I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17322

A Library in the Palm of Your Hand? A Randomized Field Experiment with Low-Income Children

Silke Anger
Bernhard Christoph
Agata Galkiewicz
Shushanik Margaryan
Frauke Peter
Malte Sandner
Thomas Siedler

SEPTEMBER 2024

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17322

A Library in the Palm of Your Hand? A Randomized Field Experiment with Low-Income Children

Silke Anger
IAB and IZA

Bernhard Christoph
IAB

Agata Galkiewicz
University of Potsdam and IAB

Shushanik Margaryan
University of Potsdam, IZA and Berlin School of Economics (BSoE)

Frauke Peter
DZHW

Malte Sandner
Technische Hochschule Nürnberg, IAB and IZA

Thomas Siedler
University of Potsdam, IZA and Berlin School of Economics (BSoE)

SEPTEMBER 2024

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

A Library in the Palm of Your Hand? A Randomized Field Experiment with Low-Income Children*

Reading comprehension is critical for academic success, yet children from disadvantaged backgrounds often engage in reading less frequently than their more advantaged peers. This study evaluates the impact of a randomized reading intervention targeting 11{12-year-olds from low-income households in Germany. As part of the intervention, children received e-book readers with access to a large digital library of age-appropriate books. The results show a signicant increase in reading engage- ment, leading to improvements in academic performance. 12 months after the start of the intervention, academic performance improved by 10.3% of a standard devia- tion, with eects persisting at 9.6% after 30 months. Additionally, the intervention positively impacted the socioemotional well-being of the children.

JEL Classification: C93, I20, I24

Keywords: randomized controlled trial, low socioeconomic status, reading comprehension, early education

Corresponding author:
Shushanik Margaryan
Universität Potsdam
Am Neuen Palais 10
14469 Potsdam
Germany
E-mail: shushanik.margaryan@uni-potsdam.de

* This version: March 2025. The experiment was preregistered in the AEA RCT registry, AEARCTR-0008561. This research was reviewed by the Institutional Review Board of the University of Bamberg. We would like to thank S. Bauernschuster, M. Bagues, E. Bettinger, E. Del Bono, L. Gonzales, H. Hermes, M. Hoekstra, A. Ichino, J. Jessen, A. Kalil, J. List, E. Nix, J. Rincke, J. Sockin, A. Uhlendorf, F. Waldinger, S. Wiederhold, N. Zinovyeva, and participants at various seminars, workshops, and conferences for comments and suggestions. We are also very grateful to U.-M. N¨utzel, T. Seibert, and U. Weber. An earlier version of this paper was circulated with results only for medium-term outcomes. This version includes analysis for longer-term outcomes too.

1 Introduction

Multiple studies document early emergence and persistence of educational achievement gaps by socioeconomic background (Boggess, 1998; Coelli et al., 2007; Cunha et al., 2006; Dustmann, 2004; Ermisch and Francesconi, 2001; Heckman, 2008). A key area in which students from lower-income backgrounds perform worse than students from higher-income backgrounds is reading. Reading skills and text comprehension are crucial, as they predict overall academic achievement and further educational attainment (Alexander et al., 1997; Bigozzi et al., 2017; Herbers et al., 2012; Hernandez, 2011; Kern and Friedman, 2009; Slavin et al., 2009; Sonnenschein et al., 2010). Moreover, they are predictive of long-term labor market success and wages (Green and Riddell, 2003; Hanushek et al., 2015).

This paper evaluates the effects of a randomized reading intervention among disadvantaged children aged 11–12 across Germany on reading engagement, academic performance, socioemotional well-being, and non-cognitive skills. The intervention consists of three key features. First, each child received an e-book reader with a one-year subscription to a digital library of about 1,000 books, allowing them to select books that align with their interests and reading abilities. Second, the devices were pre-activated, set up with password-protected child mode, and delivered directly to the children's homes, reducing barriers to use and limiting access to age-appropriate content only. Third, we provided book recommendations via postcards in collaboration with *Stiftung Lesen*, a charitable foundation that promotes reading, offering guidance similar to that which children from more educated backgrounds typically receive. Additionally to these features, our geographically representative and well-dispersed sample of disadvantaged households mitigates concerns about spillover effects.

Our analysis, conducted 12 months and 30 months after the start of the intervention, shows that children who received the e-book reader are significantly more likely to engage in reading. Specifically, after 12 months, treated children are 18 percentage points more likely to have read an e-book in the last four weeks compared to those in the control

group—a 150% increase from the 12% baseline rate in the control group. This increase in e-book reading does not come at the expense of reading printed books; instead, we observe a nine percentage point (11%) rise in overall reading. Our findings also show that the treatment group engages in reading more frequently, with children in this group being more likely to read on at least two days a week compared to the control children. The positive effect on overall reading engagement remains significant 30 months after the start of the experiment.

The reading intervention also positively impacts children's academic performance, measured through a summary index of reading comprehension test, and German and math grades in the annual school report. 12 months after the start of the intervention, we find an increase in the academic performance index of 10% of a standard deviation. A closer examination of the individual components of this index reveals that the improvements are driven by higher scores in the reading comprehension test and math grades. Specifically, we find a 13% of a standard deviation increase in reading comprehension and a nine percentage point increase in the probability of receiving a good or very good math grade. Importantly, 30 months after the start of the intervention, improvements in reading and academic performance remained significant for treated children, with a 14% of a standard deviation increase in reading comprehension and a 10 percentage point increase in the probability of receiving a good or very good math grade.

To explore the positive impact on math grades, we included four math problems requiring varying amounts of reading in the survey conducted 30 months after the start of the intervention. Our analysis shows that treated children are significantly more likely to correctly solve the problems with higher reading demands, while no statistical difference was found for the other tasks. This suggests that the observed improvement in math grades is likely attributable to better reading comprehension.

Additionally, we assess the impact of the intervention on children's socioemotional well-being using the Strengths and Difficulties Questionnaire (SDQ), a behavioral screening

tool (Goodman, 1997). Socioemotional well-being and skills are closely linked to student test performance and are strong mediators of long-term educational attainment and labor market outcomes (Heckman and Rubinstein, 2001; Cunha et al., 2010; Chetty et al., 2014; Keilow et al., 2019; Attanasio et al., 2020; Gray-Lobe et al., 2023; Del Bono et al., 2024). Our analysis indicates that treated children score 15% of a standard deviation lower on the SDQ after 12 months, indicating a positive shift in their socioemotional well-being. This improvement persists, increasing to 22% of a standard deviation 30 months after the intervention began.

We explore the heterogeneity in the impacts of our intervention by sex, migration background, reading comprehension, reading engagement, and the home learning environment (all measured at baseline). Our analysis shows that the effect of the treatment on reading engagement, academic performance, and SDQ scores is not statistically different across the various groups, indicating that the intervention is broadly effective.

Our paper contributes to the literature in several ways. First, it adds to the body of work on early childhood educational interventions, with a focus on alleviating constraints to reading (Heckman, 2006; García et al., 2020; Fryer and Levitt, 2006; Borman et al., 2007; Chetty et al., 2011; Machin and McNally, 2008). Disparities in access to educational resources are frequently cited as key contributors to the socioeconomic status (SES) gap in reading skills, which are found to emerge as early as 18 months of age (Fernald et al., 2013). The home literacy environment is especially critical, as children from low-SES households typically have limited access to educational materials, early literacy experiences, and rich language exposure (Romeo et al., 2022). Access is further challenged by the fact that the nearest public library for children in our sample is, on average, 4.5 kilometers (2.8 miles) away. Our intervention directly addresses these constraints by providing e-book readers with access to a comprehensive library of age-appropriate books at no cost, significantly reducing both the marginal costs and logistical barriers to obtaining reading materials.

Second, our study contributes to the literature on technology in education by providing

new insights into how specific digital tools can support learning outcomes. Existing research on the use of computers in educational settings often shows that their impact on academic performance is limited or negative, primarily due to potential distractions (Beuermann et al., 2015; Cristia et al., 2017; Fairlie and Robinson, 2013; Fairlie and Kalil, 2017; Leuven et al., 2007; Malamud and Pop-Eleches, 2011).[1] In contrast, e-book readers present a unique advantage by minimizing distractions, as they have no browser access, games, or apps, but do offer immediate access to a broad range of books. This focus on a distraction-free digital tool provides a clearer picture of how technology can positively impact educational outcomes. To the best of our knowledge, this study is the first large-scale randomized controlled trial to assess the impact of e-book readers as a tool to enhance reading outcomes, offering new insights into the potential of digital resources in educational interventions.[2]

Third, our intervention targets children directly. While parenting style and parental time investments are important factors in the human capital production function, our strategy circumvents parental input, focusing instead on the child's direct interaction with educational content. Research suggests that after the age of ten, a child's own time investment becomes more critical than the time investment of others (Del Boca et al., 2014). By focusing on the child's direct interaction with the e-book readers, we aim to emphasize independence in learning, potentially moderating the influence of parental background.

Fourth, the large selection of books available through the subscription alleviates concerns about selecting books that are appropriate for children's interests and reading levels. We view this flexibility as crucial, as it increases the likelihood of a meaningful match between a child's reading ability and the book's readability.[3] The literature on direct reading

[1] For an extensive overview, see Bulman and Fairlie (2016) and Escueta et al. (2020).

[2] Although a few studies have explored the use of e-book readers in classrooms with teacher guidance (Hashim and Vongkulluksn, 2018; Long and Szabo, 2016; Akbar et al., 2015), these studies are largely descriptive, with small sample sizes, and lack the rigor of randomized controlled trials.

[3] For example, when a child starts reading a book and realizes that the text is too easy, too difficult, or uninteresting, they can switch to another book within seconds with just a few clicks.

interventions, such as book and textbook provision, suggests that their effectiveness hinges on how well the books provided meet students' specific needs (Holden, 2016; Kremer et al., 2013; Abeberese et al., 2014; Guryan et al., 2016; Glewwe et al., 2009).

Fifth, we add to the literature studying non-cognitive effects of educational interventions (Gray-Lobe et al., 2023; Alan and Ertac, 2018). Our findings show that reading has significant positive effects on children's SDQ score. More generally, our results suggest cross-productivity between reading and other academic and socioemotional skills, highlighting how investments in one area of education can lead to broader benefits.

Our study presents a scalable and cost-effective intervention, providing e-book readers and access to age-appropriate digital libraries, which can be implemented across diverse educational settings. In numerous countries, there is widespread concern about poor school outcomes and growing achievement gaps early in life. Reading literacy, in particular, fell sharply between 2015 and 2022 in many EU member states (OECD, 2023).[4] Similarly, the 2024 Nation's Report Card reveals continued declines in reading scores for American students (National Assessment of Educational Progress, 2025). We observe that the present intervention results in a significant increase in reading engagement among children, positively influencing their reading skills, numeracy, and socioemotional well-being. These foundational skills are not only crucial in their own right, but are also essential for the development of higher-level cognitive abilities and have implications for further learning and academic success.

The rest of the paper is organized as follows: Section 2 describes the experimental design and our intervention, while Section 3 outlines the empirical strategy. Section 4 presents the main results, as well as potential heterogeneity and robustness checks. In the robustness section, we specifically discuss aspects of external validity, experimenter demand effects, and multiple-hypotheses testing. Section 5 presents additional findings, including a cost-benefit analysis. Finally, Section 6 concludes.

[4]From 2018 to 2022, the average reading performance across 35 OECD countries declined by ten points. This is comparable to the learning equivalent of around three-quarters of a school year.

2 Experimental Design

2.1 Sample

To draw our sample for the experimental intervention, we rely on data from the Social Security Records of the German Federal Employment Agency. In contract to recruiting the children at schools, this allows us to precisely target children with a certain household income level across Germany. When selecting our experimental population, we focus on low-income and welfare-dependent households.[5] We conducted the first baseline survey in November 2020. The study collected an array of relevant socioeconomic and academic variables, including a reading comprehension test, school grades, and educational aspirations. We also conducted surveys with the parents, allowing us to link children and their parents both at baseline and in the subsequent follow-up surveys.

The survey data has the additional advantage that the participants (and their schools) are spread across different federal states in Germany, which alleviates concerns about spillover effects between the treatment and the control group. Due to the fact that the children are dispersed across Germany, we can also address an important aspect of external validity: our results are not limited to a small regional area or a single federal state or municipality but include children from socially disadvantaged families in all 16 federal states.[6]

Our experimental sample consists of 1,000 children aged 11–12. The children are assigned to the equal-sized treatment and control groups using a stratified randomization. We stratify our sample based on the following variables: (i) an indicator of receiving welfare, (ii) sex of the child, (iii) school grade of the child, (iv) migration background, and (v) median split of baseline reading comprehension.

The focus on children aged 11–12 is motivated by research showing that this age marks a

[5]For the purposes of this study, we define low-income households as those below the 60th percentile of the income equivalence distribution. In selecting the sample, we excluded children with siblings in the same grade and those who did not complete our baseline reading comprehension test.

[6]Figure 1 shows the geographic spread of the treatment and control group, and online Appendix Table A1 presents the sample shares by federal state.

shift toward greater independence in cognitive and behavioral processes. Studies indicate that after the age of ten, a child's personal time commitment becomes more important (Del Boca et al., 2014). According to the stage model of reading by Chall (1983), children at this age are more likely to have the necessary decoding skills to read books on their own, without needing help from teachers or parents.[7] In contrast, during the early stages of childhood, the formation of a child's human capital is significantly influenced by parental involvement. Research has consistently shown that parents with a lower socioeconomic background tend to allocate fewer resources to their children's education, contributing to the well-documented achievement gaps across various skills (Fryer and Levitt, 2004; Guryan et al., 2008; Caucutt et al., 2020). A recent reading intervention aimed at enhancing parental engagement in low socioeconomic households among children around the age of four has shown mixed effects on literacy skills (Kalil et al., 2023). Our study, however, intentionally targets children who have reached the age at which they are capable of independently using e-book reader devices. In so doing, we aim to mitigate the reliance on parental involvement and school quality in children's learning and human capital development.

Table 1 shows the summary statistics for relevant pre-treatment variables for the treatment and control groups. In both groups, roughly 50% of participants are girls. The average grades in German and math are 2.4 and 2.3, on a scale from 1 (very good) to 5 (poor), and thus close to the average in Germany (Lettau, 2021).[8] A total of 78% of parents have at least an intermediate school-leaving qualification and about 80% were born in Germany. The share of households with very low socioeconomic status, defined as those receiving welfare, is just above 50% as a result of our oversampling. As the p-values in column (4) show, the randomization was successful and the baseline characteristics are balanced between the treatment and the control groups.

[7]She refers to this stage as when readers "start on the long course of reading to "learn the new"—new knowledge, information, thoughts and experiments" (Chall, 1983: p. 20).

[8]Children can also receive the grade 6 (failing), however this rarely happens in this age group.

2.2 Intervention

One year after the baseline data collection, in November 2021, the households assigned to the treatment group received a letter announcing that they would receive an e-book reader with a free subscription to more than 1,000 age-appropriate books for an entire year. Parents and children who did not wish to receive the device could opt out. Of the 500 households assigned to the treatment group, 42 opted-out.[9] In mid-December 2021, households that did not opt-out received the pre-activated e-book readers.[10]

Our intervention incorporates several key features, which we outline below. First, each device comes with a pre-activated one-year subscription to a digital library containing approximately 1,000 books. Thus, children are granted considerable autonomy to select books that match their individual interests and proficiency levels. We view this freedom as critical, as it increases the likelihood of a meaningful match between children's reading abilities and book readability.

Second, the devices are pre-configured before dispatch. Our team created individual profiles for each child and enabled a password-protected child mode, ensuring that neither children nor their parents can deactivate the child protection settings. This configuration restricts the use of the e-book readers to age-appropriate books only.[11] This further reduces the barriers to using e-book readers, marking the present intervention as an effort to lower the obstacles to starting and maintaining regular reading habits.

Third, we sent postcards with book recommendations throughout the first four months of the intervention. The first was mailed in the first week of January 2022 and the last in the last week of April 2022. To ensure age-appropriate recommendations, we collaborated

[9]The original letter (translated from German) is displayed in the online Appendix. We did not replace households that opted out. Table 2 shows the characteristics of the households that opted out compared to those that did not. The table shows that the households that opted-out do not significantly differ from those that did not.

[10]By early 2022, most pandemic-related school closures had been lifted, and the majority of schools had resumed in-person teaching, with students attending classes on-site.

[11]While parents could theoretically reset the devices to factory settings, this would result in the loss of the book subscription. Additionally, parents were informed that the e-book reader is intended exclusively for their child's use.

with *Stiftung Lesen*, the leading German charitable foundation that promotes reading. Treated children received two types of postcards. These were illustrated postcards including either the name of the recommended book and a short description, or the name of the book and a link to a video produced by *Stiftung Lesen*. In these videos, members of the *Stiftung Lesen* team stage a conversation about the book content. We host the videos on a secure platform, where only the treated children could access them.[12] The intention here is to offer book recommendations to children in the treatment group, providing them with inputs akin to those children from more educated backgrounds receive from their parents. For children from less educated families, like those in our sample, such parental guidance is likely less common.[13] Additionally, with the devices providing access to 1,000 books, the reading tips could make it easier for the participants to select their first books.

