

DISCUSSION PAPER SERIES

DP14006

**SOCIAL ORIGINS, SHARED BOOK
READING AND LANGUAGE SKILLS IN
EARLY CHILDHOOD: EVIDENCE FROM
AN INFORMATION EXPERIMENT**

Carlo Barone, Denis Fougere and Clement Pin

LABOUR ECONOMICS



SOCIAL ORIGINS, SHARED BOOK READING AND LANGUAGE SKILLS IN EARLY CHILDHOOD: EVIDENCE FROM AN INFORMATION EXPERIMENT

Carlo Barone, Denis Fougere and Clement Pin

Discussion Paper DP14006
Published 17 September 2019
Submitted 06 September 2019

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programme in **LABOUR ECONOMICS**. Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as an educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: Carlo Barone, Denis Fougere and Clement Pin

SOCIAL ORIGINS, SHARED BOOK READING AND LANGUAGE SKILLS IN EARLY CHILDHOOD: EVIDENCE FROM AN INFORMATION EXPERIMENT

Abstract

Shared book reading between parents and children is often regarded as a significant mediator of social inequalities in early skill development processes. We argue that socially biased gaps between parents in access to information about the benefits of this activity for school success contribute to inequalities between children in access to this activity and in their language development. We test this hypothesis with a large-scale field experiment assessing the causal impact of an information intervention targeting parents of pre-schoolers on both the frequency of shared book reading and the receptive vocabulary of children. Results indicate that low-educated parents are more reactive to this information intervention, with significant effects on the language development of their children. We conclude that information barriers on the potential of informal learning activities at home contribute to social inequalities in early childhood, and that removing these barriers is a cost-effective way to reduce these inequalities.

JEL Classification: I21, I24, J13, C93

Keywords: Early Childhood, Language Skills, parental reading, field experiment

Carlo Barone - carlo.barone@sciencespo.fr
OSC & LIEPP (Sciences Po Paris)

Denis Fougere - denis.fougere@sciencespo.fr
CNRS, OSC and LIEPP (Sciences Po, Paris), IZA (Bonn) and CEPR

Clement Pin - clement.pin@sciencespo.fr
LIEPP (Sciences Po Paris)

Acknowledgements

This study benefited from funding from the LIEPP (Laboratory for the interdisciplinary evaluation of public policies). The publishing house Gallimard Jeunesse gave us all the books used in the first year of this study for free and Nathan did the same in the second year. We thank especially Nathalie Marec-Breton for her advice and training concerning the use of standardized tests. We thank Madame Gaga (Académie de Paris) for her generous support for the access to the field. The experimental protocol was registered in the Social Science Registry before running the experiment. The authors acknowledge support from the French National Research Agency's (ANR) "Investissements d'Avenir" grants ANR-11-LABX-0091 (LIEPP).

Social origins, shared book reading and language skills in early childhood: evidence from an information experiment

Authors: Carlo Barone (Sciences Po Paris), Denis Fougère (CNRS, Sciences Po Paris, CEPR and IZA), Clément Pin (Sciences Po Paris)¹

Keywords: early childhood, language skills, parental reading, field experiment

JEL Codes: I21, I24, J13, C93

Abstract

Shared book reading between parents and children is often regarded as a significant mediator of social inequalities in early skill development processes. We argue that socially biased gaps between parents in access to information about the benefits of this activity for school success contribute to inequalities between children in access to this activity and in their language development. We test this hypothesis with a large-scale field experiment assessing the causal impact of an information intervention targeting parents of pre-schoolers on both the frequency of shared book reading and the receptive vocabulary of children. Results indicate that low-educated parents are more reactive to this information intervention, with significant effects on the language development of their children. We conclude that information barriers on the potential of informal learning activities at home contribute to social inequalities in early childhood, and that removing these barriers is a cost-effective way to reduce these inequalities.

¹ This study benefited from funding from the LIEPP (Laboratory for the interdisciplinary evaluation of public policies). The publishing house Gallimard Jeunesse gave us all the books used in the first year of this study for free and Nathan did the same in the second year. We thank especially Nathalie Marec-Breton for her advice and training concerning the use of standardized tests. We thank Madame Gaga (Académie de Paris) for her generous support for the access to the field. The experimental protocol was registered in the Social Science Registry before running the experiment. The authors acknowledge support from the French National Research Agency's (ANR) "Investissements d'Avenir" grants ANR-11-LABX-0091 (LIEPP).

1. Introduction

Sociologists and economists have paid increasing attention to social inequalities in early skill development. They have consistently reported that children access the school system with unequal skill endowments, which strongly correlate with their social origins (Blossfeld et al., 2017). These skill gaps persist or even increase in primary and secondary education, thus driving inequalities in academic performance and educational attainment, employment and income (Dämmrich, Triventi, 2018; Heckman et al., 2013; Potter and Roksa, 2013; Klosterman et al., 2011). These early inequalities are particularly consequential for life chances: neuroscientific studies document that the human brain is highly malleable in the preschool years (Kolb et al., 2013) and educational research indicates that learning is a cumulative process, whereby early skills lay the foundations for developing more advanced skills (Doepke and Zibbotti, 2019; Klosterman et al., 2011).

The literature on emergent literacy indicates that early language skills (ELS) are particularly consequential for school success, since they are instrumental to learning processes across virtually all domains (Whitehurst and Lonigan, 1998; Hulme and Snowling, 2015; NELP, 2008). Moreover, these studies show that the development of reading and writing skills starts well before formal instruction: vocabulary, phonological awareness and knowledge of print, the main precursors of literacy skills (NELP, 2008), are acquired since the preschool years (*ibidem*). It is also well-documented that social origins have particularly strong impacts on these ELS (Lay See et al. 2014; Myrberg and Rosen, 2009). However, the mechanisms driving social inequalities in ELS have yet to be fully spelled out.

Parent-child interactions shape the development of these ELS. More educated parents spend more time with their children, particularly more time in cognitive-enhancing activities (Dotti-Sani and Treas, 2016). They actively foster their children's skills by incorporating organised learning activities

in their daily routines, as well as opportunities and materials for informal learning at home (Lareau, 2011; Oecd, 2012). Shared Book Reading (SBR) with children is perhaps the informal learning activity that has received most attention in education research. Family background is a strong predictor of the frequency of SBR, which in turn predicts student skills and academic performances: SBR is thus regarded as a potential mediator of social inequalities in ELS (Notten and Kraaykamp, 2013; Klosterman et al., 2011; Park, 2008).

SBR may be regarded as an indicator of parental cultural capital. In this cultural reproduction perspective (Bourdieu 1979), SBR is conceived as an activity that is deeply embedded in a broad constellation of attitudes and highbrow cultural practices characterising the *habitus* of the upper classes, and it is therefore hardly amenable to change. Research on cultural capital indicates that status-related signals of highbrow cultural participation of the parents indeed correlate with children's educational attainment (Andersen and Hansen, 2012). However, they are found to be less relevant to school success than activities of direct cognitive stimulation of children, such as SBR (Notten and Kraaykamp, 2013; De Graaf et al., 2000). To the extent that the latter foster ELS regardless of their high-status connotation, as reported by recent research (*ibidem*; Park, 2008), they may be accessible to and beneficial for low-educated families, despite their lower familiarity with highbrow culture.

This work presents the results of a field experiment developed to assess whether providing parents with information on the benefits of SBR for school success impacts on the frequency of this activity and on the language skills of children from high- and low-educated families. Our purpose is twofold: assessing the role of information barriers as a potential mechanism of social inequality in early skill development, and assessing whether information interventions promoting SBR can foster language skills and reduce the related social inequalities.

