

Success for All in England: Results from the Third Year of a National Evaluation

Technical Report

Louise Tracey Institute for Effective Education, University of York
Bette Chambers Institute for Effective Education, University of York
Robert Slavin Institute for Effective Education, University of York, and
Center for Research and Reform in Education, Johns Hopkins University
Pam Hanley Institute for Effective Education, University of York
Alan Cheung Department of Educational Administration and Policy, The Chinese
University of Hong Kong

July 2013



Abstract

This paper reports the third year findings of a longitudinal evaluation in England of Success for All (SFA), a comprehensive literacy programme. Eighteen SFA schools across England and 18 control schools, matched on prior achievement and demographics, were included in this quasi-experimental study. The results of hierarchical linear modelling analysis reveal a statistically significant positive school-level effect for SFA schools compared to control schools on standardised reading measures of word-level and decoding skills, and there were directionally positive but nonsignificant school-level effects on measures of comprehension and fluency. Practical and policy implications of these findings are discussed, particularly as they relate to recent English government policies encouraging schools to implement research-proven approaches.

Corresponding author: Louise Tracey, Institute for Effective Education, University of York, Heslington, York, UK YO10 5DD



Introduction

Currently, 25 per cent of children in the UK live in poverty, and the percentage is growing (Department of Work and Pensions, 2009; Wilkinson & Pickett, 2010). The chances of these children succeeding in school and in life are much smaller than are those of their more advantaged peers (Mongon & Chapman, 2008; Strand, 2008). Inequality starts early, and every year thousands of children enter school with educational differences already apparent (Lee & Burkham, 2002; Hills et al., 2010). This educational gap often continues to increase during a child's school career, and the longer they persist, the harder they are to close (Goodman et al., 2009).

Early literacy

One of the keys to later educational attainment is early literacy acquisition. Strong literacy achievement can lead to children being able to access the whole curriculum, and the earlier they do so the more beneficial it is. In contrast, early literacy problems can hinder children's knowledge and development, with long-term consequences for their educational outcomes (Lesnick, George, Smithgall, & Gwynne, 2010).

For many years research on beginning reading has supported the use of explicit, systematic phonics teaching (National Reading Panel, 2000). In particular the Rose Review (2006) called for phonics to be taught systematically in schools. Synthetic or systematic phonics involves teaching the discrete sounds that letters make, in order to help children blend or 'sound out' new words.

The importance placed on synthetic phonics in England is reflected in current policy. For example, following the Rose Review, the English Primary Framework and the associated phonics resources *Letters and Sounds* (Department for Education and Skills, 2007) emphasise synthetic phonics practices. *Letters and Sounds* outlines systematic phonics teaching in daily, briskly paced lessons to facilitate early literacy acquisition. In 2011, a national phonics assessment to be administered at the end of Year 1 (when children are typically 6 years old) was piloted and then rolled out nationwide in 2012. The government's intention was to incentivize schools to better support pupils to master basic phonics skills in Year 1. Only 40 per cent of children tested achieved the recommended level on the first assessment administered in Summer, 2012.

Education policy makers in England (and elsewhere) are beginning to encourage schools to implement programmes and practices with strong evidence of effectiveness in helping children living in poverty to succeed in school (Allen, 2011; DCSF, 2009). For example, they have instituted a policy of allocating a "pupil premium" to schools based on the number of children who are eligible for Free School Meals due to poverty. Schools are encouraged to adopt proven practices with this



additional funding. A systematic review of reading programmes suggested that the most successful programmes were those that included a broad curriculum that included systematic phonics as well as teaching in vocabulary and comprehension (Slavin et al., 2009). The Allen report (2011) listed programmes with a record of effectiveness. One of the few programmes listed in the top category was Success for All, based mainly on evidence from evaluations conducted in the US.