Our intervention thus leverages several features to address multiple barriers to reading engagement in a highly scalable manner. Thus the current study assesses the overall impacts of our bundled treatment described above.

2.3 Timeline

Figure 2 provides an overview of the intervention and data collection timeline. Baseline information was collected in fall 2020, and the intervention began in late December 2021. Further data was collected through three follow-up online surveys. The first follow-up was conducted in summer 2022, approximately six months after the intervention started, with responses from children and one parent. The second follow-up took place between February and April 2023, shortly after the library subscription expired. The endline survey (follow-up III) was conducted in summer 2024, about 30 months after the start of the e-book reader intervention. Importantly, although the library subscription ended after one year (December 2022), children kept the e-reader devices.

[12]Online Appendix Figure A1 shows the exact mailing days of the postcards and displays an example of both types of postcards. The videos can be accessed under `vimeo.com/user/156363831/folder/6212899`.

[13]To ensure continued contact with the control group children, we sent them an illustrated postcard thanking them for participating in the study, as well as a greeting card for the New Year and summer holidays.

In the first follow-up, response rates were 76% for the treatment group and 74% for the control group, with no statistically significant differences between the groups, as shown in Table 3. Similarly, the second follow-up recorded response rates of 73% for the baseline treatment group and 71% for the control group, again with no significant differences. However, response rates declined slightly in the third follow-up, particularly in the control group, resulting in statistically significant differences in attrition between the treatment and control groups.[14] To assess whether attrition differed by key characteristics, we present results on selective attrition in online Appendix Table A2. This analysis shows that attrition was not selective with respect to baseline frequency of reading, reading comprehension, or school grades.

3 Empirical Strategy

We estimate and report both intention-to-treat (ITT) and treatment-on-the-treated (TOT) effects, with the ITT effects being our preferred specification. The equation for estimating the ITT is as follows:

$$Y_i = \alpha + \beta treat_i + \gamma' strata_i + \mu' \mathbf{X_i} + \epsilon_i \tag{1}$$

where Y_i is the outcome of interest for child i, $treat_i$ indicates the initial assignment to the treatment or control group, including those who opted out, $strata_i$ stands for a vector of fixed effects of randomization strata, year and month of birth fixed effects, and ϵ_i is the error term. To ensure our standard errors are accurate and not influenced by limited sampling, we report p-values from randomization inference (Young, 2019). The randomization test works by permuting the treatment assignments to construct the

[14] To encourage participation in the follow-up surveys, children were sent an invitation by mail with a prepaid 10 euro voucher redeemable at online or onsite retail stores. Upon survey completion, they received an additional 10 euro voucher. Non-participating children received a reminder letter after four weeks. After eight weeks, a second reminder was sent, which included an increased participation incentive of 30 euros. Parents of non-participating children were also contacted by phone to encourage survey participation.

permutation distribution of the test statistic under the null hypothesis of no treatment effect. The p-value is then calculated as the fraction of permuted statistics the same as or more extreme than the observed statistic.

Although our primary focus is on the ITT parameter, we also estimate the TOT effect, using the initial randomization assignment as an instrument for treatment take-up (Angrist et al., 1996). In the context of this experiment, we define treatment take-up as the usage of the e-reader devices at the extensive margin. To operationalize this measure, we rely on a survey question asking children whether they use the device. If a treated child reports using the device either alone or with their parents, we code treatment take-up as one.[15] The first and second stage equations are as follows:

$$T_i = \tau + \delta treat_i + \lambda' strata_i + \mu' \mathbf{X_i} + \eta_i \tag{2}$$

$$Y_i = \nu + \theta \hat{T}_i + \gamma' strata_i + \mu' \mathbf{X_i} + v_i \tag{3}$$

where T_i is the take-up, and all other variables are defined as above. The identifying assumption is that treatment assignment does not affect those individuals who were assigned to treatment but did not participate. Given that our sample is drawn from across Germany, it is highly unlikely that there were spillovers between treated children. However, it is possible that the treatment assignment variable may be influenced by an experimenter demand effect. For example, some children might report that they use the e-book reader, even if they do not actually do so in practice. As a result, we place greater emphasis on the ITT estimates, which are our preferred findings because they are less susceptible to experimenter demand bias and are the more relevant policy parameter.

[15]This information was collected during the first follow-up in summer 2022. Children were asked who typically uses the e-book reader, with response options including: (i) Me, (ii) Me, with my parents, (iii) my parents, (iv) my siblings, (v) my friends, and (vi) others. The majority of children reported using the device by themselves. Take-up is coded as one if the children reported using the device alone or with their parents, and zero otherwise. The TOT estimations include slightly fewer observations, as we exclude opt-outs and respondents who did not answer this question.

3.1 Outcome variables

Our primary outcomes of interest are in four domains: reading engagement, academic performance, socioemotional well-being, and self-efficacy. In the domain of reading engagement, we examine both the number of books read and the frequency of reading. Specifically, we consider whether children reported reading at least one e-book, printed book, or either type of book in the past four weeks. Additionally, we assess whether children typically read on two or more days per week.[16] In our pre-registration, we specified that we planned to examine weekly reading frequencies and average reading duration. However, the latter—assessed by means of an open-ended question—resulted in substantial missing data and numerous implausible responses. Consequently, we do not report results on daily reading duration.

To analyze the impact on academic performance, we estimate the effects on reading comprehension and school grades in German and math, as pre-registered. To assess the reading comprehension of the participating children, we use a standard reading test developed by the Institute of Educational Quality Improvement (IQB). This test involved reading an age-appropriate text and answering several single- and multiple-choice questions. To account for the changing age of the children, we administered a different test in each survey wave, adjusting the difficulty level accordingly. The answers were used to create standardized scores, facilitating comparison across different test versions. For our primary analysis, we recode the (self-reported) grades from the school annual report into a binary indicator, where a grade of one ("very good") or two ("good") is coded as one, and all other grades are coded as zero.[17] In keeping with our pre-registration, we also analyze the effect of the treatment on children's well-being (as measured through Strengths and Difficulties Questionnaire), self-efficacy (grit and locus of control), and educational aspirations.

[16]The original phrasing of the questions, translated from German, is presented in the online Appendix.

[17]Our wording of "very good" and "good" grades thus follows the classification commonly used in schools.

The SDQ is a behavioral and emotional screening tool, covering the domains of emotions, conduct, attention and inattention, and peer relationships. Higher scores indicate more problems and lower regulation skills. We expect reading to affect the SDQ score of children for multiple reasons. First, reading can improve children's self-regulation skills, which are directly related to the SDQ conduct and attention subdomains. Struggles with reading can lead to feelings of inadequacy and frustration, which can undermine a child's belief in their ability to succeed, translating into emotional and behavioral difficulties (Mak and Fancourt, 2020). Second, reading intervention can improve cognitive and language skills, which have a positive impact on the SDQ emotions and attention subdomains. For example, improved language skills help children express their emotions more effectively, reducing emotional distress (Eisenberg et al., 2005; Roberts et al., 2015). Third, reading offers children a relaxing and enjoyable activity, while engaging with complex stories and novels can enhance their ability to take others' perspectives, positively influencing the SDQ peer relationships subdomain.

We thus have 11 outcomes, which may lead to a multiple-inference problem. We employ several approaches to address this issue. The first is to reduce the number of tests conducted, by aggregating multiple outcomes into indices. Indices also reduce noise, thereby enhancing signal detection and the precision of estimates. We construct two indices for reading engagement and for academic performance. In the main table, we report results for a summary index that averages the equally weighted z-scores of its components, with each measure adjusted so that higher values represent better outcomes (Kling et al., 2007). In the robustness section, we report the effect on an inverse-covariance weighted index following Anderson (2008). In addition to p-values, we also report the sharpened false-discovery rate (FDR) q-values (Benjamini et al., 2006; Anderson, 2008). The FDR controls the expected proportion of falsely rejected hypotheses (type I errors). In the robustness section, we further address possible multiple-inference problems by reporting step-down adjusted p-values, p-values from bootstrap-based procedures, and p-values from vector-based resampling methods (Westfall and Young, 1989; Romano and Wolf,

2016; List et al., 2019).

4 Results

4.1 Main results

Table 4 reports the estimated intention-to-treat effects on the two indices (*reading engagement* and *academic performance*) and on the individual outcome measures, including p-values from randomized inference[18], and false discovery rate q-values (share of expected false positives). Estimated treatment-on-the-treated (TOT) effects are presented in online Appendix Table A3.

Our analysis reveals a positive and significant change in the reading engagement index among the treated children, both 12 months and 30 months after the intervention began. The ITT point estimate in the first row of Table 4 indicates that providing e-book readers increased children's reading engagement by 10.5% of a standard deviation 12 months after the program's start (column 1). This impact is precisely estimated, as shown by the randomization inference p-value ($p = 0.000$). Although the effect on reading engagement 30 months later is somewhat smaller—around 6% of a standard deviation—it remains statistically significant ($p = 0.016$). This sustained impact is worth noting, considering the subscription to the digital library expired after the first year, though the children were allowed to keep their e-reader devices.

We then turn to analyze the effect on the individual components of the reading engagement index, starting with whether a child reports having read at least one e-book in the past four weeks. The coefficient in column 1 of Table 4 shows that children in the treatment group are 18 percentage points more likely to have read at least one e-book—an increase of approximately 150%. This result is not surprising, as only 12% of children in the control group reported reading e-books. In the longer term, the likelihood of reading e-

[18]We do not report the standard p-values. We note that these tend to be smaller than the p-values from the randomized inference for most outcomes.

books remains positively and significantly higher for treated children, with a 14 percentage point (88%) increase ($p = 0.000$).

To assess whether the provision of e-book readers leads to an overall increase in reading rather than simply replacing printed books with e-books, we analyze the effect on both the number of printed books and the total number of books read. The third and fourth rows in Table 4 display the results of our analysis. 12 months after the start of the program, our findings indicate no evidence for substitution. On the contrary, we observe that children in the intervention group are 6.4 percentage points (9%) more likely to report reading a printed book in the last four weeks compared to those in the control group ($p = 0.040$). In total, the treated children have a nine percentage point (11%) higher likelihood of having read at least one book in either format within the last four weeks. This suggests that the provision of e-book readers has a positive impact on reading habits, leading to a noticeable increase in overall book consumption among the treated children. In the longer term, the positive difference with respect to printed books between the treatment and the control group fades out.

To measure reading engagement, also in terms of frequency, we analyze the impact on the variable capturing how often children read in a typical week. The outcome in row five is equal to one if the child reads on at least two days in a typical week, and zero otherwise. Our estimated coefficient shows an increased likelihood of 8.5 percentage points (approximately 11%). This positive difference in reading frequency persists over time, constituting a 9 percentage point increase (14.7%) 30 months after the start of the program. It is worth noting that while reading frequency increased among treated children, the share of children in the control group who reported reading at least two days per week declined from 74% in follow-up II to 61% in follow-up III.

Next, we study the impact on academic performance. The ITT estimates on the academic achievement index suggest that treated children experienced an improvement of around 10% of a standard deviation 12 months after the initiation of the program. This

positive effect persists 30 months after the start of the experiment, remaining statistically significant at the 10% level ($p = 0.072$).

We present the effects on the components of the index next. The results suggest an increase in treated children's reading comprehension of about 13% of a standard deviation, both in the medium and longer term, which is significant at the 10% level ($p = 0.080$ in both specifications). The q-value of the effect 30 months later indicates that with a false discovery rate threshold of 10%, the coefficient would not be considered statistically significant.

Our analysis now turns to evaluating the impact of the intervention on children's grades in German and math. The final grades are based on school report cards recorded in surveys 12 months and 30 months after the start of the intervention. We focus on binary outcomes, indicating whether a child received a very good or good grade, while results for discrete grades are presented in the robustness section. The analysis reveals no statistically significant effect on German grades, while math grades among treated children show significant improvements in both the medium and longer term. Specifically, the probability of receiving a good or very good grade in math increased by 8.6 percentage points after 12 months ($p = 0.020$) and by 10 percentage points after 30 months ($p = 0.024$), compared to the control group mean of 47% and 40% respectively.[19]

Several studies, including ours, find that educational interventions tend to have a positive impact on math grades, but not on language grades. For example, Bettinger (2012) shows that financial incentives improve math scores, while English scores remain unaffected. Similarly, Bergman (2021) finds that providing parents with information about missed assignments boosts math test scores, but has no effect on English scores. To explore the positive effects on math grades in our context, we asked children at endline (follow-up III) to solve four math problems—two of which were text-heavy, while the other

[19]In unreported regressions, we also controlled for an indicator for whether the child is attending the academic school track (Gymnasium) versus any other school track, as achieving a good or very good math grade is typically more challenging in the academic track. This does not affect our main results for school grades.

two involved minimal text. This design allows us to assess whether the improvement in math performance was driven by better reading comprehension, enabling students to understand and solve the problems more efficiently. If this were the case, we would expect a larger performance gap between treated children and those in the control group with regard to the text-heavy tasks, compared to tasks with minimal text. Figure 3 plots the estimated coefficients and their respective confidence intervals from a regression analogous to equation 1, where the outcome variables are math scores. As shown in Figure 3, we observe a significant positive treatment effect on math tasks with a large amount of reading, while there are no significant differences in performance on math tasks that require minimal reading between the treatment and control groups. The ITT estimate on the math tasks with a large amount of reading suggest that treated children experienced an improvement of nearly 20% of a standard deviation. This pattern suggests that the improvement in math grades is attributable to the treatment group's ability to comprehend text-heavy math problems rather than to improvements in core math skills.

Next, we present the results on the well-being of the children, using the SDQ. We find a statistically significant decrease in the SDQ scores, implying an improvement in socioemotional well-being. The ITT point estimate for the SDQ score suggests a reduction in the score of 15% of a standard deviation 12 months after the reading intervention was initiated ($p = 0.064$). The effect on socioemotional well-being strongly persists over time, with an even larger absolute magnitude of the effect size, showing a decline in the SDQ score of 22% of standard deviation after 30 months ($p = 0.004$). These findings suggest that the treated children experience an improvement in their socioemotional well-being. The present results primarily rely on the children's self-assessments, as parental assessments of children's non-cognitive skills may be directly influenced by the parents' own skills (Del Bono et al., 2020). For comparison, we also analyze the effect of the treatment on the SDQ as reported by parents. Our results indicate a statistically significant decrease in the SDQ score based on parental assessments as well.[20] These findings relate

[20]When analyzing the parents' assessment of their child's SDQ, we find a reduction of -0.170 of a standard deviation in the medium term, with a p-value of 0.036.

to earlier research emphasizing that early investments in children's development can yield compounding benefits over time (Cunha and Heckman, 2007; Heckman, 2008).

As pre-registered, we also analyze the impact on children's grit and locus of control. It is possible that the reading experience and improvement in reading skills could positively influence children's confidence in their abilities and perseverance. However, we do not find evidence that the reading intervention affects these outcomes either 12 months or 30 months after the intervention began. In line with our pre-registration, we also examine the effect on educational aspirations. The variable in the last row of Table 4 equals one if the child aspires to achieve the academic school track qualification, and zero otherwise. The results, however, indicate no effect on the likelihood of aspiring to this higher track in either the medium or long term.

We present the TOT effects in online Appendix Table A3. The findings here confirm the main ITT results. As expected, the TOT effects are larger in magnitude, suggesting that the effect on children who did indeed use the devices is stronger. The estimates suggest an increase in reading engagement index of 7–13% of a standard deviation, an improvement in academic performance of 11–13% of a standard deviation, and a decrease in the SDQ score of 13–18% of a standard deviation. The TOT effects also show a statistically significant increase of 16% of a standard deviation in reading comprehension of the treated children in both follow-ups.

Overall, the results in this section show that the e-book reading intervention increased children's reading engagement and improved their academic performance both in the medium and longer term, and has a positive effect on children's well-being.

