Implementing Educational Interventions at Scale

Simon Calmar Andersen & Ulrik Hvidman

Department of Political Science

TrygFonden's Centre for Child Research

Aarhus University

May 25, 2020

Abstract

Although many educational programs have demonstrated the potential to increase student learning, few examples of successful scaling exist. We study the scalability of a parent-aimed reading program that has shown promising results in an experiment within a local government. Using a nationwide experiment among the full population of 2nd-grade children in Danish public schools (n=51,312), we find that the program is less effective at large scale. We provide evidence on potential explanations for the lack of scalability, which suggests that implementation fidelity is the most important barrier to successfully scaling this type of educational interventions.

Keywords: implementation; education; experiments.

Introduction

Parents play an important role in the educational development of their children. An increasing number of field experiments have demonstrated that interventions aimed at involving parents in supporting their children's learning can have positive effects (e.g., Andersen & Nielsen, 2016; Bergman, 2019; Bergman & Chan, 2019; Bergman & Rogers, 2017; Doss et al., 2019; York et al., 2019). The evidence on the effectiveness of these early interventions suggests that there could be substantial returns from investing in parent-targeted educational programs at a large scale. However, results from small-scale experiments may not generalize to a large scale (Al-Ubaydli et al., 2017). An important next step for the research on early interventions is therefore to systematically study the barriers to scalability.

First, subjects who experience the largest effects of the interventions may be more likely to select into experimental studies (Heckman, 1992, 2020; Heckman & Smith, 1995). Second, results from small-scaled, published trials may not replicate because, conditional on being published and showing statistically significant effects, there is a relatively high probability that the trials overestimate the effect sizes (Gelman & Carlin, 2014). Third, changing the scale of the intervention may also change factors on the supply side (Duflo et al., 2008). As labor is often a key input in the implementation of education interventions, the scaling of the implementation infrastructure may be particularly important to their effectiveness.[1] Unfortunately, the use of science to study these "scalability" problems is at such an early stage that little is known about the relative importance of these problems (Czibor et al.,

[1]If governments do not scale up the number of people working on implementing the intervention proportional to the number of clients, managerial oversight is spread out more thinly within a given implementation process, which may explain that program quality decline (Muralidharan & Niehaus, 2017). Furthermore, people implementing the program may vary in their skills. Thus, even if implementation costs are held fixed per client served, scarcity in the supply of highly skilled and devoted professionals is likely to occur as programs scale (Davis et al., 2017). Research in developing countries has shown that nation-wide implementation tends to reduce effect sizes. In a meta-analysis Vivalt (2019) shows that effects of randomized controlled trials tend to be smaller when the intervention is implemented by governments rather than non-governmental organizations (NGOs). Bold et al. (2018) randomized whether the same intervention was implemented by an NGO or by the national government in Kenya. While they found positive effects when the NGO implemented the policy, effects were small and not statistically significant when implemented by the government.

2019).

To study these scalability problems, we fielded an experiment of a reading intervention called READ. The national government in Denmark invited by email a random sample of the full population of 1,142 schools with 2nd-grade students to participate in READ. The intervention had previously demonstrated positive effects on 2nd-grade students' reading and writing skills in a field experiment run by a local government (Andersen & Nielsen, 2016). We use administrative data collected independently of the experiment to measure the effect on students' reading skills as well as the socio-economic characteristics of the full population of schools and families. These features make the trial a Natural Field Experiment in the typology of Harrison and List (2004), meaning that it was conducted in the natural environment in which it would be used if implemented nationally, that there was no self-selection into the trial, and that subjects were unaware that researchers evaluated the effects of the program.

Results show no statistically significant average treatment effect when the intervention is scaled up, and the estimate is close to zero and precise enough to rule out an effect of similar magnitude as in the original trial. To understand this finding, we examine the three scalability problems. First, the features of the natural field experiment allow us to study selection into treatment and compare participants in the original, local trial to the national trial. Interestingly, in the national trial, the groups of schools and parents that opted into the program were not much different in terms of socio-economic status and previous test scores in comparison to those who opted out. Moreover, treatment effects did not seem to be heterogeneous with respect to socio-economic background. Since the original, local study was also conducted among a relatively large sample of students with variation in both socio-economic background and ethnicity, selection into the original trial does not appear to explain the scalability problem.

To study the replicability, we evaluated the effect of the program when the local government, which ran the first READ trial, subsequently put the program into operation. We use

a difference-in-differences design to compare schools within the local government that chose either to adopt or not adopt the READ program for a new cohort of students. We find effects that are of similar size as in the original, experimental trial. These results suggest that the program has the potential to improve student learning and that the lack of scalability was not merely due to a statistical artifact.

Instead, implementation fidelity (i.e., the extent to which the program is implemented as intended) appears to be an important boundary condition for the effectiveness of the program. Specifically, the way that the local government implemented the program differed from the nationwide implementation in important ways (e.g., administrative support from the municipality). Data from a smartphone app—in which invited parents could sign up—provide a behavioral measure of implementation fidelity. These data produce two sets of findings. First, the local government succeeded in making twice as many parents sign up in the app (24 % in the nationwide experiment, and 48 % in the local government intervention). Second, in the national trial, the program had a positive effect on student learning among schools in which a relatively large proportion of parents downloaded the app, a finding that is robust to various specifications. Thus, our results suggest that the implementation process is crucial if parent-aimed educational interventions should be taken to scale.

The remainder of the article is organized as follows. In the second section, we review and analyze existing research on parent-aimed and other educational interventions.The third section presents the results of the large-scale randomized trial at the national level in Denmark. The fourth section studies potential explanations of scalability problems. The final section concludes.

Parent-Aimed Programs at Scale

Evidence on the Effectiveness of Parent-aimed programs

Given that family investments and resources matter for children's skill development, one important question is whether and how governments can support and encourage parent engagement. A growing body of experimental studies supports the notion that parent-directed interventions can increase parental involvement in their children's learning and that they have the potential to improve child learning. Although parent-aimed interventions often try to enhance ability, knowledge, and motivation, initiatives differ in their focus on providing additional learning material (i.e., resources) and merely providing information to families.

One set of parent-aimed programs provide resources such as books or tablets to families in order to encourage them to enact learning activities at home. A number of parent-aimed programs have been successful in encouraging parents to read or do math with their children—often in close collaboration with their teachers. In a randomized controlled trial including 284 immigrant children from 61 child care centers, Jakobsen and Andersen (2013) found that providing families with children's books and games (and collaboration with pre-school teachers on the development of their children) increased the language test scores of children with low-educated mothers as assessed by their pre-school teachers. In the study of the READ program, on which we build the current study, Andersen and Nielsen (2016) found that the program led to an increase in students' standardized test scores in reading of 0.12-0.26 standard deviations. The family-aimed treatment consisted of books and information on how to use dialogue-based reading. 1,587 students from 28 schools (within one local government) participated in the randomized controlled trial.[2] Berkowitz et al. (2015) administered a tablet app to parents that help them do math activities at home with their

[2]Mayer et al. (2018) use behavioral interventions to more than double the time parents spend on reading with their children. They include 169 parents from 8 preschool programs, but they do not present any test on child outcomes.

1st-grade children. The study included 22 schools and 587 families and found that the more parents used the app, the better the child performed in a math test administered by trained researchers.

Taken together, these studies suggest that when parents are encouraged to read or do math with their children at home, there is a potential for improving children's learning. In contrast, Guryan et al. (2008) did not detect a positive effect of a reading intervention in their experimental study, even though the intervention resembles the other programs previously described. In their study, books were sent to students during the summer to encourage them to keep reading during the school break. Parents were invited to an afterschool family literacy event where they learned about the program. Including 5,319 students from 59 schools, this study is relatively large, which may be an indication of the challenges in scaling up programs that provide books and other resources to the families and facilitates collaboration between schools and families.

Another group of parent-aimed interventions focus on school-to-parent communication. Since the cost of sending information to parents is often lower than providing parents with learning materials, it is generally cheaper to scale up information-based interventions. In a study of 1,031 parents from one school district, York et al. (2019) found that sending parents text messages with advice on how to support the development of their children improved the pre-school children's early literacy by about 0.11 standard deviations. In another experiment, Bergman (2019) studied the impacts of emails, text messages, and phone calls from teachers to parents with information about missed assignments and grades. The effect among the 462 participating students (all from one school) was an increase in the grade point average (GPA) of about 0.20 standard deviations. In a related study, Bergman and Chan (2019) used automated text messages to scale up the intervention (i.e., sending parents information about missed assignments and grades). In a sample of 22 schools (1,137 students), they found positive effects on GPA but did not find significant effects on state-administered test scores.

In a randomized controlled trial among 6,976 students in 12 schools, Bergman and Rogers (2017) use automated text message alerts to inform parents if their child had a missing assignment, a class absence, or a low average course grade. They found effects of information on GPA of about .06 of the standard deviations of the control group at baseline. In a study, Rogers and Feller (2018) sent parents of 28,080 12^{th}-grade students (from one large school district) information about their children's school absences. The most effective treatment in the experiment reduced absences by 1.1 day, a reduction of 6.5 percent compared to the control group. However, they were not able to detect a statistically significant effect on test scores.

Scalability of Educational Programs

To more systematically examine the relationship between scale and the effectiveness of education interventions, we conducted a meta analysis of randomized controlled trials that examine the effect on standardized test outcomes.[3] In Figure 1, we plot the number of participants in the studies on a logarithm scale against the estimated, standardized effect sizes. Three patterns in Figure 1 are worth emphasizing. First, the larger studies tend to produce smaller effects on standardized student outcomes. For studies with around 1,000 or more participants, the average effect size is close to zero. Although the interventions and study samples differ in several respects, this relationship could indicate that there are some challenges in scaling up such educational interventions. Second, the scale of the program evaluations have typically been rather small. Third, there is variation in scale across type of intervention. Many evaluations of parent-aimed programs have included less than 300 participants, whereas the educational interventions included have been tested at larger scales,

[3]Our analysis is based on three data sources. First, we included the studies from Noble et al. (2019), who systematically review parent-aimed programs targeting shared book reading. Second, we included studies from Lortie-Forgues and Inglis (2019), who study all interventions commissioned by the Education Endowment Foundation (EEF) in the UK and the National Center for Education Evaluation and Regional Assistance (NCEE) in the US . Third, we supplemented these two meta-analyses with a systematic search for parent-aimed education interventions. A more detailed description of the search strategy and a full list of all publications included in our meta-analysis can be found in Appendix G.

though seldom with more than 10,000 individuals.

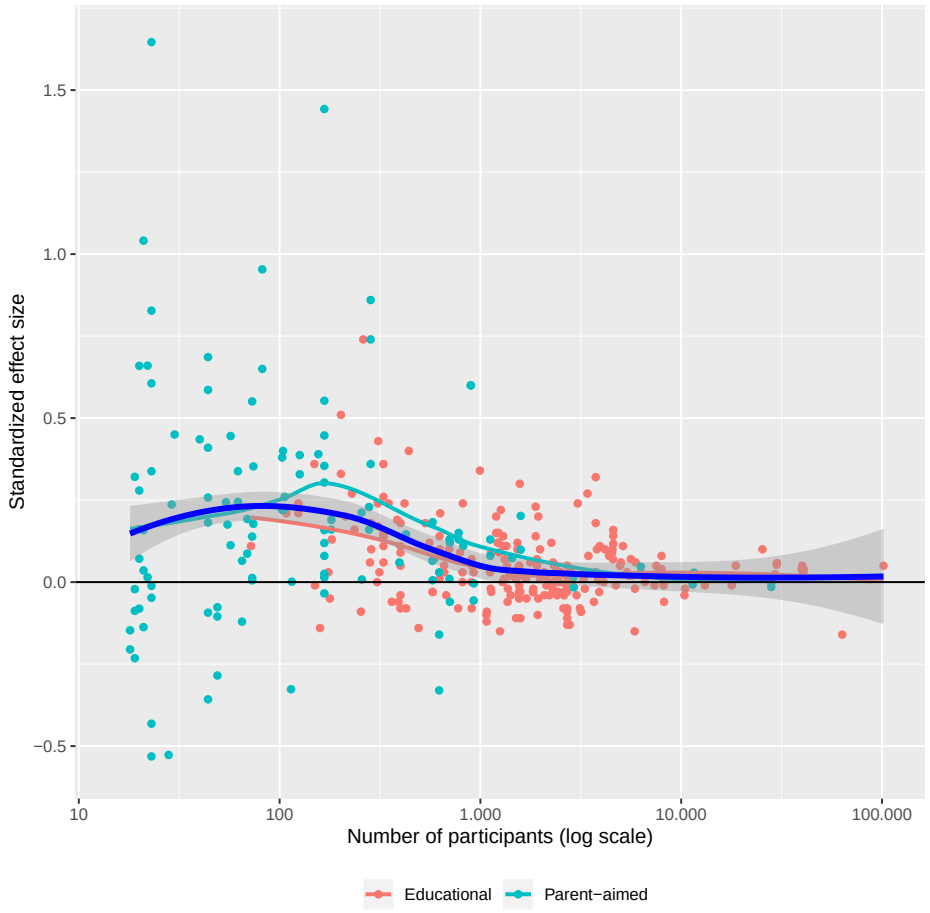


Figure 1: Number of participants (log scale) against effect sizes in randomized controlled trials of parent-aimed and general educational interventions.

Note: Red dots designate general educational interventions commissioned by the Educational Endowment Foundation and the National Center for Educational Evaluation and Regional Assistance (Lortie-Forgues & Inglis, 2019, see Appendix G for details). Green dots indicate parent-aimed interventions (Noble et al., 2019, and own review, see Appendix G). The blue line and the shaded area represent local polynomial regression with 95 % confidence intervals for all observations. The red line and the green line are local polynomail regressions based on the general educational and the parent-aimed interventions respectively.

Overall, the meta-analysis demonstrates that there are few successful examples of introducing parent-aimed educational programs at scale. Moreover, the negative relationship between the size of the study sample and the effect size may suggest that scaling educational programs is difficult. In sum, we see a growing body of evidence that small-scale parent-aimed interventions provide a promising strategy for delivering improvement in stu-

dent learning. However, there is little empirical evidence as to what are the most important barriers for scale-up.

A Full-Scale Natural Field Experiment

The READ Program

In collaboration with the Ministry of Education in Denmark, we conducted an experiment of the READ program among the full population of Danish public schools with 2nd-grade students.

READ was developed by a team of educational researchers at VIA University College and TrygFonden's Centre for Child Research in collaboration with the local government of Aarhus (Aarhus is the second-largest city in Denmark with approximately 350,000 inhabitants). The program aims at improving 2nd-grade students' literacy skills. As part of the program, families receive four books and information on how to find other reading material at the library, at the school, or in newspapers. Parents are also provided with a booklet and access to an online video (all information was translated into ten languages).

The booklet and the video underpin three components. First, the information emphasizes a growth theory of abilities by explaining to parents that their child's literacy skills can be improved regardless of its current level (Dweck, 1999, 2006). Second, the material encourages parents to take a constructive, mastery-oriented approach supporting the child's autonomous engagement with the books (Moorman & Pomerantz, 2010; Pomerantz et al., 2007). Third, the parents are encouraged not to correct their child if it reads incorrectly, unless it affects the child's understanding of the text (Haimovitz & Dweck, 2016).

The Original, Local Experiment

A previous study in the local government of Aarhus yielded encouraging results (Andersen & Nielsen, 2016). The randomized controlled trial included 1,587 children in 72 classrooms

from 28 schools. There was considerable variation in terms of immigrant and non-immigrant, high- and low-educated, and high- and low-income parents. Based on cluster randomization at the classroom level, the 1,587 children were assigned to treatment (i.e., READ) or control (i.e., treatment as usual). The duration of the program was 16 weeks, and the average costs per child approximately DKK 500 (USD 76). The implementation included a social reward component. Specifically, to encourage the child's effort, parents and children could use a logbook to note every time they read. The logbook thereby endorsed the child's effort rather than performance or results (i.e., speed and accuracy). When the child had read ten times, she could bring the logbook to her teacher, and the class would get a sticker. The class with the most stickers received a reward.

The READ treatment improved standardized test scores in reading significantly with an estimated effect size of 0.26 standard deviations after two months and 0.12 standard deviations after seven months. The treatment also improved children's expressive language skills as measured by a writing test by 0.16 standard deviations (for a thorough description of the intervention and the results, see Andersen & Nielsen, 2016).

Experiment at a Nationwide Scale

To study the effectiveness of READ at a nationwide scale, we worked with the Danish Ministry of Education to randomly assign all Danish public schools to receive the READ program or to a treatment as usual control condition. In Denmark, most children are enrolled in basic education in the summer of the year they turn six. Danish basic schooling covers a preschool year and nine years of compulsory education. Although parents can choose to enroll their children in a self-governing school or educate them at home, most children attend a public school (in 2017, 79 %). Schools are governed by 98 local governments—comparable to school districts in the US—but the national government (i.e., the Ministry of Education) formulates the general rules and can initiate policies for all public schools.

Population

The Danish Ministry of Education provided a list that included all public schools with 2nd-grade students in the school year 2017/2018.[4] As the City of Aarhus was implementing the READ program among a subgroup of their schools simultaneously with the national study, we excluded all schools from the City of Aarhus from the randomization (N = 46).[5] We end up with a sample of 1,142 public schools. Figure 2 presents the enrollment and participant flow of the experiment.

[4]We excluded schools without 2nd-grade students and self-governing schools not governed by a local government.

[5]In (section 3.1), we describe the observational replication study that evaluated the new implementation by the local government of Aarhus.

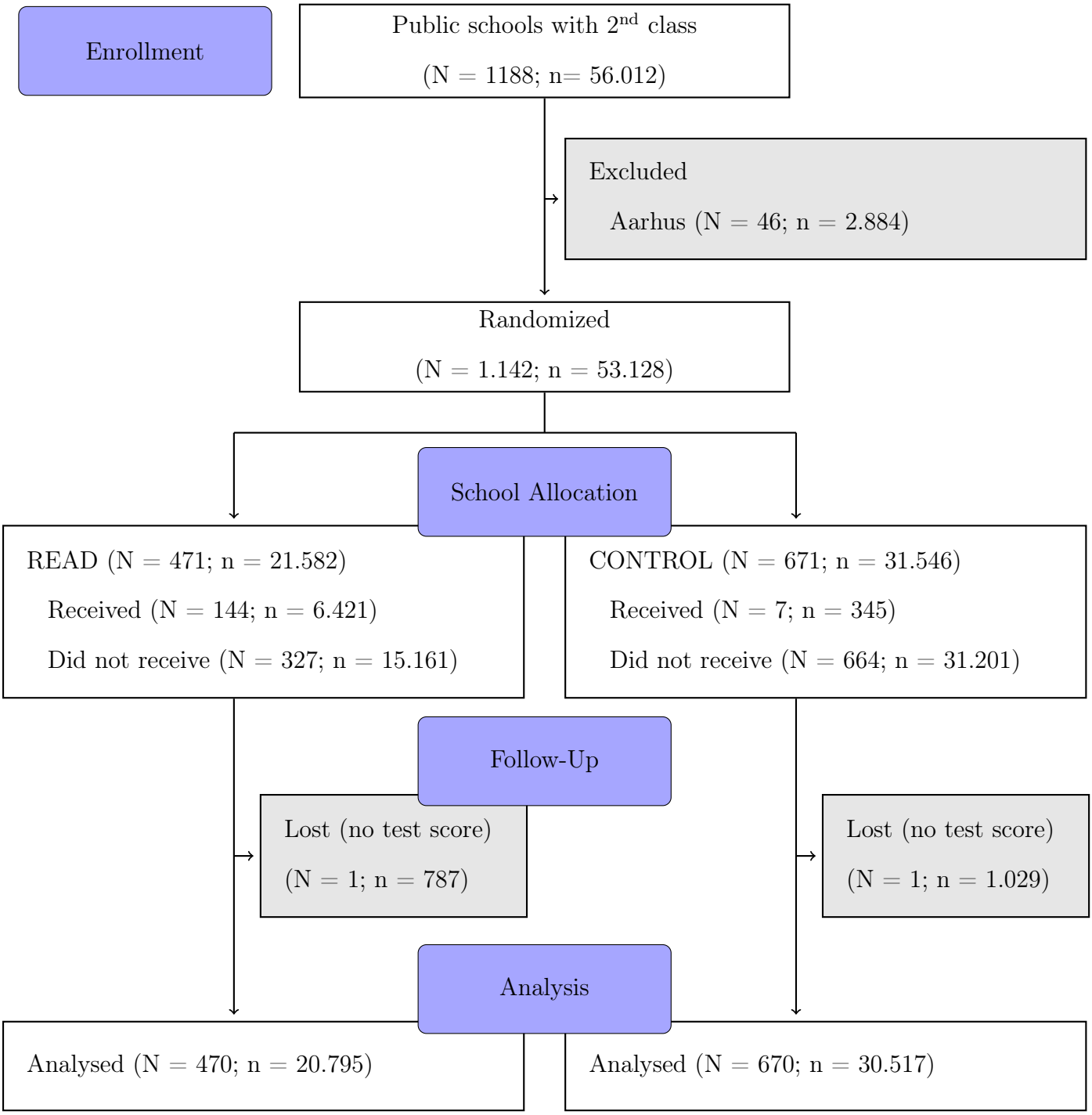


Figure 2: Participant flow diagram of study 1

N: Number of schools. n: Number of students

Randomization

The experiment was set up as an encouragement design (see Angrist & Pischke, 2009) in which schools were invited to participate in the READ program. Using a cluster-randomized design, we randomly assigned students in 1,142 schools to two conditions. The first group (our control group) did not receive an invitation to receive the READ program. The second group (READ assignment schools) received an official e-mail from The Ministry of Education with an invitation to receive READ for all 2nd-graders at the school. The local governments were informed about the project and that some of their schools had been invited.