Success for All

Success for All (SFA) is a whole-school reform programme, started in the US in 1987. The programme's design is based on a model that posits that substantially enhancing success in schools in deprived areas depends on a multidimensional intervention approach. This includes providing extensive professional development, effective teaching strategies, emphasising co-operative learning, and school-wide structures focusing on school leadership, parent involvement, and attendance, which are expected to jointly enhance reading performance and other outcomes (see Slavin, Madden, Chambers, & Haxby, 2009 for a complete description). SFA began to be used in the UK in 1997, and it now serves more than 100 schools in England, Scotland, and Wales. One previous small-scale study found positive effects of SFA in the UK (Hopkins, Youngman, Harris & Wordsworth, 1999), although another study found positive outcomes in Year 1, but mixed outcomes in Year 2 (Tymms & Merrell, 2001). However, in a policy context demanding more rigorous evidence for interventions intended for use in the UK, there remained a need for a large-scale evaluation of the approach in schools as they exist today.

SFA includes reading materials that are colourful, engaging and well-organized, in line with the findings of the Rose Review (2006) and the US National Reading Panel (2000). It includes the use of synthetic phonics in Reception (age 4-5) and Year 1 (age 5-6), real children's literature, vocabulary development, and comprehension activities. In Years 2-6 (when children are aged 6-11), the emphasis is on comprehension strategies, vocabulary, and writing, which pupils do in co-operative groups.

During daily 90-minute reading periods, pupils are regrouped across age lines so that each reading class contains pupils all at one reading level. Every eight weeks children's reading is formally assessed and if children are performing ahead of their reading class they are moved up to the next level class. If children struggle to keep up with their classmates, they are assigned to either individual or small-group tutoring to help them catch up.

A strong emphasis on teaching in SFA is the use of co-operative learning. The co-operative learning strategies employed have children learning in mixed-ability pairs or groups of four, in which positive interdependence and individual accountability are key. This means that teams are



rewarded for each team member's learning. Each pupil is held accountable for his or her own learning and for helping group mates learn as well. Research has long suggested that co-operative learning has many benefits across the curriculum, including in literacy development (Stevens, 2003; Law, 2008; Slavin, 1995, 2009). Co-operative learning has been shown to increase pupil motivation, and the brisk pacing and in-built routines ensure that time is used effectively by both pupils and teachers. It is designed to give pupils opportunities to try out their understandings in a safe environment, to receive immediate feedback, and to "learn by teaching" in describing their current state of knowledge to a peer.

SFA is a complex intervention that requires considerable initial training and in-class coaching and support. This training is provided by a registered charity called Success for All-UK (SFA-UK), established in 1997. Each school has an appointed facilitator who ensures that teachers have what they need to implement the programme and co-ordinates with the training staff from SFA-UK to provide ongoing support. Schools receive at least six days of on-site support in their first year of using SFA.

Although its basic structure is the same as its US version, SFA has been substantially adapted to the language, culture, and standards of England, Scotland, and Wales. It is aligned with the *Letters and Sounds* requirements of England's Department for Education and emphasises the same curricular elements, focusing on the systematic teaching of phonemic awareness and phonics, as well as vocabulary and comprehension. It has a fast-paced and structured approach to teaching, intended to ensure that pupils have solid reading skills by the end of Key Stage 1 (when they are 7 years old).

Previous research on SFA

More than 40 empirical studies have shown positive effects of SFA on a variety of measures of pupil achievement (Slavin et al., 2009; Slavin, Lake, Davis, & Madden, 2011). The cumulative evidence from these studies shows positive effects of SFA on a variety of measures of pupil achievement, as well as on assignments to special education, retention, and other outcomes (Borman, Hewes, Overman, & Brown, 2003).

In particular, a large cluster randomised controlled trial in the US involved a three-year longitudinal sample of children who participated in the SFA or control condition from kindergarten through to the end of the second grade (Borman et al., 2007). Hierarchical linear model (HLM) analysis revealed statistically significant school-level effects of assignment to SFA on literacy outcomes. In addition, a synthesis of 23 studies of SFA found a mean effect size of +0.29 for pupils in general



and +0.52 for pupils in the lowest 25 per cent of their classes at pretest (Slavin, Lake, Davis, & Madden, 2011).