4.2 Heterogeneity and mechanisms

In this section, we examine the potential differential impact of the e-book reader intervention across various baseline characteristics and the possible mechanisms for the positive effects of our reading intervention. For the former, we estimate a modified version of the

intention-to-treat model (equation 1), as follows:

$$Y_i = \alpha + \beta_1 treat_i + \beta_2 group_i + \beta_3 treat_i \times group_i + \gamma' strata_i + \mu' \mathbf{X_i} + \epsilon_i \qquad (4)$$

where $group_i$ is an indicator equal to one if a child is in the given group, and zero otherwise. The rest of the variables are defined as in equation (1).[21] We report heterogeneous effects for the following predetermined (baseline) dichotomous variables: female, migration background, low reading score, low reading engagement, and low home resources. The dummy variable "migration background" is equal to one if parents do not speak German at home, and zero otherwise. The variable "low reading score" is equal to one if children's reading score at baseline is below the median score, and zero otherwise. The variable "low reading engagement" is equal to one if children report that they read less often than once a week, and zero otherwise. Lastly, the dummy variable "low home resources" is equal to one if children's home environment score is below the median, and zero otherwise. The home resources score is predicted by the first principal component based on the following information (reported by children) at baseline: whether children have (i) their own desk for studying; (ii) their own room; (iii) an e-book reader that they can use; (iv) books that are helpful for homework; (v) (fast) internet at home; (vi) their own computer. We interpret a low home environment score as an indicator of a home that is not very conducive to learning.

We study the effect heterogeneity across these variables because each of these dimensions can significantly influence reading engagement and academic performance and possibly interact with the treatment. For example, the effect might differ by sex, as girls tend to perform better in reading (Logan and Johnston, 2010). Anderson (2008) analyzed early educational programs (Abecedarian, Perry, Early Training Project) and finds that the interventions provided significant short- and long-term benefits for girls, but not for boys. Children with a migration background and with low levels of reading skills

[21] In cases where we examine heterogeneity in the baseline characteristics used for stratification, we do not include an additional fixed effect for that group.

or reading engagement at baseline, might benefit more from the intervention. Beyond individual characteristics, a substantial body of literature suggests that the family and home environment play a crucial role in children's development (Angrist and Lavy, 2009; Bettinger, 2012; Guryan et al., 2014; Bergman et al., 2018), and these environmental factors may interact with intervention effectiveness.

Figure 4 plots the estimated coefficient $\hat{\beta}_3$ from equation 4 and the corresponding 90% and 95% percent confidence intervals for the outcome reading engagement index measured 12 months and 30 months after the start of the intervention. The results in both panels indicate no significant heterogeneous effects. In the top panel, the estimated interaction effects are all relatively small, close to zero, and not precisely estimated. The point estimates in the bottom panel of Figure 4 provide some suggestive evidence that girls, children with a migration background and those with lower levels of home resources might benefit slightly more from the reading intervention in the longer run. However, the confidence intervals are relatively large and our study may lack sufficient power to estimate heterogeneous effects precisely.[22] Overall, while we do not find strong evidence of significant heterogeneous effects, we cannot completely rule out the possibility that certain groups may benefit more from the intervention, although the differences do not seem substantial.

We further present exploratory evidence on the channels behind our findings. One potential mechanism for increased reading engagement through our intervention is the inherent appeal of an electronic device and advantage of easy, instant, and low-cost access to a large variety of books. Hence, we explore children's satisfaction with the devices. We surveyed children in the treatment group about their satisfaction with the e-book readers and their enjoyment in using the device. The distribution of their responses is illustrated in Figure 5. The descriptive analysis indicates a high overall level of satisfaction with the e-book

[22]Online Appendix Figures A2 and A3 display the corresponding results for the outcomes academic performance index and SDQ score, respectively, where we also do not find significant heterogeneity. In unreported regressions, we also looked at heterogeneity by parent's education and employment, but again find no significant heterogeneity.

readers, an increase in reading frequency since receiving the device, and a notable 60% of children reporting greater enjoyment in reading than before. These findings provide suggestive evidence that e-book readers can enhance the enjoyment of reading.[23] In fact, at endline more than half of the children reported that they continued using the devices, even though the prepaid subscription has expired.

We also conduct a mediation analysis following the approach of Heckman and Pinto (2015) and report the details in the online appendix. We decompose the treatment effects across our main outcomes—reading engagement, academic performance, SDQ score, and present the results in Figure B2 in the online appendix. The variables we measure do not appear to explain a substantial portion of the effect on reading engagement. However, we are able to explain 20% of the effect of academic performance 12 months after the intervention began, and 60% after 30 months, with reading engagement and self-concept being the most important factors. In contrast, while we can explain 40% of the effect on SDQ score 12 months post-intervention, the contribution of mediators to the SDQ effect becomes significantly smaller 30 months later.

Overall, the devices themselves appear to play a role, with children reporting high satisfaction and increased enjoyment of reading, suggesting that the e-book readers may help sustain engagement. Reading engagement primarily accounts for the improvements in academic performance, while its contribution to SDQ effects is observed only in the medium term.

[23] In the first follow-up survey, both control and treatment group participants were asked whether they had read any of a list of ten books. Six of these books were actively recommended to the treatment group via postcards and videos, while the control group received no such recommendations. The remaining four books were not mentioned to either group. In online Appendix Figure A4, we show coefficients from ten separate regressions, each testing whether a child reported reading a specific book. We find a statistically significant increase in the likelihood of the treated children reading the recommended books.

5 Additional evidence

5.1 Robustness

We assess the robustness of our main findings through several sensitivity checks. First, we address potential experimenter demand effects. Second, we introduce additional control variables into our main regression specification. Third, we present results from Poisson regressions and ordered probit models for discrete outcome variables that were dichotomized in the main analysis. Fourth, we report findings using an alternative summary measure: the inverse-covariance-weighted summary index (Anderson, 2008). Additionally, we provide p-values adjusted for multiple hypothesis testing (Westfall and Young, 1989; Romano and Wolf, 2016; List et al., 2019) and discuss the stable unit treatment value assumption (SUTVA) along with various aspects of external validity.

Experimenter demand effect

Self-reported reading behavior, while informative, may be influenced by experimenter demand effects, whereby participants adjust their responses to align with perceived expectations. Ideally, objective metadata directly collected from the reading devices would provide a more reliable and unbiased measure of reading behavior. However, since the control group did not receive e-reader devices, we rely on self-reported survey data for comparing the reading behavior between treated and control children. To assess the extent of potential misreporting in the self-reported data, we compare it to the objective usage data stored in the online dashboard of the reading devices of the treated children. Unfortunately, several technical issues resulted in incomplete dashboard data, such that we are only able to perform this comparison for a sub-sample of participants. Figure 6 provides a descriptive cross-validation of children's survey responses with data collected from the dashboard. The y-axis displays the number of books read in the last four weeks based on recorded (objective) reading behavior. The x-axis displays the number of e-books the children reported having read during the last four weeks in the survey. The figure shows

a positive and strong relationships between both measures, with the subjective measure of the number of e-books read being a rather lower bound. We interpret this as first suggestive evidence that a large experimenter demand effect in children's self-reported reading behavior is rather unlikely.[24]

To further address concerns about experimenter demand effects, we provide additional evidence on children's reading engagement, as reported by both children and parents. In the first follow-up we asked identical questions to both children and parents, allowing us to assess the reliability of children's responses and minimize concerns about demand effects. Table 5 summarizes the results for reading engagement, reported separately for children and parents. The binary variables are defined as described in section 3. The estimates indicate that there was a significant positive impact on reading engagement as reported by both parents and children, with point estimates aligning closely. For instance, both children and parents report an increase in the likelihood of having read an e-book in the past four weeks of just over 40 percentage points, with randomized inference p-values of 0.000. This close alignment of children's and parents' responses provides reassuring evidence of the reliability of children's survey data.

Additional control variables

Next, we introduce additional control variables into our main regression specification. In Table 6, we successively add control variables, which we refer to as Set 1, Set 2, and Set 3. All additional explanatory variables are measured at baseline, prior to the intervention. Set 1 includes the reading comprehension score, grades in math and German, and grit. Moving to Set 2, we extend the control variables to include indicators for whether the children have their own desk, their own room, their own e-book reader, access to the internet, and whether the primary language spoken at home is German.

[24]We were able to retrieve reliable information for 80 unique accounts. To ensure consistency, we matched the survey dates with the dashboard data so that the last four weeks refer to the same time period in both sources. It is important to note that the number of e-books recorded in the dashboard is often higher than the self-reported number in the survey, as the dashboard tracks books opened, regardless of whether they were finished. In contrast, the survey specifically asks children about the books they had finished reading.

Set 3 further includes fixed effects for federal state of residence, parent's employment status, and parent's education level. We deal with missing values for control variables by imputing them with sample averages and accounting for them using dummy variables in our analysis. Table 6 presents the ITT estimates and the permuted p-values. Overall, the table shows that our estimates are not sensitive to the inclusion of additional control variables. However, the effect on reading comprehension in the third follow-up looses its significance, although the magnitude of the estimate does not change much.

Poisson and ordered probit

Table 7 presents the results from Poisson and ordered probit models, where we analyze the discrete variables without converting them into binary form. We use Poisson regressions to examine the effect of the treatment on the number of (e-)books read in the past four weeks and the number of days children report reading during a typical week. The Poisson regression results align with our main findings from Table 4. For instance, the ITT estimates indicate a 74% increase in the number of e-books read in the last four weeks and a 9.3% increase in reading frequency. These results also show that the positive effects on reading engagement persist 30 months after the start of the experiment.

To analyze grades in math and German, we use ordered probit models. We reverse the grading, such that higher numbers mean better grades. Table 7 reports the results. Again, the findings confirm no significant effect on German language grades, but a positive impact on math grades.

Inverse-covariance-weighted summary index and multiple hypothesis testing

An alternative way to aggregate multiple outcome variables is by using the inverse-covariance-weighted summary index (Anderson, 2008). This approach combines multiple related outcomes by weighting each standardized outcome by the inverse of its variance-covariance matrix, effectively giving less weight to outcomes that are highly correlated and more weight to outcomes that are less correlated. Table 8 presents the estimates for

the indices reading engagement and academic performance. A comparison of the ITT results in Tables 4 and 8 shows that the weighted index also suggests a positive impact on both reading engagement and academic performance, though with larger effect sizes. We prioritize the results in Table 4, as they provide more conservative estimates.

To further examine potential issues of multiple hypothesis testing, we present three alternative testing methods in Table 9. We start by reporting step-down adjusted p-values in line with Romano and Wolf (2016). Second, following List et al. (2019) we report p-values from bootstrap-based procedures. Third, the last two columns in Table 9 report p-values for multiple hypothesis testing using vector-based resampling methods (Westfall and Young, 1989). Together with the q-values from false-discovery rates in Table 4 and the results from both summary outcome indices (Kling et al., 2007; Anderson, 2008), findings provide a comprehensive picture of multiple hypothesis testing.

Overall, the variables measuring number of (e-)books, reading frequency, and good math grades remain robustly significant across all methods of correction in both the short and longer term. SDQ scores (socioemotional well-being) also show consistent significance, strengthening over time. Although multiple hypothesis testing corrections push the p-values for reading comprehension beyond conventional thresholds, the unadjusted p-values and FDR q-values indicate a positive effect. Given that MHT corrections can be overly conservative, especially in settings with small sample sizes and correlated outcomes, we do not rule out an impact on reading comprehension. German grades, grit, locus of control, and educational aspirations remain consistently insignificant across all methods and surveys.

The stable unit treatment value assumption (SUTVA)

A typical concern in educational interventions is the potential violation of the stable unit treatment value assumption (SUTVA). This assumption would be violated if the control students also benefited from the reading intervention. Since randomization of the treatment occurred at the household level, it is highly unlikely that treatment and control

students were in the same class or school. Further, as shown in Figure 1, the experimental participants are geographically dispersed across different regions of Germany and it is therefore very unlikely that treatment and control children know or interact with one another. The average geographical distance between children in the treatment group and children in the control group is slightly more than 300 kilometers (186 miles).

External validity

Our study focuses specifically on low-income children. While this might limit the generalizability of our findings to other income groups, we believe that low-income children represent a particularly important demographic group to study because they have limited access to educational resources, early literacy, and language experiences. Consequently, these children are likely to benefit most from such interventions, rendering our results especially meaningful within this context (Agostinelli and Wiswall, 2025).

Several features of our study enhance the external validity of our findings within the context of low-income families. First, we use registry data as an effective tool for obtaining a representative sample. We invited children and their parents to participate by mail, using data from the German Federal Employment Agency—a method that improves representativeness, as noted by List et al. (2023). Second, our sample includes children from across Germany (see Figure 1), ensuring that our findings are not confined to a single region or federal state, but rather pertain to children from families in all 16 federal states. Third, as List et al. (2023) emphasize, it is important to consider whether a study's setting is appropriate for generalizing its results. We argue that our intervention minimizes disruption to children's daily routines, as they are able to read books at home—the setting where reading for pleasure most often occurs (**?**).

5.2 Comparisons with other reading interventions and cost-benefit analysis

In this section, we compare our findings with other randomized reading interventions. We primarily focus on RCTs involving children of a similar age. Additionally, we provide

cost-benefit analyses and briefly discuss them in the context of other educational studies.

Our study is most closely related to interventions in which children are directly provided with reading resources. One strand of this literature is book giveaway programs, which often target pre-school children and many of which are not causally evaluated (e.g. Dolly Parton's Imagination Library). Among studies is this area that aim at estimating causal effects, Abeberese et al. (2014) examine a 31-day reading intervention in schools in the Philippines. The key components of the program were the provision of age-appropriate books to schools, teachers being trained on how to incorporate reading, and teachers being supported in initiating reading in the classroom. The authors find an immediate increase in reading skills of 0.13 standard deviations among 4th graders. The effect decreases to 0.06 standard deviations three months after the program.

Kim and Guryan (2010) study a more structured summer reading intervention (READS) for low-income Latino children in California. At the end of 4th grade, children were randomly assigned to three different groups. In the first treatment group, children received ten books of their choice during the summer vacation. In the second treatment group, children also received ten books of their choice and they and their parents were invited to attend three literacy events. Children in the third group constitute the comparison group. The authors find significant positive effects on self-reported reading behavior, but no significant positive effect on reading achievement. In a meta study of summer reading interventions, Kim and Quinn (2013) report an 11–22% of a standard deviation effect, with larger benefits for children from low-SES backgrounds.[25]

Overall, the findings in our study are comparable in magnitude to studies that find positive and significant impacts, while being less resource-intensive and independent of external

[25]Among the most effective, yet expensive interventions are tutoring interventions, which are quite different from ours. Tutoring interventions, such as a web-based intelligent tutoring system (ITSS) used by 4th and 5th graders in the U.S., have shown effectiveness. Wijekumar et al. (2012) report that students using the ITSS for 30-45 minutes weekly over six months improved reading comprehension by 0.1 standard deviations. A follow-up study by Wijekumar et al. (2014) found a 0.2 standard deviation improvement for 5th graders. Other interventions, such as a computerized reading program (Rouse and Krueger, 2004), financial incentives (Bettinger, 2012), and after-school programs (Goux et al., 2017), show no impact on academic outcomes.

support from teachers, parents, or tutors.

Cost-benefit analysis

At the time of our reading intervention, the cost of an e-book reader, including the one-year book subscription, was around 100 euros. The cost of the production of video recommendations amounts to approximately 16,000 euros, and there were additional costs for the design of the postcards, postage, and student assistants' working hours for pre-configuration. We calculate the total cost of the e-book reader intervention at about 130,000 euros, or 260 euros per e-book reader.

For the benefit analysis, we build on Hanushek and Woessmann (2008) who summarize that several studies consistently find that a one standard deviation increase in test scores is linked to a roughly 12% increase in earnings. Based on our point estimate of an increase of 0.11 standard deviations in reading comprehension, this corresponds to an increase in earnings of around 1.3%.[26] We then calculated a present discounted value of the life long earnings gains of our reading intervention of around 5,150 euros (based on the minimum wage in Germany in 2024) and 10,135 euros (based on the average monthly gross earnings of 4,323 euros as of April 2023).[27] Overall, this implies benefit-cost ratios of about 20 and 39, respectively. These compare favorably with other early childhood educational intervention programs. Obviously, the high ratio in our setting results from the very low costs of e-book reader provision compared to other educational interventions (i.e., face-to-face tutoring programs, reducing class sizes, longer school days).

6 Conclusion

This paper provides insights into the impact of a scalable randomized reading intervention, which provides e-book readers and access to a comprehensive digital library to 11–12-year-

[26] This is very likely a lower bound, as we are not accounting for the positive effects on math, socioemotional well-being, and potential benefits for society as a whole.

[27] We calculated the present discounted value of an increase in annual earnings of 1.3% with a discount rate of 5% over a time window of 30 years.

old children from disadvantaged backgrounds. By offering e-book readers with access to a large digital library, we observe significant improvements in reading habits, comprehension, math grades, and socioemotional well-being, both 12 months and 30 months after the start of the intervention. The flexibility and minimal distraction offered by e-books make them a promising tool for improving educational outcomes, especially for children from low-income backgrounds who face significant barriers to accessing learning resources. Our findings underscore that interventions during early adolescence, when children are gaining greater autonomy in their learning processes can generate significant returns.

Schools and policymakers face numerous challenges, including teacher shortages, and disparities in school performance linked to socioeconomic background. Our study offers valuable insights into how reading habits and academic performance among disadvantaged children can be positively impacted, complementing the efforts and investments made within schools. Providing children with access to reading resources at home is both cost-effective and scalable, offering an easy-to-implement solution that can significantly support their learning outside the classroom.