Field experiments are an increasingly popular method among researchers interested in the role of information barriers for educational attainment and the related social inequalities. However, most of these studies focus on access to Higher Education: a recent systematic review reports that information interventions alone tend to be ineffective at this level, if they are not complemented by activities fostering academic skills and/or providing direct support with application procedures (Herbaut and Geven, 2019). Removing information barriers may not be enough at a stage of educational careers where students differ considerably in their school performances and have been already tracked into different curricula. Instead information may be more consequential for school success at earlier stages, as suggested for instance by information experiments targeting secondary track choices (Keller, 2019; Barone et al., 2018). If preschool years are a critical period for skill development, it is worth assessing whether providing information on the potential of informal learning activities impacts on early skill development, as suggested by two recent field experiments on parenting (York and Loeb, 2018; Sylva et al., 2008).

2. Social inequalities in shared book reading: theories and hypotheses

The benefits of SBR for ELS are a matter of consensus among educational experts (Scarborough, Dobrich, 2012; Oecd, 2012). Storybooks are much richer in vocabulary than any competing activity in the home environment, and they mobilise different language registers than daily conversations (Dickinson et al., 2012). SBR is expected to foster not only children's vocabulary, but also their knowledge of print conventions and attention span, as well as more complex narrative and reasoning skills, mobilised by the child while listening to the story (Partridge 2004). Furthermore, SBR in early childhood correlates with reading habits and reading enjoyment in adolescence, which predict academic performance (Baker et al., 1997).

Unsurprisingly, books, newspaper articles and blogs on parenting systematically recommend SBR as a highly beneficial practice. Because highly educated parents read books and newspapers more often and are more interested in expert recommendations on parenting (Radey and Randolph, 2012), they enjoy facilitated access to information about the benefits of SBR for school success. Moreover, if this information circulates within friendship networks via informal conversations, the socio-economic segregation of these networks acts as a multiplier of these information inequalities. Hence, low-educated parents may display a lower SBR frequency partly because they are less informed about the benefits of SBR. As already mentioned, survey research consistently reports that parental education correlates with SBR frequency, but little is known about the underlying mechanisms and, in particular, about the role of knowledge of the benefits of SBR.

The relationship between SBR frequency and ELS is likely to be mediated by the ‘quality’ of parent-child interactions around books². A more interactive reading style, where parents engage children to talk about the story, is regarded as highly beneficial for ELS because children are thus induced to actively use the words of the books (NELP, 2008). Vocabulary development is fostered also if parents take the time to explain unknown words and use images to facilitate the comprehension and retention of new words, rather than simply reading the text. Expert recommendations on SBR styles, which encourage parents to interact with their children around book contents, are abundant (Oecd, 2012; Partridge, 2004).

Unfortunately, social gaps in this qualitative dimension of SBR activities are much less explored. The limited existing evidence suggests that such gaps exist, but are of moderate intensity (Vandermaas-Peller et al., 2009): more educated parents tend to be less directive when proposing and managing

² The term ‘quality’ may have a normative connotation, which tends to equate ‘good’ parenting with middle class norms on parenting. We refer to SBR quality in a more descriptive sense: the quantity of time invested in this activity has different impacts on ELS depending on *how* parents read to their children.

this activity. Upper class norms involve less 'authoritarian' and more interactive styles of parenting (Doepke and Zibbotti, 2019; Auger et al., 2014). However, in practice, SBR is often a unidirectional activity where parents read and children passively listen, regardless of parental education. Research based on video recordings reveals a recurrent pattern across social groups: parents focus on the text, while children are attracted to the book images and stay quite disconnected from verbal contents (Evans and Saint-Aubin, 2005). Practical knowledge on the best strategies to foster interactions around book contents may be an information barrier that is socially widespread. Interactive styles of SBR are quite infrequent across social groups.

Information is of course only one of the potential determinants of SBR, which is in turn only one of the determinants of ELS. In particular, the cultural resources, reading habits and reading enjoyment of the parents, as well as their school involvement, correlate with both SBR and ELS (Notten and Kraaykamp, 2013; Klosterman et al., 2011; Oecd, 2012), in line with the predictions of cultural reproduction theory and of the family investment model (Bradley and Corwyn, 2004). Family income and composition, parental distress and employment conditions are also relevant predictors, reflecting the interplay between economic and social resources of the home environment (Oecd, 2012; Karrass et al., 2003). SBR practices are responsive also to children's characteristics, such as SBR enjoyment and readings skills (*ibidem*). Finally, contextual predictors identified in the literature involve the socio-economic composition and educational resources of neighbourhoods and schools (Minh et al., 2017; Myrberg and Rosen, 2009).

From a methodological perspective, all these factors are potential confounders biasing estimates of the causal relations between information barriers, SBR and ELS in observational studies. Access to information about the benefits of SBR for school achievement is not randomly distributed: it reflects a socially biased distribution of resources and attitudes that can simultaneously affect also SBR and ELS. Controlling for all these confounders with survey data is extremely difficult. From a policy

perspective, these factors are strongly inertial constraints to any intervention aimed at fostering SBR and ELS: a lack of cultural, economic and social resources in the family environment, a poor motivation toward SBR of parents and children, as well as adverse contextual factors, may hinder receptivity to information interventions on the benefits of SBR for children's academic success.

Information experiments represent a promising research strategy in both respects. The randomisation constitutes two groups of *ex ante* equivalent families: by providing information on the benefits of SBR to only one of these groups and then assessing children's ELS, it is possible to isolate the causal effects of information barriers. At the same time, the results of these studies provide policy indications on the potential of information interventions designed to foster SBR.

Overall, our theoretical arguments imply that awareness of the importance of SBR for children's ELS is unequally distributed among social groups: low-educated parents have less access to expert sources and informal networks where this information circulates more often. Therefore, we hypothesise that providing low-educated parents with information on the benefits of SBR enhances their SBR frequency (H1a); we expect that providing information on SBR enhances SBR frequency less among high-educated parents (H1b), who are less exposed to information barriers. The contrasting scenario is that inertial constraints associated with lower cultural, social and economic resources of low-educated families hinder their receptivity to information interventions.

Moreover, we hypothesise that providing information on SBR enhances ELS among pupils from low-educated families (H2a), thanks to the increased SBR frequency. Finally, we expect that providing information on SBR enhances ELS also among pupils from high-educated families (H2b), which should be more receptive to information inputs on the quality of SBR. This is because interactive styles of SBR are beneficial to ELS: we have argued that parents often lack practical knowledge to

enact them regardless of their social position, but that high-educated parents are more inclined toward less directive parenting styles.

3. Experimental evidence on shared book reading interventions

Educational research has increasingly used randomised experiments to assess whether ELS can be fostered by parenting interventions promoting higher SBR frequency and quality. Brochures, videos or face-to-face training are employed to convey information on the benefits of SBR and to provide tips and suggestions for an effective reading, together with the provision of free books for children. A recent meta-analysis has identified 30 randomised experiments on SBR interventions carried out in the past three decades (Barone et al., 2019); 19 of them assessed a specific SBR intervention methodology called dialogic reading, and they systematically reported evidence of positive impacts on ELS. However, only two of the other 11 experiments reported statistically significant, positive impacts on ELS. Hence, the experimental evidence challenges the widespread consensus that ELS can be fostered by SBR interventions: this conclusion does not apply to SBR interventions in general, but only to dialogic reading. Unfortunately, children of high-educated parents are found to benefit more from dialogic reading interventions (Mol et al., 2008). Hence, the positive impacts of these interventions on the overall level of skills come at the price of increasing social gaps.