As noted earlier, the first major evaluation of the adapted programme in the UK found positive impacts on literacy outcomes (Hopkins, Youngman, Harris, & Wordsworth, 1999). Tymms and Merrell (2001) found positive effects in Year 1 and mixed effects at the end of Year 2. Using outcome estimates from the most rigorous studies with UK cost estimates, the Dartington Social Research Unit estimated that each pound invested in SFA yields £14.78 pounds of benefits to the individuals and society (<http://dartington.org.uk/projects/investing-in-children>). However, earlier UK studies involved a small number of schools, allowing for the possibility that school characteristics could explain observed differences. Also, these studies were conducted in the 1990s. The study reported here is the first large-scale, comprehensive evaluation of SFA.

Method

Research design

This quasi-experimental study involved 20 schools already implementing SFA and 20 schools matched to the SFA schools in terms of prior attainment and demographics, to provide a comparison cohort. As the intervention affects the whole school, a long term evaluation was considered to be the most appropriate approach to gain an assessment of the impact on pupils' reading achievement over time. The large number of schools involved in the study enables the use of appropriate statistical methods for clustered (school-level) designs, with adequate statistical power to detect true differences. It also allows for an evaluation of the programme as it is actually used in England as opposed to a small study that might provide more implementation support than a typical implementation.

Sample

Schools were recruited in Spring, 2008 by researchers from the Institute for Effective Education (IEE) at the University of York using a list provided by SFA-UK. Once 20 SFA schools were recruited, researchers started to recruit control schools whose overall characteristics matched those of the SFA schools. Key matching characteristics were:

- Percentage of children achieving Level 4 or above on the Key Stage 2 literacy SATs, for the three years prior to the intervention school adopting SFA;
- Percentage of pupils in receipt of (or eligible for) Free School Meals (FSM);
- Percentage of pupils with English as an Additional Language (EAL).



All schools (control and intervention) agreed to allow for individual testing of their children and to allow observers access to SFA/literacy classes of the appropriate year group. In autumn 2008 this was the children entering Reception classes. Children were pretested in September, 2008 and then were post-tested in June-July, 2009 at the end of Reception, and then at the end of Year 1 and Year 2. Control schools were provided with a financial reward of £2,000 per year for participating in the study to compensate for the potential disruption to the school during the assessment period.

In addition to receiving head teacher consent, parental information and opt-out forms were sent to the schools to be distributed to all children entering Reception that September. All children in Reception who had permission to participate were then individually pretested and post-tested. Each year, teachers were to continue with their normal classroom practices – whether that was SFA or any other method for teaching literacy (e.g., *Letters and Sounds*, *Jolly Phonics*, *Read-Write Inc.*). The SFA schools had been involved with SFA for between one and eight years. Trainers from SFA-UK made their normal implementation visits to each school throughout the year.

Table 1 shows the characteristics of the final sample of schools by the end of the academic year 2010-2011. Over the years, four schools had dropped out of the study, two control and two experimental. The experimental and control schools involved in the original baseline were well matched on all characteristics except the percentage of children with special educational needs. However, because of attrition, by the end of the study there was a significantly higher percentage of children with English as an Additional Language in the SFA schools (46% vs 24% for the control schools). SFA schools also had non-significantly higher levels of children eligible for Free School Meals (43% vs. 36%).

The sample includes schools from a range of regional contexts throughout England representing the national reach of the programme, with a relatively high percentage of children eligible for Free School Meals (approximately 40%).



Table 1. Comparison of baseline characteristics for Success for All (intervention) schools and control schools

Variable	Condition	N	Mean	Standard Deviation (SD)	t
Pretest BPVS	Intervention	18	37.88	5.75	0.607
	Control	18	39.00	5.03	
KS2 English SATs †	Intervention	17 ¹	53.24	13.11	-0.829
	Control	14 ¹	57.21	13.51	
Female	Intervention	18	49.93	9.07	0.352
	Control	18	48.73	11.34	
Enrolment	Intervention	18	334.5	136.68	0.622
	Control	18	307.3	125.3	
% pupils eligible for Free School Meals	Intervention	18	43.41	19.05	1.403
	Control	18	35.74	13.17	
% of SEN pupils	Intervention	17 ²	11.94	5.27	1.341
	Control	16 ²	9.74	4.05	
% of pupils with EAL	Intervention	18	45.87	34.26	2.052*
	Control	18	23.91	29.79	

*p < .05

† Percentage of pupils achieving Level 4 or above for the three years prior to the intervention school in each matched pair adopting the SFA programme.

