In Clark, Margaret (Ed.). (2018). Teaching initial literacy: Policies, evidence and ideology. Birmingham, UK: Glendale Education.

Is Synthetic Phonics Working in England?
A Comment on the 'Teaching to Teach' Literacy Report

Jeff McQuillan, Senior Research Associate
Center for Educational Development, Los Angeles, CA, USA

Introduction

Since the publication of the 'Rose Report' (Rose, 2006) more than a decade ago, schools in England have been required to use 'synthetic phonics' (Wyse and Goswani, 2008) to teach children to read. Machin, McNally, and Viarengo (2016) analysed data from a large sample of English schoolchildren to assess the effectiveness of a phonics pilot program and early implementation of the new phonics mandate. Their paper, entitled 'Teaching to Teach Literacy', concluded that synthetic phonics was indeed effective compared to other methods of teaching reading in 'closing the gaps' between students who 'start out with disadvantages . . . compared to others' (p. 3).

I argue that the researchers' results do not support this claim. Both experimental studies and Machin *et al*.'s analysis show that phonics instruction has a modest effect on initial literacy levels, but little to no impact on reading achievement in later grades.

Analysis

Machin and colleagues examined the test scores of two different cohorts, each of which contained a group of students who were taught reading with synthetic phonics and a group who were not. Students were assessed at three different points:

- 'Foundation Stage' (at age 5, after one year of instruction),
- 'Key Stage 1' (at age 7), and
- 'Key Stage 2' (at age 11).

The Foundation Stage Profile (FSP) included several assessment scales, one of which was 'linking sounds and letters'. The FSP and Key Stage 1 assessments were done by the child's teacher, while Key Stage 2 was scored externally.

Machin *et al*. reported the *effect size* comparisons for the phonics groups and the non-phonics groups. Effect sizes allow us to see the magnitude of the impact an intervention has on test scores in a standardized unit, which is normally reported as the number (or fraction thereof) of standard deviations that separates the two groups (Hunter and Schmidt, 2004). This is especially important in studies such as this one, where the very large sample sizes (approaching 500,000) can cause even small score differences to be statistically significant. Effect sizes help us determine whether those differences are meaningful in practice.

Opinions differ on how to interpret effect sizes, but one widely used rule of thumb is that proposed by Cohen (1988): an effect size of 0.2 is 'small,' 0.5 is 'medium,' and 0.8 is 'large.' The U.S. Department of Education's *What Works Clearinghouse Handbook* (What Works Clearinghouse, 2014) recommended that effect sizes should be at least 0.25 to be considered 'substantively important' for education research (p. 23).

Synthetic Phonics Cohort Groups

I summarize Machin *et al*.'s effect size data in Table 1, broken down by student characteristics, including those who have English as an Additional Language (EAL) and children who were eligible for 'free meals' (all taken from their Table 6, p. 32). A positive effect size indicates that the phonics students scored higher than the non-phonics students.

Table 1

Effect Size Differences for Phonics and Non-Phonics Students in Machin et al. (2016)

	Cohort 1			**Cohort 2**		
	Age 5 (Foundation Stage)	Age 7 (Key Stage 1)	Age 11 (Key Stage 2)	Age 5 (Foundation Stage)	Age 7 (Key Stage 1)	Age 11 (Key Stage 2)
Native Speakers	0.225	0.052	-0.045	0.211	0.061	0.001
EALs	0.567	0.134	0.045	0.201	0.113	0.068
Non-Free School Meals	0.306	0.042	-0.061	0.182	0.104	0.042
Free School Meals	0.29	0.135	0.064	0.207	0.136	0.062
EALs and Free School Meals	0.3	0.216	0.181	0.221	0.195	0.099

There are two conclusions we can draw from these data, both consistent with previous, more rigorously designed studies.

1. Phonics Teaching Has a Moderate Impact Initially

Phonics training has a moderate impact during the first year of instruction. Effect sizes in favour of phonics were mostly in the small-to-medium range (0.2 to 0.5) at Age 5, although in Cohort 1, the effect size for EAL students was larger (0.567). However, we have no way of knowing how instruction affected children's *comprehension* scores versus their performance on decoding tasks, since these are combined into a single measure. The distinction is important: experimental research has found that phonics instruction boosts scores on isolated 'skills' tests, but has a much smaller impact on measures of reading comprehension (Garan, 2001; Krashen, 2009).