We would, however, like to acknowledge several limitations of our study. First, our setting does not allow us to separate the independent effect of providing children with e-book readers from the impact of book recommendations and the advantage of access to an extensive library. A multi-arm treatment design to isolate these mechanisms would have required substantially larger sample sizes to achieve adequate statistical power, exceeding the financial and logistical constraints of our research budget. Second, while our study finds positive and lasting effects of an e-book reading intervention on children's reading engagement, academic performance, and socioemotional well-being, we cannot provide information on whether and how the content of the books affect children's outcomes. For example, recent work by Adukia et al. (2023) examines the representation of identity (e.g., gender, race, skin color) in children's books using novel computer vision methods and identifies various dimensions of inequality in representation within these books, which may influence children's learning, norms, and beliefs.

Unlike many early childhood interventions that demonstrate skill attenuation over time, our treatment group maintained significant gains relative to the control group 30 months post-intervention, showing no evidence of mean reversion. This persistence suggests that the intervention may have successfully cultivated sustainable reading habits. Studying the long-term returns of the e-reader provision intervention offers promising avenues for future research.

References

Abeberese, Ama Baafra, Todd J. Kumler, and Leigh L. Linden, "Improving Reading Skills by Encouraging Children to Read in School: A Randomized Evaluation of the Sa Aklat Sisikat Reading Program in the Philippines," *Journal of Human Resources*, 2014, *49* (3), 611–633.

Adukia, Anjali, Alex Eble, Emileigh Harrison, Hakizumwami Birali Runesha, and Teodora Szasz, "What We Teach About Race and Gender: Representation in Images and Text of Children's Books," *The Quarterly Journal of Economics*, 2023, *138* (4), 2225–2285.

Agostinelli, Francesco and Matthew Wiswall, "Estimating the technology of children's skill formation," *Journal of Political Economy*, 2025, *133* (3), 000–000.

Akbar, S. R., A. Hanan Taqi, Abdulmohsin A. Dashti, and Taiba M. Sadeq, "Does E-Reading Enhance Reading Fluency?," *English Language Teaching*, 2015, *8* (5), 195–207.

Alan, Sule and Seda Ertac, "Fostering Patience in the Classroom: Results from Randomized Educational Intervention," *Journal of Political Economy*, 2018, *126* (5), 1865–1911.

Alexander, Karl L., Doris R. Entwisle, and Carrie S. Horsey, "From First Grade Forward: Early Foundations of High School Dropout," *Sociology of Education*, 1997, *70* (2), 87–107.

Anderson, Michael L., "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American Statistical Association*, 2008, *103* (484), 1481–1495.

Angrist, Joshua and Victor Lavy, "The Effects of High Stakes High School Achieve-

ment Awards: Evidence from a Randomized Trial," *American Economic Review*, 2009, *99* (4), 1384–1414.

Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin, "Identification of Causal Effects Using Instrumental Variables," *Journal of the American Statistical Association*, 1996, *91* (434), 444–455.

Attanasio, Orazio, Richard Blundell, Gabriella Conti, and Giacomo Mason, "Inequality in Socio-Emotional Skills: A Cross-Cohort Comparison," *Journal of Public Economics*, 2020, *191*, 104171.

Benjamini, Yoav, Abba M Krieger, and Daniel Yekutieli, "Adaptive Linear Step-Up Procedures that Control the False Discovery Rate," *Biometrika*, 2006, *93* (3), 491–507.

Bergman, Peter, "Parent-Child Information Frictions and Human Capital Investment: Evidence from a Field Experiment," *Journal of Political Economy*, 2021, *129* (1), 286–322.

_ , Chana Edmond-Verley, and Nicole Notario-Risk, "Parent Skills and Information Asymmetries: Experimental Evidence from Home Visits and Text Messages in Middle and High Schools," *Economics of Education Review*, 2018, *66*, 92–103.

Bettinger, Eric P., "Paying to Learn: The Effect of Financial Incentives on Elementary School Test Scores," *The Review of Economics and Statistics*, 2012, *94* (3), 686–698.

Beuermann, Diether W., Julian Cristia, Santiago Cueto, Ofer Malamud, and Yyannu Cruz-Aguayo, "One Laptop per Child at Home: Short-Term Impacts from a Randomized Experiment in Peru," *American Economic Journal: Applied Economics*, 2015, *7* (2), 53–80.

Bigozzi, Lucia, Christian Tarchi, Linda Vagnoli, Elena Valente, and Giuliana Pinto, "Reading Fluency As a Predictor of School Outcomes across Grades 4–9," *Frontiers in Psychology*, 2017, *8*.

Boca, Daniela Del, Christopher Flinn, and Matthew Wiswall, "Household Choices and Child Development," *The Review of Economic Studies*, 2014, *81* (1), 137–185.

Boggess, Scott, "Family Structure, Economic Status, and Educational Attainment," *Journal of Population Economics*, 1998, *11*, 205–222.

Bono, Emilia Del, Ben Etheridge, and Paul Garcia, "The Economic Value of Childhood Socio-Emotional Skills," *Institute for Social and Economic Research, ISER Working Paper Series*, 2024, (2024-01).

_ , Josh Kinsler, and Ronni Pavan, "Skill Formation and the Trouble with Child Non-Cognitive Skill Measures," Working Paper 13713, IZA Discussion Papers 2020.

Borman, Geoffrey D., Robert E. Slavin, Alan C. K. Cheung, Anne M. Chamberlain, Nancy A. Madden, and Bette Chambers, "Final Reading Outcomes of the National Randomized Field Trial of Success for All," *American Educational Research Journal*, 2007, *44* (3), 701–731.

Bulman, G. and R. W. Fairlie, "Technology and Education: Computers, Software, and the Internet," in Eric A. Hanushek, Stephen Machin, and Ludger Woessmann, eds., *Handbook of the Economics of Education*, Vol. 5, Elsevier, 2016, pp. 239–280.

Caucutt, Elizabeth M, Lance Lochner, Joseph Mullins, and Youngmin Park, "Child skill production: Accounting for parental and market-based time and goods investments," Working Paper w27838, National Bureau of Economic Research 2020.

Chall, Jeanne Sternlicht, *Stages of reading development*, New York: McGraw-Hil, 1983.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff, "Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood," *American Economic Review*, September 2014, *104* (9), 2633–2679.

_ , **John N Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan**, "How does your kindergarten classroom affect your earnings? Evidence from Project STAR," *The Quarterly journal of economics*, 2011, *126* (4), 1593–1660.

Coelli, Michael B., David A. Green, and William P. Warburton, "Breaking the Cycle? The Effect of Education on Welfare Receipt among Children of Welfare Recipients," *Journal of Public Economics*, 2007, *91* (7–8), 1369–1398.

Cristia, Julian, Pablo Ibarrarán, Santiago Cueto, Ana Santiago, and Eugenio Severín, "Technology and Child Development: Evidence from the One Laptop per Child Program," *American Economic Journal: Applied Economics*, 2017, *9* (3), 295–320.

Cunha, Flavio and James Heckman, "The technology of skill formation," *American Economic Review*, 2007, *97* (2), 31–47.

_ , **James J Heckman, and Susanne M Schennach**, "Estimating the Technology of Cognitive and Noncognitive Skill Formation," *Econometrica*, 2010, *78* (3), 883–931.

_ , **James J. Heckman, Lance Lochner, and Dimitriy V. Masterov**, "Chapter 12 Interpreting the Evidence on Life Cycle Skill Formation," in E. Hanushek and F. Welch, eds., *Handbook of the Economics of Education*, Vol. 1, Elsevier, 2006, pp. 697–812.

Dustmann, Christian, "Parental Background, Secondary School Track Choice, and Wages," *Oxford Economic Papers*, 2004, *56* (2), 209–230.

Eisenberg, Nancy, Adrienne Sadovsky, and Tracy L. Spinrad, "Associations of Emotion-Related Regulation with Language Skills, Emotion Knowledge, and Academic Outcomes," *New Directions for Child and Adolescent Development*, 2005, (109), 109–118.

Ermisch, John F. and Marco Francesconi, "Family Structure and Children's Achievements," *Journal of Population Economics*, 2001, *14*, 249–270.

Escueta, Maya, Andre Joshua Nickow, Philip Oreopoulos, and Vincent Quan, "Upgrading Education with Technology: Insights from Experimental Research," *Journal of Economic Literature*, 2020, *58* (4), 897–996.

Fairlie, Robert W. and Ariel Kalil, "The Effects of Computers on Children's Social Development and School Participation: Evidence from a Randomized Control Experiment," *Economics of Education Review*, 2017, *57*, 10–19.

_ **and Jonathan Robinson**, "Experimental Evidence on the Effects of Home Computers on Academic Achievement among Schoolchildren," *American Economic Journal: Applied Economics*, 2013, *5* (3), 211–240.

Fernald, Anne, Virginia A. Marchman, and Adriana Weisleder, "SES Differences in Language Processing Skill and Vocabulary are Evident at 18 Months," *Developmental Science*, 2013, *16* (2), 234–248.

Fryer, Roland G. and Steven D. Levitt, "Understanding the Black-White Test Score Gap in the First Two Years of School," *The Review of Economics and Statistics*, 2004, *86* (2), 447–464.

García, Jorge Luis, James J. Heckman, Duncan Ermini Leaf, and María José Prados, "Quantifying the Life-Cycle Benefits of an Influential Early-Childhood Program," *Journal of Political Economy*, July 2020, *128* (7), 2502–2541.

Glewwe, Paul, Michael Kremer, and Sylvie Moulin, "Many Children Left Behind? Textbooks and Test Scores in Kenya," *American Economic Journal: Applied Economics*, 2009, *1* (1), 112–135.

Goodman, Robert, "The Strengths and Difficulties Questionnaire: A Research Note," *Journal of Child Psychology and Psychiatry*, 1997, *38* (5), 581–586.

Goux, Dominique, Marc Gurgand, and Eric Maurin, "Reading Enjoyment and Reading Skills: Lessons from an Experiment with First Grade Children," *Labour Economics*, 2017, *45*, 17–25.

Gray-Lobe, Guthrie, Parag A Pathak, and Christopher R Walters, "The Long-Term Effects of Universal Preschool in Boston," *The Quarterly Journal of Economics*, 2023, *138* (1), 363–411.

Green, David A and W Craig Riddell, "Literacy and Earnings: An Investigation of the Interaction of Cognitive and Unobserved Skills in Earnings Generation," *Labour Economics*, 2003, *10* (2), 165–184.

Guryan, Jonathan, Erik Hurst, and Melissa Kearney, "Parental Education and Parental Time with Children," *Journal of Economic Perspectives*, 2008, *22* (3), 23–46.

_ , **James S Kim, and David M Quinn**, "Does reading during the summer build reading skills? Evidence from a randomized experiment in 463 classrooms," Working paper w20689, National Bureau of Economic Research 2014.

_ , **James S. Kim, and Kyung H. Park**, "Motivation and Incentives in Education: Evidence from a Summer Reading Experiment," *Economics of Education Review*, 2016, *55*, 1–20.

Hanushek, Eric A. and Ludger Woessmann, "The Role of Cognitive Skills in Economic Development," *Journal of Economic Literature*, 2008, *46* (3), 607–668.

_ , **Guido Schwerdt, Simon Wiederhold, and Ludger Woessmann**, "Returns to Skills around the World: Evidence from PIAAC," *European Economic Review*, 2015, *73*, 103–130.

Hashim, Ayesha K. and Vanessa W. Vongkulluksn, "E-Reader Apps and Reading Engagement: A Descriptive Case Study," *Computers & Education*, 2018, *125*, 358–375.

Heckman, James J., "Skill Formation and the Economics of Investing in Disadvantaged Children," *Science*, 2006.

_ , "Schools, Skills, and Synapses," *Economic Inquiry*, 2008, *46* (3), 289–324.

_ **and Rodrigo Pinto**, "Econometric Mediation Analyses: Identifying the Sources of Treatment Effects from Experimentally Estimated Production Technologies with Unmeasured and Mismeasured Inputs," *Econometric Reviews*, 2015, *34* (1-2), 6–31.

_ **and Yona Rubinstein**, "The Importance of Noncognitive Skills: Lessons from the GED Testing Program," *American Economic Review*, 2001, *91* (2), 145–149.

Herbers, Janette E., J. J. Cutuli, Laura M. Supkoff, David Heistad, Chi-Keung Chan, Elizabeth Hinz, and Ann S. Masten, "Early Reading Skills and Academic Achievement Trajectories of Students Facing Poverty, Homelessness, and High Residential Mobility," *Educational Researcher*, 2012, *41* (9), 366–374.

Hernandez, Donald J., "Double Jeopardy: How Third-Grade Reading Skills and Poverty Influence High School Graduation," Technical Report 2011.

Holden, Kristian L., "Buy the Book? Evidence on the Effect of Textbook Funding on School-Level Achievement," *American Economic Journal: Applied Economics*, 2016, *8* (4), 100–127.

Imai, Kosuke, Luke Keele, and Dustin Tingley, "A General Approach to Causal Mediation Analysis," *Psychological Methods*, 2010, *15* (4), 309–334.

Jr, Roland G. Fryer and Steven D. Levitt, "The Black-White Test Score Gap Through Third Grade," *American Law and Economics Review*, July 2006, *8* (2), 249–281.

Kalil, Ariel, Haoxuan Liu, Susan Mayer, Derek Rury, and Rohen Shah, "Nudging or Nagging? Conflicting Effects of Behavioral Tools," *University of Chicago, Becker Friedman Institute for Economics Working Paper*, 2023, (2023-02).

Keilow, Maria, Hans Henrik Sievertsen, Janni Niclasen, and Carsten Obel, "The Strengths and Difficulties Questionnaire and Standardized Academic Tests: Reliability across Respondent Type and Age," *PLoS One*, 2019, *14* (7), e0220193.

Kern, Margaret L. and Howard S. Friedman, "Early Educational Milestones as Predictors of Lifelong Academic Achievement, Midlife Adjustment, and Longevity," *Journal of Applied Developmental Psychology*, 2009, *30* (4), 419–430.

Kim, James S. and David M. Quinn, "The Effects of Summer Reading on Low-Income Children's Literacy Achievement From Kindergarten to Grade 8: A Meta-Analysis of Classroom and Home Interventions," *Review of Educational Research*, 2013, *83* (3), 386–431.

_ **and Jonathan Guryan**, "The Efficacy of a Voluntary Summer Book Reading Intervention for Low-Income Latino Children from Language Minority Families," *Journal of Educational Psychology*, 2010, *102* (1), 20–31.

Kling, Jeffrey R, Jeffrey B Liebman, and Lawrence F Katz, "Experimental Analysis of Neighborhood Effects," *Econometrica*, 2007, *75* (1), 83–119.

Kremer, Michael, Conner Brannen, and Rachel Glennerster, "The Challenge of Education and Learning in the Developing World," *Science*, 2013, *340* (6130), 297–300.

Lettau, Jacqueline, "The Impact of Children's Academic Competencies and School Grades on their Life Satisfaction: What Really Matters?," *Child Indicators Research*, 2021, *14* (6), 2171–2195.

Leuven, Edwin, Mikael Lindahl, Hessel Oosterbeek, and Dinand Webbink, "The Effect of Extra Funding for Disadvantaged Pupils on Achievement," *The Review of Economics and Statistics*, 2007, *89* (4), 721–736.

List, John A., Azeem M. Shaikh, and Yang Xu, "Multiple Hypothesis Testing in Experimental Economics," *Experimental Economics*, 2019, *22* (4), 773–793.

_ **, Ragan Petrie, and Anya Samek**, "How Experiments with Children Inform Economics," *Journal of Economic Literature*, 2023, *61* (2), 504–564.

Logan, Sarah and Rhona Johnston, "Investigating Gender Differences in Reading," *Educational Review*, 2010, *62* (2), 175–187.

Long, Deanna and Susan Szabo, "E-Readers and the Effects on Students' Reading Motivation, Attitude and Comprehension during Guided Reading," *Cogent Education*, 2016, *3* (1), 1197818.

Machin, Stephen and Sandra McNally, "The literacy hour," *Journal of Public Economics*, 2008, *92* (5-6), 1441–1462.

Mak, Hei Wan and Daisy Fancourt, "Longitudinal Associations between Reading for Pleasure and Child Maladjustment: Results from a Propensity Score Matching Analysis," *Social Science & Medicine*, 2020, *253*, 112971.

Malamud, Ofer and Cristian Pop-Eleches, "Home Computer Use and the Development of Human Capital," *The Quarterly Journal of Economics*, 2011, *126* (2), 987–1027.

National Assessment of Educational Progress, "The Nation's Report Card: Reading Results from 2024," March 2025.