¹ One intervention and four control schools were lower schools (ie took pupils up to age 9) during the three year period when the average KS2 English SATs were calculated for this study (see † above) so we did not have KS2 results for the specified timeframe.

² One intervention and two control schools were lower schools at the beginning of the study so percentage of SEN pupils was not available from the published Performance Tables.

Measures

The pretest, undertaken on entry to Reception (September 2008), was the British Picture Vocabulary Scale 2nd Edition (BPVS-II). This is a measure of receptive vocabulary and is an English adaptation of the Peabody Picture Vocabulary Scale. Children are told a word and then asked to point to one of four pictures that represents that word. The BPVS-II was normed on a national sample of children in the UK and has a Cronbach's Alpha of 0.93 and a split-half reliability of 0.86 (Dun, Dunn, Whetton, & Burley, 1997).



During June/July 2011, the same children at the end of Year 2 were post-tested using the Word Identification and Word Attack scales of the Woodcock Reading Mastery Tests-Revised (WRMT). The Word Identification scale measures the child's ability to read isolated words and the Word Attack scale assesses children's ability to decode and 'sound out' nonsense words. The WRMT was normed on a US national sample of children and the internal reliability coefficients for the two scales used were 0.97, and 0.87, respectively (Woodcock, McGrew, & Mather, 2001).

In addition, assessors administered the York Assessment of Reading Comprehension (YARC), a standardised measure of accuracy, reading rate and comprehension. The YARC was normed on a UK national sample of children. The reliability coefficients for the three subtests for Year 2 children were 0.87, 0.95 and 0.62, respectively. All assessors were hired, trained, and supervised by researchers at the IEE. Assessors were not made aware of schools' treatment conditions.

Intervention schools were visited by SFA consultants and researchers to assess implementation fidelity and control schools were visited by researchers to observe general literacy practices and assess to what extent key elements of the SFA programme (in particular co-operative learning) were being practiced.

During the evaluation, four schools left the study - two intervention and two control schools, making assessment of those children impossible. This was primarily due to a change in head teacher or direction of the school. Given the longitudinal nature of the study, a number of children left due to the high rate of movement between schools often found in vulnerable communities. As this was a longitudinal study and the analyses completed at the school level, subjects were not replaced with new pupils.

Analyses

This quasi-experimental cluster evaluation was analysed using a hierarchical linear model (HLM) with school as the unit of analysis (Raudenbush & Bryk, 2002). All multi-level models were estimated using the HLM software's restricted maximum likelihood (REML) estimation procedure (Raudenbush, Bryk, Cheong, & Congdon, 2000). Pretests (BPVS) were used as covariates. This multilevel approach is the optimal design for school-based interventions. It addresses the effects of pupils being clustered within schools, and it is well aligned with the theory of how this educational intervention works: as a co-ordinated, systemic initiative delivered by school-level elements acting in concert. Multilevel analysis greatly reduces statistical power, requiring many more schools than would be needed in individual-level analysis, but individual-level tests of statistical significance, which assume that the outcome for an individual is independent of that for any other pupil, are inappropriate for evaluations of school-level interventions.



Using HLM, one may simultaneously model both pupil and school-level sources of variability in the outcome (Raudenbush & Bryk, 2002). Specifically, we developed 2-level hierarchical models that nested pupils within schools. The fully specified level 1, or within-school model, nested pupils within schools. The linear model for this level of the analysis is written as

$$Y_{ij} = \beta_{0j} + r_{ij}$$

which represents the summer post-test achievement for pupil i in school j regressed on the level-1 residual variance, r_{ij} .