Out of the 1,142 schools, 471 were invited, and 671 did not receive an invitation. Figure 3 provides a timeline of the implementation of the READ program. Invitations were sent out during the period from September 6 to September 27 2016.[6] Schools were reminded by e-mail about the program approximately one week after the invitations were sent out. A consultant also contacted the schools to inform them about the program and remind them about the decision to participate.

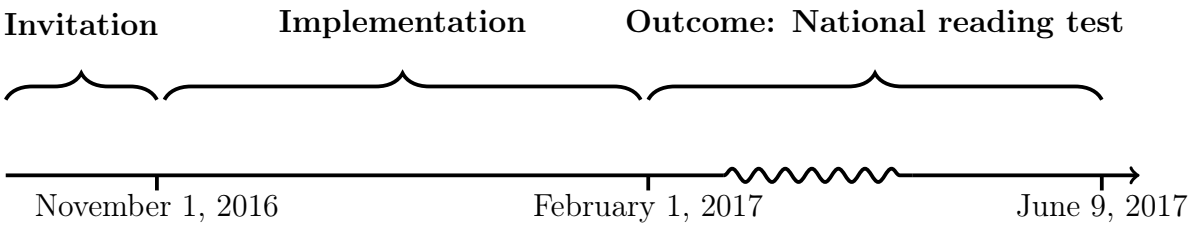


Figure 3: Timeline

The READ program was provided to all schools that accepted the invitation from the Ministry of Education. The schools were responsible for the distribution of the bags with the READ material to the families. Thus, the role of teachers was to distribute the READ material in the classroom and encourage students and their families to engage in the program. As a token of appreciation for their effort, schools were offered DKK 2,500 (USD 375) for

[6]To not exceed the budget, the 471 invitations were sent out in three waves.

participating.[7] The bags contained the original READ material, including four books to get parents started as well as information on how to find other reading material at the library. They also received a booklet and access to a video that underpin the three learning components of READ. In addition to the original READ program, parents at participating schools were also provided log-in information to a mobile application. The READ app contained video material with information on the importance of reading with children, and families could also use the app to register their reading activities and track their development (Appendix E, Figures E.3a and E.3b show screen shots from the app).

One advantage of the encouragement design using the full population of schools is that it allows us to implement the program in a way that mimics a real-world setting. Thus, schools implemented the program without any researcher involvement, and the study itself should therefore not affect the participants' behavior. The school cluster randomization also has the benefit of reducing potential spillover effects between the treatment and control groups, which might easily occur within schools.

To study the implementation of the program, two experiments were embedded in the trial. The first tested two slightly different versions of the invitation letter. The second randomized participating schools to one of two versions of the program (i.e., one set of schools received a social reward for reading, whereas the other set of schools did not receive this incentive). We found no strong effect of either of these variants. In the following analyses, we therefore pool these subconditions (for further details, see Appendix E).[8]

[7] Previous research suggests that offering schools compensation for participating is an effective tool to increase compliance (Andersen & Hvidman, 2020).

[8] A power analysis conducted prior to the experiment showed that a total of 65 schools in each treatment arm would allow us to detect an effect size of .15 standard deviations in test scores. This power calculation is conditional on a power of .80, a significance level of .05, an intra-school correlation of .10, and baseline data accounting for .20 of the variation in the outcome. To provide sufficient power to be able to detect an effect size of .15 standard deviations in the embedded experiment that tested two versions of the program, we aimed for 130 schools to accept the invitation.

Outcome

The main outcome is standardized test scores in 2nd-grade national reading tests. As of 2010, all students in public schools are tested in ten mandatory tests during basic education in grades 2 through 8. As the test result for each student is confidential and only known by the student's subject-specific teacher, tests are relatively low stakes. From 2015, however, test results at the school level are used as a soft accountability tool with potentially higher stakes for schools.

The 2nd-grade reading test constitutes a good outcome measure for three reasons. First, the tests are IT-based and performed by the students in class on a computer, and the scoring procedure is standardized (i.e., the score is automatically generated within the test system). The standardized procedure ensures that teachers and students cannot manipulate the test result, and the scoring of the tests is thereby blinded to the schools and students' treatment status in the trial. Second, tests have been shown to be a strong predictor of later-stage educational outcomes (Beuchert & Nandrup, 2018). Third, the reading test is divided into three subtests: "Language Comprehension," "Decoding," and "Reading Comprehension," which allow us to study different parts of literacy.[9]

Data

The Danish administrative registers allow us to track the full population of all 2nd-grade students in a public school. We match each student to the school that they were enrolled in when the program was implemented (September 6, 2016). Our main data consists of 53,128 students in our population of 1,142 schools.

We merge our main data with additional data sources. First, the Ministry of Education provided the student-level test data on the 2nd-grade tests in reading. Second, the data are

[9]The underlying psychometric model for the test is a Rasch model. The test score for each of the three subtests is measured on a logic scale from -7 to 7 (for further details, see Beuchert & Nandrup, 2018). We standardize each of the logit scores with mean zero and a standard deviation of one. To compute the overall score, our main outcome, we take the mean of the three standardized subtest scores and, subsequently, standardize this average score.

linked with records from Statistics Denmark containing detailed information on the children, including the child's ethnicity, gender, and age as well as on their families (e.g., the parents' length of education). The parental characteristics are measured in 2014 (two years prior to the intervention). Third, we are able to track whether parents download and register the READ application that was part of the program to measure parents' adoption of the program.

Figure 2 illustrates the experimental design and present data on compliance with the experimental protocol and attrition. Attrition, which is usually a main threat to the internal validity in experiments, is limited because the 2nd-grade reading tests are compulsory. We observe test scores on 96.6 % of the students and, importantly, there is no evidence of systematic differences in attrition between treated and controls. We define our analytical sample as the 51,312 students for which we observe the test scores.

Descriptive Statistics and Balance

In Table 1, we provide descriptive statistics for our analytical sample across experimental conditions. Because of the random assignment of READ invitations, there should be no systematic differences in the distribution of covariates between schools assigned to the treatment and control conditions. Column (3) compares invited and non-invited schools on the full set of pre-determined student and school covariates obtained from the registers. All differences are substantially small with no tests significant at the 5-% significance level. Three of the 18 tests are significant at the 10-% level.

Table 1: Differences in mean between invited and non-invited on background characteristics

	(1) Non-invited	(2) Invited	(3) 1-2
Student level			
Child is a boy	0.52	0.52	0.00
Child's age (2016)	8.09	8.09	-0.00
Child immigrant	0.10	0.12	-0.01^{+}
Mother compulsory education (2014)	0.14	0.14	-0.01
Mother upper secondary education (2014)	0.05	0.05	0.00
Mother vocational education (2014)	0.31	0.31	0.00
Mother short-cycle education (2014)	0.05	0.05	0.00^{+}
Mother medium-cycle education (2014)	0.27	0.27	-0.00
Mother long-cycle education (2014)	0.13	0.13	0.00
Father compulsory education (2014)	0.16	0.17	-0.01
Father upper secondary education (2014)	0.06	0.05	0.00^{+}
Father vocational education (2014)	0.41	0.41	-0.00
Father short-cycle education (2014)	0.08	0.08	0.00
Father medium-cycle education (2014)	0.14	0.13	0.00
Father long-cycle education (2014)	0.13	0.13	0.00
Missing test score (2017)	0.03	0.04	-0.00
School level			
School size[1]	47.01	45.82	1.19
Average test score (2016)[2]	-0.02	-0.05	0.03
Students	31546	21582	53128
Schools	671	471	1142

Notes: $+$ $p < 0.1$; * $p < 0.05$; ** $p < 0.01$. Standard errors clustered on the school level.[1] Number of students in second grade. [2]Standardized using the mean and the standard deviation from the national sample in 2017.

Estimation

The random assignment of invitations allows us to recover an unbiased estimate of the intent-to-treat effect of inviting schools to participate in READ by comparing the test scores among treatment groups (i.e., comparing the outcome between the "control group" and the "READ assignment group"). Consider the following equation:

$$y_{is} = \beta_0 + \delta INVITE_s + \boldsymbol{X'_{is}\beta_1} + \boldsymbol{Z'_s\beta_2} + u_{is} \tag{1}$$

where y_{is} is the standardized test score for student i in school s, $INVITE_s$ is an indicator that equals one for schools assigned to the READ program,$\boldsymbol{X_{is}}$ is a vector of student covariates, $\boldsymbol{Z_s}$ is a vector of school covariates, and u_{is} is a student-level error term. δ captures the reduced-form effect—that is, the difference in test scores between schools assigned to READ and the control group.

Not all schools that were assigned to the READ accepted the invitation (see Figure 2). Out of the 471 invited schools, 144 schools chose to participate (corresponding to 30.6 %). Moreover, seven of the 671 schools in the control group ended up receiving the READ program for various reasons.[10] To recover the effect of receiving the READ program, we use an instrumental variable (IV) approach in which we use the randomly assigned invitation to READ as an instrument for the school adopting READ. As invitations are randomly assigned, the instrument should be unrelated to unobserved outcome-relevant factors. Moreover, given that the invitation itself does not affect test scores, the instrument also satisfies the exclusion restriction. Under these assumptions, and given that there is a first stage, the IV approach allows us to estimate the local average treatment effect (LATE) of receiving the READ treatment as opposed to not receiving any treatment for those schools that complied with

[10]One reason for this non-compliance was that some schools are nested in administrative partnerships and, therefore, share an e-mail address. Thus, some invitations were forwarded to schools in the control group. If interested in participating, these schools were allowed to receive the READ program.

the assignment to treatment. To be specific, the first-stage equation can be written as follows:

$$READ_s = \alpha_0 + \lambda INVITE_s + \boldsymbol{X'_{is}\alpha_1} + \boldsymbol{Z'_s\alpha_2} + e_{is} \tag{2}$$

where $READ_s$ is a dummy for the school adopting the READ program. In a model without covariates, the LATE is the ratio of the reduced form estimate to the first-stage estimate ($\gamma = \frac{\delta}{\lambda}$). We use 2SLS to estimate the LATE effect. To take the nested structure of students in schools into account, we estimate all models with cluster-robust standard errors at the school level.

The Effect of READ on Student Achievement

Table 2 presents the main results on students' reading skills. Model 1 shows the reduced form effects of assigning schools to the READ program on student test scores. Model 2 shows the same model with the full set of covariates, $\boldsymbol{X_{is}}$ and $\boldsymbol{Z_s}$, included (see Equation 1). In both models, the effect is small in magnitude and not statistically distinguishable from zero. Model 3 presents first-stage estimates from Equation 2 and shows that the invitation increased the probability of participation by 28.6 percentage points compared to the non-invited group. Models 4 and 5 present two-stage least squares (2SLS) estimates of Equation 2. As the reduced-form estimates are close to zero, the 2SLS estimates are also rather small and insignificant.

Table 3 presents reduced-form estimates for the three subdomains "Language comprehension," "Decoding," and "Text comprehension," separately. For all domains, the estimates are small and statistically insignificant.

Table 2: Main results on students' reading skills

Outcome:	(1) Learning	(2) Learning	(3) Adoption	(4) Learning	(5) Learnig
	Reduced form	Reduced form	First Stage	IV	IV
Invited	-0.014	0.002	0.286**		
	(0.024)	(0.020)	(0.024)		
Participating				-0.048	0.008
				(0.084)	(0.071)
Mean of control	0.006	0.006	0.011	0.006	0.006
Observations	51312	51312	51312	51312	51312
Schools (clusters)	1140	1140	1140	1140	1140
Adjusted R-squared	0.000	0.139	0.178	0.000	0.139
Covariates	No	Yes	No	No	Yes

Notes: Columns (1), (2), and (3) are estimated with OLS. Columns (4) and (5) are estimated with 2SLS. Standard errors clustered on the school level in parentheses. + $p < 0.1$; * $p < 0.05$; ** $p < 0.01$. The full list of the included covariates is reported in Table C.1.

Table 3: Effects on the three sub domains (Reduced form)

Outcome:	(1) Language comprehension	(2) Decoding	(3) Text comprehension
Invited	-0.002	0.006	0.003
	(0.017)	(0.020)	(0.020)
Mean of control	0.008	0.003	0.004
Observations	51312	51312	51312
Schools (clusters)	1140	1140	1140
Adjusted R-squared	0.117	0.107	0.110
Covariates	Yes	Yes	Yes

Notes: Models estimated with OLS. Standard errors clustered on the school level in parentheses. + $p < 0.1$; * $p < 0.05$; ** $p < 0.01$. The full list of the included covariates is reported in Table C.1.

Scalability

To understand the small effects in the national trial compared to the original, local trial, we examine the three types of scale-up challenges: selection into the trial (Heckman, 1992, 2020; Heckman & Smith, 1995), replicability (Gelman & Carlin, 2014), and implementation (Duflo et al., 2008).

Selection into the trial

We begin by examining who adopted the READ program in the national study and in the original, local study. The detailed Danish register data provide a unique opportunity to study the selection into treatment in the experiments. Table 4 compares the schools that decided to participate to the non-participants on observed characteristics of both the families and schools. Column (1) presents descriptive statistics on the total population of 2^{nd}-grade students in public schools in Denmark in 2016. Column (2) shows descriptive statistics on those who selected into the large-scale trial run by the national government, and column (3) shows the difference between those who opted in and the full population. Although there are some differences between the READ adopters and the non-adopters, these are not large in magnitude, and there is no clear pattern of positive or negative selection. For example, whereas fathers with long-cycle education are somewhat over-represented among the READ adopters, so are fathers with vocational education. Moreover, there is no evidence of neither positive nor negative selection based on previous school performance in standardized reading tests.

Column (4) presents descriptive statistics on the students who participated in the initial test of READ in the local government. Column (5) tests for differences between the population in 2016 and participants in the original, local trial. Although many of these differences are statistically significant, most of them are small in absolute terms. However, there are some differences deserving attention. The initial READ trial had 21 % immigrants compared

Table 4: Differences in mean between participants and non-participants on background characteristics

	(1) National population	(2) National participating	(3) 1-2	(4) READ 1.0	(5) 1-4
Student level					
Child is a boy	0.52	0.53	-0.01	0.51	0.01
Child's age (2016/2013)[1]	8.09	8.09	0.00	8.12	-0.03**
Child immigrant	0.11	0.12	-0.02	0.21	-0.10**
Mother compulsory education (2014/2011)	0.14	0.15	-0.02+	0.18	-0.04
Mother upper secondary education (2014/2011)	0.05	0.05	0.01+	0.08	-0.02**
Mother vocational education (2014/2011)	0.31	0.32	-0.01	0.23	0.08**
Mother short-cycle education (2014/2011)	0.05	0.05	0.00	0.04	0.01**
Mother medium-cycle education (2014/2011)	0.27	0.27	0.00	0.23	0.04**
Mother long-cycle education (2014/2011)	0.13	0.11	0.02+	0.20	-0.07*
Father compulsory education (2014/2011)	0.16	0.17	-0.01	0.15	0.01
Father upper secondary education (2014/2011)	0.05	0.05	0.01*	0.07	-0.01+
Father vocational education (2014/2011)	0.41	0.43	-0.03*	0.26	0.15**
Father short-cycle education (2014/2011)	0.08	0.08	-0.00	0.07	0.01
Father medium-cycle education (2014/2011)	0.14	0.13	0.01	0.17	-0.03**
Father long-cycle education (2014/2011)	0.13	0.11	0.03*	0.21	-0.08**
Missing test score (2017/2014)	0.03	0.04	-0.01	0.06	-0.02
School level					
School size (2016/2013)[2]	46.52	44.81	1.98	58.78	-12.26*
Average test score (2016/2013)[3]	-0.03	-0.09	0.07*	-0.15	0.12
Students	53128	6766	53128	1587	54715
Schools	1142	151	1142	27	1169

Notes: + p < 0.1; * p < 0.05; ** p < 0.01. Standard errors clustered on the school level. Year of measurement in parentheses (National study/READ 1.0). [1]Year at the beginning of the interventions. [2] Number of students in second grade. [3]Average test score the year before implementation of READ. Standardized using the mean and the standard deviation from the national sample in 2017.

to 11 % in the full nationwide population. Moreover, 26 % of the fathers had a vocational education in the original trial, compared to 40 % in the full population. At the school level, we note that average test scores of previous cohorts were .12 standard deviations lower in the first READ compared to the population.

These differences in participation could potentially explain differences in the results when the program is scaled up. However, differences in participation would only affect the estimated effects if there are also heterogeneous treatment effects for the groups that are over- or underrepresented. Table 5 examines the same subgroups that were studied in the first READ study, that is, parental education and ethnicity (Andersen & Nielsen, 2016). Thus, we split the sample by ethnic background and whether the mother has a college education. There is no evidence that the effects are different across these subgroups (or other subgroups

Table 5: Treatment effects for subgroups (OLS)

	(1)	(2)	(3)	(4)
Subgroup:	Mother low education	Mother high education	Danish background	Immigrant background
Invited	-0.000	0.001	0.005	-0.029
	(0.024)	(0.021)	(0.020)	(0.043)
P-value	0.869		0.463	
Mean of control	-0.164	0.301	0.059	-0.500
Observations	27807	20543	46172	5140
Schools (clusters)	1139	1127	1140	870
Adjusted R-squared	0.077	0.064	0.115	0.101
Covariates	Yes	Yes	Yes	Yes

Notes: "P-value" provides p-values for the null hypothesis that the point estimates are the same for the two respective subsamples. Standard errors clustered on the school level in parentheses. + $p < 0.1$; * $p < 0.05$; ** $p < 0.01$. The full list of the included covariates is reported in Table C.1.

in Table 4) in the nationwide trial.

In the original, local trial, there was a tendency for more disadvantaged subgroups to experience larger effects of the intervention (Andersen & Nielsen, 2016). However, if there are no effects for any of the groups in the nationwide trial, then the scale-up challenge is unlikely a result of different types of students.

Replicability: An Observational Study

During the same school year as the nationwide implementation, the local government in the City of Aarhus implemented the READ program again among a new cohort of 2nd-grade students. Putting the program into operation, a subgroup of schools chose to participate in the program. The participating schools received the same material as the participating schools in the nationwide implementation. This was a natural field study in the sense that schools were not asked to participate in any data collection for the purpose of research. We use this observational replication to examine whether effects replicate when implemented at the same scale and in the same environment as in the original trial, but put into operation by the local government.

Design and Estimation

Participating schools in Aarhus received the READ intervention in January 2017. As schools were not randomly assigned to the program, systematic differences may be seen between local READ schools and non-READ schools that could explain differences in outcomes even in the absence of the READ program. To identify the impact of the READ program in the local government, we apply a difference-in-differences (DiD) design in which we compare trends in student learning before and after the implementation of the READ program in the local government schools to two control groups.First, we compare the local treatment schools to a national control group—that is, schools in the other municipalities that were not invited and did not participate in the national READ program ("National controls"). Second, as a local control group, we compare the local treatment schools to the remaining schools in the local government that did not participate in READ ("Local controls").

To be specific, we estimate the following DiD model:

$$\begin{aligned} y_{ist} = \alpha &+ \beta_1 National\_Controls_s + \beta_2 Local\_Controls_s + \beta_3 Post\_treatment_t \\ &+ \delta_1 (National\_Controls_s \times Post\_treatment_t) \\ &+ \delta_2 (Local\_Controls_s \times Post\_treatment_t) + u_{ist}, \end{aligned} \tag{3}$$

where $National\_Controls_s$ and $Local\_Controls_s$ are dummy variables indicating control group schools at the national and local levels, respectively. $Post\_treatment_t$ is a dummy indicating the school year after the implementation of READ. The interaction terms $National\_Controls_s \times Post\_treatment_t$ and $Local\_Controls_s \times Post\_treatment_t$ indicate control schools in the year after READ implementation. Under the assumption of common trends in the absence of treatment, the coefficients δ_1 and δ_2 capture the effect of *not* being assigned to the READ intervention on the reading outcome, y_{ist}.

As the national reading tests were changed in 2015—and test scores therefore not comparable before/after 2015—we include data as of 2015. To not confound the analysis by

changes in school composition, we use the balanced panel of schools for which we have test score data and consistent school identifiers for the years 2015, 2016, and 2017. To enable comparisons in effect size to the nationwide experiment, we standardize the test scores based on the national mean and standard deviation of the 2016 population (reported in Table A.1).

Results

Figure 4 presents the results visually, and Table 6 presents results of the DiD estimation and a formal test. The figure shows results for the total composite reading score as well as for each of the three subdomains. Pre-trends from 2014 to 2015 are rather similar for the local control group and the local READ group, whereas there is some divergence between the READ schools and the national control group. After the intervention period, the local treament school experienced an upward increase in test scores, which neither the local control schools nor the national control schools experienced. This is especially pronounced for the Text comprehension subscale.