These results are highly informative from a social stratification perspective. Dialogic reading is a SBR technique where the adult and the child switch roles so that the child learns to become the storyteller with the support of the adult, who acts as a listener and questioner (Whitehurst and Lonigan, 1998). This intervention approach thus focuses on the qualitative, interactive dimension of SBR. The experimental evidence on its efficacy thus supports the argument that the relation between SBR frequency and ELS is contingent on the quality of parent-child interactions around

books. Moreover, the fact that children from high-educated families benefit more from dialogic reading seems to support our previous theoretical arguments. First, this result suggests that also in these families routine SBR practices are not much interactive. Second, the differentiated impact of dialogic reading suggests that high-educated parents are more inclined to incorporate interactive reading styles. Third, the failures of other (non-dialogic) SBR interventions may illustrate the importance of the above-discussed inertial constraints hindering SBR, particularly in disadvantaged families.

However, these experimental results should be taken with caution, for at least three reasons (Barone et al., 2019). Lack of statistical power is a major concern, since the sample sizes of these studies are very small (106 cases on average). Moreover, none of these experiments employed random sampling: respondents were selected *ad hoc* (for instance, via newspapers announcements or among library users). Hence, the external validity of these studies is poor. Virtually all previous experiments were carried out in the US or in other Anglo-Saxon countries. Finally, these sampling methods and sample sizes prevent systematic comparisons of treatment impacts across social groups, which can be tentatively inferred only *ex post* via meta-analyses comparing studies based on different target populations. We describe below how we tried to overcome these limitations in our experimental design.

4.1 The experimental design

Our study started in the school year 2016-2017, when we collected information on the family background, SBR activities and ELS of a first sample of 853 children. Data collection (pre-test in January 2017) started before the information intervention, which lasted four months and was followed by a post-test in June 2017, where we updated this information. In the school year 2017-

2018, we carried out a follow-up (December 2017), as well as a replication experiment on a new sample of 934 children. Sections 4.2 and 4.3 present the first experiment; section 4.4 illustrates the replication experiment.

4.2 Targeting, sampling and randomisation

We carried out the study in the city of Paris and we targeted the districts that display a significant social mix (12th, 18th, 19th and 20th *arrondissements*), the rest of the city being virtually monopolised by the upper class. Following the same logic, we screened out a few upper-class enclaves in these districts (such as Montmartre, located in the 18th district) by targeting the so-called priority education schools (whose French acronyms are *REP* and *REP+*). The resulting population of schools falls below the average levels of parental income and education of the city of Paris, but displays a substantial presence of high-educated households (see tab. 1) owing to the gentrification process characterizing these areas. We thus have enough variability to carry out systematic comparisons across social groups.

From the list of eligible kindergarten, we drew a random sample of 25 schools, with probability of selection proportional to school size. We presented our study to the school principals and we screened out three schools that had ongoing books loan programmes³. Only five schools out of 22 refused to participate in the project, and they were replaced by randomly drawing additional schools. Access to the field was considerably facilitated by the partnership with the school authority of the education department of the city of Paris (*Académie de Paris*).

³ Students of the control group could not be involved in any book loan activity in order to avoid treatment replacement bias.

Within each school, we randomly assigned half of the classes to the treatment and half to the control status (cluster randomisation with blocking across schools). Then, we communicated the list of treated and control classes to the school principals; no school (or class) dropped out of the project. We had agreed with the schools that children assigned to the control status would receive a delayed treatment in the following school year. This agreement restricted the time window of our study, but it was necessary to ensure participation of the schools. Moreover, a within-school randomisation entails higher risks of treatment contamination than a between-school randomisation, but the latter was unfeasible, since our statistical power calculations indicated that its minimum detectable effect size was too high. Hence, the estimates of treatment effects should be regarded as conservative, although we believe that treatment contamination was rather marginal also thanks to the delayed treatment⁴.

In France kindergarten are attended by children aged 3 to 5. We targeted children aged 4 so that those assigned to the control group would still be in kindergarten the following year to receive the delayed treatment. We did not target children aged 3 to facilitate the administration of the vocabulary test described in the next section: at this age, some children may fail to understand the test procedure. If a school had multi-age classes (*classes multi-niveaux*), only children aged 4 were involved in the study. We did not have any case of eligible siblings attending the same school.

Teachers distributed a parental consent form, which was signed by 96% of the parents (this value is constant across treated and control classes). Among the pupils with signed consent form, 3% were

⁴ In the project agreement signed by the school principals, they had committed to comply with the results of the randomisation and teachers had no incentive to contaminate the treatment: they knew that children in control classes would receive the delayed treatment. Treated and control pupils could play together in the courtyard, but for the vocabulary test described below we chose words (such as wrench or sledge) that are unlikely to occur in this context. Parents are the main source of potential contamination, but interactions between parents of different classes, which in kindergartens usually take place a few minutes before opening or closing hours, are sporadic and superficial in comparison to the systematic, intensive and prolonged treatment described in the next section.

absent when we collected the data, thus resulting in a final sample size of 853 children at the pre-test⁵.

We did not screen out families where parents could not speak any French (2.7% of our sample). Such a screening is common in the field experiments reviewed in section 3, but it would be perceived as discriminatory in the context of a school-based intervention. For similar reasons, we did not formally screen out students with learning disability, but these students were screened out *de facto*, either because their parents did not sign the consent form, or because it was impossible to administer the vocabulary test to them.

4.3 Data collection

The primary outcome of the experiment is children's receptive vocabulary. This is the most common outcome in SBR interventions, because it is a strong predictor of reading comprehension, writing and academic achievement (NELP, 2008). Moreover, reliable measures of receptive vocabulary are available for children aged 4. For the pre-test we used the PPVT (Peabody Picture Vocabulary Test, Dunn & Dunn, 2007), adapted in French as EVIP (*Echelle de Vocabulaire en Images Peabody*). In this test, the interviewer assesses if a child understands the meaning of a given word by naming this word, showing four pictures and asking the child to point to the relevant one. The scale has 170 items, ranked in order of increasing difficulty; respondents start the test at different points according to their age, and they stop when they make six errors out of a sequence of eight words. For children aged 4, the test starts with item number 20 and typically ends after 30-40 words. PPVT is the most common test of receptive vocabulary used in SBR experiments and, more generally, in

⁵ The post-test was administered to 829 of these 853 children.

educational research (Fryer and Levitt, 2006). Its validity and reliability have been extensively documented (*ibidem*) and PPVT is available in two parallel forms, which is useful for test-retest longitudinal designs of experimental studies. We used form A for the pre-test and form B for the post-test.

However, a significant limitation of the PPVT is its limited sensitivity to *changes* in children's vocabulary (Marulis and Neuman, 2010). For children aged 4, the final score typically depends on their knowledge of only 30-40 words. If these words are not used in the books distributed in SBR interventions, the PPVT cannot record any treatment impact by construction. This is why several studies rely instead on author-created measures, which use the same methodology as the PPVT, but target the specific words of the books (*ibidem*). However, these context-sensitive measures cannot be any longer interpreted as measures of general vocabulary. Hence, for the post-test we decided to employ both types of measures: we complemented the standard PPVT with a test of 16 words used in the books distributed to the treated classes ('PPVT-C'). To validate this measure, we ran a factor analysis using the two parallel forms of the PPVT and the adapted PPVT-C. We found strong evidence that PPVT-C measures the same latent construct as the canonical forms, and their Cronbach's alpha (0.72) is satisfactory⁶. Interviewers were blind to the experimental status of children.