At level 2 of the model, we estimated SFA treatment effects on the mean post-test achievement outcome in school j . We included a school-level covariate, the school mean BPVS pretest score, to help reduce the unexplained variance in the outcome and to improve the power and precision of our treatment effect estimates. The fully specified level 2 model is written as

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(\text{Mean BPVS Pretest})_j + \gamma_{02}(\text{SFA})_j + u_{0j}$$

where the mean post-test intercept for school j , β_{0j} is regressed on the school-level mean BPVS score, the SFA treatment indicator, plus a residual, u_{0j} .

Results

Empirical analyses of reading achievement

For each of the achievement post-test outcomes a series of multi-level models was specified. The analyses for the WRMT subtests are shown in Table 2 and those for the YARC are shown in Table 3. These show a school-level significant effect on the WRMT Word Identification (ES = +0.20, $p < 0.03$) and Word Attack (ES = +0.25, $p < 0.01$). On the three YARC scales, effects were directionally positive but not statistically significant, with effect sizes as follows: Rate (ES = +0.11, n.s.), Comprehension (ES = +0.06, n.s.), and Accuracy (ES = +0.12, n.s.).



Table 2. Multilevel Models Predicting Pupil and School-Level Literacy Outcomes: Woodcock

Level 1 model: $Y_{ij} = \beta_{0j} + r_{ij}$

Level 2 model: $\beta_{0j} = \gamma_{00} + \gamma_{01}(\text{Mean BPVS})_j + \gamma_{02}(\text{SFA})_j + u_{0j}$

<i>Fixed Effect</i>	Word ID T=415, C=471			Word Attack T=415, C=471		
	Effect	SE	t	Effect	SE	t
School mean achievement						
Intercept	0.01	0.04	0.19	0.01	0.04	0.11
Mean BPVS pretest	0.02	0.10	0.16	-0.07	0.08	-0.90
SFA assignment	+0.20*	0.09	2.33	+0.25**	0.08	3.16
<i>Random Effect</i>	Estimate	χ^2	Df	Estimate	χ^2	Df
School mean achievement	0.18	59.15	33	0.15	49.68	33
Within-school variation	0.98			0.98		

* $p < .05$

** $p < .01$.

Note: BPVS = British Vocabulary Picture Test; SFA = Success for All.

Table 3. SFA vs. Control Multilevel Models Predicting School-Level Literacy Outcomes: York Assessment of Reading Comprehension

<i>Fixed Effect</i>	YARC Rate Ability T=356, C=381			YARC Comprehension T=407, C=461			YARC Accuracy T=412, C=468		
	Effect	SE	t	Effect	SE	t	Effect	SE	t
School mean achievement									
Intercept	-0.03	0.05	-0.55	0.03	0.04	0.73	0.01	0.04	0.35
Mean BPVS pretest	-0.08	0.12	-0.70	0.45**	0.07	6.41	0.16	0.08	1.89
SFA assignment	+0.11	0.14	1.09	+0.06	0.08	0.69	+0.12	0.08	1.46
<i>Random Effect</i>	Estimate	χ^2	Df	Estimate	χ^2	Df	Estimate	χ^2	Df
School mean achievement	0.24	74.99	33	0.17	57.47	33	0.16	52.67	33
Within-school variation	0.97			0.97			0.98		

NOTE: YARC = York Assessment of Reading Comprehension.0

BPVS = British Vocabulary Picture Test. SFA = Success for All.

* $p < .05$

** $p < .01$



Implementation observations

Observations were designed to enable researchers to assess the fidelity with which the SFA programme was being implemented and to determine if critical components of the SFA teaching process were being implemented in control schools. They included factors relating to literacy teaching, co-operative learning, and assessment. These observations were in addition to the regular, routine visits by SFA advisors. During the 2010-11 academic year, 10 of the intervention schools were visited once and eight were visited twice during the school year. Nine out of 18 control schools were visited once each during the school year.

For both the intervention and control schools, where possible, observers attempted to observe more than one class or literacy group within each school, covering a range of ability levels within Year 2. Within SFA schools, teachers usually taught groups of pupils from different year levels but all one reading level. Pupils changed classes at reading time to make this possible as the SFA model prescribes. In control schools, teachers generally taught several reading groups within the year-level class they taught all day.