Table 6, model 1 suggests that the program increased student test scores by 0.18 standard deviations compared to the local comparison group and 0.14 compared to the national control group (both statistically significant at a 10-% level). Although the difference in pre-trend slopes between the the treatment group and the national controls question the validity of the common-trends assumption, the similarity in the two DiD estimates across the two control groups is reassuring. Models 2-4 show that the coefficients are positive and substantially large in magnitude across all subdomains—but largest for "Text comprehension."[11] Although slightly smaller in magnitude, the pattern in the effect estimates across subdomains is rather similar to the findings in the original, local trial, in which effect sizes were estimated to be .19 for "Language Comprehension," .23 for "Decoding," and .27 for "Text comprehension" (see Andersen & Nielsen, 2016, Table 1).

Even though the identification of the causal effect is not as credible in the DiD design

[11]Table D.1-D.4 present robustness results and show that the estimates are rather similar across specifications with and without school fixed effects and with student and school covariates.

as in a randomized controlled trial, the fact that a second study finds effect sizes across subdomains that are consistent with the first provides some evidence of the effectiveness of the program. Thus, the positive effects in the observational study suggest that problems due to statistical inference may not be the main reason that the READ program was not effective at improving student learning in the nationwide trial.

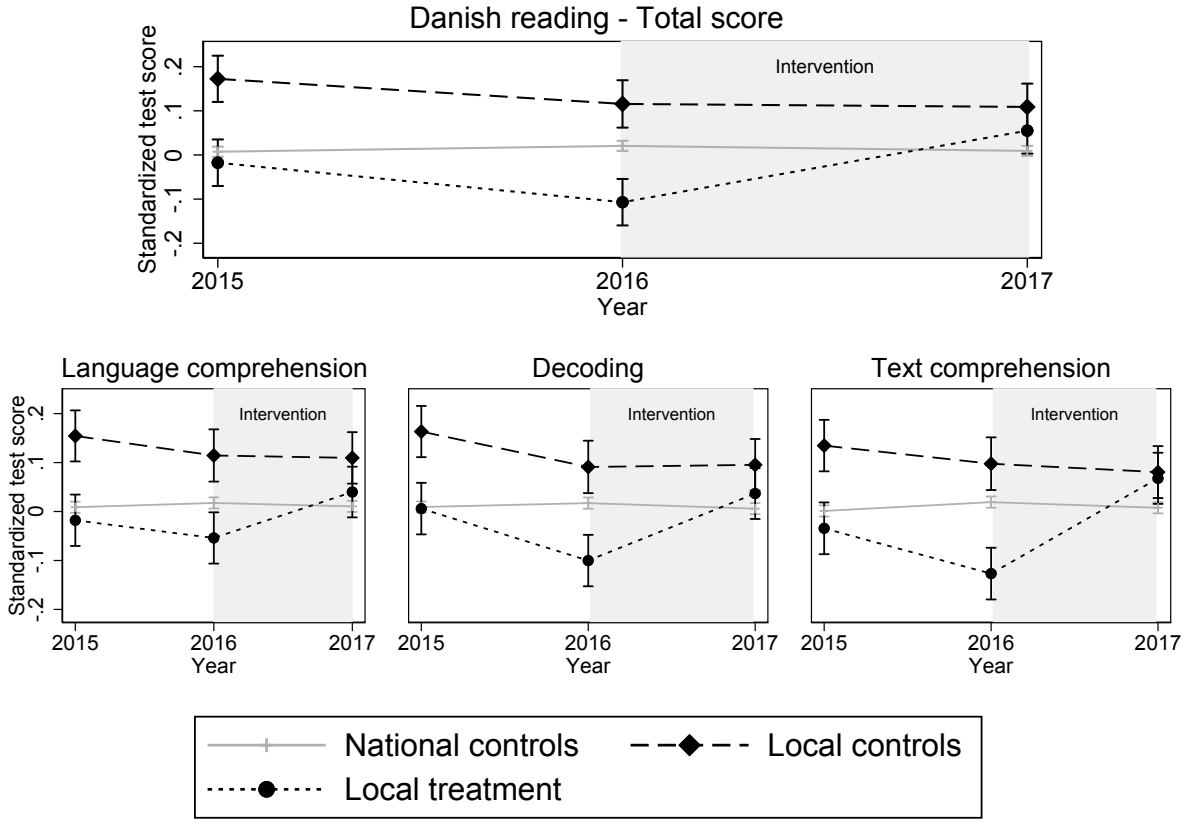


Figure 4: Effect of READ in observational replication study on total, composite reading test score and three subscales

Table 6: Difference-in-difference estimates on students' reading skills. Total score and the three subdomains (OLS)

Outcome:	(1) Total score	(2) Language comprehension	(3) Decoding	(4) Text comprehension
Post treatment	0.133^{+}	0.092^{+}	0.097	0.161*
	(0.070)	(0.054)	(0.073)	(0.064)
National controls X Post treatment	-0.139^{+}	-0.096^{+}	-0.104	-0.163*
	(0.071)	(0.056)	(0.074)	(0.066)
Local controls X Post treatment	-0.176^{+}	-0.123	-0.138	-0.201*
	(0.091)	(0.081)	(0.088)	(0.085)
Mean of control	-0.062	-0.036	-0.047	-0.081
Observations	96965	96965	96965	96965
Schools (clusters)	699	699	699	699
Adjusted R-squared	0.000	0.000	0.000	0.000
Fixed Effects	Yes	Yes	Yes	Yes

Notes: Standard errors clustered on the school level in parentheses. $+$ p $<$ 0.1; * p $<$ 0.05; ** p $<$ 0.01. Different specifications reported in Appendix D.

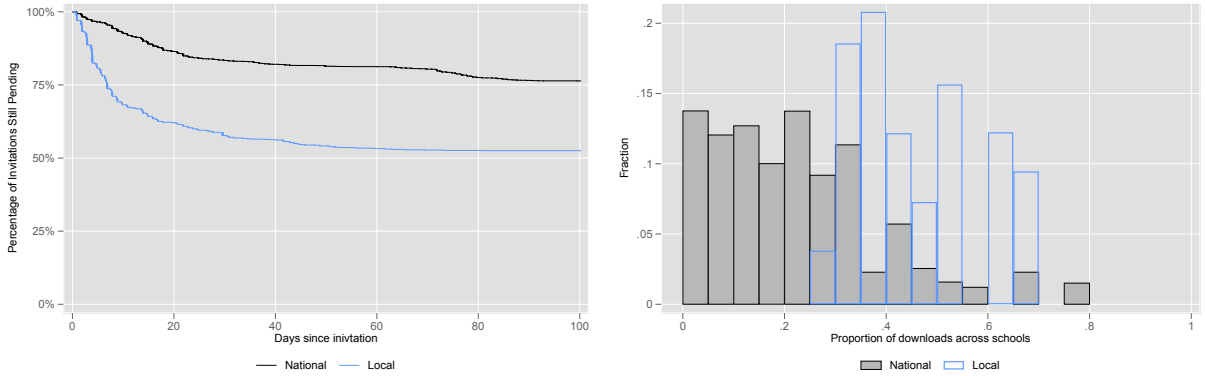
Implementation

An alternative explanation for not finding an average effect in the national, large-scale experiment of similar magnitude as in the original, local study could be poor program implementation. To explore the implementation fidelity, we analyse the extent to which the program is implemented. The READ smartphone app generated data that we can use to assess whether parents took up the program and thereby provides an indication of the quality of the implementation of the program.[12]

Implementation fidelity in the national and local settings

To compare the level of implementation fidelity in the local replication study to the national government study, Figure 5 presents data on the app use. Figure 5a (left panel) shows the proportion of pending users that had not signed up in the app over time. About 48 % of the children's families that were assigned to the program in the local government of

[12]Out of the 151 READ schools that chose to participate in the national trial, 109 schools provided access to student identifiers that enable us to match the student with the administrative records provided by Statistics Denmark. In contrast, we were able to match students at all participating schools in the local replication. Given that the schools in the national trial that did not provide access were less likely to implement the program subsequently, a higher take-up in the local replication than in the national trial would be a lower bound.



(a) Pending downloads across families **(b)** Downloads across schools

Figure 5: Implementation: App users.

Notes: Panel (a) plots survival curves that show the distribution of families that downloaded the app in the national sample and the Aarhus sample. Each survival curve plots the percentage of app user names still pending versus the number of days elapsed since the families received the invitation to download the app. The Aarhus sample consists of all families at schools that participated in READ. Among these families, 47.6 percent downloaded the READ app within 100 days. The national sample consists of all families at schools that accepted to receive the READ program. Panel (b) shows the distribution of the proportion of downloads at the school level across setting.

Aarhus registered in the app, whereas only about 24 % registered among the families in the national study. Figure 5b (right panel) compares the distribution of the proportion of users across schools in the two studies. The distribution is moved to the right in the Aarhus study. Moreover, a relatively large proportion of schools have no students signing up in the nationwide implementation.

We can use qualitative information on how the program was run in the national trial and in the local replication study to help clarify the large implementation gap. Interestingly, there were noticeable differences in the implementation procedure between the local program and the national program. First, more resources were allocated to the implementation of the program in the local setting than in the national setting. Specifically, in Aarhus, a team of internal consultants worked on the implementation of the intervention, whereas there was little administrative support at the national level. Second, the people implementing the programs may vary in their skills and the effort that they invested. For example, the

Aarhus team held information meetings with the schools and communicated continuously with the schools about the program, whereas the national implementations included merely information sent out about the program in the beginning. These differences in the way the program was run locally and nationally may provide one explanation for the profound differences that we observe in the implementation fidelity, which could drive the differences in the overall effectiveness of the program that we observe in the two studies.

Implementation fidelity and child outcomes

To understand the importance of implementation fidelity for the effectiveness of the program, we study the extent to which implementation fidelity is predictive of the impact of the program on student learning in the national trial.

Figure 6 reports the effects of the READ program conditional on the proportion of parents at the school who have downloaded the app. The graph plots the marginal effect based on a linear model and a histogram of the proportion of adopters at the school. To relax the linear functional form assumption, we also present results from a binning estimator that splits schools into tertiles based on their the percentage of parents who downloaded the app and estimates the effect within each bin separately.

The linear model suggests that the effect of the READ program increases with implementation fidelity. The binning estimator shows a significant positive effect among the top tertile, whereas the effect is not significantly different from zero for the remaining groups. This finding is robust to several specification checks. First, the result is not sensitive to whether we include the full set of school or student covariates (see Appendix F, Tabel F.2). Second, the statistically significant finding for top-adopters holds whether we devide bins by the median (Appendix F, Tabel F.3, Model 1) or by quartiles (Appendix F, Tabel F.3, Model 2). Third, the estimates are very similar irrespective of whether we include the non-compliers among the non-invited schools (Appendix F, Tabel F.3, Model 3). Moreover, the three groups (as measured by the tertiles) are rather balanced on covariates with little

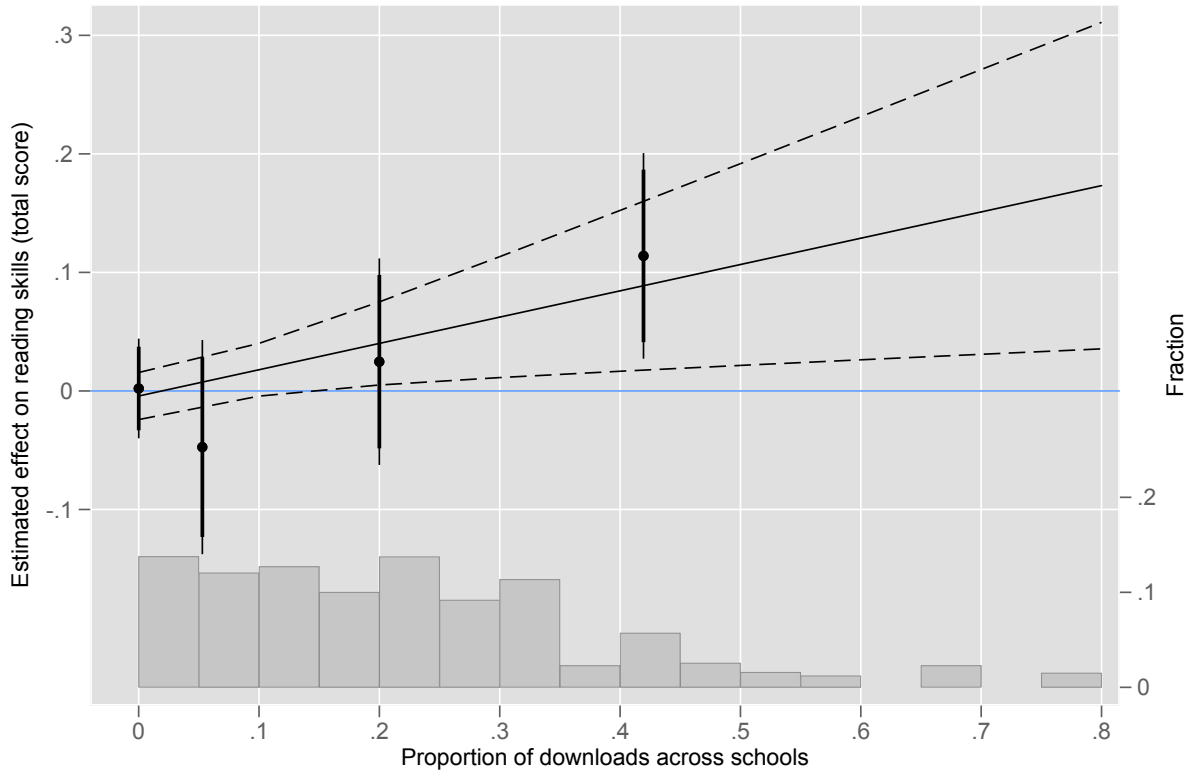


Figure 6: Estimated effect of intervention at nationwide scale

Notes: Dots and vertical lines show coefficients and 90% and 95% confidence intervals based on invited schools that did not receive the intervention and three levels (tertiles) of downloads among invited schools that received the intervention. Black line and dotted lines show estimated effect size and 95% confidence intervals based on a continuous measures of proportion of downloads across schools excluding non-compliers from the control group. Bars at the bottom show fraction of families in schools with different proportions of downloads (similar to Figure 5b).

evidence of the level of implementation fidelity being systematically related to school and student characteristics (see Appendix F, Tabel F.1). Importantly, prior school performance does not explain the degree to which the program is being implemented. Although schools may differ systematically on outcome-relevant factors that are unobserved, these findings suggest that the effect of the READ program is heterogeneous with respect to downloads of the app.

Overall, these supplementary analyses suggest that the way the program is implemented is particularly important to successfully scaling educational interventions.[13]

Conclusion

Various explanations for lack of scalability of experimental findings have been discussed in the literature. To study the scalability of parent-aimed interventions, we fielded an experiment as part of the scale-up of a reading program that had shown promising results in a local setting. The findings from the large-scale, national study suggest that the reading program was not effective at improving student learning to the same extent as in the original, local setting. The coefficients in the national implementation were small in magnitude and not statistically significant. However, we did not find much evidence of differential selection into the national program by schools compared to the full population. Schools participating in the local READ program in the original, local trial differed from the broader population on some parameters, such as the share of immigrants. Yet, we did not find much evidence of heterogeneous effects across subgroups in the national trial, which suggests that representativeness of the original sample was not the main cause of the unsuccessful scaling.

The observational replication study run by the same local government as the original

[13]We embedded two experiments aimed at increasing implementation at the school and family level. The first used info-graphics in the invitation letter to increase schools' acceptance of the invitation. The second embedded experiment used social rewards to motivate the students to read more frequently. Both experiments are described in more detail in Appendix E. Results (also presented in the appendix) show that the treatments were not strong enough to enhance implementation fidelity. Therefore, we cannot use the experiments to further test the effect of implementation fidelity on child outcomes.

study found effects that were rather similar to those of the original study. The replication of effects from the original, local trial suggests that the challenge of scaling may not be driven by problems of replicability.

The third scalability problem, implementation fidelity, is difficult to study because factors related to the implementation of the intervention are numerous and the effect of each of them therefore difficult to separate. However, behavioral data on how many families signed into the READ app provides one objective measure of differences in the level of implementation fidelity. Even though parents may have used the READ program without using the app, comparing data on app use between the local and the national studies gives an indication of differences in implementation. These data show that a much larger proportion of parents signed into the app in the local government program than in the national program. Supplementary information on the process suggests that the implementation fidelity was substantially higher when the program was run on a small, local scale with very devoted personnel than at a nationwide scale with less resources allocated to the implementation. At the national level, students at schools with relatively high levels of implementation experienced the most positive effects of the program, an effect that persists even after controlling for a large set of covariates at both the family and school level.

These findings are a strong indication that the situation in the local government programs was not representative of the way the national government implemented the program. Future research should study such implementation factors systematically since they seem to be crucial for gaining the potential benefits of scaling parent-aimed education programs.

References

Al-Ubaydli, O., List, J. A., & Suskind, D. L. (2017). What Can We Learn from Experiments? Understanding the Threats to the Scalability of Experimental Results. *American Economic Review*, *107*(5), 282–286. https://doi.org/10.1257/aer.p20171115

Andersen, S. C., & Hvidman, U. (2020). Can Reminders and Incentives Improve Implementation Within Government? Evidence from a Field Experiment. *Journal of Public Administration Research and Theory*. https://doi.org/10.1093/jopart/muaa022

Andersen, S. C., & Nielsen, H. S. (2016). Reading intervention with a growth mindset approach improves children's skills. *Proceedings of the National Academy of Sciences*, *113*(43), 12111–12113. https://doi.org/10.1073/pnas.1607946113

Angrist, J. D., & Pischke, J. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton, Princeton University Press.

Bergman, P. (2019). Parent-Child Information Frictions and Human Capital Investment: Evidence from a Field Experiment. *Journal of Political Economy*, *forthcoming*.

Bergman, P., & Chan, E. W. (2019). Leveraging Parents: The Impact of High-Frequency Information on Student Achievement. *Journal of Human Resources*, *forthcoming*.

Bergman, P., & Rogers, T. (2017). *The Impact of Defaults on Technology Adoption, and its Underappreciation by Policymakers* (Working Paper No. 6721). CESifo Working Paper Series. https://doi.org/10.3386/w25451

Berkowitz, T., Schaeffer, M. W., Maloney, E. A., Peterson, L., Gregor, C., Levine, S. C., & Beilock, S. L. (2015). Math at home adds up to achievement in school. *Science*, *350*(6257), 196–198. https://doi.org/10.1126/science.aac7427

Beuchert, L. V., & Nandrup, A. B. (2018). The Danish national tests at a glance. *Nationaloekonomisk Tidsskrift*, *1*, 1–37.

Bold, T., Kimenyi, M., Mwabu, G., Ng'ang'a, A., & Sandefur, J. (2018). Experimental evidence on scaling up education reforms in Kenya. *Journal of Public Economics*, *168*, 1–20. https://doi.org/10.1016/j.jpubeco.2018.08.007

Czibor, E., Jimenez-Gomez, D., & List, J. A. (2019). The dozen things experimental economists should do (more of). *Southern Economic Journal*, *86*(2), https://onlinelibrary.wiley.com/doi/pdf/10 371–432. https://doi.org/10.1002/soej.12392

Davis, J. M., Guryan, J., Hallberg, K., & Ludwig, J. (2017). *The economics of scale-up* (Working Paper No. 23925). National Bureau of Economic Research. https://doi.org/10.3386/w23925

Doss, C., Fahle, E. M., Loeb, S., & York, B. N. (2019). More than just a nudge: Supporting kindergarten parents with differentiated and personalized text messages. *Journal of Human Resources*, *54*(3), http://jhr.uwpress.org/content/54/3/567.full.pdf+html, 567–603. https://doi.org/10.3368/jhr.54.3.0317-8637R

Duflo, E., Glennerster, R., & Kremer, M. (2008). Using randomization in development economics research: A toolkit. In *T. schultz and john strauss, eds., handbook of development economics*. Amsterdam, New York, North Holland.

Dweck, C. S. (1999). *Self-theories: Their Role in Motivation, Personality, and Development*. Psychology Press.

Dweck, C. S. (2006). *Mindset: The New Psychology of Success* (Reprint edition). New York, Random House.

Gelman, A., & Carlin, J. (2014). Beyond power calculations: Assessing type s (sign) and type m (magnitude) errors [PMID: 26186114]. *Perspectives on Psychological Science*, *9*(6), https://doi.org/10.1177/1745691614551642, 641–651. https://doi.org/10.1177/1745691614551642

Guryan, J., Hurst, E., & Kearney, M. (2008). Parental Education and Parental Time with Children. *Journal of Economic Perspectives*, *22*(3), 23–46. https://doi.org/10.1257/jep.22.3.23

Haimovitz, K., & Dweck, C. S. (2016). Parents' Views of Failure Predict Children's Fixed and Growth Intelligence Mind-Sets. *Psychological Science*, *27*(6), 859–869. https://doi.org/10.1177/0956797616639727

Harrison, G. W., & List, J. A. (2004). Field Experiments. *Journal of Economic Literature*, *42*(4), 1009–1055. https://doi.org/10.1257/0022051043004577

Heckman, J. J. (1992). Randomization and social policy evaluation (C. Manski & I. Garfinkel, Eds.). In C. Manski & I. Garfinkel (Eds.), *Evaluating welfare training programs*. Cambridge, MA, Harvard University Press.