Roberts, Garrett J., Michael Solis, Stephen Ciullo, John W. McKenna, and Sharon Vaughn, "Reading Interventions With Behavioral and Social Skill Outcomes: A Synthesis of Research," *Behavior Modification*, 2015, *39* (1), 8–42.

Romano, Joseph P. and Michael Wolf, "Efficient Computation of Adjusted p-values for Resampling-Based Stepdown Multiple Testing," *Statistics & Probability Letters*, 2016, *113*, 38–40.

Romeo, Rachel R., Lili Uchida, and Joanna A. Christodoulou, "Socioeconomic Status and Reading Outcomes: Neurobiological and Behavioral Correlates," *New Directions for Child and Adolescent Development*, 2022, *2022* (183-184), 57–70.

Rouse, Cecilia Elena and Alan B. Krueger, "Putting Computerized Instruction to the Test: A Randomized Evaluation of a "Scientifically Based" Reading Program," *Economics of Education Review*, 2004, *23* (4), 323–338.

Slavin, Robert E., Cynthia Lake, Bette Chambers, Alan Cheung, and Susan Davis, "Effective Reading Programs for the Elementary Grades: A Best-Evidence

Synthesis," *Review of Educational Research*, 2009, *79* (4), 1391–1466.

Sonnenschein, Susan, Laura M. Stapleton, and Amy Benson, "The Relation Between the Type and Amount of Instruction and Growth in Children's Reading Competencies," *American Educational Research Journal*, 2010, *47* (2), 358–389.

Westfall, Peter H. and S. Stanley Young, "p Value Adjustments for Multiple Tests in Multivariate Binomial Models," *Journal of the American Statistical Association*, 1989, *84* (407), 780–786.

Wijekumar, Kausalai, Bonnie J. F. Meyer, Pui-Wa Lei, Yu-Chu Lin, Lori A. Johnson, James A. Spielvogel, Kathryn M. Shurmatz, Melissa Ray, and Michael Cook, "Multisite Randomized Controlled Trial Examining Intelligent Tutoring of Structure Strategy for Fifth-Grade Readers," *Journal of Research on Educational Effectiveness*, 2014, *7* (4), 331–357.

Wijekumar, Kausalai Kay, Bonnie J. F. Meyer, and Puiwa Lei, "Large-Scale Randomized Controlled Trial with 4th Graders using Intelligent Tutoring of the Structure Strategy to Improve Nonfiction Reading Comprehension," *Educational Technology Research and Development*, 2012, *60*, 987–1013.

Young, Alwyn, "Channeling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results," *The Quarterly Journal of Economics*, 2019, *134* (2), 557–598.

Figures and Tables

Figure 1: Geographic spread of the sample

Notes: This figure illustrates the geographic distribution of treatment and control groups across Germany. The map highlights the federal states, with dots representing the number of sampled units in the treatment group and triangles representing the control group. The scale of treatment and control groups is further categorized by size at the county level. Treatment units are represented as circles, with three categories: small circles (1–5 units), medium circles (6–10 units), and large circles (11–14 units). Similarly, control units are depicted with triangles, with small (1–5 units), medium (6–10 units), and large (11–17 units) triangle symbols.

Figure 2: Intervention and data collection timeline

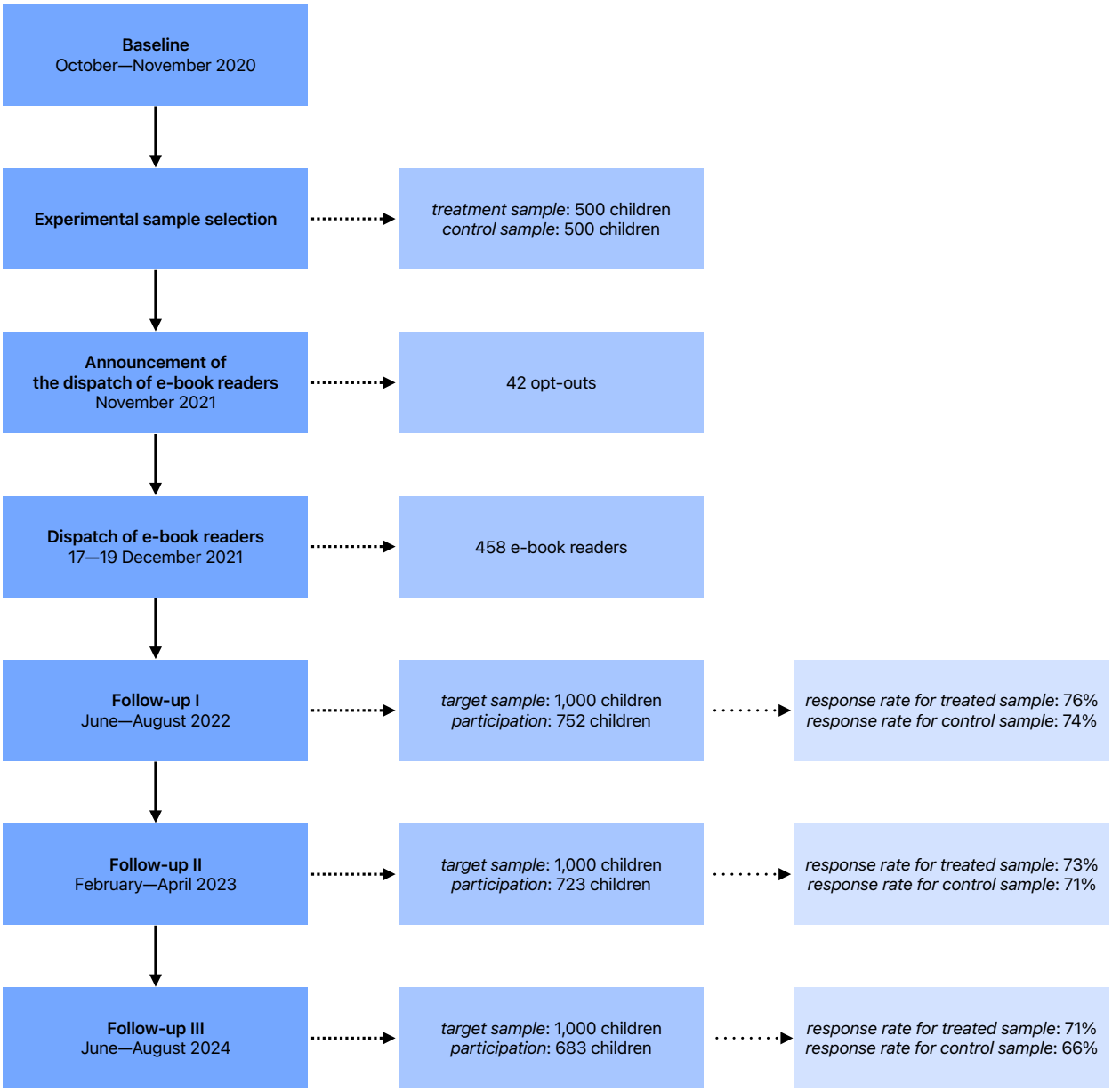
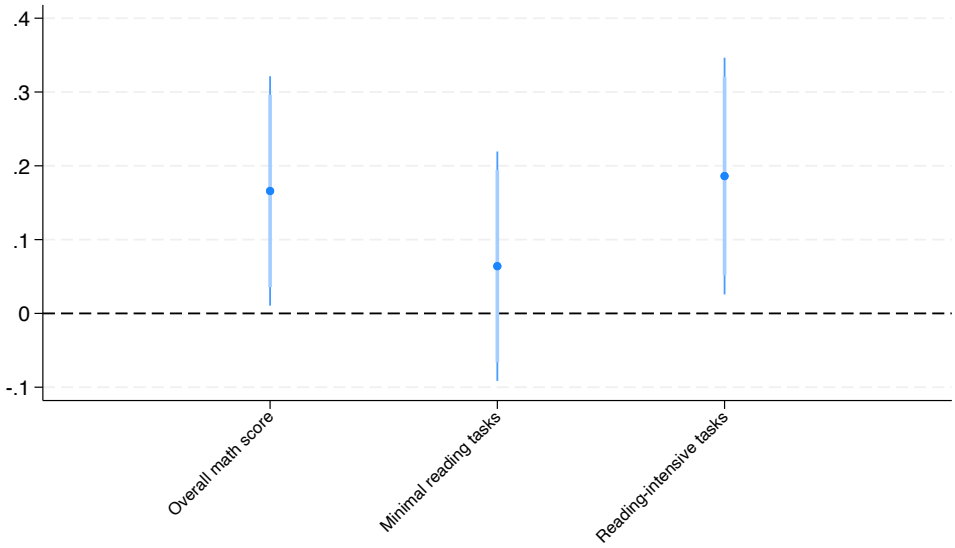


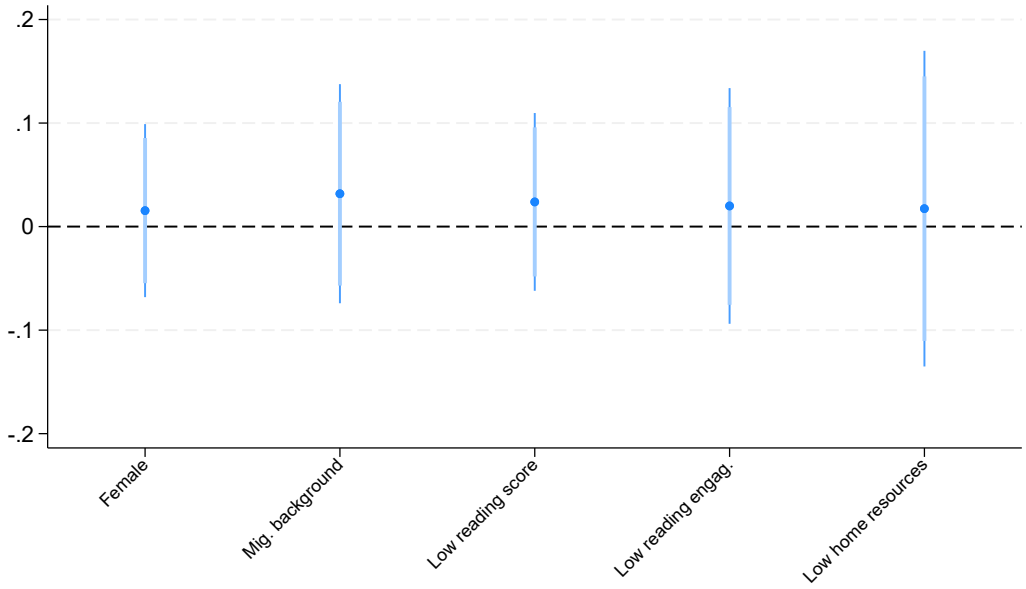
Figure 3: The effect of the treatment on math score by type of math problem



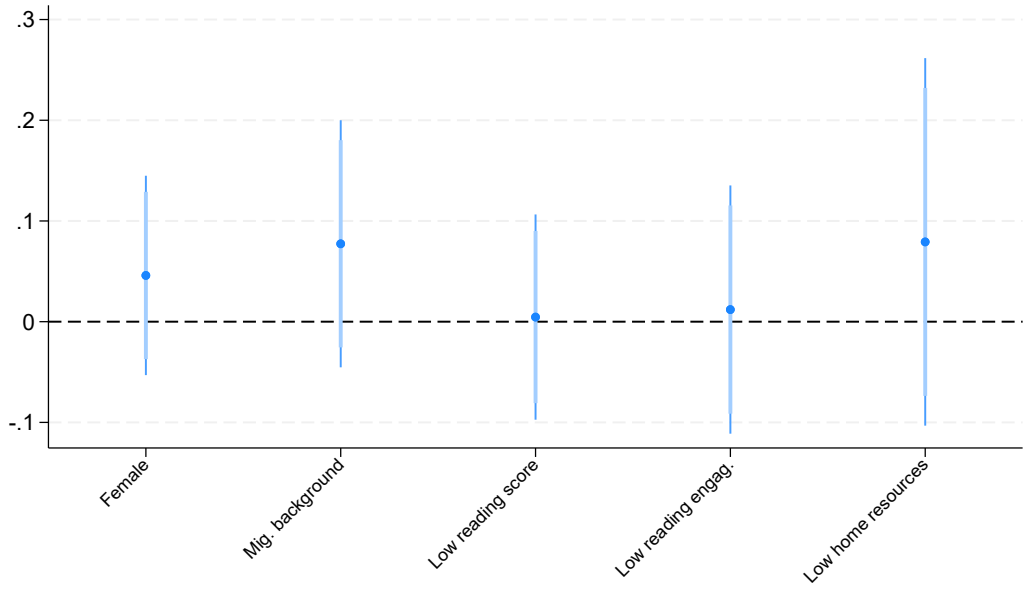
Notes: The figure plots the estimated coefficient $\hat{\beta}$ from a model similar to equation 1 with corresponding 90% and 95% confidence intervals. The outcome variable is children's math score in math problems measured 30 months after the start of the RCT (follow-up III). We differentiate between two types of tasks: minimal reading and reading-intensive. Each point estimate comes from a separate regression. The first point estimate (confidence intervals) displays the treatment effect on the overall score (average score for all four math problems), the second point estimate (confidence intervals) is for math tasks with minimal reading, and the final point estimates (confidence intervals) are for reading-intensive math tasks.

Figure 4: Heterogeneous effects of the e-book reading intervention on the reading engagement index

(a) Follow-up II

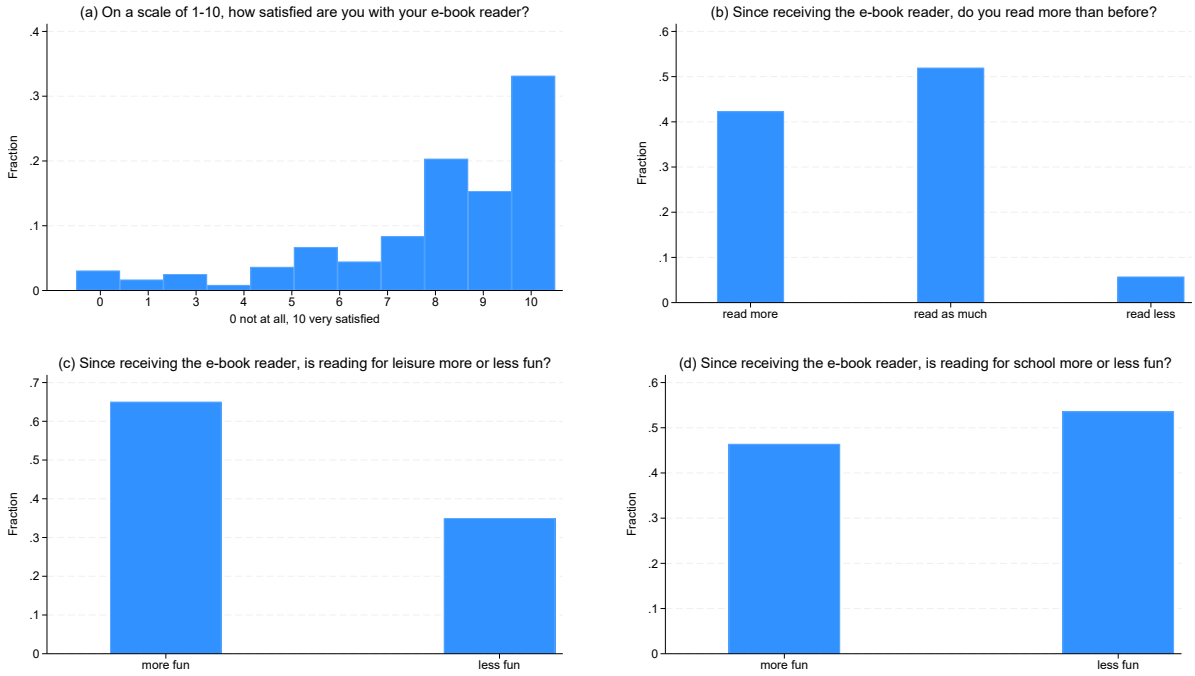


(b) Follow-up III



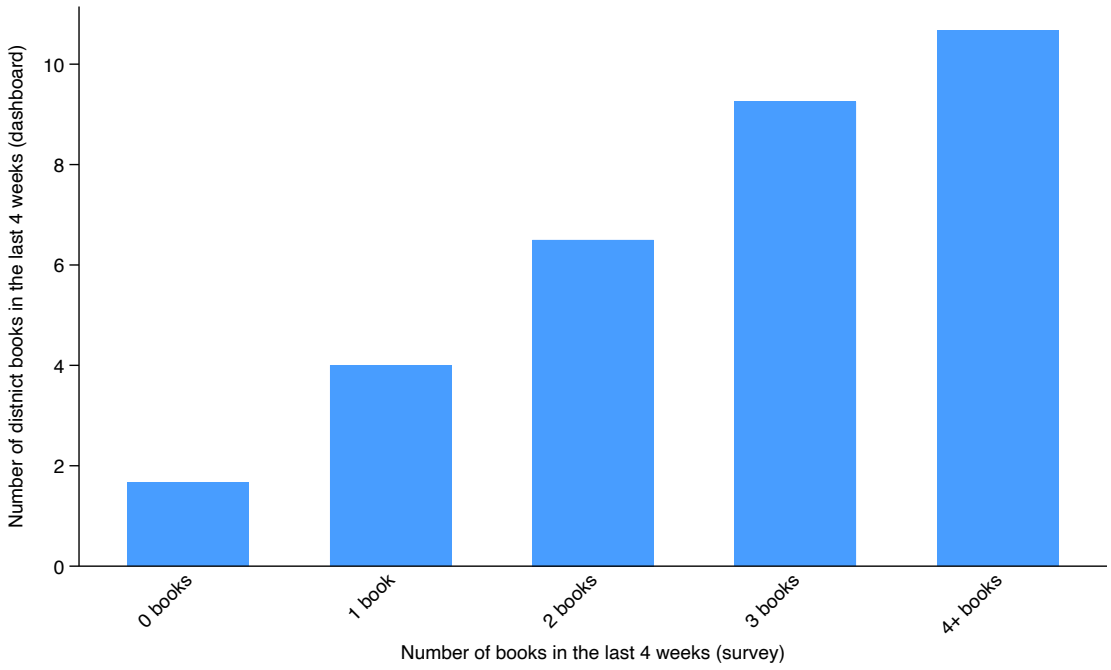
Notes: The figures plot the estimated $\hat{\beta}_3$ coefficients from equation 4 with corresponding 90% and 95% confidence intervals. Each point estimate comes from a separate regression. All stratifying variables are measured at baseline. Follow-up II: 12 months after the start of RCT. Follow-up III (endline): 30 months after the start of RCT.