Phonics. In all the schools visited who had a phonics lesson observed, researchers saw some form of teaching in synthetic phonics in use - teaching letter sounds and then blending these sounds together to read whole words. '*Letters and Sounds*' (Department for Education and Skills, 2007) was the most widely used system in control schools, although other programmes were also mentioned by teachers or recognized by observers, including *Jolly Phonics* and *Read, Write Inc.*

Co-operative learning. A key feature of the SFA programme is co-operative learning. Partner work was observed in many control as well as intervention schools. At its most limited, this involved sharing of resources within pairs. At its most sophisticated it involved partner reading – a strategy observed in most SFA schools, whereby children work together and take turns reading and summarising text. The most popular form of partner work in both control and intervention schools involved partner talk, the sharing of ideas, formalised in the SFA schools by such strategies as 'Think-Pair-Share'. In most SFA schools and some control schools the partnerships were formally assigned by the teacher, who directed children to work together in specific pairs.

SFA implementation fidelity. It is important to note that there was variation in implementation of the SFA programme among the intervention schools. Project researchers rated the implementation of the SFA Year 2 teachers on a scale of 0-3 - zero being 'No fidelity to the programme or co-operative learning in place' and a score of 3 meaning 'materials and routines are followed with fidelity and co-operative learning is embedded within the school culture'. Given the nature of SFA, with its cross-class, cross-year reading classes grouped by reading level, this is an important



indicator of overall fidelity to the programme. An average fidelity score was calculated across reading classes within schools and, where more than one visit was made, an average was then obtained across visits. This resulted in an implementation fidelity score for each school on a scale of 0-3, with 0 being very weak implementation fidelity and 3 being very high.

Of the 18 intervention schools, ten schools received a rating of 3, seven were rated 2, and one was rated 1. The key element that varied among schools was the extent to which co-operative learning was followed with consistency. Additional issues elicited from informal interviews after observations in SFA schools included:

- Lessons in some schools were reduced from 90 minutes per day, five days a week to less than 90 minutes a day and/or less than five days a week;
- Other reading schemes were used within the classroom within the literacy time in a few schools, thereby diluting the 'whole programme' approach advocated by SFA.
- Sometimes the organisation of the schools did not facilitate the mixing of year groups within a school.

This latter point was compounded by the fact that the Year 2 children were due to take the end of Key Stage 1 SATs during the summer of 2011. This meant that sometimes the mixed year group approach advocated by SFA was not followed through at this stage because of the importance to schools of ensuring that children reached their required 'competences' and the felt need to provide additional coaching to Year 2 children who were less likely to meet those requirements.

Discussion

These results indicate educationally significant results for the word level reading assessments. There are positive effect sizes for all the assessments undertaken but only significant effects for the Word ID and Word Attack subtests of the WRMT.

There may have been additional beneficial outcomes of the SFA programme that we did not measure. For example, one teacher indicated that behaviour tended to improve with the co-operative learning framework.

These outcomes of this evaluation of SFA are noteworthy for several reasons. First, we were able to obtain the co-operation of a sufficient number of SFA and control schools to provide an acceptable level of statistical power to detect school-level effects within a multilevel model framework. SFA and control samples were reasonably well matched on a variety of baseline characteristics, including demographics and BPVS pretest scores.



Second, there were positive effect sizes demonstrated for the intervention schools on the beginning reading skills of word identification and decoding. Despite the strong focus on phonics and decoding in English schools in recent years, SFA schools still greatly improved phonics and word-level outcomes in comparison to otherwise similar control schools.

Thirdly, the treatment fidelity and SFA implementation quality seemed reasonably good. There were similarities between the intervention and control schools in that they all were teaching some form of synthetic phonics, yet co-operative learning strategies were generally absent within control schools.