Heckman, J. J. (2020). *Randomization and social policy evaluation revisited* (Working Paper No. 2020-001). Human Capital and Economic Opportunity Global Working Group. https://hceconomics.uchicago.edu/research/working-paper/randomization-and-social-policy-evaluation-revisited

Heckman, J. J., & Smith, J. A. (1995). Assessing the Case for Social Experiments. *Journal of Economic Perspectives*, *9*(2), 85–110. https://doi.org/10.1257/jep.9.2.85

Jakobsen, M., & Andersen, S. C. (2013). Coproduction and Equity in Public Service Delivery. *Public Administration Review*, *73*(5), 704–713. https://doi.org/10.1111/puar.12094

Koch, A., Nafziger, J., & Nielsen, H. S. (2015). Behavioral economics of education. *Journal of Economic Behavior & Organization*, *115*, 3–17. https://doi.org/10.1016/j.jebo.2014.09.005

Lavecchia, A. M., Liu, H., & Oreopoulos, P. (2016). Chapter 1 - Behavioral Economics of Education: Progress and Possibilities (E. A. Hanushek, S. Machin, & L. Woessmann, Eds.). In E. A. Hanushek, S. Machin, & L. Woessmann (Eds.), *Handbook of the Economics of Education*. Elsevier. https://doi.org/10.1016/B978-0-444-63459-7.00001-4

Lortie-Forgues, H., & Inglis, M. (2019). Rigorous large-scale educational rcts are often uninformative : Should we be concerned? *Educational Researcher*. http://eprints.whiterose.ac.uk/141754/

Mayer, S. E., Kalil, A., Oreopoulos, P., & Gallegos, S. (2018). Using Behavioral Insights to Increase Parental Engagement: The Parents and Children Together Intervention. *Journal of Human Resources*, 0617–8835R. https://doi.org/10.3368/jhr.54.4.0617.8835R

Moorman, E. A., & Pomerantz, E. M. (2010). Ability mindsets influence the quality of mothers' involvement in children's learning: An experimental investigation. *Developmental Psychology*, *46*(5), 1354–1362. https://doi.org/10.1037/a0020376

Mullainathan, S., & Shafir, E. (2013). *Scarcity: Why Having Too Little Means So Much*. New York, Times books.

Muralidharan, K., & Niehaus, P. (2017). Experimentation at Scale. *Journal of Economic Perspectives*, *31*(4), 103–124. https://doi.org/10.1257/jep.31.4.103

Noble, C., Sala, G., Peter, M., Lingwood, J., Rowland, C., Gobet, F., & Pine, J. (2019). The impact of shared book reading on children's language skills: A meta-analysis. *Educational Research Review*, *28*, 100290. https://doi.org/https://doi.org/10.1016/j.edurev.2019.100290

Pomerantz, E. M., Moorman, E. A., & Litwack, S. D. (2007). The How, Whom, and Why of Parents' Involvement in Children's Academic Lives: More Is Not Always Better. *Review of Educational Research*, *77*(3), 373–410. https://doi.org/10.3102/003465430305567

Rogers, T., & Feller, A. (2018). Reducing student absences at scale by targeting parents' misbeliefs. *Nature Human Behaviour*, *2*(5), 335. https://doi.org/10.1038/s41562-018-0328-1

Vivalt, E. (2019). *How Much Can We Generalize from Impact Evaluations?* (Working Paper).

York, B. N., Loeb, S., & Doss, C. (2019). One step at a time: The effects of an early literacy text-messaging program for parents of preschoolers. *Journal of Human Resources*, *54*(3), http://jhr.uwpress.org/content/54/3/537.full.pdf+html, 537–566. https://doi.org/10.3368/jhr.54.3.0517-8756R

A. Descriptive Statistics. National sample.

Table A.1: Descriptive Statistics. National sample.

	Mean	SD	N
Student level			
Child is a boy (2016)	0.52	0.50	53128
Child's age (2016)	8.09	0.34	53128
Child immigrant (2016)	0.11	0.31	53128
No. of children in family (2016)	2.32	0.90	53128
Mother's total income (1000 kr.) (2014)	260.60	189.73	51959
Mother's age (2014)	38.71	4.97	51976
Father's total income (1000 kr.) (2014)	382.37	319.98	51218
Father's age (2014)	41.27	5.79	51025
Child living with both parents (2016)	0.71	0.45	53128
Child living with single parent (2016)	0.20	0.40	53128
Child living with parent in new relationship or not living with own parents (2016)	0.08	0.28	53128
Mother compulsory education (2014)	0.14	0.35	51700
Mother upper secondary education (2014)	0.05	0.23	51700
Mother vocational education (2014)	0.31	0.46	51700
Mother short-cycle education (2014)	0.05	0.22	51700
Mother medium-cycle education (2014)	0.27	0.45	51700
Mother long-cycle education (2014)	0.13	0.34	51700
Mother outside labor market (2014)	0.17	0.38	51850
Mother unemployed (2014)	0.04	0.20	51849
Mother employed (2014)	0.79	0.41	51874
Father compulsory education (2014)	0.16	0.37	50715
Father upper secondary education (2014)	0.05	0.22	50715
Father vocational education (2014)	0.41	0.49	50715
Father short-cycle education (2014)	0.08	0.27	50715
Father medium-cycle education (2014)	0.14	0.34	50715
Father long-cycle education (2014)	0.13	0.34	50715
Father outside labor market (2014)	0.09	0.29	50984
Father unemployed (2014)	0.03	0.17	50984
Father employed (2014)	0.88	0.33	50997
Missing on Mother's education	0.06	0.25	53128
Missing on Mother's employment status	0.02	0.15	53128
Missing on Mother's total income (1000 kr.)	0.02	0.15	53128
Missing on Mother's age	0.02	0.15	53128
Missing on Father's education	0.08	0.27	53128
Missing on Father's employment status	0.04	0.19	53128
Missing on Father's total income (1000 kr.)	0.04	0.19	53128
Missing on Father's age	0.04	0.19	53128
Missing test score	0.03	0.18	53128
Danish reading - Total score (2017)	0.00	1.00	51312
School level			
School size[1]	46.52	25.83	1142
Average test score 2016[2]	-0.03	0.41	1131
Number of Schools			1142

Notes: Mean, standard deviation and number of non-missing observations. [1] Number of students in second grade. [2]Standardized using the mean and the standard deviation from the national sample in 2017.

B. Balance: National sample and local treatment

Table B.1: Differences in mean between full cohort and local treatment on background characteristics

	(1) National population	(2) Local treatment	(3) 1-2
Student level			
Child is a boy	0.52	0.52	-0.00
Child's age (2016)	8.09	8.10	-0.01
Child immigrant	0.11	0.20	-0.09**
Mother compulsory education (2014)	0.14	0.13	0.01
Mother upper secondary education (2014)	0.05	0.06	-0.00
Mother vocational education (2014)	0.31	0.18	0.12**
Mother short-cycle education (2014)	0.05	0.05	-0.00
Mother medium-cycle education (2014)	0.27	0.27	0.01
Mother long-cycle education (2014)	0.13	0.27	-0.14**
Father compulsory education (2014)	0.16	0.15	0.01
Father upper secondary education (2014)	0.05	0.06	-0.01
Father vocational education (2014)	0.41	0.22	0.19**
Father short-cycle education (2014)	0.08	0.08	0.00
Father medium-cycle education (2014)	0.14	0.18	-0.05**
Father long-cycle education (2014)	0.13	0.27	-0.14**
Missing test score (2017)	0.03	0.04	-0.00
School level			
School size[1]	46.52	63.91	-17.39**
Average test score (2016)[2]	-0.03	-0.21	0.19*
Students	53128	1470	54598
Schools	1142	23	1165

Notes: + p < 0.1; * p < 0.05; ** p < 0.01. Standard errors clustered on the school level. [1] Number of students in second grade. [2] Number of students in second grade. [2]Average test score the year before implementation of READ. Standardized using the mean and the standard deviation from the national sample in 2017.

C. Background Characteristics and Standardized Test Scores.

Table C.1: Background Characteristics and Standardized Test Scores.

	(1) OLS	
Child is a boy	-0.266**	(0.009)
Child's age (2016) = 7	-	-
Child's age = 8	0.044	(0.037)
Child's age = 9	-0.168**	(0.041)
Child's age > 9	-0.358*	(0.144)
Child immigrant	-0.282**	(0.023)
Child living with single parent (2016)	-0.078**	(0.012)
Child living with parent in new relationship or not living with own parents (2016)	-0.050**	(0.017)
No. of children in family (2016)	-0.033**	(0.005)
Mother upper secondary education (2014)	0.261**	(0.023)
Mother vocational education (2014)	0.113**	(0.016)
Mother short-cycle education (2014)	0.287**	(0.023)
Mother medium-cycle education (2014)	0.341**	(0.018)
Mother long-cycle education (2014)	0.478**	(0.021)
Mother unemployed (2014)	-0.034	(0.027)
Mother employed (2014)	0.048**	(0.016)
Mother's total income (1000 kr.) (2014)	0.000**	(0.000)
Mother's ag (2014)	0.003*	(0.001)
Father upper secondary education (2014)	0.263**	(0.022)
Father vocational education (2014)	0.126**	(0.014)
Father short-cycle education (2014)	0.259**	(0.020)
Father medium-cycle education (2014)	0.316**	(0.017)
Father long-cycle education (2014)	0.409**	(0.019)
Father unemployed (2014)	0.011	(0.030)
Father employed (2014)	0.064**	(0.017)
Father's total income (1000 kr.) (2014)	0.000**	(0.000)
Father's age (2014)	-0.001	(0.001)
Missing on Mother's total income (1000 kr.)	-0.229	(0.141)
Missing on Mother's age	-0.183	(0.137)
Missing on Father's total income (1000 kr.)	-0.053	(0.083)
Missing on Father's age	-0.095	(0.082)
Constant	-0.388**	(0.060)
Students	51312	
Schools (clusters)	1140	
Adjusted R-squared	0.139	

Notes: Standard errors clustered on the school level in parentheses. + $p < 0.1$; * $p < 0.05$; ** $p < 0.01$.

D. Specifications

Table D.1: Total score

	(1)	(2)	(3)
Not READ	0.076		
	(0.055)		
Not READ Aarhus	0.207**		
	(0.073)		
Post READ	0.117	0.133+	0.103
	(0.073)	(0.070)	(0.071)
Not READ X Post READ	-0.122	-0.139+	-0.113
	(0.074)	(0.071)	(0.072)
Not READ Aarhus X Post READ	-0.153	-0.176+	-0.153+
	(0.095)	(0.091)	(0.091)
Mean of control	-0.062	-0.062	-0.062
Observations	96965	96965	96965
Schools (clusters)	699	699	699
Adjusted R-squared	0.001	0.000	0.116
Fixed Effects	No	Yes	Yes
Covariates	No	No	Yes

Notes: Standard errors clustered on the school level in parentheses. + p < 0.1; * p < 0.05; ** p < 0.01. The full list of the included covariates is reported in Table C.1.

Table D.2: Language comprehension

	(1)	(2)	(3)
Not READ	0.049		
	(0.048)		
Not READ Aarhus	0.171**		
	(0.064)		
Post READ	0.076	0.092+	0.068
	(0.057)	(0.054)	(0.053)
Not READ X Post READ	-0.078	-0.096+	-0.071
	(0.058)	(0.056)	(0.054)
Not READ Aarhus X Post READ	-0.101	-0.123	-0.102
	(0.083)	(0.081)	(0.079)
Mean of control	-0.036	-0.036	-0.036
Observations	96965	96965	96965
Schools (clusters)	699	699	699
Adjusted R-squared	0.001	0.000	0.092
Fixed Effects	No	Yes	Yes
Covariates	No	No	Yes

Notes: Standard errors clustered on the school level in parentheses. + $p < 0.1$; * $p < 0.05$; ** $p < 0.01$. The full list of the included covariates is reported in Table C.1.

Table D.3: Decoding

	(1)	(2)	(3)
Not READ	0.060		
	(0.053)		
Not READ Aarhus	0.175*		
	(0.072)		
Post READ	0.084	0.097	0.070
	(0.076)	(0.073)	(0.074)
Not READ X Post READ	-0.091	-0.104	-0.085
	(0.077)	(0.074)	(0.075)
Not READ Aarhus X Post READ	-0.117	-0.138	-0.119
	(0.092)	(0.088)	(0.088)
Mean of control	-0.047	-0.047	-0.047
Observations	96965	96965	96965
Schools (clusters)	699	699	699
Adjusted R-squared	0.001	0.000	0.092
Fixed Effects	No	Yes	Yes
Covariates	No	No	Yes

Notes: Standard errors clustered on the school level in parentheses. + $p < 0.1$; * $p < 0.05$; ** $p < 0.01$. The full list of the included covariates is reported in Table C.1.

Table D.4: Text comprehension

	(1)	(2)	(3)
Not READ	0.091^{+}		
	(0.050)		
Not READ Aarhus	0.197**		
	(0.064)		
Post READ	0.149*	0.161*	0.133*
	(0.067)	(0.064)	(0.067)
Not READ X Post READ	-0.151*	-0.163*	-0.142*
	(0.068)	(0.066)	(0.068)
Not READ Aarhus X Post READ	-0.185*	-0.201*	-0.179*
	(0.088)	(0.085)	(0.086)
Mean of control	-0.081	-0.081	-0.081
Observations	96965	96965	96965
Schools (clusters)	699	699	699
Adjusted R-squared	0.001	0.000	0.093
Fixed Effects	No	Yes	Yes
Covariates	No	No	Yes

Notes: Standard errors clustered on the school level in parentheses. $+$ p $<$ 0.1; $*$ p $<$ 0.05; $**$ p $<$ 0.01. The full list of the included covariates is reported in Table C.1.

E. Supplementary Materials

Two experiments were embedded in the trial. First, as illustrated in Figure E.1 two variants of the invitation letter were send to the schools. Second, among schools that accepted the invitations, two variance of the READ program was tested. Below we describe each of these embedded experiments. Figure E.1 contains less information than Figure 2. The purpose is to illustrate the two embedded experiments.

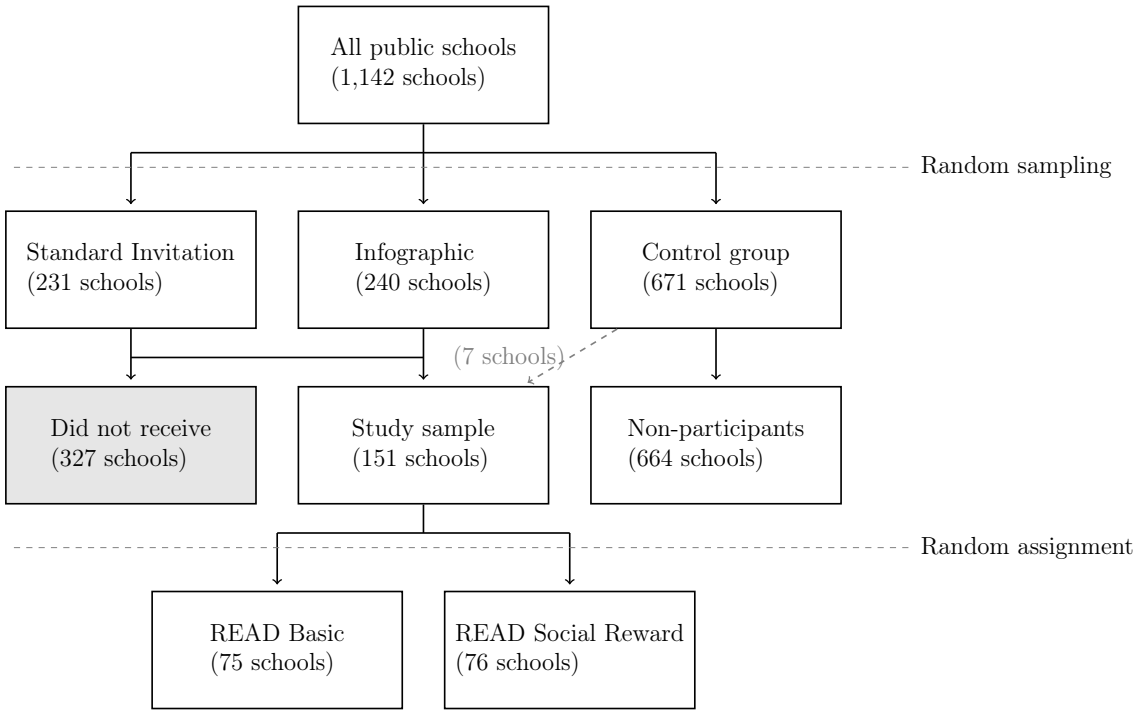
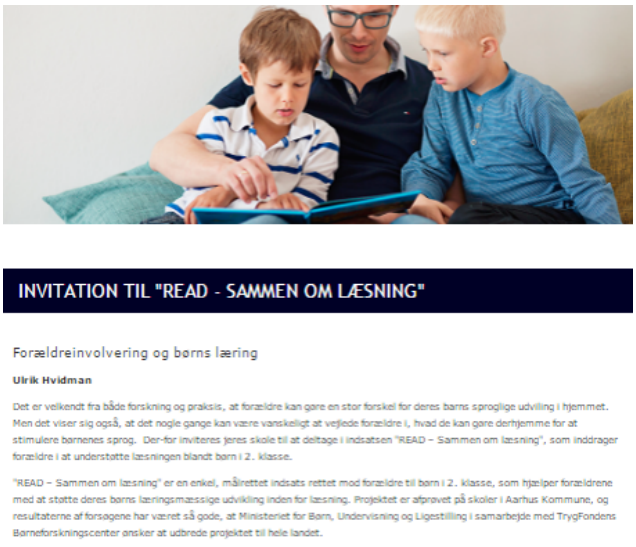


Figure E.1: Design of two embedded experiments

Invitations

To study how to encourage schools to adopt the program, we randomly assigned schools to one of two versions of the invitation letter. Both groups received an invitation from The Ministry of Education describing the program and its effects. However, since effects from randomized controlled trials may be difficult to convey to persons without a background in research, the one group of schools were assigned to an infographic illustrating the effect of the intervention as estimated in the first randomized controlled-trial in 2014. Apart from infographic, the invitations were identical.

Figure E.2 panel (a) shows the main invitation that all schools receive. Panel (b) shows the infographic that was randomly assigned to halfe of the schools in the invitation group.



INVITATION TIL "READ - SAMMEN OM LÆSNING"

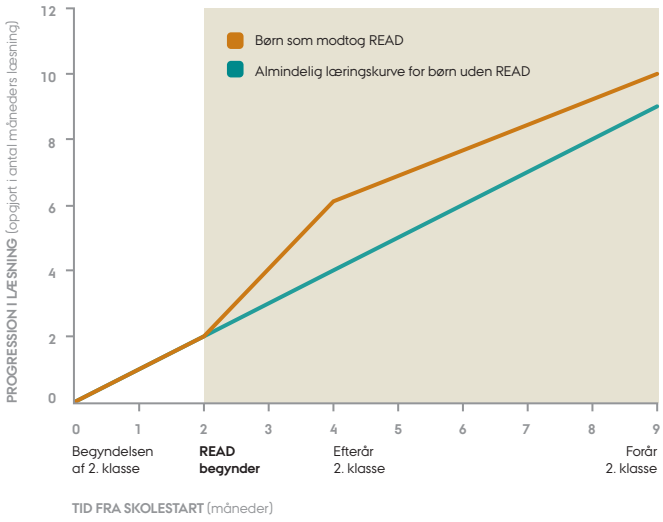
Forældreinvolvering og børns læring

Ulrik Hvidman

Det er velkendt fra både forskning og praksis, at forældre kan gøre en stor forskel for deres barns sproglige udvikling i hjemmet. Men det viser sig også, at det nogle gange kan være vanskeligt at vejlede forældre i, hvad de kan gøre derhjemme for at stimulere børnenes sprog. Der-for inviteres jeres skole til at deltage i indsatsen "READ – Sammen om læsning", som inddrager forældre i at understøtte læsningen blandt børn i 2. klasse.

"READ – Sammen om læsning" er en enkel, målrettet indsats rettet mod forældre til børn i 2. klasse, som hjælper forældrene med at støtte deres børns læringsmæssige udvikling inden for læsning. Projektet er afprøvet på skoler i Aarhus Kommune, og resultaterne af forsøgene har været så gode, at Ministeriet for Børn, Undervisning og Ligestilling i samarbejde med TrygFondens Børneforskningscenter ønsker at udbrede projektet til hele landet.

(a) Standard Invitation



Note til figur: Effektevalueringen er gennemført af Aarhus Universitet i samarbejde med Børn og Unge i Aarhus og VIA University College. Den stiplede sorte linje viser den gennemsnitlige udvikling i læsefærdigheder for elever i løbet af de første 9 måneder i 2. klasse, som ikke modtog READ. Den røde linje viser den gennemsnitlige udvikling i tilsvarende periode for børn i 2. klasse, som modtog READ to måneder inde i skoleåret (indførelsen af READ er angivet med vertikale stiplede linje).