Figure 5: Self-reported satisfaction with e-book readers, reading frequency, and enjoyment of reading



Notes: Figure (a) on the top left depicts treated children's satisfaction with their e-book readers on a scale from 0 to 10. Figure (b) on the top right plots the share of treated children based on whether they read more, as much, or less since receiving the e-book reader. Figures (c) and (d) at the bottom show the share of treated children who report that reading, for pleasure or school, respectively, has been more or less enjoyable since receiving the e-book reader.

Figure 6: Cross-validation of survey responses



Notes: The y-axis shows the number of e-books read in the last four weeks as (objectively) recorded in the dashboard. The x-axis displays the number of e-books read as reported by treated children in the survey.

Table 1: Balance of covariates at baseline (November 2020)

	Control	Treatment	Difference	p-value
	(1)	(2)	(3)	(4)
Female	0.52	0.51	0.01	(0.85)
School year	4.43	4.43	-0.00	(0.98)
Year of birth	2010.07	2010.13	-0.05	(0.24)
Month of birth	6.53	6.38	0.15	(0.50)
Reading score	-0.00	0.00	-0.01	(0.92)
German grade	2.39	2.34	0.05	(0.33)
Math grade	2.33	2.28	0.05	(0.38)
Parents' school leaving qualification	0.78	0.79	-0.01	(0.69)
Parents born in Germany	0.79	0.79	0.00	(0.88)
Very low socioeconomic status	0.53	0.54	-0.00	(0.95)
Number of observations	500	500		

Notes: German and math are the respective school grades, measured on a scale of 1–5, where 1 is the best possible grade. Parental school leaving qualification is a binary variable taking on the value 1 if the parent has at least a middle school (Realschule) qualification, and zero otherwise. Very low socioeconomic status refers to students who come from households receiving welfare.

Table 2: Balance of the covariates at baseline between opt-outs and those who received e-book readers

	Received e-book reader	Opted out	Difference	p-value
	(1)	(2)	(3)	(4)
Female	0.52	0.43	0.09	(0.25)
School year	4.43	4.43	0.00	(1.00)
Year of birth	2010.13	2010.07	0.06	(0.61)
Month of birth	6.37	6.48	-0.11	(0.85)
Reading score	0.02	-0.21	0.24	(0.15)
German grade	2.33	2.40	-0.07	(0.61)
Math grade	2.28	2.33	-0.05	(0.75)
Parents' school leaving qualification	0.79	0.81	-0.02	(0.75)
Parents born in Germany	0.80	0.71	0.08	(0.21)
Very low socioeconomic status	0.54	0.50	0.04	(0.63)
Number of observations	458	42		

Notes: German and math are the respective school grades, measured on a scale of 1–5, where one is the best possible grade. Parental school leaving qualification is a binary variable taking the value one if the parent has at least an intermediate school-leaving qualification (Realschule), and zero otherwise. Very low socioeconomic status refers to children who come from households receiving welfare benefit.

Table 3: Response rates in follow-up surveys

	Treatment	Control	Difference	p-value
	(1)	(2)	(3)	(4)
Response rate:				
Follow-up I	0.76	0.74	-0.02	(0.42)
Follow-up II	0.73	0.71	-0.02	(0.44)
Follow-up III (endline)	0.71	0.66	-0.05	(0.07)
Responded all surveys	0.61	0.59	-0.02	(0.45)

Notes: Response rates are calculated based on the original sample of 1,000 participants. The last row of the table presents the share of children who participated in all three follow-up surveys. Follow-up I: 6 months after the start of RCT. Follow-up II: 12 months after the start of RCT. Follow-up III: 30 months after the start of RCT.

Table 4: Effect of e-book reading intervention on main outcomes

	Follow-up II: 12 months after the start of RCT			**Follow-up III: 30 months after the start of RCT**		
	Intention-to-treat (ITT)	p-value from randomized inference	FDR q-value	Intention-to-treat (ITT)	p-value from randomized inference	FDR q-value
Reading engagement (index)	0.105	0.000		0.063	0.016	
At least one e-book	0.177	0.000	0.022	0.136	0.000	0.001
At least one printed book	0.064	0.040	0.066	0.017	0.636	0.613
At least one book	0.089	0.000	0.009	0.010	0.768	0.732
At least two days/week	0.085	0.000	0.022	0.091	0.008	0.043
Academic performance (index)	0.103	0.036		0.096	0.072	
Reading comprehension	0.129	0.080	0.09	0.137	0.080	0.120
Good math grade	0.086	0.020	0.063	0.102	0.024	0.043
Good German grade	0.023	0.590	0.285	0.025	0.582	0.613
SDQ score	-0.147	0.064	0.077	-0.218	0.004	0.027
Grit	-0.055	0.450	0.264	-0.044	0.592	0.380
Locus of control	0.089	0.282	0.155	-0.050	0.568	0.613
School aspirations	-0.018	0.646	0.337	0.049	0.300	0.613

Notes: This table shows the ITT estimates of the e-book treatment on the two summary indices (*reading engagement* and *academic performance*) and on the 11 separate outcome measures. The variables in the index *reading engagement* are binary, taking the value one if the child has read at least one (e-)book in the last four weeks, or reported reading on at least two days in a typical week. Reading comprehension is a standardized score of the reading tests. Math and German grades are dichotomized to take the value one if the student has a good or very good grade (1 or 2 on the German grading scale), and zero otherwise. Components of SDQ are standardized with a mean of zero and a standard deviation of one. All regressions control for strata (randomization block), year and month of birth fixed effects. Only observations with observed outcomes are included. The number of observations in regressions varies between 712 to 645, depending on item non-response. The table reports p-values from randomization inference (Young, 2019) and sharpened false-discovery rate (FDR) q-values.

Table 5: Reading engagement (follow-up I)

	ITT	p-value of randomized inference	Number of obs.
	Reported by children		
At least one e-book	0.433	0.000	734
At least one printed book	0.002	0.990	742
At least one book	0.061	0.018	733
At least two days/week	0.056	0.014	741
	Reported by parents		
At least one e-book	0.412	0.000	705
At least one printed book	-0.029	0.310	719
At least one book	0.065	0.006	703
At least two days/week	0.064	0.002	723

Notes: This table shows the impact of e-book readers on reading engagement at follow-up I (six months after the start of the experiment), as reported by the children and their parents. All variables are binary. The variable *(At least one e-book (printed book))* takes the value one if the child has read at least one e-book (printed book) in the last four weeks, and zero otherwise. The variable *At least one book* takes the value one if the child has read at least one book in the last four weeks, regardless of the format, and zero otherwise. The variable *At least two days/week* takes the value one if the child reads on at least two days a week, and zero otherwise. Randomization inference p-values are reported next to the coefficients. All regressions control for strata (randomization block), year and month of birth fixed effects. Only observations with observed outcomes are included. The table reports p-values from randomization inference (Young, 2019).

Table 6: Robustness test: Additional control variables

	Set 1		Set 2		Set 3	
	ITT	p-value	ITT	p-value	ITT	p-value
Panel A. Follow-up II						
Reading engagement (index)	0.104	0.000	0.110	0.000	0.110	0.000
At least one e-book	0.177	0.000	0.183	0.000	0.183	0.000
At least one printed book	0.060	0.038	0.063	0.052	0.062	0.040
At least one book	0.084	0.006	0.090	0.000	0.092	0.000
At least two days/week	0.089	0.004	0.096	0.004	0.099	0.004
Academic performance (index)	0.107	0.032	0.106	0.020	0.110	0.018
Reading comprehension	0.138	0.078	0.136	0.092	0.142	0.084
Good math grade	0.078	0.066	0.079	0.056	0.080	0.044
Good German grade	0.027	0.564	0.030	0.470	0.028	0.512
SDQ score	-0.155	0.058	-0.138	0.090	-0.133	0.086
Grit	-0.059	0.438	-0.052	0.524	-0.050	0.512
Locus of control	0.101	0.232	0.115	0.184	0.115	0.198
School aspirations	0.001	0.978	0.006	0.872	0.004	0.936
Panel B. Follow-up III						
Reading engagement (index)	0.050	0.060	0.049	0.076	0.049	0.078
At least one e-book	0.130	0.000	0.130	0.002	0.129	0.000
At least one printed book	0.006	0.882	0.007	0.888	0.008	0.814
At least one book	-0.004	0.930	-0.003	0.948	-0.003	0.914
At least two days/week	0.073	0.062	0.069	0.070	0.068	0.084
Academic performance (index)	0.086	0.094	0.077	0.158	0.081	0.134
Reading comprehension	0.109	0.184	0.097	0.214	0.100	0.226
Good math grade	0.099	0.024	0.095	0.048	0.096	0.018
Good German grade	0.023	0.600	0.019	0.652	0.014	0.766
SDQ score	-0.211	0.006	-0.208	0.012	-0.205	0.004
Grit	-0.017	0.818	-0.002	0.936	-0.002	0.986
Locus of control	-0.070	0.410	-0.064	0.448	-0.081	0.350
School aspirations	0.041	0.316	0.047	0.286	0.051	0.212

Notes: This table reports estimates when adding further explanatory variables (all measured at baseline, prior to the intervention). Set 1 comprises controls for reading competence score, grades in math and German, and grit. Set 2 adds whether the children have their own desk, their own room, their own e-book reader, access to the internet, and whether the primary language spoken at home is German. Set 3 adds fixed effects for federal state of residence, parent's employment status, and parent's educational qualifications. The table reports p-values from randomization inference (Young, 2019). Follow-up II: 12 months after the start of RCT. Follow-up III (endline): 30 months after the start of RCT.

Table 7: Poisson and ordered probit regressions

	Follow-up II			**Follow-up III**		
Poisson	Intention-to-treat (ITT)	100× $(exp(\beta)-1)/$	p-value	Intention-to-treat (ITT)	100× $(exp(\beta)-1)$	p-value
Number of e-books[a]	0.553	73.907	0.008	0.789	120.188	0.000
Number of printed books[a]	0.072	7.482	0.330	-0.047	-4.606	0.569
Number of books[a]	0.175	19.143	0.025	0.119	12.671	0.155
Number of days read[b]	0.089	9.280	0.043	0.130	13.932	0.032
Ordered probit		$dy/\ dx$			$dy/\ dx$	
Math grade	0.176	0.036	0.055	0.228	-0.004	0.008
German grade	0.075	0.013	0.410	0.078	-0.000	0.368

Notes: The top panel reports Poisson regression coefficients and the change in the outcome in percent. [a] Number of books read in the last four weeks. [b] Number of days read in a typical week. The bottom panel reports the ordered probit coefficients, and the marginal effects evaluated for the outcomes is to the best grade. The grading is reversed such that higher values mean better grades. The table reports p-values from randomization inference (Young, 2019). Follow-up II: 12 months after the start of RCT. Follow-up III (endline): 30 months after the start of RCT.

Table 8: Effect of e-book reading intervention on inverse-covariance-weighted summary indices

	Intention-to-treat (ITT)	p-value from randomized inference
Reading engagement (index)		
Follow-up II	0.385	0.000
Follow-up III (endline)	0.305	0.002
Academic performance (index)		
Follow-up II	0.157	0.034
Follow-up III (endline)	0.166	0.024

Notes: The table reports treatment effects (ITT) on the inverse-covariance-weighted summary indices *reading engagement* and *academic performance* (Anderson, 2008). The table reports p-values from randomization inference (Young, 2019). Each coefficient and p-value comes from a different regression. Follow-up II: 12 months after the start of RCT. Follow-up III (endline): 30 months after the start of RCT.

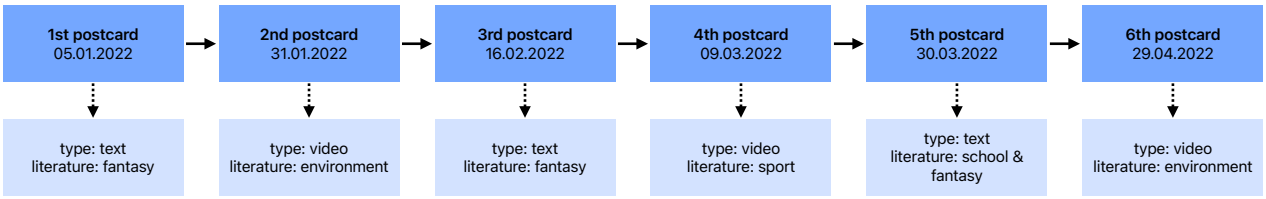
Table 9: Family-wise error: Multiple hypothesis testing

	Unadjusted		Romano & Wolf		List et al.		Westfall & Young	
	Follow-up II	Follow-up III	Follow-up II	Follow-up III	Follow-up II	Follow-up III	Follow-up II	Follow-up III
At least one e-book	**0.000**	**0.016**	**0.001**	**0.001**	**0.001**	**0.001**	**0.000**	**0.001**
At least one printed book	**0.040**	0.636	**0.036**	0.742	**0.040**	0.797	**0.022**	0.786
At least one book	**0.000**	0.768	**0.007**	0.766	**0.001**	0.788	**0.006**	0.790
At least two days/week	**0.000**	**0.008**	**0.019**	**0.041**	**0.015**	**0.034**	**0.013**	**0.043**
Reading comprehension	**0.080**	**0.080**	0.163	0.168	0.140	0.185	0.178	0.157
Good math grade	**0.020**	**0.024**	**0.082**	**0.055**	**0.085**	**0.051**	0.102	**0.053**
Good German grade	0.590	0.582	0.544	0.543	0.553	0.552	0.582	0.576
SDQ score	**0.064**	**0.004**	0.204	**0.022**	0.202	**0.017**	0.197	**0.017**
Grit	0.450	0.592	0.591	0.810	0.724	0.605	0.588	0.847
Locus of control	0.282	0.568	0.951	0.810	0.946	0.807	0.956	0.847
School aspirations	0.646	0.300	0.709	0.810	0.624	0.802	0.730	0.847

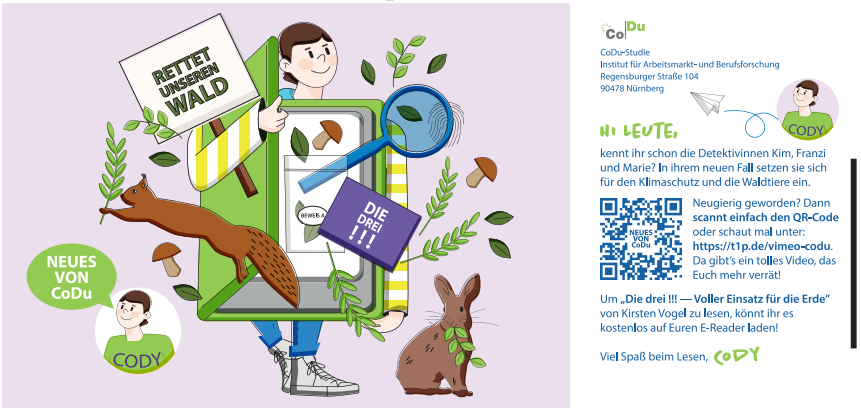
Notes: Columns 1–2 present the unadjusted p-values from the main Table 4. Columns 3–4 report step-down adjusted p-values (Romano and Wolf, 2016). Columns 5–6 report p-values from bootstrap-based procedures (List et al., 2019), and columns 7–8 report p-values for multiple hypothesis testing using vector-based resampling methods (Westfall and Young, 1989). p-values below 0.1 are highlighted in bold. Follow-up II: 12 months after the start of RCT. Follow-up III (endline): 30 months after the start of RCT.