The pattern of the Year 2 treatment effects appears to be consistent with previous US studies, the SFA programme theory, and more general research and theory on the development of young children's emergent literacy skills. We found effects of both statistical and educational significance. The logic model behind the programme is consistent with more general theories of how young children develop as emergent readers (Snow, Burns, & Griffin, 1998). Specifically, powerful decoding strategies and phonemic awareness, as stressed by the SFA Reception and Year 1 programme, are key building blocks upon which children can develop a broader range of skills. However, comprehension results, although generally positive, were not statistically significant. The largest cluster randomised evaluation of SFA in the US, by Borman et al. (2007), found positive effects of SFA in word-level outcomes in kindergarten (Year 1) and first grade (Year 2), but comprehension effects did not appear until second or third grades. If the current evaluation were continued for another year, it might also find significant comprehension effects as Borman et al. (2007) did.

This aside, implementation was not perfect and this has always been an issue with large-scale programme evaluation. There was variability in the level of implementation among SFA schools, although the majority of schools implemented the programme with at least moderate fidelity (10 out of 18). Yet they didn't always do so for 90 minutes, five days a week.

The findings of this experiment are important for policy and practice. Overall, this study responds to the doubts that have been raised about the viability and appropriateness of large-scale evaluations in school settings (Cook & Payne, 2002). As a large, lengthy field evaluation, rather than a relatively small, brief experiment, the results of this study have strong external validity and relevance for policy and practice. This project ties together two central themes of educational research and policy: the scale-up, or replication, of school-based interventions and the development of high-quality evidence of their causal effects. These outcomes have established



that large-scale quasi-experiments involving replicable school-based interventions are both possible and desirable in the UK.

If a complex intervention such as SFA can be implemented with success in 18 primary schools in deprived areas across England, then there is no reason it could not be successful on a broader scale. As government policies now provide schools in England more autonomy and less top-down prescription, more schools should have opportunities to choose among effective and replicable interventions to provide educators with tools capable of narrowing the achievement gap between high and low income populations. With the movement towards evidence-based education in England, hopefully more programmes such as SFA will be created, evaluated and disseminated to help them do so.

Acknowledgements

The authors would like to thank Marilyn Jones-Hill, Beth Fowkes and Caroline Harvey of Success for All-UK for their support of the schools in the study and Louise Elliott at the Institute for Effective Education, University of York for her data management and invaluable contribution to the project.

Please note:

Robert Slavin and Bette Chambers were involved in the development of the SFA programme and serve on the Board of Directors for the Success for All-UK registered charity.

This research was supported by funding from the Bowland Trust. However, any opinions expressed are those of the authors and do not necessarily represent Bowland Trust positions or policies.



References

- Allen, G. (2011). *Early intervention: Smart investment, massive savings*. Independent review on early intervention delivery. London, UK: Cabinet Office.
- Borman, G.D., Hewes, G.M., Overman, L.T., & Brown, S. (2003). Comprehensive school reform and achievement: A meta-analysis. *Review of Educational Research, 73* (2), 125-230.
- Borman, G.D., Slavin, R.E., Cheung, A., Chamberlain, A., Madden, N.A., & Chambers, B. (2007). Final reading outcomes of the national randomized field trial of Success for All. *American Educational Research Journal, 44* (3), 701-731.
- Cook, T. D. & M. R. Payne. (2002). Objecting to the objections to using random assignment in educational research. In F. Mosteller and R. F. Boruch (Eds.), *Evidence matters: Randomized trials in education research*. Washington, D.C.: Brookings Institution.
- Department for Education and Skills. (2007). *Letters and sounds: Principles and practice of high quality phonics six-phase teaching programme*. London: DfES.
- Department for Children, Families and Schools. (2009). *Analysis and evidence strategy 2009-10: Delivering evidence based policy*. London: DCFS.
- Department for Work and Pensions (2009) *Households below average income: an analysis of the income distribution 1994/95–2007/08*, London: DWP (available at http://statistics.dwp.gov.uk/asd/hbai/hbai2008/pdf_files/full_hbai09.pdf, accessed 14 October 2010).
- Dun, L. M., Dunn L. M., Whetton, C., & Burley, J. (1997). *The British Picture Vocabulary Scale* (2nd Edition), London, UK: GL Assessment.
- Goodman, A., Sibieta, L., and Washbrook, E. (2009). *Inequalities in educational outcomes among children aged 3 to 16*. Final report for the National Equality Panel.
- Hills, J., Brewer, M., Jenkins, S., Lister, R., Lupton, R., Machin, S., Mills, C., Modood, T., Rees, T., Riddell, S. (2010). *An anatomy of economic inequality in the UK: Report of the National Equality Panel*, Government Equalities Office, London.
- Hopkins, D. Youngman, M. Harris, A., & Wordsworth, J. (1999). Evaluation of the initial effects and implementation of Success for All in England. *Journal of Research in Reading, 22*(3), 257-270.