(b) Infographic

Figure E.2: Invitation email. All schools received the standard invitation. Half of the schools were randomly assigned to also receive the infographic.

Table E.1 shows that the two experimental groups in the embedded invitation experiment

were balanced on major baseline characteristics.

Table E.1: Balance Invitation experiment

	(1) Standard Invitation		(2) Infographic		(3) 1-2
Average test score 2016	-0.05	(0.44)	-0.05	(0.40)	0.00
Mother high education	0.39	(0.17)	0.38	(0.16)	0.01
Child immigrant	0.11	(0.14)	0.09	(0.12)	0.02^{+}
School size	46.29	(26.43)	45.38	(25.53)	0.91
Observations (Schools)	231		240		471

Notes: Standard errors in parentheses. + p < 0.1; * p < 0.05; ** p < 0.01.

Table E.2 shows that the Infographic invitation did not increase participation in the READ program significantly.

Table E.2: Effect of Infographic on participation in the READ program

	(1)	(2)
Infographic	-0.063	-0.070
	(0.042)	(0.043)
Mean of control	0.338	0.338
Observations (Schools)	471	468
Adjusted R-squared	0.003	0.001
Covariates	No	Yes

Notes: Standard errors in parentheses. + p < 0.1; * p < 0.05; ** p < 0.01. The full list of the included covariates is reported in Table C.1.

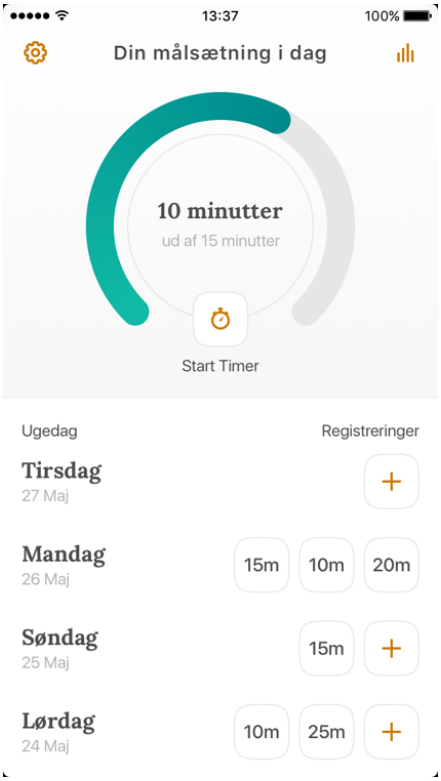
Social Rewards

Behavioral barriers may constitute a key challenge to the effectiveness of educational interventions that target families. The basic notion behind parent-aimed interventions is that parents will build a better learning environment at home. However, a rapidly growing research literature in behavioral social science has focused on understanding why people often fail to do things they know they should do. Even parents who know what steps to take to significantly improve their children's abilities may fail to take these steps because of behavioral factors (for reviews of behavioral economics of education, see Koch et al., 2015; Lavecchia et al., 2016). One explanation may be that parents experience a present bias or lack self-control. As people often discount future outcomes relative to immediate outcomes, it is hard for parents to invest time and effort today for a return on their child's human capital that might show up years later. Moreover, parents may lack self-control and perseverance in their busy everyday lives. As a result, many programs stop at the good intentions because of scarcity of time, energy, and persistence among participants (Mullainathan & Shafir, 2013). In a study of a school information system that provided information to parents, Bergman (2019) found that less than half of the families ever used the system—and that non-users were typically low-income families and families of low-achieving students. The same constraints may be true for interventions that provide resources to parents and try to encourage them to read with their children. Some have proposed approaches to mitigate these behavioral barriers. In a study of the use of a reading application, a treatment group was exposed to three different behavioral tools (i.e., a commitment device, text message reminders, and a social reward). The study suggests that behavioral tools were effective as they increased the usage of the reading application by 1 standard deviation (Mayer et al., 2018).

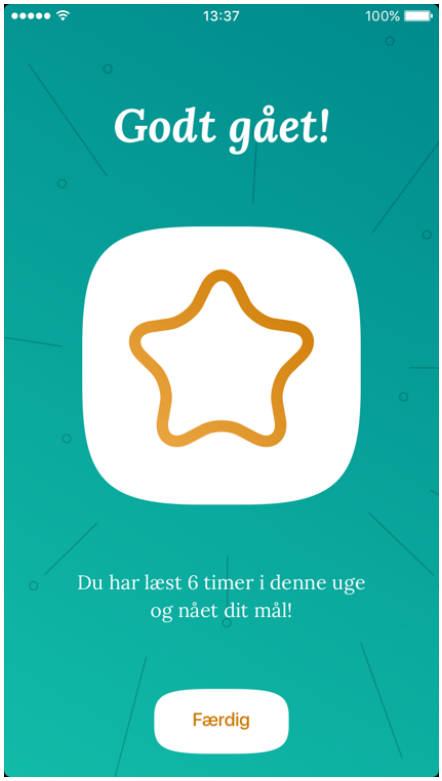
To examine the influence of such behavioral barriers, we randomly assigned participating schools to two versions of the READ program. By making small deviations from the basic READ program, we can test the effect of these modifications. We experimented with social rewards designed to shift preferences by increasing the utility of the current behavior. Social

rewards have been proposed as an effective strategy for changing behaviors. In a study by Mayer et al., 2018, a treatment group was exposed to three different behavioral tools (i.e., a commitment device, text message reminders, and a social reward). Although this study suggests that behavioral tools are effective, it is difficult to disentangle the effects of the different interventions as the treatment also included an information intervention apart from the behavioral tools that increased their engagement with their children. In the original READ program, some teachers decided to use a logbook in which families could note every time the child read (as previously described). The logbook endorsed child effort, not performance or results (not the speed or accuracy of the reading). When the children had read ten times, they could bring the logbook to their schoolteacher, and the class would receive a sticker. The class with the most stickers received a prize. In the original trial, use of the logbook was not randomized but selected by teachers. To test the additional effect of this social reward entailed by the logbook competition, we randomly assigned schools that accepted to receive READ to one of two conditions: READ Basic and READ Social Reward. Parents in the READ Social Reward group were provided with the same material as READ Basic, but also with the logbook. As in the original trial, when the children had read ten times, they could bring the logbooks to their schoolteacher, and the class would receive a sticker. The class with most stickers at the school received a prize; a gift card to a reading store worth 10,000 DKK (USD 1,500). This experiment enabled us to test the effect of the social rewards component.

Figure E.3 shows screenshots from the app, where parents could register every time they had read with their child.



(a) Screenshot (I) from READ app



(b) Screenshot (II) from READ app

Figure E.3: Smartphone app

Table E.3 shows that schools assigned to either READ Basic or READ Social Reward were balanced at baseline.

Table E.3: Baseline balance of READ Social Reward relative to READ Basic

	(1) Basic	(2) Social Reward	(3) 1-2
Student level			
Child is a boy	0.53	0.52	0.01
Child's age (2016)	8.08	8.10	-0.02^{+}
Child immigrant	0.12	0.12	-0.00
Mother compulsory education (2014)	0.15	0.15	0.00
Mother upper secondary education (2014)	0.05	0.05	0.00
Mother vocational education (2014)	0.32	0.33	-0.01
Mother short-cycle education (2014)	0.05	0.05	-0.01
Mother medium-cycle education (2014)	0.27	0.27	0.00
Mother long-cycle education (2014)	0.12	0.11	0.01
Father compulsory education (2014)	0.17	0.18	-0.01
Father upper secondary education (2014)	0.05	0.04	0.01
Father vocational education (2014)	0.43	0.44	-0.00
Father short-cycle education (2014)	0.08	0.08	-0.00
Father medium-cycle education (2014)	0.13	0.13	-0.00
Father long-cycle education (2014)	0.12	0.10	0.02
Missing test score (2017)	0.03	0.04	-0.01
School level			
School size[1]	45.00	44.62	0.38
Average test score (2016)[2]	-0.08	-0.10	0.02
Students	3375	3391	6766
Schools	75	76	151

Notes: $+$ p $<$ 0.1; * p $<$ 0.05; ** p $<$ 0.01. Standard errors clustered on the school level. [1] Number of students in second grade. [2]Standardized using the mean and the standard deviation from the national sample in 2017.

Table E.4 shows that READ Social Reward did not change the number of app downloads compared to READ Basic.

Table E.4: Effect of READ Social Reward relative to READ Basic on the number of app downloads

	(1)	(2)
Social Reward	0.040	0.043
	(0.037)	(0.036)
Mean of control	0.228	0.228
Observations	4804	4804
Schools (clusters)	110	110
Adjusted R-squared	0.002	0.032
Covariates	No	Yes

Notes: Standard errors clustered on the school level in parentheses. + $p < 0.1$; * $p < 0.05$; ** $p < 0.01$. The full list of the included covariates is reported in Table C.1.

F. Robustness of Implementation Analyses

Table F.1: Baseline balance

	(1) Invite, not part	(2) Lower third	(3) Midlle third	(4) Upper third	(5) 1-2	(6) 1-3	(7) 1-4
Student level							
Child is a boy	0.51	0.52	0.53	0.53	-0.01	-0.02^{+}	-0.02
Child's age (2016)	8.09	8.08	8.09	8.09	0.01	0.00	0.01
Child immigrant	0.11	0.17	0.10	0.09	-0.06*	0.01	0.03
Mother compulsory education (2014)	0.14	0.17	0.14	0.14	-0.03^{+}	-0.00	0.01
Mother upper seconday education (2014)	0.06	0.05	0.05	0.05	0.01	0.01	0.00
Mother vocational education (2014)	0.30	0.30	0.32	0.34	-0.00	-0.01	-0.04^{+}
Mother short-cycle education (2014)	0.05	0.04	0.05	0.05	0.01^{+}	0.00	-0.00
Mother medium-cycle education (2014)	0.27	0.25	0.28	0.29	0.02	-0.01	-0.01
Mother long-cycle education (2014)	0.13	0.13	0.12	0.10	0.01	0.01	0.03
Father compulsory education (2014)	0.17	0.18	0.17	0.15	-0.01	-0.00	0.01
Father upper seconday education (2014)	0.05	0.05	0.05	0.05	0.00	0.01	0.00
Father vocational education (2014)	0.40	0.41	0.42	0.46	-0.01	-0.03	-0.07*
Father short-cycle education (2014)	0.08	0.07	0.08	0.08	0.00	-0.01	-0.01
Father medium-cycle education (2014)	0.14	0.12	0.13	0.13	0.02	0.01	0.00
Father long-cycle education (2014)	0.14	0.12	0.11	0.10	0.01	0.03	0.04*
Missing test score (2017)	0.04	0.05	0.03	0.03	-0.02^{+}	0.01	0.00
School level							
School size[1]	46.36	43.23	46.59	43.89	3.13	-0.23	2.47
Average test score (2016)[2]	-0.04	-0.13	-0.03	-0.06	0.09	-0.01	0.02
Students	15161	2075	2283	2063	17236	17444	17224
Schools	327	48	49	47	375	376	374

Notes: $+ p < 0.1$; * $p < 0.05$; ** $p < 0.01$. Standard errors clustered on the school level. [1] Number of students in second grade. [2]Standardized using the mean and the standard deviation from the national sample in 2017.

Table F.2: Effect of READ by Level of Implementation (Downloads). Models with and without covariates included.

	Categorical variable			Continous variable		
	(1)	(2)	(3)	(4)	(5)	(6)
Invite, not part	-0.012	-0.004	0.002			
	(0.027)	(0.023)	(0.021)			
Lower third	-0.128*	-0.072	-0.047			
	(0.060)	(0.051)	(0.046)			
Midlle third	-0.004	0.023	0.025			
	(0.052)	(0.046)	(0.044)			
Upper third	0.074+	0.099*	0.114*			
	(0.043)	(0.043)	(0.044)			
Proportion of downloads across schools				0.121	0.201*	0.222*
				(0.092)	(0.087)	(0.091)
Girl		ref.	ref.		ref.	ref.
Child is a boy=1		-0.267**	-0.268**		-0.267**	-0.268**
		(0.009)	(0.009)		(0.009)	(0.009)
Child age 7 years		ref.	ref.		ref.	ref.
8		0.043	0.045		0.047	0.049
		(0.037)	(0.037)		(0.037)	(0.037)
9		-0.170**	-0.179**		-0.167**	-0.175**
		(0.041)	(0.041)		(0.041)	(0.041)
10 +		-0.358*	-0.406**		-0.358*	-0.403**
		(0.145)	(0.147)		(0.145)	(0.147)
Child immigrant=1		-0.279**	-0.254**		-0.280**	-0.255**
		(0.023)	(0.023)		(0.023)	(0.023)
Child lives with both parents		ref.	ref.		ref.	ref.
Child living with single parent		-0.078**	-0.072**		-0.078**	-0.072**
		(0.012)	(0.013)		(0.012)	(0.013)
Child living with parent in new relationship or not living with own parents		-0.051**	-0.053**		-0.049**	-0.051**
		(0.017)	(0.017)		(0.017)	(0.017)
No. of children in family		-0.033**	-0.031**		-0.034**	-0.032**
		(0.005)	(0.006)		(0.006)	(0.006)
Mother compulsory education		ref.	ref.		ref.	ref.
Mother upper seconday education (2014)		0.261**	0.251**		0.263**	0.253**
		(0.023)	(0.023)		(0.023)	(0.024)
Mother vocational education (2014)		0.112**	0.110**		0.115**	0.112**
		(0.016)	(0.016)		(0.016)	(0.016)
Mother short-cycle education (2014)		0.286**	0.275**		0.291**	0.279**
		(0.023)	(0.023)		(0.023)	(0.023)
Mother medium-cycle education (2014)		0.340**	0.331**		0.344**	0.334**
		(0.018)	(0.018)		(0.018)	(0.018)
Mother long-cycle education (2014)		0.478**	0.457**		0.480**	0.459**
		(0.021)	(0.021)		(0.021)	(0.021)
Mother outside labor market		ref.	ref.		ref.	ref.
Mother unemployed		-0.034	-0.030		-0.033	-0.028
		(0.027)	(0.027)		(0.027)	(0.027)
Mother employed		0.048**	0.043**		0.050**	0.046**
		(0.016)	(0.016)		(0.016)	(0.016)
Mother's total income (1000 kr.)		0.000**	0.000**		0.000**	0.000**
		(0.000)	(0.000)		(0.000)	(0.000)
Mother's age in 2014, y		0.003*	0.002+		0.003*	0.003*
		(0.001)	(0.001)		(0.001)	(0.001)

Continues next page.

Tabel F.2 continued

Father compulsory education		ref.	ref.		ref.	ref.
Father upper seconday education (2014)		0.262**	0.246**		0.263**	0.248**
		(0.022)	(0.022)		(0.022)	(0.022)
Father vocational education (2014)		0.125**	0.119**		0.126**	0.120**
		(0.014)	(0.014)		(0.014)	(0.014)
Father short-cycle education (2014)		0.258**	0.245**		0.256**	0.243**
		(0.020)	(0.020)		(0.020)	(0.020)
Father medium-cycle education (2014)		0.316**	0.300**		0.317**	0.302**
		(0.017)	(0.017)		(0.017)	(0.017)
Father long-cycle education (2014)		0.409**	0.382**		0.409**	0.384**
		(0.019)	(0.019)		(0.019)	(0.019)
Father outside labor market		ref.	ref.		ref.	ref.
Father unemployed		0.011	0.014		0.018	0.021
		(0.030)	(0.030)		(0.030)	(0.030)
Father employed		0.063**	0.058**		0.065**	0.059**
		(0.017)	(0.017)		(0.017)	(0.017)
Father's total income (1000 kr.)		0.000**	0.000*		0.000**	0.000*
		(0.000)	(0.000)		(0.000)	(0.000)
Father's age in 2014, y		-0.001	-0.001		-0.001	-0.001
		(0.001)	(0.001)		(0.001)	(0.001)
Missing on Mother's education (6 categories)		0.138**	0.134**		0.141**	0.138**
		(0.029)	(0.029)		(0.029)	(0.029)
Missing on Mother's employment status (3 categories)		0.174+	0.126		0.172+	0.127
		(0.093)	(0.092)		(0.093)	(0.093)
Missing on Mother's total income (1000 kr.)		-0.237+	-0.199		-0.230	-0.196
		(0.141)	(0.139)		(0.142)	(0.140)
Missing on Mother's age in 2014, y		-0.179	-0.192		-0.171	-0.183
		(0.136)	(0.136)		(0.138)	(0.137)
Missing on Father's education (6 categories)		0.125**	0.104**		0.127**	0.106**
		(0.029)	(0.029)		(0.029)	(0.029)
Missing on Father's employment status (3 categories)		0.173+	0.180*		0.172+	0.179*
		(0.088)	(0.086)		(0.089)	(0.086)
Missing on Father's total income (1000 kr.)		-0.056	-0.074		-0.053	-0.070
		(0.083)	(0.077)		(0.084)	(0.078)
Missing on Father's age in 2014, y		-0.093	-0.086		-0.095	-0.088
		(0.081)	(0.081)		(0.082)	(0.081)
School size			-0.000			-0.000
			(0.000)			(0.000)
School average test score 2016			0.285**			0.283**
			(0.028)			(0.027)
Constant	0.006	-0.387**	-0.322**	-0.002	-0.406**	-0.339**
	(0.016)	(0.060)	(0.064)	(0.013)	(0.060)	(0.064)
Observations	51312	51312	51030	50980	50980	50698
Clusters (Schools/Municipalities)	1140	1140	1130	1133	1133	1123
Adjusted R-squared	0.001	0.139	0.150	0.000	0.140	0.150

Notes: Standard errors clustered on the school level in parentheses. + p < 0.1; * p < 0.05; ** p < 0.01.

Table F.3: Effect of READ by Level of Implementation (Downloads). Proportion of downloads in 2 and 4 categories, and continuous variable with non-compliers included.

	(1) 2 categories	(2) 4 categories	(3) Continous, incl. non-compliers	(4) Continous, excl. no student identifiers
Invite, not part	0.002			
	(0.021)			
Lower half	-0.038			
	(0.039)			
Upper half	0.099**			
	(0.036)			
Control group, not invited		ref.		
Invite, not part		0.002		
		(0.021)		
1st quarter		-0.099+		
		(0.051)		
2nd quarter		0.013		
		(0.054)		
3rd quarter		0.094+		
		(0.054)		
4th quarter		0.104*		
		(0.042)		
Proportion of downloads across schools			0.196*	0.224*
			(0.087)	(0.101)
Constant	-0.322**	-0.321**	-0.325**	-0.343**
	(0.064)	(0.064)	(0.064)	(0.065)
Observations	51030	51030	51030	49258
Schools (clusters)	1130	1130	1130	1087
Adjusted R-squared	0.150	0.150	0.150	0.149
Student covariats	YES	YES	YES	YES
School covariates	YES	YES	YES	YES

Notes: Standard errors clustered on the school level in parentheses. + $p < 0.1$; * $p < 0.05$; ** $p < 0.01$. Model 3 for includes non-compliers, i.e. seven schools that were not invited but participated. Model 4 excludes schools that participated, but did not grant access to student identifiers on use of app. The full list of the included covariates is reported in Table F.2.

G. Meta-analysis of Educational Interventions

The meta-analysis is based on data from two studies. First, studies of educational programs commissioned by EEF and NCEE are based on data from Lortie-Forgues and Inglis (2019). These are categorized as educational interventions. Second, studies specifically focusing on parent-aimed interventions are based on data from a systematic review of shared book reading by Noble et al. (2019). Following Noble et al. (2019), we exclude studies with effect sizes greater than three standard deviations, and we include only randomized controlled trials with standardized test outcomes in the meta-analysis.

We supplement the data on parent-aimed interventions with a systematic literature search that broadens the search criteria. The systematic search was based on the same search string as Noble et al. (2019) with the following addition: "caregiver*", "parent*", or "home" combined with "reading", "training", "education", "information", "implement*", "intervention", "achievement", "engagement", "text messag*", or "provid* knowledge". The systematic search was limited to studies with the words "random*", "causal", "experiment", or "impact" in the following journals: *Science, Proceedings of the National Academy of Sciences, Nature Human Behaviour, Journal of Economic Perspectives, Economics Letters, Journal of Human Resources, Economics of Education Review, Journal of Policy Analysis and Management, Journal of Public Administration Research and Theory, Public Administration Review, Annals of The American Academy of Political and Social Science, and Early Child Development and Care.*

Tabel G.1 lists all the studies included in our meta-analysis. A full list of references is provided below the table.