We exploited the parental consent form to survey parents on their highest level of education, the language spoken with their children, the frequency of SBR at home, as well as the enjoyment of parents and of children for this activity. We could ask only few questions, because the consent form had to be filled up when parents dropped their children at school. Therefore, we could not collect reliable information on parents' occupation or income. Parents' education is the most common

⁶ The results are reported in appendix 2.

indicator of family background used in the literature on SBR, a possible indication that socio-cultural barriers are regarded as more relevant than economic hurdles. Having agreed to deliver a delayed treatment to control students, we could not observe long-term impacts. However, we managed to negotiate with the schools the administration of a follow-up in December 2017, that is, six months after the end of the intervention.

4.4 The replication study in the second year

During summer 2017, we obtained funding for a second experiment on SBR and we decided to make a replication study. Hence, we targeted children aged 4 attending priority education schools in the same districts of Paris. We applied the same sampling and randomisation designs, as well as the same rules of access to the field. Six schools out of 24 refused to participate in the second year and had to be replaced. 96% of the parents signed the consent form.

We closely replicated also the data collection design, with two differences. First, due to budget constraints, we tested children only in the short-term assessment, using only the adapted PPVT-C. As documented in section 5, the results of the first year indicated that the canonical forms of PPVT were insensitive to treatment impacts, for the reasons discussed above. Second, the specific words tested in the adapted PPVT-C were different from those of the first year, because we had distributed different books (see footnote 1).

4.5 Treatment design

In the first year, the SBR intervention was implemented for four months (between February and May 2017) and involved four components:

- a) parents received two books per week during these four months. The research team had selected 18 book titles and delivered to the teachers of treated classes a number of books twice as large as the size of their classes. Each child was left free to choose two books every week. This was important to capture children's interest in the books and bring them to ask their parents to read them. Teachers were asked to monitor that children chose a variety of titles;
- b) during the first six weeks, parents received one brochure per week that explained why SBR is considered by experts as highly beneficial to school achievement. These brochures also provided several tips for an effective and enjoyable reading time: practical suggestions to foster interactive reading styles were stressed. Every week treated children went back home with a coloured project bag containing the two books and the weekly brochure;
- c) in the second month of this intervention, a team of four interviewers hired and trained by the research team called the parents of treated children, asked for informal feedbacks on the project and delivered a short, standardised message, which reiterated the key messages of the brochures concerning the benefits of SBR and of interactive reading styles. These phone calls could be made in five languages (French, Arab, Chinese, Spanish, English) in order to reach also allophone families (see the online supplementary materials, Appendix 1);
- d) during the last six weeks, parents received one text message per week reiterating the main messages of the intervention.

The contents of this SBR intervention were organised around the twofold objective of raising the frequency and the quality of SBR activities. On one side, in order to motivate parents to read regularly, we stressed that SBR involves high benefits for children and low costs for their parents. We illustrated the positive effects of this activity on cognitive, language and social development, as well as its emotional value for parent-child relations. We stressed the limited time investment

demanded and we provided detailed information on public libraries in these neighbourhoods and on opportunities to buy cheap books. On the other side, information materials explained to the parents how to propose this activity to their children, how to 'play the story' using voice and gestures, and how to encourage children to talk about the story, to use the new words that they had learnt and to reinvent the story. These suggestions were based on a core set of frequent recommendations found in the grey literature on parenting. We exploited also the materials of a preparatory study carried out in the target neighbourhoods by means of qualitative interviews with parents, teachers and school principals.

The treatment of the second year was a close replication of the one of the first year, but it was more diluted over time: having a less tight time schedule, the book loan could last for six months (instead of four) and we borrowed one book per week (instead of two). We report a detailed description of treatment contents in Appendix 1.

It must be noted that the treatment provided not only information to parents, but also books for free. We considered the option of employing a factorial design to isolate the specific impact of information, but we ruled out this option, which demands a much larger sample size. Moreover, as discussed above (section 4.3), selecting the books that parents would read was necessary to develop a context-specific vocabulary test, but this was feasible only by freely providing these books to the families. Hence, caution is warranted when interpreting the mechanisms driving treatment impacts as purely informational. This said, economic constraints are unlikely to be significant hurdles to SBR: books for children can be freely borrowed in the widespread network of libraries available in these

neighbourhoods. Of course, families are not always aware of this opportunity: this is another information barrier that we tried to remove⁷.

5.1 Equivalence *ex ante* and correlational results

In what follows we present results based on the merged data for the two school years. In Appendix 3, we report separate descriptives and treatment impacts for each year. Since we did not detect any statistically significant difference between the estimates for the two years, the replication was successful. Therefore, we have merged the data to enhance statistical power and to focus on substantive results.

TABLE 1 HERE

In Table 1 we report descriptive statistics on treated and control pupils at the pre-test. Pointwise estimates for the two groups are close and differences are not statistically significant⁸. Importantly, the average vocabulary scores were identical *ex ante*. Hence, the randomisation worked well. In about one third of the families, parents speak only French with their children, an indication of the high presence of first- and second-generation immigrants in Paris. However, parents who speak only a language other than French with their children are a tiny minority (2.3%). In about half of the

⁷ The experimental protocol of this study was registered in the Social Science Registry before running the experiment.

⁸ The only minor exception is the category ‘somewhat’ of the variable ‘child enjoyment’ (parental self-reports). Controlling for this variable in the models presented below does not affect our conclusions (results available upon request). The statistical significance of differences between the two groups is tested by means of regression models which incorporate dummy covariates for experimental status and school year.

families (49%) of the sample, at least one parent has a tertiary degree; the share of the tertiary-educated among adults aged 25-34 is 44.7% in France.

Parents report that SBR is a common and frequent home activity: in four households out of ten, it is practiced on a daily basis and in 20% of them on a weekly basis. However, these averages conceal strong variations by parental education. For instance, SBR on a daily basis is practiced by 64% of tertiary-educated parents, but only by 14% of low-educated parents (not reported). Similarly, the average high levels of parents' and children's enjoyment of SBR reported in Table 1, which indicates that almost half of parents and children enjoy this activity to a great extent, are socially stratified: in low-educated households, 40% of the children and 36% of the parents enjoy this activity; the corresponding values for high-educated families are 65% and 63%, respectively. Importantly, these are self-reported data which may be subject to social desirability bias.

TABLE 2 HERE

In Table 2, we report correlational evidence that SBR mediates social gaps in ELS using the data of the pre-test for both groups. Model 1 refers to the total effect of parental education, controlling for language spoken at home, gender and age of the child. The standardised coefficient for tertiary-educated parents (0.62) confirms that parental education is a strong predictor of language skills. In model 2, we add SBR frequency, which displays a strong association with children's vocabulary: the coefficient for the category 'reading every day' is 0.73 (reference category: reading rarely or never). Interestingly, we detect a marked reduction of the total effect of parental education from 0.62 in model 1 to 0.36 in model 2 (-42%), suggesting that SBR is a significant mediator of socio-economic gaps in ELS. Compared with previous analyses, which were based on retrospective parental reports (Parks 2008), this mediation effect looks particularly strong. However, as discussed above, the effect of SBR on ELS is clearly not causal. In the next section, we present the experimental results.

5.2 Experimental evidence

The vocabulary score at the post-test is our primary outcome, while SBR frequency at the post-test is our main instrumental outcome. We cannot reliably observe any treatment impact on the quality of SBR activities: this would demand direct observations in the home environment, which would be highly intrusive. Hence, we use parents' and children's enjoyment of SBR as proxies. Each outcome at the post-test is regressed on the same variable at the pre-test, a dummy indicating assignment to treated or control classes (intention-to-treat estimators) and a dummy for school year.

For vocabulary scores (EVIP B and C), we use OLS regression models and report standardised coefficients. For the other outcomes, we use binomial logistic regression and report average marginal effects. Standard errors are clustered at the school level.