- Law, Y. K. (2008). Effects of cooperative learning on second graders' learning from text, *Educational Psychology: An International Journal of Experimental Educational Psychology*, 28(5), 567-582.
- Lee, V.E., & Burkham, D.T. (2002). *Inequality at the starting gate: Social background differences in achievement as children begin school*, Economic Policy Institute, Washington.
- Lesnick, J., George, R., Smithgall, C., & Gwynne, J. (2010). *Reading on grade level in third grade: How is it related to high school performance and college enrollment?* Baltimore, MD: Annie E. Casey Foundation.
- Mongon, D. and Chapman, C. (2008) *Successful leadership for promoting the achievement of white working class pupils*, Nottingham: National College of School Leadership (available at www.nationalcollege.org.uk/index/docinfo.htm?id=17406, accessed 18 October 2010).
- 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*. Rockville, MD: National Institute of Child Health and Human Development.
- Raudenbush, S. W. & Bryk, A. A. (2002). *Hierarchical linear models* (2nd Ed.). Thousand Oaks, CA: Sage.
- Raudenbush, S. W., Bryk, A. S., Cheung, Y., & Gordon, R. T. (2000). *HLM5: Hierarchical linear and nonlinear modeling*. Chicago: Scientific Software International.
- Rose, J. (2006). *Independent review of the teaching of early reading*. London: Department of Education and Skills.
- Slavin, R. E. (1995). *Cooperative learning: Theory, research, and practice* (2nd Ed.). Boston: Allyn & Bacon.
- Slavin, R. E. (2009). Cooperative learning. In G. McCulloch & D. Crook (Eds.), *International Encyclopedia of Education*. Abington, UK: Routledge.
- Slavin, R. E., Lake, C., Chambers, B., Cheung, A., & Davis, S. (2009). Effective reading programs for the elementary grades: A best-evidence synthesis. *Review of Educational Research*, 79(4), 1391-1465.
- Slavin, R. E., Lake, C., Davis, S., & Madden, N. (2011). Effective programs for struggling readers: A best-evidence synthesis. *Educational Research Review*, 6, 1-26.
- Slavin, R.E., Madden, N.A., Chambers, B., & Haxby, B. (2009). *Two million children: Success for All*. Thousand Oaks, CA: Corwin.



- Snow, C. E., Burns, M. S., & Griffin, P. (Eds.) (1998). *Preventing reading difficulties in young children*. Washington, DC: Committee on the Prevention of Reading Difficulties in Young Children, National Research Council.
- Stevens, R.J. (2003). Student Team Reading and Writing: A cooperative learning approach to middle school literacy instruction, *Educational Research and Evaluation*, 9(2), 137-160.
- Strand, S. (2008) Minority ethnic pupils in the Longitudinal Study of Young People in England: Extension report on performance in public examinations at age 16 (DCSF research report 029), London: DCSF (available at www.education.gov.uk/research/data/uploadfiles/dcsf-rr029.pdf, accessed 18 October 2010).
- Tymms, P., & Merrell, C. (2001). *Success for All quantitative evaluation report: Second year*. Durham, UK: Curriculum Evaluation and Management Centre, Durham University.
- Wilkinson, R., & Pickett, K. (2010). *The spirit level: Why more equal societies almost always do better*, London, UK: Penguin.
- Woodcock, R. W., Mc Grew, K. S., & Mather, N. (2001). *Woodcock-Johnson III Tests of Achievement*. Itasca, IL: Riverside Publishing.