Table G.1: Analysed sample size (n) and effect size (ES) of studies included in Figure 1

Intervention type	Study	n	ES
Parent-aimed	Vaquero (2014)	18	-0.15
Parent-aimed	Vaquero (2014)	18	-0.21

Table G.1 continued

Intervention type	Study	n	ES
Parent-aimed	Woods (2017)	19	-0.09
Parent-aimed	Woods (2017)	19	-0.23
Parent-aimed	Woods (2017)	19	-0.02
Parent-aimed	Woods (2017)	19	0.32
Parent-aimed	Lonigan and Whitehurst (1998)	20	-0.08
Parent-aimed	Lonigan and Whitehurst (1998)	20	0.16
Parent-aimed	Lonigan and Whitehurst (1998)	20	0.07
Parent-aimed	Vaquero (2014)	20	0.66
Parent-aimed	Vaquero (2014)	20	0.28
Parent-aimed	Lonigan and Whitehurst (1998)	21	0.16
Parent-aimed	Lonigan and Whitehurst (1998)	21	0.04
Parent-aimed	Lonigan and Whitehurst (1998)	21	1.04
Parent-aimed	Reese et al. (2010)	21	-0.14
Parent-aimed	Reese et al. (2010)	22	0.01
Parent-aimed	Zevenbergen et al. (2018)	22	0.66
Parent-aimed	Lonigan and Whitehurst (1998)	23	-0.05
Parent-aimed	Lonigan and Whitehurst (1998)	23	0.83
Parent-aimed	Lonigan and Whitehurst (1998)	23	0.61
Parent-aimed	Lonigan and Whitehurst (1998)	23	-0.01
Parent-aimed	Lonigan and Whitehurst (1998)	23	0.34
Parent-aimed	Lonigan and Whitehurst (1998)	23	1.65
Parent-aimed	Reese et al. (2010)	23	-0.53
Parent-aimed	Reese et al. (2010)	23	-0.43
Parent-aimed	Sheets and Buyer (1999)	28	-0.53

Table G.1 continued

Intervention type	Study	n	ES
Parent-aimed	Sheets and Buyer (1999)	29	0.24
Parent-aimed	Zevenbergen et al. (2018)	30	0.45
Parent-aimed	Kotaman (2013)	40	0.44
Parent-aimed	Whitehurst and Others (1994)	44	0.69
Parent-aimed	Whitehurst and Others (1994)	44	0.41
Parent-aimed	Whitehurst and Others (1994)	44	0.59
Parent-aimed	Whitehurst and Others (1994)	44	-0.36
Parent-aimed	Whitehurst and Others (1994)	44	0.26
Parent-aimed	Whitehurst and Others (1994)	44	0.18
Parent-aimed	Whitehurst and Others (1994)	44	-0.09
Parent-aimed	Fielding-Barnsley and Purdie (2003)	49	-0.08
Parent-aimed	Fielding-Barnsley and Purdie (2003)	49	-0.28
Parent-aimed	Fielding-Barnsley and Purdie (2003)	49	-0.10
Parent-aimed	Wing-Yin Chow and McBride-Chang (2003)	54	0.24
Parent-aimed	Wing-Yin Chow and McBride-Chang (2003)	55	0.18
Parent-aimed	Wing-Yin Chow and McBride-Chang (2003)	57	0.45
Parent-aimed	Wing-Yin Chow and McBride-Chang (2003)	57	0.11
Parent-aimed	Levin and Aram (2012)	62	0.34
Parent-aimed	Levin and Aram (2012)	62	0.24
Parent-aimed	Levin and Aram (2012)	65	0.06
Parent-aimed	Levin and Aram (2012)	65	-0.12
Parent-aimed	Levin and Aram (2012)	69	0.19
Parent-aimed	Levin and Aram (2012)	69	0.09
Educational	Biggart, O'Hare, et al. (2015)	72	0.11

Table G.1 continued

Intervention type	Study	n	ES
Parent-aimed	Chow et al. (2008)	73	0.01
Parent-aimed	Chow et al. (2008)	73	0.01
Parent-aimed	Chow et al. (2008)	73	0.55
Parent-aimed	Chow et al. (2008)	73	0.14
Parent-aimed	Chow et al. (2008)	74	0.18
Parent-aimed	Chow et al. (2008)	74	0.35
Parent-aimed	Murray et al. (2016)	82	0.95
Parent-aimed	Murray et al. (2016)	82	5.26
Parent-aimed	Murray et al. (2016)	82	0.65
Parent-aimed	Neville et al. (2013)	103	0.22
Parent-aimed	Neville et al. (2013)	103	0.38
Parent-aimed	Neville et al. (2013)	104	0.22
Parent-aimed	Neville et al. (2013)	104	0.40
Parent-aimed	Hadeed (2011)	106	0.26
Educational	*Catch Up® Numeracy* (2014)	108	0.21
Parent-aimed	Huebner (2000)	114	-0.33
Parent-aimed	Huebner (2000)	115	0.00
Educational	D. Torgerson, Torgerson, Jefferson, et al. (2014)	124	0.21
Educational	D. Torgerson, Torgerson, Jefferson, et al. (2014)	124	0.24
Parent-aimed	Chacko et al. (2018)	126	0.33
Parent-aimed	Chacko et al. (2018)	126	0.39
Educational	Lord et al. (2015)	149	0.36
Educational	Styles et al. (2014)	150	-0.01
Parent-aimed	Hadeed (2011)	156	0.39

Table G.1 continued

Intervention type	Study	n	ES
Educational	Styles et al. (2014)	159	-0.14
Parent-aimed	Whitehurst et al. (1994)	167	0.08
Parent-aimed	Whitehurst et al. (1994)	167	-0.03
Parent-aimed	Whitehurst et al. (1994)	167	0.12
Parent-aimed	Whitehurst et al. (1994)	167	0.30
Parent-aimed	Whitehurst et al. (1994)	167	0.45
Parent-aimed	Whitehurst et al. (1994)	167	1.44
Parent-aimed	Whitehurst et al. (1994)	167	0.01
Parent-aimed	Whitehurst et al. (1994)	167	0.03
Parent-aimed	Whitehurst et al. (1994)	167	0.16
Parent-aimed	Whitehurst et al. (1994)	167	0.55
Parent-aimed	Whitehurst et al. (1994)	167	0.35
Educational	Styles, Clarkson, et al. (n.d.)	175	0.03
Educational	King and Kasim (2015)	178	-0.05
Parent-aimed	Rozek et al. (2017)	181	0.16
Parent-aimed	Rozek et al. (2017)	181	0.19
Educational	Maxwell, Connolly, Demack, O'Hare, Stevens, Clague, and Stiell (2014)	182	0.13
Educational	Sibieta (2016)	202	0.51
Educational	Sibieta (2016)	202	0.33
Educational	Styles and Bradshaw (2015)	213	0.20
Educational	Sibieta et al. (2016)	229	0.27
Educational	Sibieta et al. (2016)	236	0.16
Educational	Gorard, Siddiqui, and See (2014b)	254	-0.09

Table G.1 continued

Intervention type	Study	n	ES
Parent-aimed	Bergman (2015)	256	0.21
Parent-aimed	Bergman (2015)	257	0.01
Educational	D. Torgerson, Torgerson, Ainsworth, et al. (2014)	261	0.74
Parent-aimed	Bergman (2015)	279	0.16
Parent-aimed	Bergman (2015)	279	0.23
Educational	Styles and Bradshaw (2015)	282	0.06
Parent-aimed	Jakobsen and Andersen (2013)	284	0.36
Parent-aimed	Jakobsen and Andersen (2013)	284	0.74
Parent-aimed	Jakobsen and Andersen (2013)	284	0.86
Educational	Rienzo et al. (2015)	286	0.18
Educational	Rienzo et al. (2015)	286	0.10
Educational	Gorard, Siddiqui, and See (2014a)	306	0.00
Educational	Gorard, See, and Siddiqui (2014)	308	0.24
Educational	Gorard, Siddiqui, and See (2014a)	310	0.17
Educational	Merrell and Kasim (2015)	310	0.43
Educational	Menzies, Kasim, et al. (2016)	314	0.03
Educational	Torgesen et al. (2007)	329	0.36
Educational	Torgesen et al. (2007)	329	0.26
Educational	Torgesen et al. (2007)	329	0.15
Educational	Torgesen et al. (2007)	329	0.14
Educational	Torgesen et al. (2007)	329	0.14
Educational	Torgesen et al. (2007)	329	0.11
Educational	Torgesen et al. (2007)	329	0.06
Educational	Gorard et al. (2015a)	349	0.24

Table G.1 continued

Intervention type	Study	n	ES
Educational	Worth et al. (2018)	362	-0.06
Educational	Gorard, Siddiqui, and See (2014b)	385	0.19
Educational	Maxwell, Connolly, Demack, O'Hare, Stevens, and Clague (2014)	391	-0.06
Parent-aimed	York et al. (2019)	395	0.06
Educational	Torgesen et al. (2007)	400	0.18
Educational	Torgesen et al. (2007)	400	0.11
Educational	Torgesen et al. (2007)	400	0.09
Educational	Torgesen et al. (2007)	400	0.05
Educational	Torgesen et al. (2007)	400	-0.04
Educational	Torgesen et al. (2007)	400	-0.08
Educational	Torgesen et al. (2007)	400	-0.08
Educational	Gorard et al. (2015b)	419	0.24
Parent-aimed	York et al. (2019)	426	0.15
Educational	Sheard et al. (2015)	427	-0.08
Educational	Heppen et al. (2011)	440	0.40
Educational	Dorsett et al. (2014)	492	-0.14
Educational	Nunes et al. (2018)	532	0.18
Educational	Rutt et al. (2015)	557	0.12
Parent-aimed	Kraft and Rogers (2015)	576	0.07
Parent-aimed	Doss et al. (2019)	578	0.01
Parent-aimed	Doss et al. (2019)	578	0.18
Educational	C. Torgerson et al. (2016)	578	-0.03
Parent-aimed	Frank (2016)	587	0.07

Table G.1 continued

Intervention type	Study	n	ES
Parent-aimed	Knight et al. (2019)	623	-0.33
Parent-aimed	Knight et al. (2019)	623	-0.16
Parent-aimed	Knight et al. (2019)	623	0.03
Educational	Black et al. (2009)	626	0.14
Educational	Black et al. (2009)	626	0.09
Educational	Black et al. (2009)	626	0.01
Educational	Robinson-Smith et al. (2018)	628	0.10
Educational	D. Torgerson, Torgerson, Mitchell, et al. (2014)	631	0.21
Educational	Tracey et al. (2016)	659	0.05
Educational	Tracey et al. (2016)	660	0.03
Educational	Buchanan, Worth, et al. (2015)	676	-0.04
Educational	Tracey et al. (2016)	679	0.08
Educational	Menzies, Kasim, et al. (2016)	699	0.10
Parent-aimed	Bergman et al. (2018)	701	0.01
Parent-aimed	Bergman et al. (2018)	701	0.13
Parent-aimed	Bergman et al. (2018)	705	-0.06
Parent-aimed	Bergman et al. (2018)	705	0.12
Educational	Buchanan, Worth, et al. (2015)	775	-0.08
Parent-aimed	Kraft and Monti-Nussbaum (2017)	779	0.13
Parent-aimed	Kraft and Monti-Nussbaum (2017)	779	0.15
Educational	Black et al. (2009)	792	0.09
Educational	Haywood et al. (2015)	814	0.03
Educational	Haywood et al. (2015)	814	0.00
Educational	D. Torgerson, Torgerson, Mitchell, et al. (2014)	817	0.24

Table G.1 continued

Intervention type	Study	n	ES
Parent-aimed	York et al. (2019)	821	0.11
Parent-aimed	Kraft and Monti-Nussbaum (2017)	896	0.60
Parent-aimed	Kraft and Monti-Nussbaum (2017)	896	0.60
Educational	Patel et al. (2017)	902	0.00
Educational	Black et al. (2009)	905	0.07
Educational	Black et al. (2009)	905	0.05
Educational	Black et al. (2009)	905	-0.08
Educational	McNally (2014)	924	0.03
Parent-aimed	Bergman and Chan (2019)	925	-0.06
Parent-aimed	Bergman and Chan (2019)	927	0.00
Educational	Rolfhus et al. (2012)	994	0.34
Educational	Menzies, Kasim, et al. (2016)	1071	-0.10
Educational	Dynarski et al. (2017)	1074	-0.12
Educational	Dynarski et al. (2017)	1077	-0.09
Educational	Humphrey et al. (2015)	1117	0.03
Parent-aimed	Bergman et al. (2018)	1120	0.08
Parent-aimed	Bergman et al. (2018)	1120	0.13
Educational	McNally (2014)	1124	-0.02
Educational	Rudd et al. (2017)	1129	0.09
Educational	Humphrey et al. (2015)	1134	-0.03
Educational	Black et al. (2009)	1144	0.09
Educational	Jay et al. (2017)	1198	0.15
Educational	Glazerman et al. (2010)	1198	0.20
Educational	Jay et al. (2017)	1223	0.12

Table G.1 continued

Intervention type	Study	n	ES
Educational	Gorard et al. (2017)	1224	0.12
Educational	Gorard et al. (2017)	1231	0.15
Educational	Jay et al. (2017)	1239	0.09
Educational	Motteram et al. (2016)	1252	-0.15
Educational	Hanley et al. (2015)	1264	0.22
Educational	Husain et al. (2016)	1290	0.00
Educational	Goodson et al. (2010)	1296	0.14
Educational	Goodson et al. (2010)	1296	0.14
Educational	Goodson et al. (2010)	1296	0.11
Educational	Husain et al. (2016)	1312	0.01
Educational	Miller, Biggart, et al. (2017)	1322	0.07
Educational	Glazerman et al. (2010)	1347	0.11
Educational	Hitchcock et al. (2011)	1355	0.05
Educational	Lloyd, Edovald, Kiss, et al. (2015)	1366	-0.06
Educational	Lloyd, Edovald, Kiss, et al. (2015)	1370	-0.02
Educational	Husain et al. (2016)	1398	0.02
Educational	Husain et al. (2016)	1414	-0.04
Parent-aimed	Berlinski et al. (2016)	1439	0.07
Educational	Rienzo et al. (2015)	1505	-0.11
Educational	Rienzo et al. (2015)	1505	0.01
Educational	Gorard et al. (2015c)	1529	0.12
Educational	Gorard et al. (2015c)	1529	0.03
Educational	Gorard et al. (2015c)	1529	0.10
Educational	Miller, Biggart, et al. (2017)	1537	0.04

Table G.1 continued

Intervention type	Study	n	ES
Educational	Bernstein et al. (2009)	1558	-0.04
Educational	Bernstein et al. (2009)	1558	-0.05
Educational	Bernstein et al. (2009)	1558	-0.03
Educational	Bernstein et al. (2009)	1558	-0.01
Educational	Judkins et al. (2008)	1560	0.05
Educational	Motteram et al. (2016)	1570	0.30
Educational	Wilkins et al. (2012)	1571	0.02
Educational	Humphrey et al. (2015)	1582	-0.11
Educational	Humphrey et al. (2015)	1582	-0.03
Parent-aimed	Andersen and Nielsen (2016)	1587	0.10
Parent-aimed	Andersen and Nielsen (2016)	1587	0.20
Educational	Glazerman et al. (2010)	1690	0.01
Educational	Garet et al. (2016)	1697	-0.05
Educational	D. Torgerson, Torgerson, Mitchell, et al. (2014)	1772	0.06
Educational	Bernstein et al. (2009)	1830	-0.03
Educational	Bernstein et al. (2009)	1830	-0.03
Educational	James-Burdumy et al. (2010)	1833	-0.02
Educational	McNally et al. (2016)	1884	0.23
Educational	McNally et al. (2016)	1884	0.14
Educational	Cordray et al. (2012)	1914	0.07
Educational	Cordray et al. (2012)	1914	0.05
Educational	Glazerman et al. (2010)	1929	-0.10
Educational	Worth et al. (2015)	1940	-0.05
Educational	Worth et al. (2015)	1942	0.20

Table G.1 continued

Intervention type	Study	n	ES
Educational	D. Torgerson, Torgerson, Mitchell, et al. (2014)	1982	0.10
Educational	Husain et al. (2018)	1985	0.01
Educational	Lord, Rabiasz, Roy, et al. (2017)	2080	0.03
Educational	Biggart, Sloan, et al. (2015)	2083	-0.04
Educational	Lord, Rabiasz, Roy, et al. (2017)	2122	0.01
Educational	Garet et al. (2011)	2132	-0.01
Educational	Wiggins, Parrao, et al. (2017)	2166	0.00
Educational	Lord, Rabiasz, Roy, et al. (2017)	2174	0.00
Educational	Lord, Rabiasz, Roy, et al. (2017)	2203	-0.04
Educational	Somers et al. (2010)	2255	0.12
Educational	Lord, Rabiasz, Roy, et al. (2017)	2291	-0.02
Educational	James-Burdumy et al. (2010)	2302	0.06
Educational	Somers et al. (2010)	2329	0.05
Educational	Lord, Rabiasz, Roy, et al. (2017)	2337	0.00
Educational	Bos et al. (2012)	2373	0.05
Educational	Greaves et al. (2017)	2379	0.01
Educational	Lord, Rabiasz, Roy, et al. (2017)	2386	0.03
Educational	James-Burdumy et al. (2010)	2395	0.00
Educational	Drummond et al. (2011)	2407	0.03
Educational	Drummond et al. (2011)	2407	-0.04
Educational	Lord, Rabiasz, Roy, et al. (2017)	2448	-0.03
Educational	Wijekumar et al. (2009)	2456	0.02
Educational	Lord, Rabiasz, Roy, et al. (2017)	2474	0.00
Educational	Humphrey, Hennessey, et al. (n.d.)	2504	0.03

Table G.1 continued

Intervention type	Study	n	ES
Educational	James-Burdumy et al. (2010)	2590	-0.08
Educational	James-Burdumy et al. (2010)	2607	-0.04
Educational	Arens et al. (2012)	2612	-0.03
Educational	James-Burdumy et al. (2010)	2681	-0.01
Educational	Lloyd, Edovald, Morris, et al. (2015)	2683	0.02
Educational	Thurston (2016)	2696	-0.03
Educational	Judkins et al. (2008)	2704	0.05
Educational	Judkins et al. (2008)	2704	0.00
Educational	Judkins et al. (2008)	2704	-0.08
Educational	Judkins et al. (2008)	2704	-0.09
Educational	Judkins et al. (2008)	2704	-0.11
Educational	Judkins et al. (2008)	2704	-0.13
Educational	Menzies, Hewitt, et al. (2016)	2784	-0.13
Educational	Lloyd, Edovald, Morris, et al. (2015)	2786	0.01
Educational	Glazerman et al. (2013)	2827	0.04
Educational	Wiggins, Sawtell, et al. (2017)	2829	0.00
Educational	Wiggins, Sawtell, et al. (2017)	2837	0.00
Parent-aimed	Cortes et al. (2018)	2920	-0.02
Parent-aimed	Cortes et al. (2018)	2920	0.02
Educational	Wiggins, Sawtell, et al. (2017)	3013	-0.04
Educational	Abe et al. (2012)	3052	0.24
Educational	Wiggins, Sawtell, et al. (2017)	3127	-0.08
Educational	Gorard et al. (2016)	3170	-0.09
Educational	Gorard et al. (2016)	3170	-0.09

Table G.1 continued

Intervention type	Study	n	ES
Educational	Glazerman et al. (2013)	3261	0.01
Educational	Garet et al. (2008)	3266	0.03
Educational	Sloan et al. (2018)	3312	-0.02
Educational	Finkelstein et al. (2011)	3415	0.27
Educational	Garet et al. (2008)	3450	0.08
Educational	Bos et al. (2012)	3456	0.03
Educational	Garet et al. (2016)	3677	-0.06
Educational	Glazerman et al. (2013)	3751	0.18
Educational	Finkelstein et al. (2011)	3752	0.32
Educational	Heller (2012)	3768	0.03
Educational	Glazerman et al. (2013)	3804	0.10
Educational	Jerrim et al. (2016)	3865	0.01
Educational	Crawford et al. (2019)	3907	-0.03
Educational	Crawford et al. (2019)	3907	0.01
Educational	Coe et al. (2011)	3931	0.11
Educational	Clark et al. (2013)	4116	0.00
Educational	Vignoles et al. (2015)	4176	0.10
Educational	Hanson et al. (2012)	4376	0.08
Educational	Crawford and Skipp (2014)	4413	0.09
Educational	Hanson et al. (2012)	4525	0.09
Educational	Corrin et al. (2012)	4546	0.10
Educational	Clark et al. (2013)	4573	0.07
Educational	Crawford et al. (2019)	4586	0.12
Educational	Crawford et al. (2019)	4586	0.16

Table G.1 continued

Intervention type	Study	n	ES
Educational	Crawford et al. (2019)	4586	0.14
Educational	Miller, Davison, et al. (2017)	4726	-0.01
Educational	Kushman et al. (2011)	4959	0.05
Educational	Corrin et al. (2012)	5011	0.06
Educational	Heller (2012)	5130	0.11
Educational	Miller, Davison, et al. (2017)	5376	0.03
Educational	Rose et al. (2017)	5462	0.02
Educational	Miller, Davison, et al. (2017)	5613	0.07
Educational	Martin et al. (2012)	5677	0.02
Educational	Cavalluzzo et al. (2012)	5863	-0.15
Educational	Hanley et al. (2016)	5882	-0.02
Educational	Jerrim et al. (2015)	5938	0.06
Parent-aimed	Bergman and Rogers (2017)	6291	0.05
Educational	Murphy et al. (2017)	6304	0.03
Educational	Murphy et al. (2017)	6437	0.02
Educational	West et al. (2017)	6596	0.00
Educational	Worth et al. (2017)	7366	-0.01
Educational	Newman et al. (2012)	7528	0.05
Educational	Bos et al. (2012)	7699	0.03
Educational	Sibieta et al. (2014)	7730	0.02
Educational	Sibieta et al. (2014)	7730	0.04
Educational	Sibieta et al. (2014)	7980	0.04
Educational	Sibieta et al. (2014)	7980	0.08
Educational	Bos et al. (2012)	8098	0.01