Table 3 reports the estimates of treatment impacts (see Appendix 4 for the full results). Column 1 indicates that the treatment has a significant, positive effect on SBR frequency: the share of parents who report reading to their children every day is enhanced by 8 percent points in the experimental group at the post-test⁹. This main effect conceals strong variations according to parental education: it is twice as large for low-educated households (+16.2%), while we detect no significant treatment impact (+2.9%) among high-educated families, where SBR frequency was high already before the intervention. The confidence intervals of the point estimates for the two groups do not overlap, indicating that the intervention has reduced social inequalities in access to SBR. If we model SBR frequency less restrictively (parents report reading at least once a week), the main effect is 14% and

⁹ Because 85% of the parents who had signed the consent form filled up both questionnaires, the analytical sample for reading practices comprises 1374 cases.

the values for low- and high-educated families are, respectively, 25.7% and 7.7%. Hence, hypotheses 1a and 1b are confirmed.

TABLE 3 HERE

Columns 2 and 3 refer to treatment impacts on parents' and children's enjoyment of SBR (as reported by the parents). As can be seen, the main effects are positive and statistically significant. The coefficients referring to high-educated families are statistically significant, while those for low-educated families do not reach statistical significance. However, the difference between the two sets of point estimates is small and not statistically significant. Hence, our expectation that the more qualitative dimension of the SBR intervention would impact more on high-educated families is not supported by the results.

Column 4 refers to treatment impacts on children's ELS, measured by a standardised test of receptive vocabulary. The effect on the standard PPVT (form B) is not statistically significant: as discussed above, this general measure of vocabulary may be weakly sensitive to treatment impacts. Indeed, the treatment impact (effect size of 0.11) on the book-specific measure of vocabulary (PPVT-C) is positive and statistically significant (column 5)¹⁰. Moreover, the heterogeneity analysis indicates that the treatment impacted positively on the ELS of children from low-educated families, in line with hypothesis 3. Hence, when low-educated parents access information on the benefits of SBR for school success, they report increasing the frequency of this activity, which results in vocabulary gains for their children. For children from high-educated households, the effect is positive, but not statistically significant. Hence, we cannot reject the null hypothesis that the treatment had no effect on this subpopulation, a finding that militates against hypothesis 2b.

¹⁰ As reported in appendix 2, a factor analysis reveals that the three vocabulary measures tap into the same underlying construct. The difference is that PPVT-C specifically tests the words used in the books of the project.

However, it should be noted that the point estimates for the two groups are similar and the confidence intervals overlap.

We have commented so far the effects on children's skills at the short-term assessment, carried out in the month following the end of the intervention. Column 6 reports the results of the follow-up, carried out six months afterwards (only in the first year of the study). We detect a statistically significant, positive effect size (0.16). Hence, the treatment impact persists six months after the conclusion of the intervention. The point estimates for the two groups overlap again, but in this mid-term assessment the point estimate for children from low-educated families (0.23) is much higher than that for children of tertiary graduates (0.05)¹¹.

Finally, let us comment on some robustness checks. The models presented in Table 3 incorporate a treatment dummy, the corresponding pre-treatment variable and a dummy for the school year. We have estimated some nested models incorporating also socio-demographic control variables (gender, age and language spoken at home), as well as simpler models, where the only predictor is the treatment dummy. Results are unchanged (see appendix 5).

Since we cannot observe whether and how parents read to their children at home, we cannot have any direct indicator of treatment compliance. However, the phone calls could reach 84% of the families: considering that each family was contacted at least four times (in different days and at different hours), we suspect that the remaining 16% of families mainly comprised parents who preferred avoiding contact with the interviewers, probably because of their weak involvement in the project (parents received a text message before the phone call). Therefore, we used successful phone contact as an indirect indicator of treatment compliance to estimate average treatment

¹¹ Due to longitudinal attrition, at the mid-term assessment we lost 21% of the children participating in the short-term assessment. In Appendix 1, we show that *ex ante* equivalence holds also for the subsample of 'survivors', that attrition is not socially patterned and that there is no statistically significant difference between short- and mid-term impacts for this subsample.

effects on the treated. As expected, treatment impacts are larger among compliers than the ITT estimates reported in table 3 (Appendix 4), but the difference is not large.

6. Concluding remarks

This work has examined the social stratification of SBR practices by means of a field experiment assessing the causal impact on SBR and ELS of information barriers concerning the beneficial effects of SBR for school success. We designed an information treatment concerning these benefits using flyers, phone calls and text messages. The experiment involved a large sample of children aged 4 attending 44 kindergarten located in the city of Paris and employed a clustered, classroom-level randomisation. As documented above, this experiment displays high internal validity, in terms of *ex ante* equivalence of the two groups and similar attrition rates. Moreover, this experiment is not based on a national sample, but it still marks a significant improvement over previous SBR experiments in terms of external validity, particularly as regards the sampling design and sample size, which allows systematic comparisons among children from low- and high-educated families.

We hypothesised that information on the cognitive-enhancing potential of SBR is unequally distributed among social groups. High-educated parents access this information more easily than the low-educated through direct access to expert sources on parenting and the indirect circulation of this information within their social networks. In line with this hypothesis, we found that the information treatment impacts on the SBR frequency of low-educated parents, but not on that of high-educated parents. Moreover, we found that the treatment has positive impacts on children's ELS, measured via standardised test scores. Importantly, these effects persist six months after the conclusion of the intervention. These positive mid-term impacts involve children of low-educated

parents, while they are negligible and not statistically significant for children of high-educated families.

The treatment impacts on SBR frequency are quite pronounced: at the pre-test, SBR on a daily basis involved 41.2% of the families, and the treatment fostered this practice by 8 percent points. However, this is a self-reported outcome, which could reflect social desirability bias. It is therefore reassuring that the treatment impacts also on children's skills, measured via standardised tests. These latter effects are not weak, at least for disadvantaged children. Following Cohen's general benchmarks for effect sizes in statistics, they should be regarded as small. However, if consider that the more context-specific benchmark of the average effect size in educational field experiments is 0.06 (Lortie-Forgues and Inglis, 2018), the effect sizes of 0.11 (short-term assessment) to 0.16 (follow-up) are not small, and those involving low-educated families (0.13 and 0.23) are comparatively strong. Moreover, these estimates could be conservative because of potential spill-over effects, although contamination between the treated and control pupils is arguably marginal (*see footnote 3*).

From a policy perspective, the magnitude of these effects is far from negligible, particularly if we consider the limited duration of the intervention and its low marginal costs (approximately 3.50 euros per child) if it is scaled up. Moreover, we carried out qualitative interviews with the teachers and school principals of treated classes after the post-test and they consistently reported that the reception of the intervention from children, parents and teachers had been very positive.

Finally, in substantive terms, these causal effects seem far from negligible considering that information constraints are only one of the several determinants of SBR practices, and that we left all other cultural, social and motivational barriers unchanged. Previous research has extensively documented the importance of these barriers, but it has paid scant attention to the role of

information as a mechanism driving social inequalities in early childhood. Our study thus makes a novel contribution by documenting that the lack knowledge on the benefits of informal learning activities at home has a genuinely causal effect on the opportunities to access these activities for children from low-educated families, with significant consequences for their language development. At the same time, a cheap, light-touch parenting intervention can contrast these information inequalities, with beneficial effects for these children.

