WORKING PAPER · NO. 2023-02

Nudging or Nagging? Conflicting Effects of Behavioral Tools

Ariel Kalil, Haoxuan Liu, Susan Mayer, Derek Rury, and Rohen Shah

JANUARY 2023

Becker Friedman Institute
FOR ECONOMICS AT UCHICAGO

5757 S. University Ave.
Chicago, IL 60637
Main: 773.702.5599
bfi.uchicago.edu

NUDGING OR NAGGING? CONFLICTING EFFECTS OF BEHAVIORAL TOOLS[*]

Ariel Kalil
Haoxuan Liu
Susan Mayer
Derek Rury
Rohen Shah[†]

January 2023

Abstract:

The gap in reading skills between low-income children and their higher income peers emerges early in life. To help boost the reading skills of low-income children, we conducted an RCT with low-income parents of young children in Chicago. The RCT aimed to increase parental reading time and child's literacy skills. Parents were randomized into 4 groups: 1) a control group, and groups that received 2) a digital library tablet, 3) a digital library tablet with reminder texts, and 4) a digital library tablet with goal-setting texts. Both reminders and goal setting text messages were designed to nudge parents to manage present bias. Relative to the digital library tablet treatment, we find that goal-setting messages led to an increase (.32 SD) in parent reading time but had no effect on literacy skills. Unexpectedly, reminders messages led to a *decrease* in literacy skills, despite no significant difference in reading time. This demonstrates that nudging might have the unintended consequence of reducing the quality of the task (reading). Another important result is that technology may help boost the reading skills of low-income children: children in the digital library tablet group increase their literacy skills by .29 SD relative to the control group. All results are robust to controlling for school fixed effects, age and in the case of literacy skills, baseline assessment scores.

JEL Codes: C93, D80, D91, I20.
Keywords: Education, Behavioral Economics, Field Experiments

[*]This study was registered on the AEA RCT Registry (AEARCTR-0005195) and received approval from the Social and Behavioral Sciences IRB from the University of Chicago (IRB19-0953). We thank William Delgado, Sebastian Gallegos, Phil Oreopoulos, and conference participants at AEA, AFE, APPAM, SEA, SABE, and ESA for helpful comments. We are very grateful to the staff at the Behavioral Insights and Parenting Lab for their invaluable effort: Ana Arellano Jimenez, Michelle Park Michelini, and Paula Rusca. We also thank the numerous research assistants without whom this work would not have been possible. This was work funded by the Valhalla Foundation.
†Corresponding author. Email: shahr@uchicago.edu. Phone: 734-578-5684. Address: Harris School of Public Policy, University of Chicago, 1307 E. 60th St., Chicago IL 60637 USA.

1. INTRODUCTION

The achievement gap in literacy between advantaged and disadvantaged children emerges before formal schooling begins and persists over the school years (Heckman, 2006; Waldfogel & Washbrook, 2011). Time-use surveys show that advantaged parents spend more time on educational activities with their children (Guryan et al., 2008; Kalil et al., 2012), and there is also evidence of a causal link between this type of parent engagement and child skills (Price, 2010; Villena-Roldan & Ríos-Aguilar, 2012). Consequently, interventions have attempted to increase parental engagement such as reading to their children. In particular, interventions that prompt parent engagement with text messages designed with a behavioral tool, or "nudge" have shown some promise, in part because of their low cost relative to programs that have a higher-touch design, such as visits to a parent's home by a trained practitioner (Mayer et al., 2019; York et al., 2019).

One difficulty with interventions aimed at increasing reading is measuring the time parents read with their children. While York et al. (2019) finds that that text message reminders led to an increase in child literacy skills, their measure for reading time is self-reported and therefore may be subject to measurement error. Mayer et al. (2019) overcomes this issue by providing parents with a tablet that contains a digital library of books and tracks each parent's reading time. However, Mayer et al. (2019) does not measure literacy skills, and therefore it is unclear whether the additional reading induced by behavioral tools leads to an increase in child literacy skills. Additionally, the text messages in the intervention in Mayer et al. (2019) reflected a bundle of behavioral tools (reminders, goal setting, peer competition), leaving it unclear which behavioral tool drives the treatment effect.

In the present study, we implement an 11-month RCT with 379 low-income parents in Chicago to study both parent-child reading time and child literacy skills. Parents were randomized into four groups: 1) a control group, 2) a group that received a tablet containing a digital library, 3) a digital library tablet group with reminder texts, and 4) a digital library tablet group with goal-setting texts. This design allows us to distinguish between two different types of behavioral tools meant to address present bias: reminders and goal setting. Additionally, this design allows us to measure the impact of receiving a digital library tablet on child literacy skills.

We find that, relative to the group that received only the digital library tablet, adding goal setting messages increased parent reading time by 50% (.32 SD), whereas adding reminder messages had no significant impact on parent reading time. Although the magnitude of impact is smaller than the results from Mayer et al. (2019), we replicate the finding that behavioral tools, delivered via text messages, increase reading time. Further, we can identify goal setting messages as the driver of this effect, relative to reminder messages.