Table G.1 continued

Intervention type	Study	n	ES
Educational	Wilkerson et al. (2012)	8182	-0.01
Educational	Wilkerson et al. (2012)	8213	-0.06
Educational	Randel et al. (2011)	9596	0.01
Educational	Institute for Effective Education (2016)	10385	-0.04
Educational	Institute for Effective Education (2016)	10449	-0.02
Educational	West et al. (2017)	13131	-0.01
Educational	Bos et al. (2012)	17837	-0.01
Educational	Bos et al. (2012)	18180	0.01
Educational	Newman et al. (2012)	18713	0.05
Educational	Speckesser et al. (2018)	25393	0.10
Parent-aimed	Rogers and Feller (2018)	28080	-0.02
Parent-aimed	Rogers and Feller (2018)	28080	0.00
Educational	Garet et al. (2017)	28492	0.01
Educational	Garet et al. (2017)	29385	0.02
Educational	Garet et al. (2017)	29874	0.05
Educational	Garet et al. (2017)	29995	0.06
Educational	Wellington et al. (2016)	39807	0.04
Educational	Wellington et al. (2016)	40037	0.05
Educational	Wellington et al. (2016)	40390	0.03
Educational	Wellington et al. (2016)	40571	0.03
Educational	Wellington et al. (2016)	40708	0.04
Educational	Wellington et al. (2016)	40847	0.02
Educational	Buchanan, Morrison, et al. (2015)	63379	-0.16
Educational	Lord, Rabiasz, and Styles (2017)	86155	0.01

Table G.1 continued

Intervention type	Study	n	ES
Educational	Lord, Rabiasz, and Styles (2017)	86742	0.01
Educational	Lord, Rabiasz, and Styles (2017)	87701	0.01
Educational	Lord, Rabiasz, and Styles (2017)	88088	0.01
Educational	Buchanan, Morrison, et al. (2015)	101772	0.05
Educational	Constantine et al. (2009)	N/A	-0.01
Educational	Constantine et al. (2009)	N/A	-0.05
Educational	Dynarski et al. (2007)	N/A	0.09
Educational	Dynarski et al. (2007)	N/A	0.01
Educational	Dynarski et al. (2007)	N/A	0.02
Educational	Dynarski et al. (2007)	N/A	0.02
Educational	Dynarski et al. (2007)	N/A	-0.01
Educational	Dynarski et al. (2007)	N/A	0.09
Educational	Dynarski et al. (2007)	N/A	0.11
Educational	Dynarski et al. (2007)	N/A	-0.03
Educational	Dynarski et al. (2007)	N/A	-0.06
Educational	Dynarski et al. (2007)	N/A	0.00
Educational	Gleason et al. (2010)	N/A	-0.07
Educational	Gleason et al. (2010)	N/A	-0.06
Educational	Agodini et al. (2010)	N/A	0.09
Educational	Agodini et al. (2010)	N/A	0.00
Educational	Agodini et al. (2010)	N/A	0.12
Educational	Agodini et al. (2010)	N/A	0.11
Educational	Agodini et al. (2010)	N/A	0.17
Educational	Agodini et al. (2010)	N/A	0.07

Table G.1 continued

Intervention type	Study	n	ES
Educational	Chiang et al. (2017)	N/A	0.39
Educational	Chiang et al. (2017)	N/A	0.36
Educational	Chiang et al. (2017)	N/A	0.27
Educational	Chiang et al. (2017)	N/A	0.04

References for studies in Table G.1

Abe, Y., Thomas, V., Sinicrope, C., & Gee, K. A. (2012). *Effects of the Pacific CHILD Professional Development Program. Final Report. NCEE 2013-4002* [Publication Title: National Center for Education Evaluation and Regional Assistance]. National Center for Education Evaluation; Regional Assistance. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED538073

Agodini, R., Harris, B., Thomas, M., Murphy, R., & Gallagher, L. (2010). *Achievement Effects of Four Early Elementary School Math Curricula: Findings for First and Second Graders. NCEE 2011-4001* [Publication Title: National Center for Education Evaluation and Regional Assistance]. National Center for Education Evaluation; Regional Assistance. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED512551

Andersen, S. C., & Nielsen, H. S. (2016). Reading intervention with a growth mindset approach improves children's skills [Publisher: National Academy of Sciences _eprint: https://www.pnas.org/content/113/43/12111.full.pdf]. *Proceedings of the National Academy of Sciences*, *113*(43), 12111–12113. https://doi.org/10.1073/pnas.1607946113

Arens, S. A., Stoker, G., Barker, J., Shebby, S., Wang, X., Cicchinelli, L. F., & Williams, J. M. (2012). *Effects of Curriculum and Teacher Professional Development on the Language Proficiency of Elementary English Language Learner Students in the Central Region. Final Report. NCEE 2012-4013* [Publication Title: National Center for Education

Evaluation and Regional Assistance]. National Center for Education Evaluation; Regional Assistance. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED530839

Bergman, P. (2015). *Parent-Child Information Frictions and Human Capital Investment: Evidence from a Field Experiment* (SSRN Scholarly Paper No. ID 2622034). Social Science Research Network. Rochester, NY. Retrieved March 21, 2020, from https://papers.ssrn.com/abstract=2622034

Bergman, P., & Chan, E. W. (2019). Leveraging Parents through Low-Cost Technology: The Impact of High-Frequency Information on Student Achievement [Publisher: University of Wisconsin Press]. *Journal of Human Resources*, 1118–9837R1.

Bergman, P., Edmond-Verley, C., & Notario-Risk, N. (2018). Parent skills and information asymmetries: Experimental evidence from home visits and text messages in middle and high schools [Publisher: Elsevier]. *Economics of Education Review*, *66*, 92–103.

Bergman, P., & Rogers, T. (2017). The impact of defaults on technology adoption, and its underappreciation by policymakers [Publisher: CESifo Working Paper Series].

Berlinski, S., Busso, M., Dinkelman, T., & Martinez, C. (2016). Reducing parent-school information gaps and improving education outcomes: Evidence from high frequency text messaging in Chile. *Unpublished Manuscript*.

Bernstein, L., Rappaport, C. D., Olsho, L., Hunt, D., & Levin, M. (2009). *Impact Evaluation of the U.S. Department of Education's Student Mentoring Program. Final Report. NCEE 2009-4047* [Publication Title: National Center for Education Evaluation and Regional Assistance]. National Center for Education Evaluation; Regional Assistance. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED504310

Biggart, A., O'Hare, L., & Miller, S. (2015). *Tutoring with Alphie: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581542

Biggart, A., Sloan, S., O'Hare, L., & Miller, S. (2015). Quest: Evaluation report and executive summary. https://doi.org/10.13140/RG.2.1.1215.1287

Black, A. R., Somers, M.-A., Doolittle, F., Unterman, R., & Grossman, J. B. (2009). *The Evaluation of Enhanced Academic Instruction in After-School Programs: Final Report. NCEE 2009-4077* [Publication Title: National Center for Education Evaluation and Regional Assistance]. National Center for Education Evaluation; Regional Assistance. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED506725

Bos, J. M., Sanchez, R. C., Tseng, F., Rayyes, N., Ortiz, L., & Sinicrope, C. (2012). *Evaluation of Quality Teaching for English Learners (QTEL) Professional Development. Final Report. NCEE 2012-4005* [Publication Title: National Center for Education Evaluation and Regional Assistance]. National Center for Education Evaluation; Regional Assistance. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED529843

Buchanan, E., Morrison, J., Walker, M., Aston, H., & Cook, R. (2015). *Tutor Trust Secondary: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581539

Buchanan, E., Worth, J., & Aston, H. (2015). *Tutor Trust Primary: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581540

Catch Up® Numeracy: Evaluation Report and Executive Summary [ISSN: ISSN- Publication Title: Education Endowment Foundation]. (2014). Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581123

Cavalluzzo, L., Lowther, D. L., Mokher, C., & Fan, X. (2012). *Effects of the Kentucky Virtual Schools' Hybrid Program for Algebra I on Grade 9 Student Math Achievement. Final Report. NCEE 2012-4020* [Publication Title: National Center for Education Evalu-

ation and Regional Assistance]. National Center for Education Evaluation; Regional Assistance. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED530803

Chacko, A., Fabiano, G. A., Doctoroff, G. L., & Fortson, B. (2018). Engaging fathers in effective parenting for preschool children using shared book reading: A randomized controlled trial [Publisher: Taylor & Francis]. *Journal of Clinical Child & Adolescent Psychology*, *47*(1), 79–93.

Chiang, H., Speroni, C., Herrmann, M., Hallgren, K., Burkander, P., & Wellington, A. (2017). *Evaluation of the Teacher Incentive Fund: Final Report on Implementation and Impacts of Pay-for-Performance across Four Years. NCEE 2018-4004* [Publication Title: National Center for Education Evaluation and Regional Assistance]. National Center for Education Evaluation; Regional Assistance. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED578857

Chow, B. W.-Y., McBride-Chang, C., Cheung, H., & Chow, C. S.-L. (2008). Dialogic reading and morphology training in Chinese children: Effects on language and literacy. [Publisher: American Psychological Association]. *Developmental Psychology*, *44*(1), 233.

Clark, M. A., Chiang, H. S., Silva, T., McConnell, S., Sonnenfeld, K., Erbe, A., & Puma, M. (2013). *The Effectiveness of Secondary Math Teachers from Teach For America and the Teaching Fellows Programs. NCEE 2013-4015* [Publication Title: National Center for Education Evaluation and Regional Assistance]. National Center for Education Evaluation; Regional Assistance. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED544171

Coe, M., Hanita, M., Nishioka, V., & Smiley, R. (2011). *An Investigation of the Impact of the 6+1 Trait Writing Model on Grade 5 Student Writing Achievement. Final Report. NCEE 2012-4010* [Publication Title: National Center for Education Evaluation and Regional Assistance]. National Center for Education Evaluation; Regional Assistance. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED527445

Constantine, J., Player, D., Silva, T., Hallgren, K., Grider, M., & Deke, J. (2009). *An Evaluation of Teachers Trained through Different Routes to Certification. Final Report. NCEE 2009-4043* [Publication Title: National Center for Education Evaluation and Regional Assistance]. National Center for Education Evaluation; Regional Assistance. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED504313

Cordray, D., Pion, G., Brandt, C., Molefe, A., & Toby, M. (2012). *The Impact of the Measures of Academic Progress (MAP) Program on Student Reading Achievement. Final Report. NCEE 2013-4000* [Publication Title: National Center for Education Evaluation and Regional Assistance]. National Center for Education Evaluation; Regional Assistance. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED537982

Corrin, W., Lindsay, J. J., Somers, M.-A., Myers, N. E., Meyers, C. V., Condon, C. A., & Smith, J. K. (2012). *Evaluation of the Content Literacy Continuum: Report on Program Impacts, Program Fidelity, and Contrast. Final Report. NCEE 2013-4001* [Publication Title: National Center for Education Evaluation and Regional Assistance]. National Center for Education Evaluation; Regional Assistance. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED538060

Cortes, K. E., Fricke, H., Loeb, S., Song, D. S., & York, B. N. (2018). Too little or too much? Actionable Advice in an Early-Childhood Text Messaging Experiment [Publisher: MIT Press]. *Education Finance and Policy*, 1–44.

Crawford, C., Edwards, A., Farquharson, C., Greaves, E., Trevelyan, G., Wallace, E., & White, C. (2019). *Magic breakfast: Evaluation report and executive summary*. Education Endowment Foundation.

Crawford, C., & Skipp, A. (2014). *LIT Programme: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581139

Dorsett, R., Rienzo, C., Rolfe, H., Burns, H., Robertson, B.-A., Thorpe, B., & Wall, K. (2014). *Mind the Gap: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581160

Doss, C., Fahle, E. M., Loeb, S., & York, B. N. (2019). More Than Just a Nudge Supporting Kindergarten Parents with Differentiated and Personalized Text Messages [Publisher: University of Wisconsin Press]. *Journal of Human Resources*, *54*(3), 567–603.

Drummond, K., Chinen, M., Duncan, T. G., Miller, H. R., Fryer, L., Zmach, C., & Culp, K. (2011). *Impact of the Thinking Reader[R] Software Program on Grade 6 Reading Vocabulary, Comprehension, Strategies, and Motivation: Final Report. NCEE 2010-4035* [Publication Title: National Center for Education Evaluation and Regional Assistance]. National Center for Education Evaluation; Regional Assistance. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED517968

Dynarski, M., Agodini, R., Heaviside, S., Novak, T., Carey, N., Campuzano, L., Means, B., Murphy, R., Penuel, W., Javitz, H., Et al. (2007). Effectiveness of reading and mathematics software products: Findings from the first student cohort.

Dynarski, M., Rui, N., Webber, A., & Gutmann, B. (2017). *Evaluation of the DC Opportunity Scholarship Program: Impacts after One Year. NCEE 2017-4022* [Publication Title: National Center for Education Evaluation and Regional Assistance]. National Center for Education Evaluation; Regional Assistance. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED573822

Fielding-Barnsley, R., & Purdie, N. (2003). Early intervention in the home for children at risk of reading failure [Publisher: Wiley Online Library]. *Support for Learning*, *18*(2), 77–82.

Finkelstein, N., Hanson, T., Huang, C.-W., Hirschman, B., & Huang, M. (2011). *Effects of Problem Based Economics on High School Economics Instruction. Final Report. NCEE 2010-4022rev* [Publication Title: National Center for Education Evaluation

and Regional Assistance]. National Center for Education Evaluation; Regional Assistance. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED533126

Frank, M. C. (2016). Comment on "Math at home adds up to achievement in school". *Science*, *351*(6278), 1161.2–1161. https://doi.org/10.1126/science.aad8008

Garet, M. S., Cronen, S., Eaton, M., Kurki, A., Ludwig, M., Jones, W., Uekawa, K., Falk, A., Bloom, H. S., Doolittle, F., Zhu, P., & Sztejnberg, L. (2008). *The Impact of Two Professional Development Interventions on Early Reading Instruction and Achievement. NCEE 2008-4030* [Publication Title: National Center for Education Evaluation and Regional Assistance]. National Center for Education Evaluation; Regional Assistance. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED502700

Garet, M. S., Heppen, J. B., Walters, K., Parkinson, J., Smith, T. M., Song, M., Garrett, R., Yang, R., & Borman, G. D. (2016). *Focusing on Mathematical Knowledge: The Impact of Content-Intensive Teacher Professional Development. NCEE 2016-4010* [Publication Title: National Center for Education Evaluation and Regional Assistance]. National Center for Education Evaluation; Regional Assistance. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED569154

Garet, M. S., Wayne, A. J., Brown, S., Rickles, J., Song, M., & Manzeske, D. (2017). *The Impact of Providing Performance Feedback to Teachers and Principals. NCEE 2018-4001* [Publication Title: National Center for Education Evaluation and Regional Assistance]. National Center for Education Evaluation; Regional Assistance. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED578873

Garet, M. S., Wayne, A. J., Stancavage, F., Taylor, J., Eaton, M., Walters, K., Song, M., Brown, S., Hurlburt, S., Zhu, P., Sepanik, S., & Doolittle, F. (2011). *Middle School Mathematics Professional Development Impact Study: Findings after the Second Year of Implementation. NCEE 2011-4024* [Publication Title: National Center for Education Evaluation and Regional Assistance]. National Center for Education Evaluation;

Regional Assistance. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED519922

Glazerman, S., Isenberg, E., Dolfin, S., Bleeker, M., Johnson, A., Grider, M., & Jacobus, M. (2010). *Impacts of Comprehensive Teacher Induction: Final Results from a Randomized Controlled Study. NCEE 2010-4027* [Publication Title: National Center for Education Evaluation and Regional Assistance]. National Center for Education Evaluation; Regional Assistance. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED565837

Glazerman, S., Protik, A., Teh, B.-r., Bruch, J., & Max, J. (2013). *Transfer Incentives for High-Performing Teachers: Final Results from a Multisite Randomized Experiment. NCEE 2014-4004* [ISSN: ISSN- Publication Title: National Center for Education Evaluation and Regional Assistance]. National Center for Education Evaluation; Regional Assistance. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED544269

Gleason, P., Clark, M., Tuttle, C. C., & Dwoyer, E. (2010). *The Evaluation of Charter School Impacts: Final Report. NCEE 2010-4029* [Publication Title: National Center for Education Evaluation and Regional Assistance]. National Center for Education Evaluation; Regional Assistance. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED510573

Goodson, B., Wolf, A., Bell, S., Turner, H., & Finney, P. B. (2010). *The Effectiveness of a Program to Accelerate Vocabulary Development in Kindergarten (VOCAB): Kindergarten Final Evaluation Report. NCEE 2010-4014* [Publication Title: National Center for Education Evaluation and Regional Assistance]. National Center for Education Evaluation; Regional Assistance. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED512900

Gorard, S., See, B. H., Siddiqui, N., Smith, E., & White, P. (2016). *Youth Social Action Trials: Youth United Evaluation Report and Executive Summary*. London. http://dro.dur.ac.uk/20113/1/20113.pdf

Gorard, S., See, B. H., & Siddiqui, N. (2014). *Switch-On Reading: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581466

Gorard, S., Siddiqui, N., & See, B. H. (2014a). *Future Foundations: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581128

Gorard, S., Siddiqui, N., & See, B. H. (2014b). *Response to Intervention: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581264

Gorard, S., Siddiqui, N., & See, B. H. (2015a). *Accelerated Reader: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581101

Gorard, S., Siddiqui, N., & See, B. H. (2015b). *Fresh Start: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581126

Gorard, S., Siddiqui, N., & See, B. H. (2015c). *Philosophy for Children: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581147

Gorard, S., Siddiqui, N., See, B. H., Smith, E., & White, P. (2017). *Children's University: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education

Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581159

Greaves, E., Sianesi, B., Sibieta, L., Amin-Smith, N., Callanan, M., & Hudson, R. (2017). *Achieve Together: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581115

Hadeed, J. (2011). The continued effects of home intervention on child development outcomes in the Kingdom of Bahrain [Publisher: Taylor & Francis]. *Early Child Development and Care*, *181*(10), 1291–1313.

Hanley, P., Bohnke, J., Slavin, B., Elliott, L., & Croudace, T. (2016). Let's Think Secondary Science: Evaluation report and executive summary [Library Catalog: eprints.hud.ac.uk Num Pages: 88 Place: London Publisher: Education Endowment Foundation]. Retrieved March 21, 2020, from http://eprints.hud.ac.uk/id/eprint/29811/1/EEF_Project_Report_Lets_Think_Secondary_Science.pdf

Hanley, P., Slavin, R., & Elliott, L. (2015). *Thinking, Doing, Talking Science: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581125

Hanson, T., Dietsch, B., & Zheng, H. (2012). *Lessons in Character Impact Evaluation. Final Report. NCEE 2012-4004* [Publication Title: National Center for Education Evaluation and Regional Assistance]. National Center for Education Evaluation; Regional Assistance. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED530370

Haywood, S., Griggs, J., Lloyd, C., Morris, S., Kiss, Z., & Skipp, A. (2015). *Creative Futures: Act, Sing, Play. Evaluation report and executive summary* (Report). NatCen Social Research. Retrieved March 21, 2020, from https://educationendowmentfoundation.org.uk/

Heller, J. I. (2012). *Effects of Making Sense of SCIENCE[TM] Professional Development on the Achievement of Middle School Students, Including English Language Learners. Final Report. NCEE 2012-4002* [Publication Title: National Center for Education Evaluation and Regional Assistance]. National Center for Education Evaluation; Regional Assistance. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED530414

Heppen, J. B., Walters, K., Clements, M., Faria, A.-M., Tobey, C., Sorensen, N., & Culp, K. (2011). *Access to Algebra I: The Effects of Online Mathematics for Grade 8 Students. NCEE 2012-4021* [Publication Title: National Center for Education Evaluation and Regional Assistance]. National Center for Education Evaluation; Regional Assistance. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED527394

Hitchcock, J., Dimino, J., Kurki, A., Wilkins, C., & Gersten, R. (2011). *The Impact of Collaborative Strategic Reading on the Reading Comprehension of Grade 5 Students in Linguistically Diverse Schools. Final Report. NCEE 2011-4001* [Publication Title: National Center for Education Evaluation and Regional Assistance]. National Center for Education Evaluation; Regional Assistance. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED517770

Huebner, C. E. (2000). Promoting toddlers' language development through community-based intervention [Publisher: Elsevier]. *Journal of Applied Developmental Psychology*, *21*(5), 513–535.