References:

- Andersen P., Hansen M., 2012, Class and cultural capital—The case of class inequality in educational performance, *European Sociological Review*, 28(5), 607–621.
- Auger A., Reich S., Penner E. (2014). The effect of baby books on mother's beliefs and reading practices. *Journal of Applied Developmental Psychology*, 35(4): 337-346.
- Barone C., Chambuleyron E., Vonnak R., Assirelli G. (2019). The effects of shared book reading interventions on children's skills: a meta-analysis of randomised controlled trials, *LIEPP Working Paper No. 19*.
- Barone C., Assirelli G., Abbiati G., Argentin G., De Luca D. (2018). Social origins, relative risk aversion and track choice: A field experiment on the role of information biases. *Acta Sociologica*, 61(4): 441–459.
- Baker L., Scher D., Mackler K. (1997). Home and family influences on motivations for reading. *Educational Psychologist*, 32(2): 69-82.
- Blossfeld HP., Kulic N., Skopek J., Triventi M. (eds) 2017. *Childcare, early education and social inequality: an international perspective*. London: Edward Elgar.
- Bourdieu P. (1979). *La distinction, critique sociale du jugement*. Paris: Les Editions de Minuit.
- Bradley, R., Corwyn R. (eds) 2004. *Family investments in children's potential: resources and parenting behaviors that promote success*, Lawrence Erlbaum Associates, Mahwah, US.
- Dämmrich J., Triventi M. (2018). The dynamics of social inequalities in cognitive-related competencies along the early life course – a comparative study. *International Journal of Educational Research*, 88: 73–84.
- De Graaf N., De Graaf P., Kraaykamp G. (2000). Parental cultural capital and educational attainment in the Netherlands: a refinement of the cultural capital perspective. *Sociology of Education*, 73(2): 92-108.
- Dickinson D., Griffith J., Michnick R., Golinkoff R., Hirsh-Pasek K (2012). How reading books fosters language development around the world. *Child Development*, Vol. 2012.

- Doepke M., Zilibotti F. (2019). *Love, money and parenting*. Princeton: Princeton University Press.
- Dotti Sani G., Treas J. (2016). Educational gradients in parents' child-care time across countries, 1965–2012. *Journal of Marriage & Family*, 78(4): 1083–1096.
- Dunn L., Dunn D. (2007). *Peabody Picture Vocabulary Test*. London: Pearson Education.
- Evans M., Saint-Aubin J. (2005). What children are looking at during shared storybook reading: evidence from eye movement monitoring. *Psychological Science*, 16(11): 913-920.
- Fryer R., Levitt S. (2006). The Black-White Test Score Gap Through Third Grade. *American Law and Economics Review*. 8(2): 249–281.
- Hoffman J., Teale W., Paciga K. (2014). Assessing vocabulary learning in early childhood. *Journal of Early Childhood Literacy*, 14(4): 459-481.
- Karrass J., Van Deventer M., Braungart-Rieker J. (2003). Predicting shared parent-child book reading in infancy. *Journal of Family Psychology*, 17(1): 134-146.
- Keller, T. (2003). Information and track choices: evidence from a randomized experiment, paper under review.
- Kloosterman R., Notten N., Tolsma J., Kraaykamp G. (2011). The effects of parental reading socialization and early school involvement on children's academic performance: a panel study of primary school pupils in the Netherlands. *European Sociological Review*, 27(3): 291-306.
- Kolb B., Mychasiuk M., Muhammad A., Gibb R. (2013). Brain plasticity in the developing brain. *Progress in brain research*, 27: 35-64.
- Heckman, J., Pinto R., Savelyev P. (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review*, 103(6): 2052–86.

- Herbaut E., Geven K. (2019). What works to reduce inequalities in Higher Education? A systematic review of the (quasi-)experimental literature on outreach and financial aid. *WB Policy Research Working Paper 8802*.
- Hulme C., Snowling M. (2015). Learning to read: what we know and what we need to understand better. *Child development perspectives*, 7(1): 1–5.
- Lareau A. (2011). *Unequal childhoods*. Berkeley: University of California Press.
- Lay See Y., Winston W., Charis M. (2014). The home literacy environment and preschool children's reading skills and interest, *Early Education and Development*, 25(6): 791-814.
- Lortie-Forgues H., Inglis M. (2018). Rigorous large-scale educational RCTs are often uninformative: should we be concerned? http://eprints.whiterose.ac.uk/141754/3/LFI_2019_manuscript_supp.pdf
- Mol S., Bus A., De Jong, M., Smeets J. (2008). Added value of dialogic parent-child book readings: a meta-analysis. *Early Education and Development*, 19(1): 7-26.
- Minh A., Muhajarine N., Janus M., Guhn M. (2017). A review of neighbourhood effects and early child development. *Health & Place*, 46: 155-174.
- Myrberg E., Rosen M. (2009). Direct and indirect effects of parents' education on reading achievement among 3rd graders in Sweden. *British Journal of Educational Psychology*, 79(4): 695-711.
- National Early Literacy Panel, 2008. *Developing Early Literacy*. <https://www.nichd.nih.gov>
- Notten B., Kraaykamp G. (2013). Parental media socialization and educational attainment: resource or disadvantage? *Research in Social Stratification and Mobility*, 28(4): 453-464.
- Oecd (2012). *Let's read them a story!* Paris: Oecd.
- Park H. (2008). Home literacy environments and children's reading performance: a comparative study of 25 countries. *Educational Research and Evaluation*, 14(6): 489-505.
- Partridge H. (2004). Helping parents make the most of shared book reading. *Early Childhood Education Journal*, 32(1): 211-244.

- Potter D., Roksa J. (2013). Accumulating advantages over time: family experiences and social class inequality in academic achievement. *Social Science Research*, 42(4): 1018–32.
- Radey M., Randolph K. (2012). Parenting sources: how do parents differ in their efforts to learn about parenting? *Family Relations*, 58(5): 536-548.
- Scarborough H., Dobrich W. (1994). On the efficacy of reading to preschoolers. *Developmental Review*, 14(3): 245–302.
- Sylva K., Scott S., Totsika V., Ereky-Stevens K., Crook K. (2008). Training parents to help their children read: a randomized control trial. *British Journal of Educational Psychology*, 78(3): 435–455.
- Vandermaas-Peller M., Nelson J., Bumpass C., Sassine B. (2009). Social contexts of development: parent-child interactions during reading and play. *Journal of Early Childhood Literacy*, 9(3): 295-317.
- Whitehurst, G., Lonigan C. (1998). Child Development and Emergent Literacy. *Child Development*, 69(3): 848-872.
- York B., Loeb S. (2018), One step at a time: the effects of an early literacy text messaging program for parents of preschoolers. *NBER Working Paper* 20659.