The impact of this behavioral messaging on literacy skills, however, reveals an unintended negative consequence of behaviorally-informed, or "nudging" interventions. Despite leading to a significant increase in reading time, the goal setting messages had no significant impact on child literacy skills relative to the digital library tablet group. Further, the reminder messages led to a significant *decrease* in literacy skills compared to the tablet group, despite no significant difference in reading time. We describe a theoretical model of skills production that is consistent with our findings for both goal setting and reminder messages. We hypothesize a "nag factor" that scales down task quality as an unintended consequence of nudging interventions. This relates to the literature on intrinsic and extrinsic motivation, where monetary incentives potentially backfire if they reduce intrinsic motivation. Nudge interventions are often described

as having high benefit-cost ratios because even small benefits outweigh the nearly-zero cost of sending a text message. Our model implies that nudges could have a potentially high cost, thereby challenging this conventional wisdom.

We also find that the digital library tablet itself caused a significant increase in literacy skills relative to the control group. The treatment group who received just the digital library tablet scored 0.30 SD higher than the control group on the literacy skills test.[2] Because low-income parents are less likely to have access to this technology than their high-income counterparts, this finding highlights the role that technology could play in raising child skill.

This work contributes to several literatures. First, this work adds to the growing literature on the hidden cost of nudges (Allcott & Kessler, 2019; Barron et al., 2022; Damgaard & Gravert, 2018). The value proposition for text message reminders is typically that they are almost costless to implement, which justifies their use even if their benefit is relatively small. Our work in contrast shows that while text message reminders might have a relatively low monetary cost, the decrease in intrinsic motivation might change the nudge to a "nag" and reduce the quality of task performance. This potential cost should be considered if behavioral tools are used to change behavior for a task where the quality of performance matters. The difference in findings for reading time versus literacy skills also highlights the limitation of using proxy outcome (such as reading) in place of an intended outcome (literacy skills).

This study also adds to the literature on human capital development, highlighting the role of technology (Escueta et al., 2020). While there has been work studying educational technology used at home in a developing or middle-income country setting (Beuermann et al., 2015; Malamud & Pop-Eleches, 2011), there is little research on this topic in a developed country

[2] If we pool the average score for all three treatment groups that received a digital library tablet the treatment effect is 0.20 SD higher than the control group after controlling for baseline scores.

setting. We provide, to our knowledge, the first experimental evidence of providing low-income families with digital literacy technology on child literacy skills. Because the technology used in our intervention is widely available, the results from this study can help policymakers make important cost-benefit decisions about how best to address the early-life gap in literacy skills between low- and high-income children.

Our work also contributes to the literature that explores heterogeneity in effects of behavioral tools. For example, Clark et al. (2020) conduct an RCT to find that goal setting messages are effective when the goal is tied to the task but not when it is tied to performance, and Cortes et al. (2021) find that sending three text messages per week to parents was more effective at increasing child literacy compared to sending either one or five messages per week.

The remainder of this paper proceeds as follows. Section 2 describes the theoretical framework. Section 3 describes our experimental design, including the sample, treatments, and outcomes. Section 4 shows the results of the experiment. Section 5 provides a discussion of the findings as well as limitations, and Section 6 concludes.

2. THEORETICAL FRAMEWORK

This section provides our theoretical framework that describes how prompting parents with behavioral tools could lead to increased child literacy skills.

2.1 Setup

In this model, parent i chooses how much time to invest in reading in the current period, denoted R_{i0}, to maximize their net benefits V, given by:

$$V(R_{i0}) = u(R_{i0}) - c(R_{i0})$$

Where $u(R_{i0})$ and $c(R_{i0})$ are twice differentiable functions such that $u''(R_{i0}) < 0$ and $c''(R_{i0}) > 0$. Specifically, $u(R_{i0})$ and $c(R_{i0})$ are the present values of the benefits and costs respectively of reading R_{i0} minutes in the current period. That is:

$$u(R_{i0}) = \sum_{t=0}^{\infty} \delta^t \cdot u_t(R_{i0})\,; \qquad c(R_{i0}) = \sum_{t=0}^{\infty} \delta^t \cdot c_t(R_{i0})$$

Where δ is the discount factor and u_t and c_t are the benefit and cost functions in period t.

2.2 Present Bias

Let us assume that the parent is present-biased and has a quasi-hyperbolic discount function (Laibson, 1997). That is, the functions above are now:

$$u(R_{i0}) = u_0(R_{i0}) + \beta \sum_{t=1}^{\infty} \delta^t \cdot u_t(R_{i0})\,; \qquad c(R_{i0}) = c_0(R_{i0}) + \beta \sum_{t=1}^{\infty} \delta^t \cdot c_t(R_{i0})$$