Online Appendix A: Additional Figures and Tables

Figure A1: Postcard mailing timeline and sample postcards with and without a video link



A postcard with a link to a video

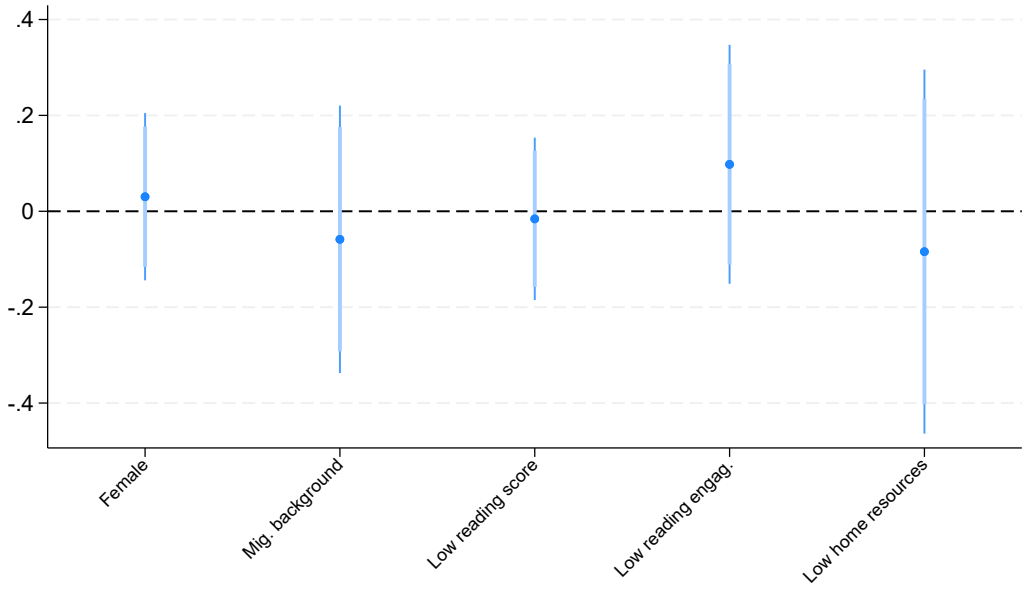


A postcard without a link to a video

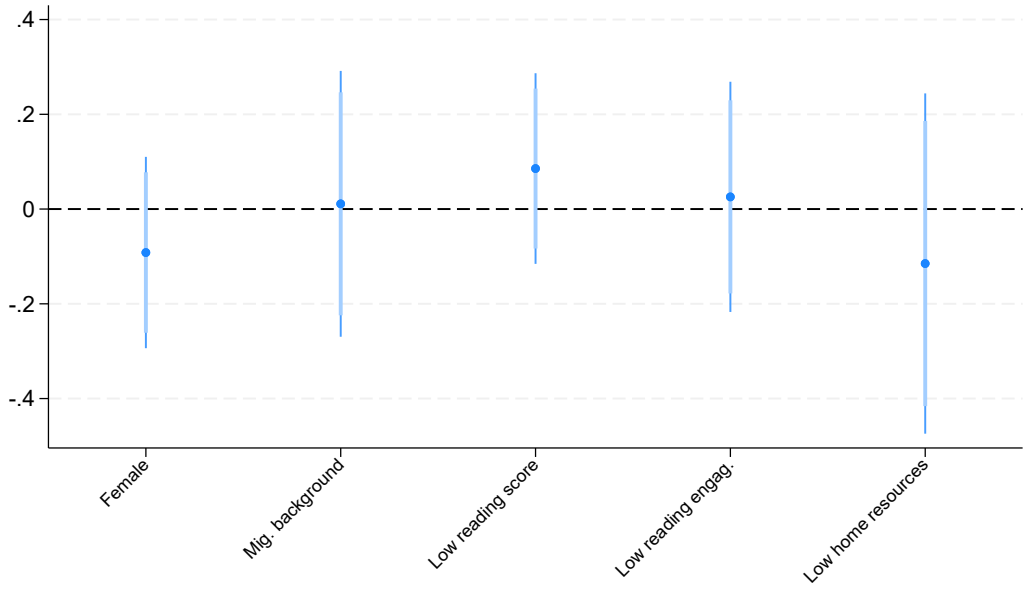
Notes: The upper panel indicates when the postcards with reading tips were sent to the treatment children, as well as the type of book (e.g., fantasy, environment, etc.) and type of recommendation (postcard with a link to a video (example shown in the middle panel) or postcard without a link to a video (example shown in the lower panel)).

Figure A2: Heterogeneous effects of the e-book reading intervention on academic performance index

(a) Follow-up II



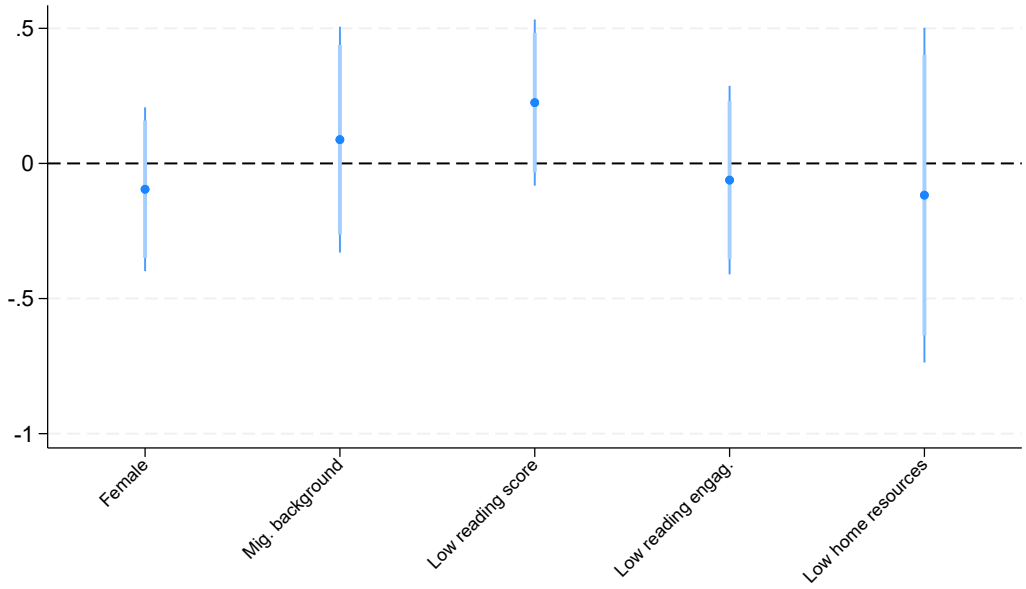
(b) Follow-up III



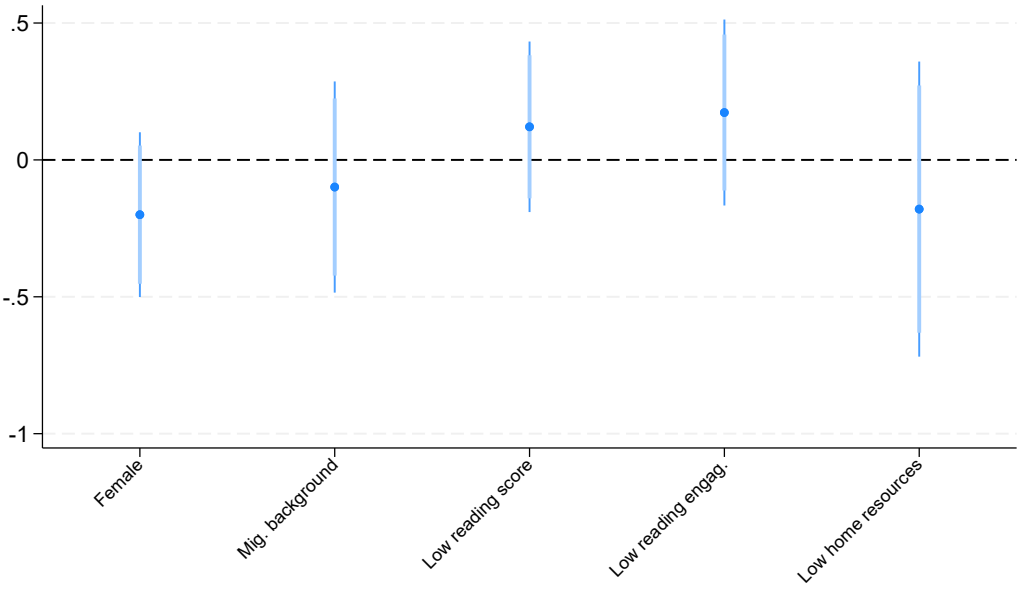
Notes: The figures plot the estimated $\hat{\beta}_3$ coefficients from equation 4 with corresponding 90% and 95% confidence intervals, respectively. Each point estimate comes from a separate regression. All stratifying variables are measured at baseline. Follow-up II: 12 months after the start of RCT. Follow-up III (endline): 30 months after the start of RCT.

Figure A3: Heterogeneous effects of the e-book reading intervention on Strengths and Difficulties Questionnaire (SDQ)

(a) Follow-up II

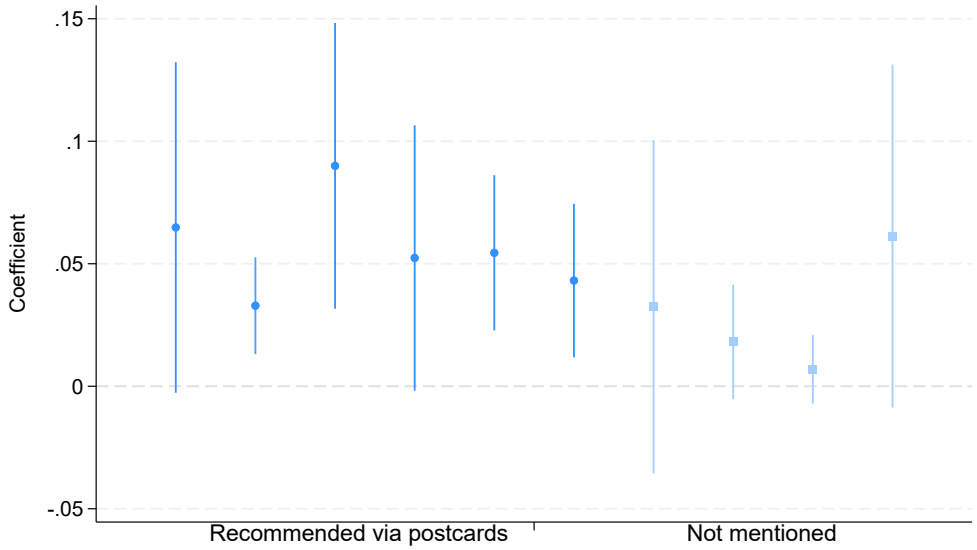


(b) Follow-up III



Notes: The figures plot the estimated $\hat{\beta}_3$ coefficients from equation 4 with corresponding 90% and 95% confidence intervals, respectively. Each point estimate comes from a separate regression. All stratifying variables are measured at baseline. Follow-up II: 12 months after the start of RCT. Follow-up III (endline): 30 months after the start of RCT.

Figure A4: The effects on reading a book title (recommended vs. not mentioned)



Notes: The figure presents the estimated coefficients (95% confidence intervals) from ten separate regressions, each examining whether the child reported reading a specific book. The dependent variable in each regression is equal to one if the child says they read the given book, and zero otherwise. In the figure, the darker blue points with confidence intervals represent the books that were recommended, while lighter ones represent those that were not. The results indicate a statistically significant increase in the likelihood of children in the treatment group reporting having read four of the six recommended books. In contrast, no such effect was observed for the books that were not explicitly recommended.
In follow-up I, participants from both the control and treatment groups were asked to answer a series of questions. Specifically, we presented them with a list of ten book titles and asked whether they had read any of them. Of these ten books, six were actively recommended to the children in the treatment group through postcards and videos from our research team. The control group did not receive these recommendations. The remaining four books were not mentioned in any of our communications to either group.

Table A1: Share of participants by federal state

State	Our sample	Germany
Baden-Wuerttemberg	14.2	13.3
Bavaria	17.0	15.8
Berlin	3.5	4.4
Brandenburg	4.4	3.0
Bremen	0.3	0.8
Hamburg	2.5	2.2
Hesse	6.6	7.5
Mecklenburg-Western Pomerania	3.0	1.9
Lower Saxony	8.9	9.6
North Rhine-Westphalia	19.1	21.6
Rhineland-Palatinate	3.8	4.9
Saarland	0.6	1.2
Saxony	6.5	4.9
Saxony-Anhalt	2.5	2.7
Schleswig-Holstein	4.2	3.5
Thuringia	2.8	2.6

Notes: Column (1) presents the share of students in the experimental sample by federal state. Column (2) presents the share of Germany's population by federal state. Source: German Statistical Office.

Table A2: Selective attrition

	Follow-up I			**Follow-up II**			**Follow-up III**		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treat	0.019	0.017	0.016	0.021	0.020	0.020	0.054*	0.055*	0.055*
	(0.497)	(0.535)	(0.553)	(0.443)	(0.459)	(0.476)	(0.063)	(0.059)	(0.055)
Treat × baseline e-books			-0.047			-0.047			0.021
			(0.115)			(0.138)			(0.475)
Treat × baseline books			0.043			0.052*			0.045
			(0.135)			(0.072)			(0.139)
Treat × baseline reading comprehension			0.020			-0.005			0.008
			(0.510)			(0.864)			(0.793)
Treat × baseline math grade			-0.011			-0.005			0.016
			(0.737)			(0.882)			(0.652)
Treat × baseline German grade			0.005			0.000			-0.022
			(0.877)			(0.991)			(0.537)
Baseline e-books		-0.005	0.015		-0.000	0.019		0.015	0.008
		(0.729)	(0.423)		(0.984)	(0.310)		(0.336)	(0.676)
Baseline books		0.002	-0.020		0.003	-0.024		0.005	-0.018
		(0.865)	(0.323)		(0.844)	(0.253)		(0.730)	(0.425)
Baseline reading comprehension		0.013	0.002		-0.013	-0.011		0.017	0.011
		(0.588)	(0.946)		(0.580)	(0.692)		(0.504)	(0.715)
Baseline math grade		0.001	0.006		0.017	0.019		0.015	0.007
		(0.944)	(0.792)		(0.333)	(0.441)		(0.412)	(0.772)
Baseline German grade		0.033*	0.031		0.034*	0.035		0.027	0.039
		(0.061)	(0.193)		(0.056)	(0.175)		(0.134)	(0.129)
Number of observations	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000

Notes: The table shows the selective attrition analysis for follow-ups I–III. All dependent variables are binary, taking the value one if the child participated in the survey, and zero otherwise. In columns (1), (4), (7) only the treatment status is included as an independent variable. In columns (2), (5), (8), additional baseline covariates are added and, in columns (3), (6), and (9), all these baseline covariates are interacted with the treatment status variable. Math and German grades are dichotomized to take the value one if the child has a very good or good grade, and zero otherwise. All specifications include strata fixed effects. Robust standard errors are in parentheses. Follow-up I: 6 months after the start of RCT. Follow-up II: 12 months after the start of RCT. Follow-up III (endline): 30 months after the start of RCT.

Table A3: Treatment-on-the-treated (TOT) effects of e-book reading intervention on main outcomes

	Follow-up II		**Follow-up III**	
	Treatment-on-the-treated (TOT)	p-value from randomized inference	Treatment-on-the-treated (TOT)	p-value from randomized inference
Reading engagement (index)	0.134	0.000	0.071	0.040
At least one ebook	0.258	0.000	0.161	0.000
At least one printed book	0.070	0.074	0.017	0.724
At least one book	0.105	0.006	0.005	0.986
At least two days/week	0.099	0.022	0.106	0.082
Academic performance (index)	0.126	0.018	0.106	0.084
Reading comprehension	0.158	0.100	0.156	0.060
Good math grade	0.104	0.020	0.130	0.008
Good German grade	0.060	0.324	0.016	0.722
SDQ score	-0.132	0.092	-0.176	0.060
Grit	-0.047	0.508	-0.022	0.736
Locus of control	0.066	0.528	-0.103	0.398
School aspirations	0.005	0.806	0.054	0.210

Notes: This table shows the TOT estimates of the e-book treatment on the two indices (*reading engagement* and *academic performance*) and on the 11 separate outcome measures. The variables in the index *reading engagement* are binary, taking the value one if the child has read at least one (e-)book in the last four weeks or reported reading on at least two days in a typical week. Reading comprehension is a standardized score of the reading test a year after the intervention. Math and German grades are dichotomized to take the value one if the student has a good or very good grade (1 or 2 on the German grading scale), and zero otherwise. Components of SDQ are standardized with a mean of zero and a standard deviation of one. All regressions control for strata (randomization block), year and month of birth fixed effects. Only observations with observed outcomes are included. The table reports p-values from randomization inference (Young, 2019) and sharpened false-discovery rate (FDR) q-values. Follow-up II: 12 months after the start of RCT. Follow-up III: 30 months after the start of RCT.