Tab. 1: Descriptive statistics on the equivalence between treated and control children before the intervention (N=1,725)

Variable, wording of the question	Category	% control	% treated	p-value difference
Gender of the child	Male	52.2	49.7	0.278
	Female	47.8	50.3	0.278
Age in months of the child	Mean	54.1	54.2	0.777
Language spoken at home “Which language do you usually speak with your child?”	Only French	36.1	35.8	0.907
	French + another language	36.7	39.4	0.312
	Only another language	3.3	2.3	0.256
	Another language + French	13.8	12.9	0.636
	Missing	10.1	9.6	0.855
Parental education “What is the highest level of education of the parents?”	Primary	5.2	4.4	0.516
	Lower secondary	13.6	13.3	0.849
	Upper secondary	22.3	23.8	0.503
	Tertiary	49.7	49.4	0.909
	Missing	9.2	9.1	0.962
Shared book reading frequency “How often do you read books to your child?”	Rarely or never	7.7	7.9	0.917
	Sometimes	24.8	21.6	0.142
	Every week	20.8	20.3	0.814
	Every day	39.8	43.8	0.190
	Missing	7.0	6.5	0.719
Child enjoyment for shared book reading “To what extent is your child interested in books?”	Not at all	0.7	1.6	0.094
	A little	15.9	17.7	0.424
	Somewhat	30.1	24.3	0.010
	To a great extent	47.7	49.6	0.489
	Missing	5.6	6.8	0.434
Parental enjoyment for shared book reading “To what extent do you enjoy reading books to your child?”	Not at all	0.6	1.7	0.024
	A little	15.9	17.0	0.601
	Somewhat	33.4	29.1	0.071
	To a great extent	43.5	44.9	0.591
	Missing	6.6	7.2	0.721
Standardised score at the pre-test (PPVT-A)	Mean	-0.006	0.006	0.829

Tab. 2 The association between parental education, shared book reading frequency and children's vocabulary in the pre-test (treated and control students, N=1,725), OLS regression coefficients

Parental education (Ref.: Lower secondary or less)	Upper secondary	0.17** (0.07)	0.10 (0.07)
	Tertiary	0.62*** (0.07)	0.36*** (0.06)
	Missing	0.14 (0.10)	0.27** (0.10)
	Reading frequency (Ref.: Rarely or never)	Sometimes per month	
	Every week		0.27*** (0.09)
	Almost every day		0.73*** (0.09)
	Missing		-0.05 (0.18)
R-squared		0.22	0.28
Robust standard errors in parentheses			
*** p<0.01, ** p<0.05, * p<0.1			

Tab. 3: Treatment impacts on primary and secondary outcomes: main effects and variations by parental education (average marginal effects for binomial logit regression and standardised coefficients for OLS regression; confidence intervals at the 95% level in brackets)

	Book reading frequency (1=every day)	Parental enjoyment (1=to a great extent)	Child enjoyment (1=to a great extent)	EVIP-B (post-test, year 1)	EVIP-C (post-test, years 1+2)	EVIP-C (follow up in year 1)
Model	Binomial logit	Binomial logit	Binomial logit	OLS	OLS	OLS
Main effects of the treatment	0.08*** [0.04; 0.12]	0.05** [0.01; 0.09]	0.07*** [0.01; 0.13]	-0.06 [-0.18; 0.06]	0.11** [0.01; 0.21]	0.16** [0.02; 0.30]
Effects of the treatment by parental education						
Tertiary degree	0.03 [-0.03; 0.09]	0.05* [-0.01; 0.11]	0.07*** [0.01; 0.13]	-0.16* [-0.32; -0.01]	0.11 [-0.03; 0.25]	0.05 [-0.15; 0.25]
Upper secondary degree or less	0.16*** [0.10; 0.22]	0.05 [-0.03; 0.13]	0.05 [-0.03; 0.13]	0.00 [-0.14; 0.14]	0.13* [-0.01; 0.27]	0.23** [0.03; 0.43]
N	1,374	1,374	1,374	829	1,611	665
R-squared	0.40	0.24	0.22	0.56	0.38	0.35

*** p<0.01, ** p<0.05, * p<0.1

ONLINE SUPPLEMENTARY MATERIALS

APPENDIX 1

Treatment contents and formats

The contents of the intervention conveyed in the brochures were articulated into three stages corresponding to different thematic domains:

Stage 1: Motivating parents to engage in SBR (weeks 1-3). Four types of messages were stressed:

1a) the *high benefits for children*: the importance of SBR for language and cognitive skills, as well as for the emotional and social development and the creativity of children. The materials stressed the connections between skill development in preschool years and school success.

1b) *emotional benefits*: SBR can be fun time to spend with children, making them feel important;

1c) SBR is a *low-effort investment*: even a few minutes per day have beneficial effects for the children, if this activity is carried out regularly;

1d) *self-efficacy*: parents can make a difference for the school success of their children thanks to SBR, regardless of their level of education.

Stage 2: Explaining how to effectively read (weeks 3-5):

2a) *how to propose SBR to children*: how to capture their interest and avoid imposing this activity;

2b) *how to play the story*: how to read a book effectively by 'playing the story' (the importance of voice & gestures); the role of images to convey the contents and emotions of the story; the importance of using images to capture the interest of the child and to explain new words;

2c) the *setting*: SBR as a ritual, that is, a special moment that should have a fixed time in the day and a fixed place; why avoiding interruptions from other family members is important;

Stage 3: Fostering interactive, enriched SBR experiences (weeks 5-8):

3a) *how to develop more interactive reading styles*: how to prompt children to talk about the story, to express their emotions, to retell the story with the support of parents or even to reinvent some its events;

3b) *promoting more opportunities for SBR*: tips to access more books and to offer more reading occasions to children (for instance, practical information on libraries);

3c) *SBR as a family affair*: adults as role models for children and the importance of promoting a broad involvement of all family members in reading activities (both parents, siblings, grandparents, etc.).

Each brochure was a coloured A4 paper with images and short texts that could be read in 2-3 minutes and that used simple sentences. The phone calls were designed to reiterate the main messages of the brochures, as we expected that some parents would not read them. The text messages were intended to summarise the core messages of this intervention and to keep up parents' awareness of their participation in the project.

The books, the information materials and the vocabulary tests were in French. If immigrant parents could not easily read French, or if they simply preferred to use their own language, we invited them to follow the book images and to tell the story in this language. We selected books adapted to this purpose (simple and short stories, lots of images and a simple vocabulary). The phone calls could be made in five languages (French, Arab, Chinese, Spanish, English) to reach allophone parents and to communicate the main messages of the brochures as well as the possibility to use a language other than French.

APPENDIX 2: Validation of EVIP-C

Tab. A1: Factor loadings, eigenvalues and explained variance of a principal-component factor analysis on EVIP-A, EVIP-B, EVIP-C (N=853)

Index	Value
Factor loadings	
EVIP-A	0.90
EVIP-B	0.90
EVIP-C	0.88
Eigenvalue 1	2.36
Eigenvalue 2	0.38
Explained variance	78%

APPENDIX 3: Comparison between years 1 and 2

Tab. A2: Descriptive statistics on the equivalence between treated and control children before the intervention, first year (N=853)

Variable, wording of the question	Category	% control	% treated	p-value difference
Gender of the child	Male	52.9	50.2	0.401
	Female	47.1	49.8	0.401
Age in months of the child	Mean	54.2	54.2	0.958
Language spoken at home “Which language do you usually speak with your child?”	Only French	35.6	33.1	0.556
	French + another language	36.5	39.9	0.437
	Only another language	4.2	2.8	0.272
	Another language + French	13.8	12.7	0.702
	Missing	9.8	11.5	0.714
Parental education “What is the highest level of education of the parents?”	Primary	5.1	4.7	0.808
	Lower secondary	14.1	13.6	0.849
	Upper secondary	22.5	24.9	0.440
	Tertiary	48.7	47.2	0.730
	Missing	9.6	9.6	0.994
Shared book reading frequency “How often do you read books to your child ?”	Rarely or never	7.3	6.6	0.702
	Sometimes	26.2	21.6	0.090
	Every week	21.6	22.8	0.676
	Every day	39.3	43.2	0.349
	Missing	5.6	5.9	0.913
Child enjoyment for shared book reading “To what extent is your child interested in books?”	Not at all	0.7	1.9	0.125
	A little	15.0	16.2	0.678
	Somewhat	28.6	24.4	0.169
	To a great extent	51.3	50.5	0.852
	Missing	4.4	7.0	0.207
Parental enjoyment for shared book reading “To what extent do you enjoy reading books to your child?”	Not at all	0.7	1.4	0.276
	A little	13.8	15.0	0.614
	Somewhat	34.0	30.1	0.290
	To a great extent	45.4	46.2	0.846
	Missing	6.1	7.3	0.607
Standardised score at the pre-test (PPVT-A)	Mean	-0.03	0.03	0.467