Where $\beta \leq 1$ represents the present-bias factor. For exposition, assume that 1) the only cost of reading occurs in the present period (so $c_t(R_{i0}) = 0$ for $t > 0$), 2) the benefit is received in a single future period T, and 3) $\delta = 1$. For ease of notation, we suppress indices on R_{i0}. The net benefit V simplifies to:

$$V(R) = \beta u_T(R) - c_0(R)$$

This results in the FOC: $\beta u_T'(R^*) = c_0'(R^*)$

2.3 Impact of Behavioral Tools

We define a behavioral tool, or a "nudge", as an intervention that increases the value of the present bias factor from β to $\beta' \in [\beta, 1]$. For exposition, we make standard functional form assumptions where $u_T(R) = \frac{R^{1-\alpha}}{1-\alpha}$ and $c_0(R) = \frac{C(R)^{1+\varepsilon}}{1+\varepsilon}$, where $\alpha, \varepsilon > 0$.

$$V(R) = \beta \cdot \frac{R^{1-\alpha}}{1-\alpha} - \frac{(C(R))^{1+\varepsilon}}{1+\varepsilon}$$

$$0 = \beta R^{-\alpha} - C(R)^{\varepsilon} C'(R)$$

$$C(R)^{\varepsilon} C'(R) = \frac{\beta}{R^{\alpha}}$$

$$R^{*} = \left(\frac{\beta}{C(R)^{\varepsilon} C'(R)} \right)^{\frac{1}{\alpha}}$$

$$\frac{\partial R^{*}}{\partial \beta} > 0$$

assuming $C, C' > 0$. This shows that an intervention that increases β will increase the parent's choice of reading time R^{*}. Note that if a parent is not present biased $(\beta = 1)$, the behavioral tool will result in no change in R^{*}.

2.4 Literacy Production Function

One aspect of the parent's utility $u_T(R)$ is the impact that the reading will have on their child's literacy skills. We define $y = f(R)$ as a literacy production function that maps a parent's reading time R onto their child's literacy skills y. Assuming $f'(R) > 0$, a behavioral tool that increases parent reading time would be expected to increase their child's literacy skills; that is:

$$\frac{\partial y}{\partial \beta} = \frac{\partial f(R)}{\partial R} \cdot \frac{\partial R}{\partial \beta} > 0$$

With this model, we see the potential benefits of behavioral tools. Later in the paper, we will extend this in the context of our experiment and describe potential nagging costs in Section 5.

3. EXPERIMENTAL DESIGN

3.1 Sample

We recruited families from 13 subsidized preschools across Chicago to participate in the study. Nine of these were Head Start early childhood centers in Chicago, and the remaining four were operated by Chicago Public Schools as part of the "Preschool for All" program. To be eligible to participate, parents had to have a child between three and four years old enrolled in a participating preschool. In addition, parents had to have a mobile phone and be willing to receive up to four text messages per week and speak either English or Spanish as their primary language. Parents were also told that they would be asked to participate in two surveys and that their child's literacy skills would be assessed twice.

Eight of the centers allowed opt-out recruitment, where all parents were automatically enrolled and received an informational flyer with a description of the study and directions for several ways they could opt out. The other five centers were opt-in, meaning that parents had to actively sign up to participate. If a parent had more than one qualifying child in the preschool, all the siblings were assigned to the same treatment group.

We ultimately recruited 617 children to participate in the study. Nine dropped out before randomization. Of the remaining 608 children, 14 were siblings of enrolled children. We dropped siblings, leaving only one child per parent, leaving us with 594 children. At the beginning of the randomization, one child was dropped because of the wrong treatment materials provided. This

left 593 children who were assigned to three treatment arms and one control arm. After the randomization, their parents could decide to drop out of the study by contacting us directly or by notifying their preschool that they no longer wanted to participate.

3.2 Treatment

The digital library is an application on an electronic tablet called "CAPER" (Children and Parents Engaged in Reading). To create the CAPER digital library, we secured a digital rights extension agreement with multiple partnerships to populate our digital library, including *Cricket Media* and *Jumpstart Book*s. This was done with the support of the University of Chicago Library. We also obtained high-quality, free Spanish books that required permission from the authors. We carefully vetted books to ensure they were age-appropriate and engaging for our target audience. The CAPER app had 207 books (149 in English and 58 in Spanish).

No other applications were available on the tablets. Outside the CAPER app, only the camera and microphone were enabled, although families could not see these icons or access them in any way. Parents and children could not use the internet or download any apps to the tablet. This is important to note, because a treatment that provides parents access to an unrestricted electronic tablet might not be expected to have the same effect if it is primarily used for applications other than the digital library. When a parent opened the CAPER app, a video and audio recording began automatically. This allows us to measure the amount of reading done using the app directly rather than relying on self-reports. Parents that received tablets kept them for the 11-month duration of the study, from December 2019 until October 2020.

Parents were randomly assigned to one of four experimental arms: (1) control; (2) digital library tablet only; (3) digital library tablet with reminder text messages; and (4) digital library

tablet only with goal setting text messages. The randomization was stratified at the school level and assigned at the individual level. While this might lead to some concerns with spillover effects, this design maximizes statistical power. Parents assigned to the control group received an activity book with crayons and stickers for their child at the beginning of the intervention as a thank you for participating. Parents in the digital library tablet only *(tablet only)* group received the tablet for 11 months. Parents in these two groups also received administrative text messages such as welcome and thank-you messages and requests to complete the surveys and assessments. Table A1.1 in Appendix 1 provides sample text messages as well as the administrative messages that were sent to parents in the control and the tablet only groups.

