are methodologically debatable.

However, a recent paper by Kraft (2020) explains that these effect size classifications were devised from clinical studies and makes a persuasive empirical case that they are not appropriate for applied educational research. Kraft proposes the following effect size interpretations: <0.05 is small; 0.05 – 0.2 is medium or moderate; and >0.2 is large. Using these interpretations, the effect sizes of using a systematic phonics program are almost all very large.

Bowers' other key criticism, aside from the relative effect sizes, is what he regards to be weak evidence directly comparing systematic phonics with 'unsystematic' phonics. Given the difficulty of classifying the comparison conditions as unsystematic phonics, whole language (with or without unsystematic phonics), balanced literacy, and rare 'no phonics' teaching, it seems reasonable and practical to do what almost all studies and meta-analyses have done – compare the presence of systematic phonics instruction with the absence of systematic phonics instruction.

The available evidence from multiple studies shows that reading instruction that includes systematic phonics is more effective than instruction that does not. The range of effect sizes is due to numerous factors, including the duration, level of systematicity, intensity, age of students, beginning level of students, group size, instructional fidelity, and the quality of classroom instruction. Nevertheless, the overall effect size is invariably and significantly positive.

What are the potential "alternatives" to systematic phonics instruction?

What are the alternative methods to systematic phonics, and what is the likelihood that they will be more effective? Bowers suggests that instruction "should focus more on the role that meaning plays in organizing spellings (via morphology) and that English spelling system (sic) makes sense once the interrelation between phonology, morphology, and etymology are considered." (p. 23).

Jeffrey Bowers' brother Peter Bowers has developed such a program – Structured Word Inquiry (Bowers & Bowers, 2008). Jeffrey Bowers has co-authored papers with Peter Bowers on the rationale for SWI (Bowers & Bowers, 2017) as well as participated in evaluations of the program (Colenbrander et al., 2018).

There is no problem with academics developing reading programs. Such reading programs would naturally be informed by the developers' understanding of the best available evidence. The problem with positing Structured Word Inquiry (SWI) as a superior alternative to systematic phonics is that there are no studies showing that SWI is effective for teaching beginning reading, either with or without the sort of comparison group that Bowers (2020) says is necessary to truly prove efficacy. Evaluations of SWI do not compare it with systematic phonics for initial instruction.

Studies of SWI show that children can benefit from instruction in morphology and etymology after one or more years of initial reading instruction that includes phonics (Bowers & Kirby, 2010; Devonshire & Fluck, 2010; Devonshire et al. 2013; Colenbrander et al., 2018). They do not provide evidence to support the argument that instruction based on morphology and etymology could or should be an alternative to systematic phonics in the initial stages of learning to read.

There is strong evidence for the inclusion of systematic phonics in initial reading instruction.

Systematic phonics has one of the largest and most consistent evidence bases in education. Synthetic phonics, which is the most systematic form of phonics instruction, has been specifically investigated in a number of randomised control trials (Christensen & Bowey, 2005; Hatcher, Hulme, & Snowling, 2004; Johnston, McGeown & Watson, 2011) and has been found to be a common factor in high performing schools (Joseph, 2019; Louden, 2015; OFSTED, 2010). After the introduction of mandatory synthetic phonics instruction in 2006 and a phonics screening check in 2012 in all English primary schools, there was an improvement in upper primary reading in national assessments and early indications of gains in international assessments (Buckingham, 2016; Machen et al., 2018; Double et al., 2019).

Synthetic phonics is strongly aligned with cognitive scientific research and models of reading that have been found to be highly predictive – the Dual Route Cascading Model (of word reading) and the Simple View of Reading (for reading comprehension) in particular (Castles, Rastle & Nation, 2018). The same cannot be said for whole language, balanced literacy, or analytic phonics.

And while there is some validity to the argument that meta-analyses provide a more accurate estimate of the effect of an intervention, there is also a good argument to be made for giving strong consideration to the findings of individual studies that investigate a higher quality version of the intervention of interest. Meta-analyses include interventions that are short in duration, with small numbers, and restricted instructional scope and depth. Emphasis should also be given to the findings of larger studies with implementations that more closely resemble what would generally be considered ideal classroom practice, such as the Clackmannanshire study (Johnston et al., 2011).

Bowers' thesis rests on the flawed argument that when held up to the highest possible standards of evidence, systematic phonics falls short. It is therefore illogical to suggest using "alternative teaching methods" that have either much weaker evidence or no evidence base whatsoever.

It is one thing to say that researchers should consider investigating unproven alternative methods, but it is irresponsible to make the same recommendation for teachers. Classroom practice should use the methods with the strongest evidence available base, and at the moment that is undeniably systematic synthetic phonics.

References

Adesope, O. O., Lavin, T., Thompson, T., & Ungerleider, C. (2011). Pedagogical strategies for teaching literacy to ESL immigrant students: a meta-analysis. *British Journal of Educational Psychology, 81*(Pt 4), 629–653. https://doi.org/10.1111/j.2044-8279.2010.02015.x

Bowers, J. S. (2020). Reconsidering the evidence that systematic phonics is more effective than alternative methods of reading instruction. *Educational Psychology Review*, Online first. https://doi.org/10.1007/s10648-019-09515-y

Bowers, S., & Bowers, P. (2008). Understanding SWI: "Structured word inquiry" or "scientific word investigation". http://www.wordworkskingston.com/WordWorks/Structured_Word_Inquiry.html

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

Bowers, P. N., & Kirby, J. R. (2010). Effects of morphological instruction on vocabulary acquisition. *Reading and Writing: An Interdisciplinary Journal, 23*(5), 515–537. https://doi.org/10.1007/s11145-009-9172-z

Brooks, G. (2015). *Dictionary of the British English Spelling System*. Open Book Publishers.