Tab. A3: Descriptive statistics on the equivalence between treated and control children before the intervention, second year (N=934)

Variable, wording of the question	Category	% control	% treated	p-value difference
Gender of the child	Male	51.5	49.3	0.489
	Female	48.5	50.7	0.489
Age in months of the child	Mean	54.0	54.1	0.652
Language spoken at home "Which language do you usually speak with your child?"	Only French	36.5	38.1	0.696
	French + another language	36.9	38.9	0.522
	Only another language	2.3	1.8	0.632
	Another language + French	13.9	13.2	0.781
	Missing	10.4	8.0	0.331
Parental education "What is the highest level of education of the parents?"	Primary	5.3	4.2	0.495
	Lower secondary	13.2	13.0	0.934
	Upper secondary	22.2	22.9	0.821
	Tertiary	50.6	51.3	0.827
	Missing	8.9	8.6	0.937
Shared book reading frequency "How often do you read books to your child?"	Rarely or never	8.1	9.0	0.653
	Sometimes	23.3	21.6	0.587
	Every week	20.1	18.2	0.367
	Every day	40.2	44.3	0.359
	Missing	8.3	7.0	0.558
Child enjoyment for shared book reading "To what extent is your child interested in books?"	Not at all	0.7	1.4	0.382
	A little	16.9	19.0	0.488
	Somewhat	31.6	24.2	0.027
	To a great extent	44.1	48.9	0.261
	Missing	6.7	6.6	0.960
Parental enjoyment for shared book reading "To what extent do you enjoy reading books to your child?"	Not at all	0.5	2.0	0.047
	A little	18.0	18.8	0.790
	Somewhat	32.8	28.3	0.127
	To a great extent	41.6	43.7	0.575
	Missing	7.2	7.2	0.991
Standardized score at the pre-test (PPVT-C)	Mean	0.18	-0.16	0.659

Tab. A4: Treatment impacts on primary and secondary outcomes: main effects and variations by parental education (average marginal effects for binomial logit regression and standardised coefficients for OLS regression), first year

	Book reading frequency (1=every day)	Parental enjoyment (1=to a great extent)	Child enjoyment (1=to a great extent)	EVIP-C
Model	Binomial logit	Binomial logit	Binomial logit	OLS
Main effects of the treatment	0.10*** (0.03)	0.02 (0.03)	0.04 (0.04)	0.11 (0.07)
Effects of the treatment by parental education				
Tertiary degree	0.04 (0.05)	0.02 (0.04)	0.04 (0.04)	0.04 (0.08)
Upper secondary degree or less	0.20*** (0.04)	0.03 (0.06)	0.00 (0.06)	0.17* (0.10)
Confidence intervals overlapping	Yes	Yes	Yes	Yes
N	665	657	662	829
R-squared	0.37	0.22	0.20	0.56

*** p<0.01, ** p<0.05, * p<0.1

Tab. A5: Treatment impacts on primary and secondary outcomes: main effects and variations by parental education (average marginal effects for binomial logit regression and standardised coefficients for OLS regression), second year

	Book reading frequency (1=every day)	Parental enjoyment (1=to a great extent)	Child enjoyment (1=to a great extent)	EVIP-C
Main effects of the treatment	0.07*** (0.03)	0.08** (0.03)	0.11*** (0.04)	0.11 (0.08)
Effects of the treatment by parental education				
Tertiary degree	0.03 (0.04)	0.06 (0.03)	0.11*** (0.04)	0.17 (0.10)
Upper secondary degree or less	0.13*** (0.05)	0.07 (0.06)	0.10 (0.07)	0.07 (0.09)
Confidence intervals overlapping	Yes	Yes	Yes	Yes
N	709	703	710	782
R-squared	0.43	0.26	0.27	0.34

*** p<0.01, ** p<0.05, * p<0.1

APPENDIX 4: Robustness checks on modeling specifications

Tab. A6: Treatment impacts on EVIP C: ITT and ATT estimates with different model specifications: main effects (standardised coefficients for OLS regression), first + second year

	ITT			ATT		
	M0	M1	M2	M0	M1	M2
Treatment : main effect	0.133* (0.071)	0.111** (0.053)	0.112** (0.051)	0.161* (0.086)	0.134** (0.064)	0.135** (0.061)
Vocabulary score at the pre-test		0.612*** (0.019)	0.543*** (0.023)		0.611*** (0.019)	0.541*** (0.023)
Female (ref : male)			0.001 (0.040)			0.003 (0.040)
Language spoken at home: French + other (ref .: only French)			-0.074 (0.049)			-0.073 (0.049)
Language spoken at home: Only other			-0.173*** (0.065)			-0.173*** (0.065)
Langue spoken at home: missing			-0.097 (0.097)			-0.098 (0.098)
Parental education: Upper secondary (ref .: Lower sec. or less)			0.036 (0.067)			0.034 (0.066)
Parental education: Tertiary			0.253*** (0.058)			0.255*** (0.058)
Parental education: missing			-0.174* (0.097)			-0.170* (0.097)
Year: second	-0.005 (0.071)	0.008 (0.053)	-0.027 (0.051)	-0.006 (0.070)	0.007 (0.053)	-0.028 (0.050)
Constant	-0.061 (0.117)	-0.071 (0.088)	-0.083 (0.112)	-0.060 (0.116)	-0.070 (0.087)	-0.083 (0.111)
Observations	1,628	1,611	1,611	1,628	1,611	1,611

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Tab. A7: Treatment impacts on parental reading frequency: ITT and ATT estimates with different model specifications: main effects (average marginal effects for logit regression), first + second year

	ITT			ATT		
	M0	M1	M2	M0	M1	M2
Treatment: main effect	0.114*** (0.032)	0.084*** (0.020)	0.087*** (0.020)	0.139*** (0.038)	0.102*** (0.025)	0.106*** (0.024)
Pre-intervention measure: Sometimes per month		0.048 (0.031)	0.026 (0.032)		0.044 (0.032)	0.021 (0.033)
Pre-intervention measure: Every week		0.219*** (0.041)	0.168*** (0.040)		0.213*** (0.042)	0.162*** (0.041)
Pre-intervention measure: Almost every day		0.728*** (0.036)	0.622*** (0.041)		0.725*** (0.037)	0.617*** (0.042)
Pre-intervention measure: missing value		0.146** (0.061)	0.180* (0.093)		0.147** (0.062)	0.179* (0.092)
Female			-0.007 (0.020)			-0.005 (0.019)
Language spoken at home: French + other (ref .: only French)			-0.063** (0.026)			-0.063** (0.026)
Language spoken at home: Only other			-0.093** (0.036)			-0.094*** (0.036)
Language spoken at home: missing			-0.073 (0.079)			-0.077 (0.079)
Parental education: Upper secondary (ref .: Lower sec. or less)			0.002 (0.033)			0.001 (0.032)
Parental education: Tertiary			0.115*** (0.037)			0.117*** (0.037)
Parental education: missing			-0.017 (0.054)			-0.011 (0.055)
Year: second	0.018 (0.032)	-0.001 (0.020)	-0.008 (0.020)	0.019 (0.032)	0.000 (0.020)	-0.007 (0.020)
Constant	0.429*** (0.052)	0.077* (0.045)	0.136** (0.053)	0.428*** (0.051)	0.080* (0.045)	0.137** (0.054)
Observations	1,374	1,374	1,374	1,374	1,374	1,374

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