Parents in both the digital library tablet with reminders (*reminders*) and digital library tablet with goal setting messages (*goal setting*) treatment groups received text messages every week for 11 months in addition to the digital library tablet. The parents in the *reminders* group received 4 text messages per week reminding them to use the CAPER app to read to their child. Examples of reminder messages are, "Does [child] have a favorite CAPER book? Read it to [him/her] tonight!" and "Don't forget to start the new week by finding time to read CAPER books to your child." Table A1.2 provides additional examples and the schedule of messages for the reminders group.

Parents in the *goal setting* group also received 4 text messages per week; these messages were focused on meeting reading goals. These texts are shown in Table A1.3. In addition to these 4 weekly messages, parents in this group received a message in the first week of the intervention assigning them a goal of reading for 10 minutes during the first week. Every subsequent Friday, we sent parents a text message informing them of the number of minutes they read using the CAPER app in the previous week. This allows parents to see whether they met their assigned

goal. Moreover, every two weeks, we gave parents a new weekly goal. The new goal was automated such that it was always 20% higher than a parent's average weekly reading time from the previous two weeks[3]. The goal-setting algorithm is based on the finding that outcomes tend to improve the most when goals are continually increased in small amounts. (Landers et al., 2017; Tondello et al., 2018).

Our behavioral messages aim to reduce present bias. Present bias arises from the tendency to put more weight on the present than on the future, which leads to procrastination in the short term and preference reversals (i.e., regret) in the future. There is a large literature studying present bias across the social sciences, showing that the type of behavioral texts we use can help individuals "bring the future to the present," thereby closing the intention-action gaps (Chabris et al., 2008; Meier & Sprenger, 2010; Rodgers et al., 2005).

3.3 Measures

The two primary outcome variables for this study are parent reading time and child literacy skills. We measured parental reading time using the minutes read in the CAPER app. Note that this is not available for the control group as parents in this group were not assigned the CAPER app and did not receive a digital library tablet.

We measured children's literacy skills at baseline and 11 months later at the end of the intervention using three subscales of the core language index of the Clinical Evaluation of Language Fundamentals (CELF-P). These subscales measure general language ability. At

[3]Goals were rounded up to the next highest multiple of 5. The minimum possible goal was 10 minutes. The 20% increase is dropped if their average time is greater than 75 minutes The maximum possible goal was 90 minutes. The goal is never lower than 80% of the previous period's goal. An example of a goal and feedback message is, "In the past two weeks you read an average of [feedback] minutes per week to [child_name] using CAPER. This week try reading [goal] minutes."

baseline, we assessed children in person by having an assessor visit the preschool and test one student after another within a four-week assessment period. The assessment at the end of the intervention was conducted remotely via a video Zoom call due to COVID-19 restrictions.

To assess children remotely, assessors phoned parents to schedule individual appointments and help parents set up the required technology. The technology included a device that had a screen that was at least 9.7 inches (a desktop, laptop, tablet) with a reliable internet connection and the capability to join a video Zoom call with the assessment page open on a web browser. Many parents did not have a computer, internet access or computer skills sufficient to use the software for the assessments. In those cases, the research team provided both equipment and technical assistance.

We also surveyed parents at baseline and at the end of the intervention. These surveys collected data on parent and child demographics so that we could assess the representativeness of the sample, balance across treatment groups, and study potential attrition from the project. At both baseline and the end of the intervention the surveys also collected data on parents' attitudes and behavior relevant to their child's skills to study how parents responded to the treatments.[4]

3.4 Attrition

While 505 families were randomized into a treatment group that required a tablet, we only distributed 479 tablets because 26 families either refused to accept or never picked up the

[4]At baseline, research assistants surveyed parents when they dropped off or picked up their child from preschool. We surveyed all 593 parents who had been randomized. We administered the follow-up survey remotely using a hybrid of online and telephone survey methods. We were able to survey 379 (64%) parents who were surveyed at the baseline. Survey data results are available upon request.

tablet from the preschool. Of the 479 tablets distributed, we retrieved 351 at the end of the intervention and used their reading data as the analytical sample for reading time.[5]

We assessed 553 children (93.25% of children who were randomized) at baseline using the CELF assessment. At the end of the intervention, we assessed 300 children (54.24% of children who were randomized) among 553 children assessed at the baseline. The smaller number of children assessed at follow-up reflects our having to switch to remote assessments due to the COVID-19 pandemic. Because of the constraints of remote assessment, it took 11 weeks to assess children at the end of the intervention. However, there was no statistically significant difference in attrition rate across treatment groups or by baseline score. Our analytic sample in this study is therefore the 300 children for whom we have both baseline and final test scores.