Buckingham, J. (2016). *Focus on phonics: why Australia should adopt the year 1 phonics screening check*. The Centre for Independent Studies. Research Report 22. https://doi.org/10.13140/RG.2.2.35472.61440

Buckingham, J. (2020). Systematic phonics belongs in evidence-based reading programs: A response to Bowers. *The Educational and Developmental Psychologist*, Online first. https://doi.org/10.1017/edp.2020.12

Camilli, G., Vargan, S., & Yurecko, M. (2003). Teaching children to read: The fragile link between science and federal education policy. *Education Policy Analysis Archives, 11*(15), 1–51. https://doi.org/10.14507/epaa.v11n15.2003

Camilli, G., Wolfe, M., & Smith, M. L. (2006). Meta-analysis and reading policy: Perspectives on teaching children to read. *The Elementary School Journal, 107*(1), 27–36. https://doi.org/10.1086/509525

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

Christensen, C. A. & Bowey J. A. (2005). The efficacy of orthographic rime, grapheme–phoneme correspondence, and implicit phonics approaches to teaching decoding skills. *Scientific Studies of Reading, 9*(4), 327–349. https://doi.org/10.1207/s1532799xssr0904_1

Cohen, J. (1969). *Statistical power analysis for the behavioral sciences*. Academic Press.

Colenbrander, D., Parsons, L., Murphy, S., Hon, Q., Bowers, J. & Davis, C. (2018). Morphological intervention for children with reading and spelling difficulties. https://www.nuffieldfoundation.org/wp-content/uploads/2019/11/Davis204217920-20Colenbrander20SSSR202018.pdf

Devonshire, V., & Fluck, M. (2010). Spelling development: Fine-tuning strategy use and capitalising on the connections between words. *Learning and Instruction, 20*(5), 361-371. https://doi.org/10.1016/j.learninstruc.2009.02.025

Devonshire, V., Morris, P., & Fluck, M. (2013). Spelling and reading development: The effect of teaching children multiple levels of representation in their orthography.

Double, K. S., McGrane, J. A., Stiff, J. C., & Hopfenbeck, T. (2019). The importance of early phonics improvements for predicting later reading comprehension. *British Educational Research Journal, 45*(6), 1220-1234. https://doi.org/10.1002/berj.3559

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

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

Garcia, J. G., & Cain, K. (2014). Decoding and reading comprehension: A meta-analysis to identify which readers and assessment characteristics influence the strength of the relationship in English. *Review of Educational Research, 84*(1), 74-111. https://doi.org/10.3102/0034654313499616

Hatcher, P. J., Hulme, C. & Snowling, M. J. (2004). Explicit phoneme training combined with phonic reading instruction helps young children at risk of reading failure. *The Journal of Child Psychology and Psychiatry, 45*(2), 338-358.

Hjetland, H. N., Lervåg, A., Lyster, S.-A. H., Hagtvet, B. E., Hulme, C., & Melby-Lervåg, M. (2019). Pathways to reading comprehension: A longitudinal study from 4 to 9 years of age. *Journal of Educational Psychology. 11*(5), 751-763. http://dx.doi.org/10.1037/edu0000321

Johnston, R., McGeown, S. & Watson, J. (2011). Long-term effects of synthetic versus analytic phonics teaching on the reading and spelling ability of 10 year old boys and girls. *Reading and Writing: An Interdisciplinary Journal, 25*(6), 1365–1384.

Joseph, B. (2019). *Overcoming the odds: A study of Australia's top performing disadvantaged schools*. Research Report 39. The Centre for Independent Studies.

Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher, 49*(4), 241-253. https://doi.org/10.3102/0013189X20912798

Louden, B. (2015). *High performing schools: What do they have in common?* Western Australia Department of Education.

Machin, S., McNally, S., & Viarengo, M. (2018). Changing how literacy is taught: Evidence on synthetic phonics. *American Economic Journal: Economic Policy, 10*, 217–241. https://doi.org/10.1257/pol.20160514

McArthur, G., Sheehan, Y., Badcock, N.A., Francis, D.A., Wang, H.C., Kohnen, S., Banales, E., Anandakumar, T., Marinus, E., Castles, A. (2018). Phonics training for English-speaking poor readers. *Cochrane Database of Systematic Reviews 2018*, Issue 11. Article No.: CD009115. DOI: 10.1002/14651858.CD009115.pub3.

Nation, K., & Hulme, C. (1997). Phonemic segmentation, not onset-rime segmentation, predicts early reading and spelling skills. *Reading Research Quarterly, 32*(2), 154-167. https://doi.org/10.1598/RRQ.32.2.2

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. Bethesda, MD: National Institute of Child Health and Human Development.

OFSTED (2010). *Reading by six: How the best schools do it*. UK Office for Standards in Education, Children's Services and Skills.

Rastle, K. (2019). The place of morphology in learning to read in English. *Cortex, 116*(July 2019), 45-64. https://doi.org/10.1016/j.cortex.2018.02.008

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. https://doi.org/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. https://doi.org/10.1177/0022219414528540

Torgerson, C. J., Brooks, G., & Hall, J. (2006). *A systematic review of the research literature on the use of phonics in the teaching of reading and spelling* (DfES Research Report 711). Department for Education and Skills, University of Sheffield.

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*, 1-31. https://doi.org/10.1080/02671522.2017.1420816

Vousden, J. I., Ellefson, M. R., Solity, J., & Chater, N. (2011). Simplifying reading: Applying the simplicity principle to reading. *Cognitive Science, 35*(1), 34-78. https://doi.org/10.1111/j.1551-6709.2010.01134.x

Dr Jennifer Buckingham is Director of Strategy and Senior Research Fellow at MultiLit. She is also the director of the Five from Five reading project.