Online Appendix B: Mediation Analysis

We perform the mediation analysis for our three main outcomes: reading engagement index, academic performance index, and SDQ score. The potential mediators we investigate are a) the child's reading self-concept, b) positive leisure activities, and c) study hours. To shed light on the mediators of academic performance and SDQ, we additionally consider the index of reading behavior as a mediator. All mediator variables are measured in follow-up I (six months after the start of the intervention), before the outcome assessments 12 and 30 months later. The mediators we consider are those that are positively and significantly impacted by the treatment.

When examining the impact of leisure activities, we categorize them into positive types and less constructive types. Positive leisure activities include computer usage, music-related activities, and sports. Conversely, less constructive activities include social media engagement and watching television. Since the treatment does not significantly affect the latter, they are excluded from our mediation analysis. We asked respondents to report the frequency of their engagement in these activities during a typical week.

Following the approach of Heckman and Pinto (2015), we assume that the treatment influences the outcomes through both direct and indirect channels. The indirect effects operate via the observed mediators discussed above, while the direct effects include all unmeasured factors. Our mediation analysis, therefore, aims to quantify the relative contribution of these indirect effects. To implement this, we augment equation 1 by incorporating the vector of mediators, denoted as M. The coefficients associated with these mediators, γ^j, where j indexes the individual mediators, are then estimated. Subsequently, we estimate the treatment effect of the intervention on each mediator separately, capturing these effects with the coefficients λ^j. The contribution of each mediator to the overall treatment effect is calculated as the ratio: $\frac{\hat{\gamma}^j \times \hat{\lambda}^j}{\hat{\beta}}$, where $\hat{\beta}$ represents the estimated treatment effect on the outcomes derived from the unconditional regression in equation (1).

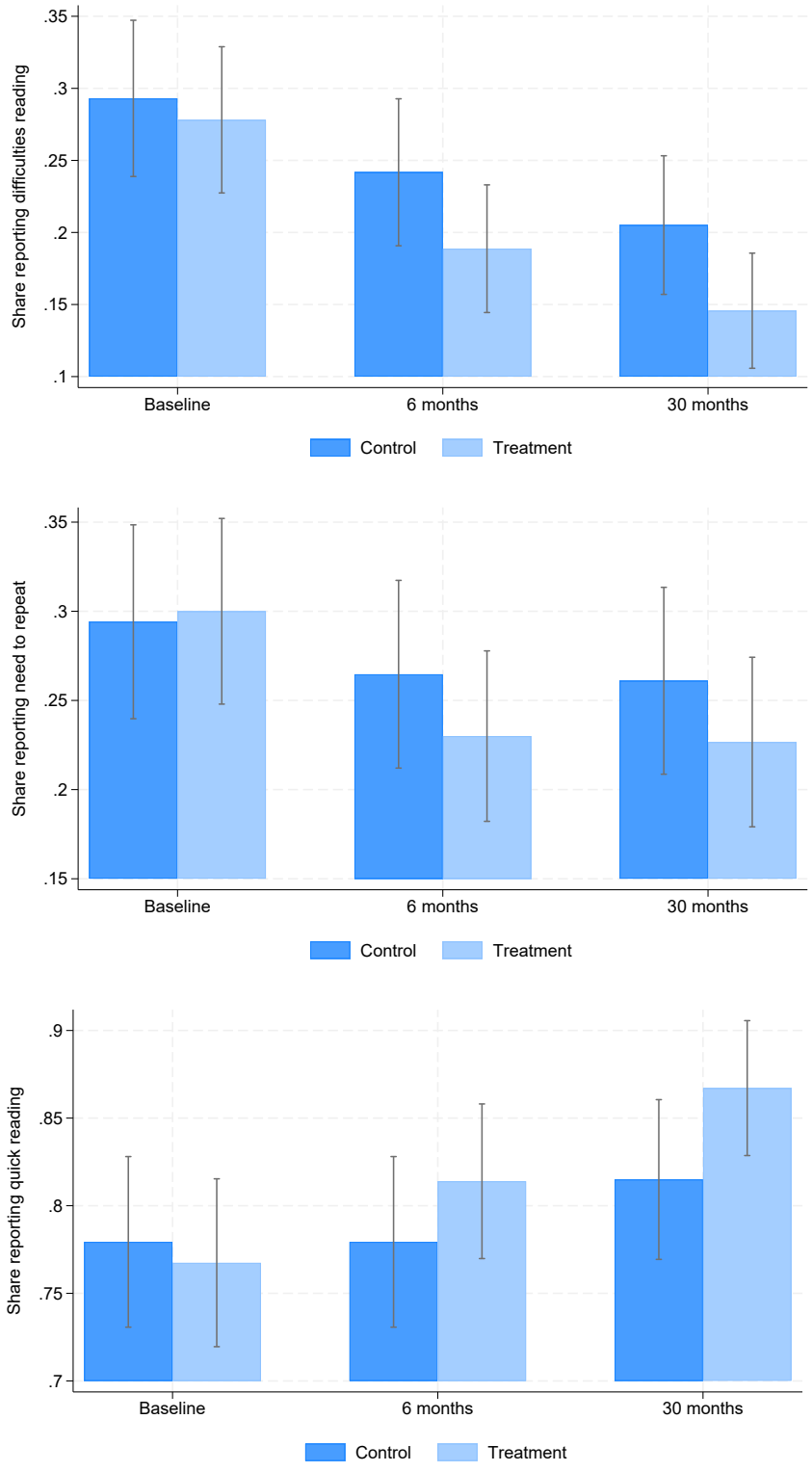
The results for outcomes measured 12 and 30 months after the start of the RCT are illus-

trated in Figure B2. The top panel examines potential mediators of reading engagement index, followed by academic achievement index, and the SDQ score. Each mediator's contribution is represented by different shades of the horizontal bar, with the gray area indicating the unexplained portion of the treatment effect.

The figures reveal that the explanatory power of the measured mediators is modest in explaining the treatment effect on the reading index. Specifically, 92% of the observed positive effect on reading behavior remains unexplained by the factors we consider 12 months after the start of the RCT. The equivalent percentage 30 months later is 85%. Next, we decompose the effect on academic performance index, where we find that that reading self-concept and reading behavior are significant mediators for academic achievement, especially in the longer term. 30 months after the start of the experiment, these factors combined explain just above 60% of the effect we find. When it comes to children's socioemotional well-being, reading behavior seems to be important in explaining our results 12 months after the start of the RCT, but in the longer-term the considered mediators do not appear to explain the effect we estimate.

We would like to note, however, that drawing causal inferences from these results requires making assumptions about the sources of variation in the mediators, as discussed by Imai et al. (2010). Since we do not have exogenous variation across specific channels and are limited to a single instance of randomization, this analysis can only provide suggestive evidence regarding the role of different mediators in explaining the observed treatment effects.

Figure B1: Self-assessed reading skills (self-concept)

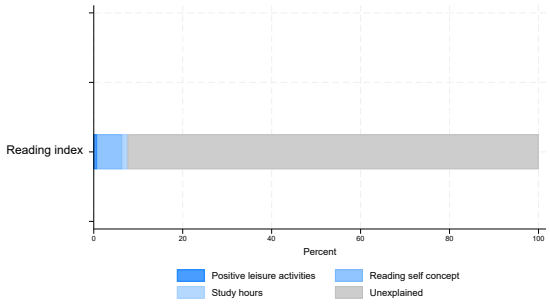


Notes: The figure presents the share (and 95% confidence intervals) of children reporting reading difficulties (top panel), the need to reread text (middle panel), and the share considering themselves fast readers (bottom panel). The results are shown for the baseline and two follow-up surveys (follow-up I: 6 months after the start of RCT) and follow-up III (30 months after the start of RCT.)

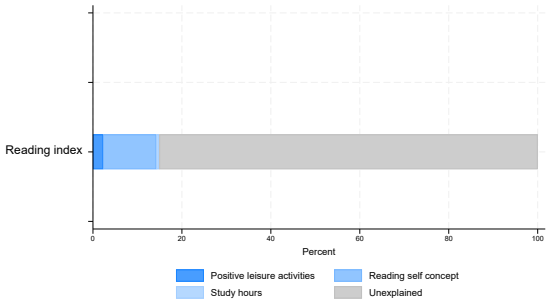
Figure B2: Mediation analysis

Reading engagement (index)

(a) Follow-up II

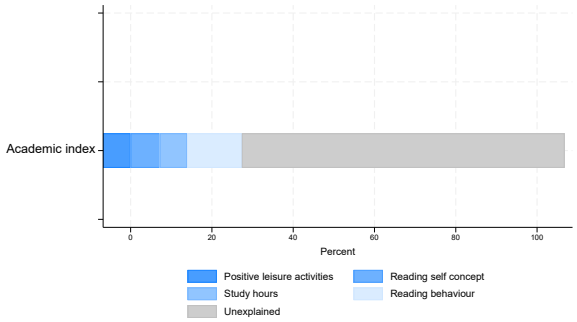


(b) Follow-up III

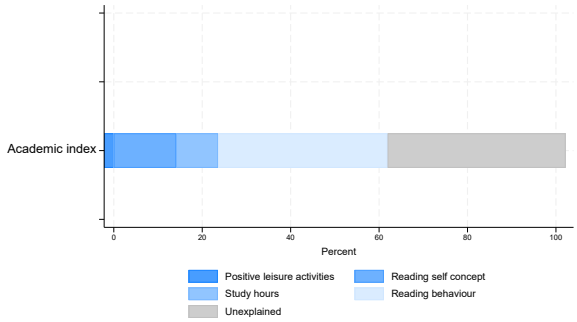


Academic performance (index)

(c) Follow-up II

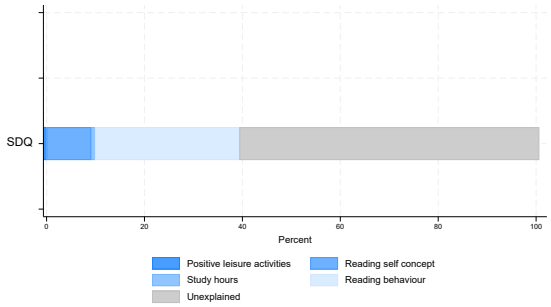


(d) Follow-up III

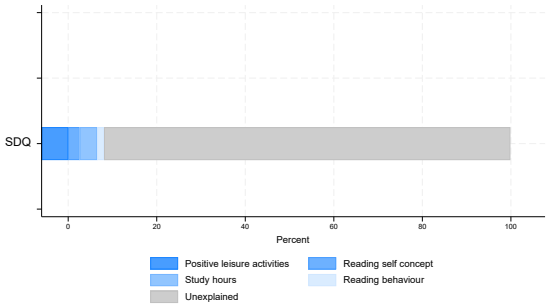


SDQ score

(e) Follow-up II



(f) Follow-up III



Notes: This figure shows the results of our mediation analysis. Panel (a) shows the decomposition of the treatment effect on the outcome reading index, panel (b) displays the decomposition for the academic achievement index, and panel (c) shows the decomposition of the treatment effect on the outcome SDQ index. Follow-up II: 12 months after the start of RCT. Follow-up III (endline): 30 months after the start of RCT.

Online Appendix C: Survey Materials and Questionnaires

Cover letter and information sheet about the e-reader

Own translation from German

Mr./Mrs. [Last Name]

Recipient's Address

City, Zip Code

Dear Mr./Mrs. [Last Name],

We would like to thank you and your child once again for participating in the first CoDu survey in fall 2020. The next survey has been delayed due to current developments and is expected to take place in mid-2022. We would be delighted if you were also to take part in further surveys.

In the meantime, we would like to provide an e-reader to [your daughter/son] who was born in [month of birth] [year of birth] and attended 4th or 5th grade in the last school year and participated in our first survey. As part of our study, children will receive an e-reader giving them access to over 1,000 popular and age-appropriate books for free for a year. You can find more information about the e-reader on the next page.

If you agree, you don't need to do anything else. The e-reader will be sent to you automatically within the next three weeks. This will not incur any costs for you.

If you do not wish to receive the e-reader, please return the detachable, anonymized reply form on the last page in the stamped addressed envelope enclosed by November 26, 2021. This will not incur any costs or disadvantages for you.

However, we hope that you and your child will take this opportunity to read fun and adventurous stories together.

Sincerely,

P.S. Below you will find additional information on the e-reader:

Information sheet on the e-reader:

- As part of our study, we are providing you and your child with an e-reader free of charge.
- Parental controls are already activated on the e-reader, so you do not need to set up anything.
- With the e-reader, your child can read over 1,000 popular and age-appropriate books for a year free of charge.
- The e-reader makes reading easy on the eyes.
- There is no access to games, the internet, or videos. No messages can be sent, and there is no advertising.
- If you have any questions or problems, please call our hotline at (0911) 179 XXXX, and we will be happy to help you.

Questionnaire excerpts

Reading

1. How much time do you usually spend reading outside of school?

   Consider books, e-books (e.g., for Kindle or Tolino), or magazines. Please provide the number of hours and minutes. If you read for less than one hour, write `0` in the first field.

   - On a school day, I usually read (outside of school) for:

     _____ hours and _____ minutes.

2. On a non-school day, I usually read (outside of school) for:

   _____ hours and _____ minutes.

3. How many days per week do you usually read outside of school?

   Consider books, e-books (e.g., for Kindle or Tolino), or magazines.

   - On 0 days/week
   - On 1 day/week
   - On 2 days/week
   - On 3 days/week
   - On 4 days/week
   - On 5 days/week
   - On 6 days/week
   - On 7 days/week

4. How many books have you read outside of school in the last 4 weeks?

   By books, we mean books in paper format and e-books (e.g., for Kindle or Tolino). Please specify the number for e-books and paper books separately.

   - Number of e-books read:
     - No e-books
     - 1 e-book
     - 2 e-books

  - 3 e-books
  - 4 e-books
  - 5 e-books
  - More than 5 e-books

- Number of paper books read:
  - No books
  - 1 book
  - 2 books
  - 3 books
  - 4 books
  - 5 books
  - More than 5 books

School aspirations

What school qualification do you aspire to achieve?

- Hauptschulabschluss (basic school-leaving certificate)
- Realschulabschluss / Mittlere Reife (intermediate school-leaving certificate)
- Abitur (university entrance qualification)
- Leave school with no qualification

Locus of control

To what extent do you agree with the following statements?

Scale: **1 = Strongly disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly agree**

- What one achieves in life is mostly a matter of fate or luck.
- Success must be earned through hard work.

Grit

To what extent do the following statements apply to you? There are no right or wrong answers!

Scale: **1 = Not at all like me, 2 = Not much like me, 3 = Somewhat like me, 4 = Mostly like me, 5 = Very much like me**

- New ideas and projects sometimes distract me from previous ones.
- Setbacks do not discourage me.
- I am passionate about an idea or project at first, but lose interest later.
- I am someone who works hard.
- I often set myself a goal but switch to a different one later.
- I find it hard to stay focused on a project that takes more than a few months.
- I finish everything I start.
- I am careful and thorough in my work.

SDQ

To what extent do the following statements apply to you? There are no right or wrong answers!

Scale: **1 = Not at all like me, 2 = Not much like me, 3 = Somewhat like me, 4 = Mostly like me, 5 = Very much like me**

Item	Not true	Some-what true	Certainly true
I try to be nice to other people. I care about their feelings.			
I am restless, I cannot stay still for long.			
I get a lot of headaches, stomach-aches, or nausea.			
I usually share with others, for example CDs, games, food.			
I get very angry and often lose my temper.			
I would rather be alone than with people of my age.			
I usually do as I am told.			
I worry a lot.			
I am helpful if someone is hurt, upset, or feeling ill.			
I am constantly fidgeting or squirming.			
I have one or more good friends.			
I fight a lot. I can make other people do what I want.			
I am often unhappy, depressed, or tearful.			
Other people my age generally like me.			
I am easily distracted, I find it difficult to concentrate.			
I am nervous in new situations. I easily lose confidence.			
I am kind to younger children.			
I am often accused of lying or cheating.			
Other children or young people pick on me or bully me.			
I often offer to help others (parents, teachers, children).			
I think before I do things.			
I take things that are not mine from home, school, or elsewhere.			
I get along better with adults than with people my own age.			
I have many fears, I am easily scared.			
I finish the work I'm doing. My attention span is good.			