4. RESULTS

4.1 Descriptive Results

Table 1 shows the descriptive statistics for demographic variables, including a balance test across treatment conditions. There is no significant difference across treatment conditions for the available variables. Additionally, in terms of race, parental education, and gender, there is no statistically significant difference between the children in our sample and that of the Head Start population based on the Head Start Family and Child Experience Survey (FACES) study in 2019

[5] See Table A2.1 for the distributions of tablets across treatment groups. Among 128 unretrieved tablets, there were 46 tablets in the goal setting group. Because these 46 tablets were cellular-enabled to get real-time data and set reading goals during the intervention, we could collect reading data from these tablets remotely at the end of the intervention. However, we are concerned that parents who did not return the tablets were systemically different from parents who did return the tablets; further, we were unable to collect data from unretrieved tablets from the tablet-only and the reminders groups. Any comparison of the reading time of (1) parents who were in the tablet-only and the reminders groups and returned the tablets with (2) parents in the goal setting group who did not return the tablet may be biased. Thus, we did not use those 46 observations from the unretrieved tablets in the goal setting group.

(Kopack Klein et al., 2021). While not conclusive, this gives us confidence that our sample might be representative of the population of low-income families enrolled in Head Start.

(Table 1 here)

Figure 1 shows the total number of minutes read by parents in each of the treatment groups over the entire 11-month duration of the study. From this we can see that there is a steady decrease in the amount of reading over time across all treatment conditions. The decrease in usage over time might be due to a "novelty" effect, as is the case with many studies where a treatment becomes less enticing to participants over time (Asensio & Delmas, 2016). Using a regression discontinuity model, we find that there was no significant change in reading immediately before and after COVID-19 school closures.

(Figure 1 here)

4.2 Regression Results for Reading Time

We first estimate the following regression to measure the treatment effect of reminder and goal setting messages on parent reading time:

$$y_i = \beta_0 + \beta_1 R_i + \beta_2 G_i + \varepsilon_i$$

where y_i represents parent i's reading time, R_i and G_i are indicators for assignment to reminders and goal setting treatment conditions respectively and ε_i is the error term. The omitted (comparison) group in this regression are those in the digital library tablet group. Note that this analysis did not include parents assigned to the control group, where no tablet was provided and therefore no reading time is available. The result of this regression is shown in Table 2.

(Table 2 here)

In the first column, we see that the average reading time for the digital library tablet group was 270 minutes. Relative to this group, the reminder group read 20 additional minutes, although this difference is not statistically significant. The goal setting group read about 134 more minutes, and the difference is statistically significant. This 50% increase in reading time caused by goal setting messages is equivalent to a 0.32 SD treatment effect. The second column includes school fixed effects and shows qualitatively similar results.[6] The third column adds the child's age to the regression and shows that the results are largely the same. Additionally, we also look at the total reading time in the first 2 months and the first 6 months of the study. As with the results from the total intervention period, only the goal setting messages cause a statistically significant increase in reading time.

[6] To study the robustness of our results on reading time, we checked whether the results hold up for alternative specifications of the outcome variable. Specifically, we consider whether the treatment affects the share of parents who read at all. Using a binary variable with a value of 1 if the parent reads a non-zero amount, we similarly find that the only the goal setting messages cause a significant increase in the share of parents who read. We also consider a binary variable with a value of 1 if the parent is in the 90th percentile of reading time, and similarly find that goal setting messages – but not reminder messages – cause an increase in the likelihood that parents reading a high amount. This implies that the goal setting messages impact parental reading at both the extensive and intensive margins.

4.3 Regression Results for Literacy Skills

We next estimate the following regression to measure the treatment effect on literacy skills:

$$L_{i,t} = \alpha_0 + \alpha_0 T_i + \alpha_2 R_i + \alpha_3 G_i + \alpha_4 L_{i,t-1} + \varepsilon_i$$

where $L_{i,t}$ represents child i's literacy skills at the end of the experiment, T_i, R_i and G_i are indicators for assignment to tablet only, reminders, and goal setting respectively, $L_{i,t-1}$ represents the baseline literacy skills and ε_i is the error term. Note that the comparison group in this regression is the control group, which did not receive a digital library tablet. The result of this regression is shown in Table 3.

(Table 3 here)

In the first column, we see that those in the tablet only group score almost 5 points higher on the CELF exam than the control group after controlling for baseline test scores. This treatment effect is 0.29 SD and is significant at the 1% level. Although the reminders group also scores higher than the control group, the difference is not statistically significant. However, the goal setting group scored a significant 0.26 SD higher than the control group. The results are qualitatively similar in Columns 2 – 4, where we include school and age fixed effects (column 2) and estimate the regression using percentile rather than raw score as the outcome (columns 3 and 4).

We also test whether the effect for the tablet only group differs from that of the reminders or goal setting group. To do so we run the regression without the control group and with the tablet only condition as the reference group. The result of this regression is shown in Table 4.

(Table 4 here)

This result shows that while goal setting messages do not cause a significant difference in literacy skills relative to the tablet only group, the reminder messages caused a 0.19 SD, or 6 percentile points, *decrease* in literacy skills relative to the tablet only group. This difference is significant at the 10% level. This result is robust to adding school fixed effects and controlling for child's age. This finding fails to support our pre-registered hypothesis that all the behavioral tools would increase literacy skills. The next section builds on the model in Section 2 to describe a hypothesis that is consistent with these findings.