The results of three different meta-analyses (Ehri, Nunes, Stahl, and Willows, 2001; McArthur *et al.*, 2011; Torgerson, Brooks, and Hall, 2006, also reported in Torgerson *et al.*, 2018) show that phonics instruction has a medium-to-large impact on reading isolated words and pseudo-word words in all three analyses, ranging from 0.38 to 0.76. But on reading comprehension tests, the effects are much smaller: 0.27 in Ehri *et al*., 0.24 in Torgerson *et al*., and 0.14 in McArthur *et al*.'s review. The estimates from Torgerson *et al.* (2006) and McArthur *et al*. were not significantly different from zero. Any adequate analysis of the effectiveness of synthetic phonics, then, must examine the impact of instruction on comprehension tests as opposed to decoding measures. Machin *et al*.'s analysis does not.

2. The Effect of Phonics Instruction Fades Quickly

As can be seen in Table 1, by age 7, when most students are probably reading independently, the difference between the children taught synthetic phonics and the controls declines sharply for all groups, and is less than 0.2 for all comparisons except EAL + free meals group (and for Cohort 1 only: 0.216). By Age 11, only one of the 10 comparisons shown in Table 1 is even greater than 0.1, and all are under 0.2.

Again, experimental studies confirm these results. Suggate (2016) conducted a meta-analysis of 16 experimental studies on the long-term impact of phonics instruction. He found that the impact of phonics instruction all but disappears a year or so after it is introduced. Suggate reports that the effect size for reading comprehension dropped from 0.47 to -0.10 from immediate post-test to 'follow-up' tests (Table 3, p. 86; mean time to delayed post-test: 11.17 months). On tests involving reading words in isolation (grouped under the 'Reading' outcome), which are strongly influenced by decoding skills, the effect size advantage for phonics dropped from 0.26 to 0.08. Similar declines are reported in Vaden-Kiernan *et al.* (2017).

Conclusion: Policy Should Be Based on Best Evidence

Machin *et al*.'s analysis on synthetic phonics was based on a 'natural experiment,' allowing them to use a very large dataset with two separate cohorts. But policy decisions should be made on the strongest evidence we have, not the weakest (Garan, 2004), and the best evidence comes from *experimental* studies. In any case, the results of the 'Teaching to Teach' Literacy study do not support the assertion that synthetic phonics is having a positive impact on the reading scores of primary schoolchildren in England. The evidence Machin *et al.* presented is consistent with experimental studies that have found intensive phonics instruction makes a modest initial impact, but has very small effects on reading achievement later.

References

Cohen, J. (1988) *Statistical Power Analysis for the Social Sciences*. Thousand Oaks, CA: Sage Publications.

Ehri, L., Nunes, S., Stahl, S., and Willows, D. (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.

Garan, E. (2001) 'Beyond the smoke and mirrors: A critique of the National Reading Panel report on phonics'. *Phi Delta Kappan 82*(7): 500-506.

Garan, E. (2004) *In Defense of Our Children: When Politics, Profit, and Education Collide.* Portsmouth: Heinemann.

Hunter, J. and Schmidt, F. (2004) *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings*. 2nd edition. Thousand Oaks, CA: Sage Publications.

Krashen, S. (2009) 'Does intensive decoding instruction contribute to reading comprehension?' *Knowledge Quest, 37*(4): 72-74.

Machin, S., McNally, S., and Viarengo, M. (2016) 'Teaching to Teach Literacy'. *Centre for Economic Performance Discussion Paper* Number 1425.

McArthur, G., Eve, P., Jones, K., Banales, E, Kohnen, S., Anadakumar, T., and Castles, A. (2011) 'Phonics training for English-speaking poor readers'. *Cochrane Database of Systematic Reviews,* 12. doi: 10.1002/14651858.CD009115.pub2

Rose, J. (2006) *Independent Review of the Teaching of Early Reading. Final Report*. Nottingham, England: Department for Education and Skills (DfES) Publications.

Suggate, S. (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.

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*. Nottingham, England: Department for Education and Skills (DfES) Publications.

Torgerson, C., Brooks, G., Gascoine, L., and 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

Vaden-Kiernan, M., Borman, G., Caverly, S., Bell, N., Sullivan, K., Ruiz de Castilla, V., Fleming, G., Rodriquez, D., Henry, C., Long, T., and Hughes Jones, D. (2017) 'Findings From a Multiyear Scale-Up Effectiveness Trial of *Open Court* Reading'. *Journal of Research on Educational Effectiveness*, 1-24. Advance online publication. doi: 10.1080/19345747.2017.1342886

What Works Clearinghouse. (2014). *Procedures and Standards Handbook* (version 3.0). Washington, DC: Institute of Education Sciences, U.S. Department of Education.

Wyse, D., and Goswani, U. (2008). 'Synthetic phonics and the teaching of reading'. *British Educational Research Journal, 34*(6): 691-710.