Humphrey, N., Hennessey, A., Ashworth, E., Frearson, K., Black, L., Petersen, K., Wo, L., Panayiotou, M., Lendrum, A., Wigelsworth, M., Birchinall, L., Squires, G., & Pampaka, M. (n.d.). *Good Behaviour Game Evaluation Report and Executive Summary*. London, Education Endowment Foundation. https://pdfs.semanticscholar.org/b2cb/43b8b8f1de6c66f81f3468820f69b78fcd41.pdf

Humphrey, N., Barlow, A., Wigelsworth, M., Lendrum, A., Pert, K., Joyce, C., Stephens, E., Wo, L., Squires, G., Woods, K., Calam, R., Harrison, M., Turner, A., & Humphrey, N. (2015). *Promoting Alternative Thinking Strategies (PATHS): Evaluation Report and*

Executive Summary [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581278

Husain, F., Wishart, R., Marshall, L., Frankenberg, S., Bussard, L., Chidley, S., Hudson, R., Votjkova, M., & Morris, S. (2018). *Family Skills Evaluation report and executive summary, Report to the Education Endowment Foundation* (Report). Education Endowment Foundation. Retrieved March 21, 2020, from https://e-space.mmu.ac.uk/620513/1/Family_Skills.pdf

Husain, F., Jabin, N., Haywood, S., Kasim, A., & Paylor, J. (2016). *Parent Academy: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://files.eric.ed.gov/fulltext/ED581143.pdf

Institute for Effective Education. (2016). *Teacher Effectiveness Enhancement Programme: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://files.eric.ed.gov/fulltext/ED581110.pdf

Jakobsen, M., & Andersen, S. C. (2013). Coproduction and equity in public service delivery [Publisher: Wiley Online Library]. *Public Administration Review*, *73*(5), 704–713.

James-Burdumy, S., Deke, J., Lugo-Gil, J., Carey, N., Hershey, A., Gersten, R., Newman-Gonchar, R., Dimino, J., Haymond, K., & Faddis, B. (2010). *Effectiveness of Selected Supplemental Reading Comprehension Interventions: Findings from Two Student Cohorts. NCEE 2010-4015* [Publication Title: National Center for Education Evaluation and Regional Assistance]. National Center for Education Evaluation; Regional Assistance. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED579176

Jay, T., Willis, B., Thomas, P., Taylor, R., Moore, N., Burnett, C., Merchant, G., & Stevens, A. (2017). Dialogic Teaching : Evaluation Report and Executive Summary [Library Catalog: shura.shu.ac.uk Num Pages: 77 Place: London, UK Publisher: Education En-

dowment Foundation]. Retrieved March 21, 2020, from https://educationendowmentfoundation.org.uk/public/files/Projects/Evaluation_Reports/Dialogic_Teaching_Evaluation_Report.pdf

Jerrim, J., Austerberry, H., Crisan, C., Ingold, A., Morgan, C., Pratt, D., Smith, C., & Wiggins, M. (2015). *Mathematics Mastery: Secondary Evaluation Report* [ISSN: ISSN-Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581187

Jerrim, J., Macmillan, L., Micklewright, J., Sawtell, M., & Wiggins, M. (2016). *Chess in Schools: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581100

Judkins, D., St. Pierre, R., Gutmann, B., Goodson, B., von Glatz, A., Hamilton, J., Webber, A., Troppe, P., & Rimdzius, T. (2008). *A Study of Classroom Literacy Interventions and Outcomes in Even Start. NCEE 2008-4028* [Publication Title: National Center for Education Evaluation and Regional Assistance]. National Center for Education Evaluation; Regional Assistance. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED502817

King, B., & Kasim, A. (2015). *Rapid Phonics: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581293

Knight, D. S., Landry, S., Zucker, T. A., Merz, E. C., Guttentag, C. L., & Taylor, H. B. (2019). Cost-Effectiveness of Early Childhood Interventions to Enhance Preschool: Evidence from a Randomized Experiment in Head Start Centers Enrolling Historically Underserved Populations [Publisher: Wiley Online Library]. *Journal of Policy Analysis and Management*, *38*(4), 891–917.

Kotaman, H. (2013). Impacts of dialogical storybook reading on young children's reading attitudes and vocabulary development [Publisher: Project Innovation]. *Reading Improvement*, *50*(4), 199–204.

Kraft, M. A., & Monti-Nussbaum, M. (2017). Can schools enable parents to prevent summer learning loss? A text-messaging field experiment to promote literacy skills [Publisher: SAGE Publications Sage CA: Los Angeles, CA]. *The ANNALS of the American Academy of Political and Social Science*, *674*(1), 85–112.

Kraft, M. A., & Rogers, T. (2015). The underutilized potential of teacher-to-parent communication: Evidence from a field experiment [Publisher: Elsevier]. *Economics of Education Review*, *47*, 49–63.

Kushman, J., Hanita, M., & Raphael, J. (2011). *An Experimental Study of the Project CRISS Reading Program on Grade 9 Reading Achievement in Rural High Schools. Final Report NCEE 2011-4007* [Publication Title: National Center for Education Evaluation and Regional Assistance]. National Center for Education Evaluation; Regional Assistance. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED518226

Levin, I., & Aram, D. (2012). Mother–child joint writing and storybook reading and their effects on kindergartners' literacy: An intervention study [Publisher: Springer]. *Reading and Writing*, *25*(1), 217–249.

Lloyd, C., Edovald, T., Morris, S. P., Skipp, A., KIss, Z., & Haywood, S. (2015). *Durham Shared Maths Project. Evaluation report and executive summary* (Report). Education Endowment Foundation. Retrieved March 21, 2020, from https://educationendowmentfoundation.org.uk/

Lloyd, C., Edovald, T., Kiss, Z., Morris, S., Skipp, A., & Ahmed, H. (2015). *Paired Reading: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581127

Lonigan, C. J., & Whitehurst, G. J. (1998). Relative efficacy of parent and teacher involvement in a shared-reading intervention for preschool children from low-income backgrounds. *Early Childhood Research Quarterly*, *13*(2), 263–290.

Lord, P., Bradshaw, S., Stevens, E., & Styles, B. (2015). *Perry Beeches Coaching Programme: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581144

Lord, P., Rabiasz, A., Roy, P., Harland, J., Styles, B., & Fowler, K. (2017). *Evidence-Based Literacy Support: The "Literacy Octopus" Trial. Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581248

Lord, P., Rabiasz, A., & Styles, B. (2017). *'Literacy Octopus' Dissemination Trial: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581230

Martin, T., Brasiel, S. J., Turner, H., & Wise, J. C. (2012). *Effects of the Connected Mathematics Project 2 (CMP2) on the Mathematics Achievement of Grade 6 Students in the Mid-Atlantic Region. Final Report. NCEE 2012-4017* [Publication Title: National Center for Education Evaluation and Regional Assistance]. National Center for Education Evaluation; Regional Assistance. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED530513

Maxwell, B., Connolly, P., Demack, S., O'Hare, L., Stevens, A., & Clague, L. (2014). *TextNow Transition Programme: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581131

Maxwell, B., Connolly, P., Demack, S., O'Hare, L., Stevens, A., Clague, L., & Stiell, B. (2014). Summer Active Reading Programme : Evaluation report and executive summary [Library Catalog: shura.shu.ac.uk Num Pages: 63 Place: London, UK Publisher: Education Endowment Foundation]. Retrieved March 21, 2020, from https://educationendowmentfoundation.org.uk/uploads/pdf/EEF_Evaluation_Report_-_Summer_Active_Reading_Programme_-_October_2014.pdf

McNally, S. (2014). *Hampshire Hundreds: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581163

McNally, S., Ruiz-Valenzuela, J., & Rolfe, H. (2016). *ABRA: Online Reading Support. Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581099

Menzies, V., Kasim, A., Kokotsaki, D., Hewitt, C., Akhter, N., Collyer, C., Younger, K., Wiggins, A., & Torgerson, C. (2016). *Hallé SHINE on Manchester Evaluation Report and Executive Summary*. London, Education Endowment Foundation. http://dro.dur.ac.uk/19386/1/19386.pdf

Menzies, V., Hewitt, C., Kokotsaki, D., Collyer, C., & Wiggins, A. (2016). Project based learning: Evaluation report and executive summary.

Merrell, C., & Kasim, A. (2015). *Butterfly Phonics: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581118

Miller, S., Biggart, A., Sloan, S., & O'Hare, L. (2017). *Success for All: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581417

Miller, S., Davison, J., Yohanis, J., Sloan, S., Gildea, A., & Thurston, A. (2017). *Texting Parents: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581121

Motteram, G., Choudry, S., Kalambouka, A., Hutcheson, G., & Barton, H. (2016). *ReflectED: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581262

Murphy, R., Weinhardt, F., Wyness, G., & Rolfe, H. (2017). *Lesson Study: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581145

Murray, L., De Pascalis, L., Tomlinson, M., Vally, Z., Dadomo, H., MacLachlan, B., Woodward, C., & Cooper, P. J. (2016). Randomized controlled trial of a book-sharing intervention in a deprived South African community: Effects on carer–infant interactions, and their relation to infant cognitive and socioemotional outcome [Publisher: Wiley Online Library]. *Journal of Child Psychology and Psychiatry*, *57*(12), 1370–1379.

Neville, H. J., Stevens, C., Pakulak, E., Bell, T. A., Fanning, J., Klein, S., & Isbell, E. (2013). Family-based training program improves brain function, cognition, and behavior in lower socioeconomic status preschoolers [Publisher: National Acad Sciences]. *Proceedings of the National Academy of Sciences*, *110*(29), 12138–12143.

Newman, D., Finney, P. B., Bell, S., Turner, H., Jaciw, A. P., Zacamy, J. L., & Gould, L. F. (2012). *Evaluation of the Effectiveness of the Alabama Math, Science, and Technology Initiative (AMSTI). Final Report. NCEE 2012-4008* [Publication Title: National Center for Education Evaluation and Regional Assistance]. National Cen-

ter for Education Evaluation; Regional Assistance. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED529587

Nunes, T., Barros, R., Evangelou, M., Strand, S., Mathers, S., & Sanders-Ellis, D. (2018). *1stClass@Number Evaluation Report and Executive Summary*. https://educationendowmentfoundati org.uk/public/files/1stClass@Number_evaluation_report.pdf

Patel, R., Jabin, N., Bussard, L., Cartagena, J., Haywood, S., & Lumpkin, M. (2017). *Switch-on Effectiveness Trial Evaluation Report and Executive Summary*. London, Education Endowment Foundation. https://dera.ioe.ac.uk/29342/1/EEF_Project_Report_Switchon_Effectiveness-1.pdf

Randel, B., Beesley, A. D., Apthorp, H., Clark, T. F., Wang, X., Cicchinelli, L. F., & Williams, J. M. (2011). *Classroom Assessment for Student Learning: Impact on Elementary School Mathematics in the Central Region. Final Report. NCEE 2011-4005* [Publication Title: National Center for Education Evaluation and Regional Assistance]. National Center for Education Evaluation; Regional Assistance. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED517969

Reese, E., Leyva, D., Sparks, A., & Grolnick, W. (2010). Maternal elaborative reminiscing increases low-income children's narrative skills relative to dialogic reading [Publisher: Taylor & Francis]. *Early Education and Development*, *21*(3), 318–342.

Rienzo, C., Rolfe, H., & Wilkinson, D. (2015). *Changing Mindsets: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581132

Robinson-Smith, L., Fairhurst, C., Stone, G., Bell, K., Elliott, L., Gascoine, L., Hallett, S., Hewitt, C., Hugill, J., Torgerson, C., Torgerson, D., Menzies, V., & Ainsworth, H. (2018). *Maths Champions Evaluation Report and Executive Summary*. London, Education Endowment Foundation. https://educationendowmentfoundation.org.uk/

public/files/Projects/Evaluation_Reports/Maths_champions_evaluation_report.pdf

Rogers, T., & Feller, A. (2018). Reducing student absences at scale by targeting parents' misbeliefs [Publisher: Nature Publishing Group]. *Nature Human Behaviour*, *2*(5), 335–342.

Rolfhus, E., Gersten, R., Clarke, B., Decker, L. E., Wilkins, C., & Dimino, J. (2012). *An Evaluation of "Number Rockets": A Tier-2 Intervention for Grade 1 Students at Risk for Difficulties in Mathematics. Final Report. NCEE 2012-4007* [Publication Title: National Center for Education Evaluation and Regional Assistance]. National Center for Education Evaluation; Regional Assistance. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED529429

Rose, J., Thomas, S., Zhang, L., Edwards, A., Augero, A., & Roney, P. (2017). *Research Learning Communities: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581267

Rozek, C. S., Svoboda, R. C., Harackiewicz, J. M., Hulleman, C. S., & Hyde, J. S. (2017). Utility-value intervention with parents increases students' STEM preparation and career pursuit [Publisher: National Acad Sciences]. *Proceedings of the National Academy of Sciences*, *114*(5), 909–914.

Rudd, P., Aguilera, A. B. V., Elliott, L., & Chambers, B. (2017). *MathsFlip: Flipped Learning. Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581151

Rutt, S., Kettlewell, K., & Bernardinelli, D. (2015). *Catch Up® Literacy: Evaluation Report and Executive Summary* [Publication Title: National Foundation for Educational Research]. National Foundation for Educational Research. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED558735

Sheard, M., Chambers, B., & Elliott, L. (2015). *Units of Sound: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581097

Sheets, G. N., & Buyer, L. S. (1999). Examining the Effects of Shared Book Reading across Age-Groups. [Publisher: ERIC].

Sibieta, L. (2016). *REACH: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581275

Sibieta, L., Greaves, E., & Sianesi, B. (2014). *Increasing Pupil Motivation: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581249

Sibieta, L., Kotecha, M., & Skipp, A. (2016). *Nuffield Early Language Intervention: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581138

Sloan, S., Gildea, A., Miller, S., & Thurston, A. (2018). *Zippy's Friends Evaluation Report and Executive Summary*. London, Education Endowment Foundation. https://pdfs.semanticscholar.org/661e/98a44e41996a462a3624a7174186cada2393.pdf

Somers, M.-A., Corrin, W., Sepanik, S., Salinger, T., Levin, J., & Zmach, C. (2010). *The Enhanced Reading Opportunities Study Final Report: The Impact of Supplemental Literacy Courses for Struggling Ninth-Grade Readers. NCEE 2010-4021* [Publication Title: National Center for Education Evaluation and Regional Assistance]. National Center for Education Evaluation; Regional Assistance. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED511811

Speckesser, S., Runge, J., Foliano, F., Bursnall, M., Hudson-Sharp, N., Rolfe, H., & Anders, J. (2018). *Embedding Formative Assessment Evaluation Report and Executive Summary*. Education Endowment Foundation. https://dera.ioe.ac.uk/32012/1/EFA_evaluation_report.pdf

Styles, B., & Bradshaw, S. (2015). *Talk for Literacy: Evaluation Report and Executive Summary* [Publication Title: National Foundation for Educational Research]. National Foundation for Educational Research. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED558605

Styles, B., Clarkson, R., & Fowler, K. (2014). *Chatterbooks: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581155

Styles, B., Clarkson, R., & Fowler, K. (n.d.). Rhythm for Reading: Evaluation Report and Executive Summary, 38. https://www.nfer.ac.uk/publications/EFTR01/EFTR01.pdf

Thurston, A. (2016). *Talk of the Town: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581531

Torgerson, C., Ainsworth, H., Buckley, H., Hampden-Thompson, G., Hewitt, C., Humphry, D., Jefferson, L., Mitchell, N., & Torgerson, D. (2016). *Affordable Online Maths Tuition: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581116

Torgerson, D., Torgerson, C., Ainsworth, H., Buckley, H., Heaps, C., Hewitt, C., & Mitchell, N. (2014). *Improving Writing Quality: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education En-

dowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581140

Torgerson, D., Torgerson, C., Jefferson, L., Buckley, H., Ainsworth, H., Heaps, C., & Mitchell, N. (2014). *Discover Summer School: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581103

Torgerson, D., Torgerson, C., Mitchell, N., Buckley, H., Ainsworth, H., Heaps, C., & Jefferson, L. (2014). *Grammar for Writing: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581133

Torgesen, J., Schirm, A., Castner, L., Vartivarian, S., Mansfield, W., Myers, D., Stancavage, F., Durno, D., Javorsky, R., & Haan, C. (2007). *National Assessment of Title I. Final Report. Volume II: Closing the Reading Gap–Findings from a Randomized Trial of Four Reading Interventions for Striving Readers. NCEE 2008-4013* [Publication Title: National Center for Education Evaluation and Regional Assistance]. National Center for Education Evaluation; Regional Assistance. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED499018

Tracey, L., Chambers, B., Bywater, T., & Elliott, L. (2016). *SPOKES: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581435

Vaquero, J. (2014). *An Exploratory Study of a Shared-Book Reading Intervention Involving Spanish-Speaking Latino Families* (PhD Thesis).

Vignoles, A., Jerrim, J., & Cowan, R. (2015). *Mathematics Mastery: Primary Evaluation Report* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581183

Wellington, A., Chiang, H., Hallgren, K., Speroni, C., Herrmann, M., & Burkander, P. (2016). *Evaluation of the Teacher Incentive Fund: Implementation and Impacts of Pay-for-Performance after Three Years. NCEE 2016-4004* [Publication Title: National Center for Education Evaluation and Regional Assistance]. National Center for Education Evaluation; Regional Assistance. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED568150

West, M., Ainscow, M., Wigelsworth, M., & Troncoso, P. (2017). *Challenge the Gap: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581102

Whitehurst, G. J., Epstein, J. N., Angell, A. L., Payne, A. C., Crone, D. A., & Fischel, J. E. (1994). Outcomes of an emergent literacy intervention in Head Start. [Publisher: American Psychological Association]. *Journal of Educational Psychology*, *86*(4), 542.

Whitehurst, G. J., & Others, A. (1994). *A Picture Book Reading Intervention in Day Care and Home for Children from Low-Income Families* (tech. rep.) [ISSN: 0012-1649 Issue: 5 Num Pages: 679-89 Publication Title: Developmental Psychology Volume: 30]. ISSN: 0012-1649 Issue: 5 Num Pages: 679-89 Publication Title: Developmental Psychology Volume: 30. Retrieved March 21, 2020, from http://search.proquest.com/docview/62721951?rfr_id=info%3Axri%2Fsid%3Aprimo

Wiggins, M., Parrao, C. G., Austerberry, H., & Ingold, A. (2017). *Foreign Language Learning in Primary School: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581156

Wiggins, M., Sawtell, M., & Jerrim, J. (2017). *Learner Response System: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581136

Wijekumar, K., Hitchcock, J., Turner, H., Lei, P., & Peck, K. (2009). *A Multisite Cluster Randomized Trial of the Effects of CompassLearning Odyssey[R] Math on the Math Achievement of Selected Grade 4 Students in the Mid-Atlantic Region. Final Report. NCEE 2009-4068* [Publication Title: National Center for Education Evaluation and Regional Assistance]. National Center for Education Evaluation; Regional Assistance. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED507314

Wilkerson, S. B., Shannon, L. C., Styers, M. K., & Grant, B.-J. (2012). *A Study of the Effectiveness of a School Improvement Intervention (Success in Sight). Final Report. NCEE 2012-4014* [Publication Title: National Center for Education Evaluation and Regional Assistance]. National Center for Education Evaluation; Regional Assistance. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED530416

Wilkins, C., Gersten, R., Decker, L. E., Grunden, L., Brasiel, S., Brunnert, K., & Jayanthi, M. (2012). *Does a Summer Reading Program Based on Lexiles Affect Reading Comprehension? Final Report. NCEE 2012-4006* [Publication Title: National Center for Education Evaluation and Regional Assistance]. National Center for Education Evaluation; Regional Assistance. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED530029

Wing-Yin Chow, B., & McBride-Chang, C. (2003). Promoting language and literacy development through parent–child reading in Hong Kong preschoolers [Publisher: Taylor & Francis]. *Early Education and Development*, *14*(2), 233–248.

Woods, L. (2017). Interactive Book Reading: Promoting Emergent Literacy Skills in Preschool Children Through a Parent Training Program.

Worth, J., Nelson, J., Harland, J., Bernardinelli, D., & Styles, B. (2018). *GraphoGame Rime Evaluation Report and Executive Summary*. London, Education Endowment Foundation. https://nfer.ac.uk/media/2497/eerr01.pdf

Worth, J., Sizmur, J., Ager, R., & Styles, B. (2015). *Improving Numeracy and Literacy: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education

Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581142

Worth, J., Sizmur, J., Walker, M., Bradshaw, S., & Styles, B. (2017). *Teacher Observation: Evaluation Report and Executive Summary* [ISSN: ISSN- Publication Title: Education Endowment Foundation]. Education Endowment Foundation. Retrieved March 21, 2020, from https://eric.ed.gov/?id=ED581113

York, B. N., Loeb, S., & Doss, C. (2019). One step at a time the effects of an early literacy text-messaging program for parents of preschoolers [Publisher: University of Wisconsin Press]. *Journal of Human Resources*, *54*(3), 537–566.

Zevenbergen, A. A., Worth, S., Dretto, D., & Travers, K. (2018). Parents' experiences in a home-based dialogic reading programme [Publisher: Routledge _eprint: https://doi.org/10.1080/030 *Early Child Development and Care*, *188*(6), 862–874. https://doi.org/10.1080/03004430.2016.1241775