5. DISCUSSION

5.1 Unintended Consequences: Nag Factor

The theoretical framework in Section 2 defined a literacy production function $y = f(R)$ that maps a parent's reading time R onto their child's literacy skills y. A parent's chosen level of reading time R_i will depend, in part, on their present bias factor β_i. We defined a nudge as an intervention that has the potential to increase β to $\beta' \in [\beta, 1]$. Because R increases with β and $f(R)$ increases in R, a nudge can therefore increase literacy skills because $f\big(R(\beta')\big) > f(R(\beta))$.

However, a downside of nudges might be that the quality of the task performed is lower, even if the quantity increases (or remains unchanged). Previous work shows that nudges may come with an annoyance cost. For example, in the charitable giving literature, Damgaard & Gravert (2018) find that potential donors give more in the short-run as a result of nudge interventions, but avoid certain types of communication about donations, ultimately leading them to either unsubscribe from a charity listserv or avoid donation drive locations. Our results are also consistent with a "crowding out" effect where extrinsic incentives may undermine intrinsic motivation (Bénabou & Tirole, 2003). Even though nudges do not provide a monetary incentive, a parent might interpret receiving a text message as a similar signal to an extrinsic incentive, which could result in reduced intrinsic quality of effort.

In this spirit, we define a "nag factor" $\eta \in [0,1]$, such that a behavioral tool also reduces quality of output, making the production function go from $y = f(R)$ down to $y = (1-\eta)f(R)$. This describes the tradeoff involved between intentional nudging and inadvertent nagging: a nudge can increase the amount of reading, but the reduced quality of reading implies that every minute spent reading has a lower marginal return on literacy skills relative to not receiving any nudges. This is illustrated in Figure 2.

(Figure 2 here)

Figure 2 depicts a case consistent with our findings for both goal setting and reminders. Point A describes the reading time R_0 and corresponding child literacy skill $f(R_0)$ for a parent who receives no nudges. Points B and C represent different types of nudges. Note that we assume here that all nudges have a common nag factor η, but not all nudges have the same

impact on the present biased factor β and therefore different nudges might change R by varying amounts.

Point B describes the impact of a nudge like our reminders, where the nudge itself fails to increase the reading time. Even though there is no change in R_0, the nag factor still causes the literacy skills to shift down to $(1-\eta)f(R_0)$. This is consistent with our reminders group having a lower level of literacy skill despite reading the same amount. Point C represents a nudge, such as the messages in our goal setting group, which does increase the reading time to R_0+R_k. Despite this increase, the literacy skills are not guaranteed to increase because of the nag factor. This nudge could be beneficial to the overall literacy skill as long as $(1-\eta)f\big(R(\beta')\big) > f(R(\beta))$, which in this example is specifically if $(1-\eta)f(R_0+R_k) > f(R_0)$. In general, we would expect this to happen if η is relatively small or if the increase in R_i is substantially large. The illustrated case depicts results consistent with our goal setting messages, where literacy skills remain unchanged despite the increased reading time.

While future research is needed to properly design an experiment that explicitly tests this hypothesis, we present some exploratory evidence for this "inefficient" reading by parents in the goal setting group. Using data on reading time minutes, we can identify what time of day parents used the CAPER app to study the effectiveness of reading time. Although there are surely individual differences among children's bedtimes in this sample, recall that the children in this study are less than four years old, on average, and that they are all enrolled in out-of-home morning preschool. We therefore hypothesize that reading to children very late at night when children are tired and inattentive (or even asleep) cannot benefit children's literacy skills. Figure 3 shows the distribution of reading by hour of day for each treatment condition.

(Figure 3 here)

Visually, the additional goal setting minutes appear to be concentrated at later hours of the day. For a more specific test we examine whether there was a significant difference in the proportion of nighttime reading by treatment status. The results are presented in Table 5 and verify that the goal setting group is significantly more likely to read during the night hours. The results are robust to different cut off points for what we consider to be "nighttime" (i.e., 8:00, 9:00, and 10:00 pm) as shown in columns 2 and 3 of Table 5.

(Table 5 here)

A higher likelihood of inefficient reading may imply that the parent is not reading primarily to benefit their child or themselves, but simply to avoid any psychic cost of not satisfying the nudge's request. One important feature of the behavior we are trying to influence (reading) is that it is not a well-defined, concrete task, as it has multiple dimensions that affect its quality (e.g., what to read, when to read, what pace and tone to use while reading, how enthusiastically to read). We find that attempts to influence reading using behavioral nudges might come with hidden costs in these dimensions.

5.2 Impact of Technology

Our experimental design also allows us to identify the impact of receiving a tablet with a digital library (CAPER) on child literacy skills. We find a substantial impact of 0.29 SD for this group. The treatment effects from the CAPER tablet compares favorably to treatment effects obtained in other interventions intended to increase children's literacy skill, even those that are

more intensive and costly (Kraft, 2020). For example, using the National Head Start Impact Study (NHSIS) data Gibbs, Ludwig, & Miller (2011) found the effect size for a year of Head Start on literacy skills between .090 and .235 standard deviations, depending on the assessment used. These estimated impacts declined over time and were not statistically significant when measured at the end of kindergarten or first grade. Clearly the cost of a year of Head Start is also much greater than the cost of a year of providing families with a tablet and digital library. More intensive programs that focus on parents also place high demands on parents' time and effort.

The type of technology used as part of this experiment is now common in the marketplace. Various tablet technologies across a range of prices can hold hundreds of digital books with engaging graphics and functionalities. The specific tablet used in CAPER has a going market rate of slightly less than $300, offering policymakers a relatively cheaper alternative to more labor-intensive interventions aimed at increased literacy skills of young children. Programs that target tablet provision based on income status could further increase the effectiveness of such a program. As a result of the COVID-19 pandemic, many school districts have already provided tablet technologies to students as part of their school curriculum. Our results from CAPER lead us to consider that additional materials for at-home use could be added to these technologies and parents could be encouraged to engage with them.

Recall that the tablet provided to treatment parents only contained the CAPER application; internet access or other applications was not possible. An unrestricted electronic tablet might counterproductively lead to more distraction, and therefore our findings should not be generalized to imply that electronic tablets with no restrictions would lead to improved literacy skills.

6. CONCLUSION

We designed and implemented an 11-month field experiment where treatment parents were randomly assigned to receive either a tablet containing a digital library (CAPER) or, along with the tablet, one of two behavioral messaging interventions – reminders and goal setting – meant to address present bias. We then assessed how our three interventions affected reading time and literacy skills. We found that compared to just receiving the tablet, being assigned to our goal setting intervention increased reading time on the tablet by 0.32 SD. Being assigned to the reminders group, however, did not lead to any significant effects on reading time.

We also found that children in the tablet only group increased their literacy skills as measured by the CELF assessment by 0.29 standard deviations. Children in the goal setting group had an increase of 0.26 standard deviations in literacy skills. Being assigned to the reminders group had no effect on children's assessment scores compared to the control group. In fact, our reminders intervention had a negative impact on literacy skills relative to receiving the tablet alone.

These results highlight the possibility that our behavioral messaging may have had a neutral or negative effect when on children's literacy skills. The extra reading our behavioral messaging induced also took place later in the evening. This may have reduced the efficiency of reading time in terms of literacy skills.

Future work using nudges to increase parental investments in early-childhood skills should consider the potential hidden costs or crowding-out effects of such efforts. Our results are consistent with a story in which we nagged instead of nudged parents to read more, crowding out intrinsic motivation to read leading to less effective reading. We also recommend that future

research should explore how technology reduces frictions to parent-child reading. Our results show how this home-based technology may serve an important complement to school-based activities for young children. The technology used as part of CAPER is also relatively cheap and offers policymakers a unique opportunity to equip parents with the materials they need to invest in their children's literacy skills at home during a crucial time in a child's development.

UNIVERSITY OF CHICAGO, HARRIS SCHOOL OF PUBLIC POLICY

REFERENCES

Allcott, H., & Kessler, J. B. (2019). The welfare effects of nudges: A case study of energy use social comparisons. *American Economic Journal: Applied Economics*, *11*(1), 236–276.

Asensio, O. I., & Delmas, M. A. (2016). The dynamics of behavior change: Evidence from energy conservation. *Journal of Economic Behavior & Organization*, *126*, 196–212.

Barron, K., Damgaard, M. T., & Gravert, C. (2022). *When do reminders work? Memory constraints and medical adherence*. University of Copenhagen. Department of Economics. The Center for Economic ….

Bénabou, R., & Tirole, J. (2003). Intrinsic and extrinsic motivation. *The Review of Economic Studies*, *70*(3), 489–520.

Beuermann, D. W., Cristia, J., Cueto, S., Malamud, O., & Cruz-Aguayo, Y. (2015). One laptop per child at home: Short-term impacts from a randomized experiment in Peru. *American Economic Journal: Applied Economics*, *7*(2), 53–80.

Chabris, C. F., Laibson, D., Morris, C. L., Schuldt, J. P., & Taubinsky, D. (2008). Individual laboratory-measured discount rates predict field behavior. *Journal of Risk and Uncertainty*, *37*(2), 237–269.

Clark, D., Gill, D., Prowse, V., & Rush, M. (2020). Using goals to motivate college students: Theory and evidence from field experiments. *Review of Economics and Statistics*, *102*(4), 648–663.

Cortes, K. E., Fricke, H., Loeb, S., Song, D. S., & York, B. N. (2021). Too little or too much? Actionable advice in an early-childhood text messaging experiment. *Education Finance and Policy*, *16*(2), 209–232.

Damgaard, M. T., & Gravert, C. (2018). The hidden costs of nudging: Experimental evidence from reminders in fundraising. *Journal of Public Economics*, *157*, 15–26.

Escueta, M., Nickow, A. J., Oreopoulos, P., & Quan, V. (2020). Upgrading education with technology: Insights from experimental research. *Journal of Economic Literature*, *58*(4), 897–996.

Gibbs, C., Ludwig, J., & Miller, D. L. (2011). *Does Head Start do any lasting good?* National Bureau of Economic Research.

Guryan, J., Hurst, E., & Kearney, M. (2008). Parental education and parental time with children. *Journal of Economic Perspectives*, *22*(3), 23–46.

Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science*, *312*(5782), 1900–1902.

Kalil, A., Ryan, R., & Corey, M. (2012). Diverging destinies: Maternal education and the developmental gradient in time with children. *Demography*, *49*(4), 1361–1383.

Kopack Klein, A., Aikens, N., Li, A., Bernstein, S., Reid, N., Dang, M., Blesson, E., Rakibullah, S., Scott, M., & Cannon, J. (2021). Descriptive Data on Head Start Children and Families from FACES 2019: Fall 2019 Data Tables and Study Design. OPRE Report 2021-77. *Administration for Children & Families*.

Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher*, *49*(4), 241–253.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *The Quarterly Journal of Economics*, *112*(2), 443–478.

Landers, R. N., Bauer, K. N., & Callan, R. C. (2017). Gamification of task performance with leaderboards: A goal setting experiment. *Computers in Human Behavior*, *71*, 508–515.

Malamud, O., & Pop-Eleches, C. (2011). Home computer use and the development of human capital. *The Quarterly Journal of Economics*, *126*(2), 987–1027.

Mayer, S. E., Kalil, A., Oreopoulos, P., & Gallegos, S. (2019). Using Behavioral Insights to Increase Parental Engagement The Parents and Children Together Intervention. *Journal of Human Resources*, *54*(4), 900–925.

Meier, S., & Sprenger, C. (2010). Present-biased preferences and credit card borrowing. *American Economic Journal: Applied Economics*, *2*(1), 193–210.

Price, J. (2010). The effect of parental time investments: Evidence from natural within-family variation. *Unpublished Manuscript, Brigham Young University*.

Rodgers, A., Corbett, T., Bramley, D., Riddell, T., Wills, M., Lin, R.-B., & Jones, M. (2005). Do u smoke after txt? Results of a randomised trial of smoking cessation using mobile phone text messaging. *Tobacco Control*, *14*(4), 255–261.

Tondello, G. F., Premsukh, H., & Nacke, L. (2018). A theory of gamification principles through goal-setting theory. *Proceedings of the 51st Hawaii International Conference on System Sciences*.

Villena-Roldan, B., & Ríos-Aguilar, C. (2012). Causal effects of maternal time-investment on children's cognitive outcomes. *Center for Applied Economics, Working Paper*, *285*.

Waldfogel, J., & Washbrook, E. (2011). Early years policy. *Child Development Research*, *2011*.

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

TABLES

Table 1: Descriptive Statistics and Balance

Child variables (N=300)	**Control** (N=43)	**Tablet** (N=97)	**Reminder** (N=78)	**Goal Setting** (N=82)	**F Statistic p-value**
Age (months)	46.1 (5.7)	45.2 (6.6)	44.1 (5.4)	45.7 (6.4)	0.25
Male	60.5%	40.2%	48.7%	47.8%	0.17
Baseline CELF	20.1 (15.0)	21.5 (15.2)	19.8 (13.2)	21.2 (14.3)	0.86
Only speak English at home	61.9%	67.4%	63.5%	74.1%	0.44
Only speak Spanish at home	11.9%	27.1%	23.0%	17.3%	0.17
Parent variables (N=379)	**Control** (N=54)	**Tablet** (N=115)	**Reminder** (N=105)	**Goal Setting** (N=105)	**F Statistic p-value**
White	28.6%	28.3%	24.1%	27.0%	0.92
Black	28.6%	26.1%	31.0%	28.1%	0.91
Hispanic	26.2%	29.3%	32.2%	28.1%	0.90
Less than HS	21.4%	12.0%	16.5%	16.7%	0.56
HS diploma/GED	23.8%	30.4%	30.6%	26.7%	0.81
BA & higher	16.7%	17.4%	14.1%	15.6%	0.94
Attrition Rate	42.0%	36.1%	45.8%	46.4%	0.19

Note: The data are from baseline CELF assessment and parent survey. The child characteristic data are limited to the sample of children who took both baseline and follow-up assessments. The parent characteristic data are limited to the sample of parents who responded to both baseline and follow-up parent surveys. The number of respondents to each question varies because not all parents answered all questions. SD is included in the parentheses. The F-Statistic p-value column represents the p-value on a joint hypothesis test with a null hypothesis of equal means across treatment conditions. *** p<.01, ** p<.05, * p<.10.

Table 2: Treatment Effect of Reminder and Goal Setting Texts on Reading Time

	(1)	(2)	(3)
	Reading time	Reading time	Reading time
Reminders	19.79	34.25	34.89
	(50.94)	(50.87)	(51.04)
Goal Setting	133.90**	145.84**	150.08**
	(61.34)	(58.30)	(58.33)
Constant	270.83***	39.47	-244.13
	(39.21)	(44.90)	(173.75)
Control Mean	270.83	270.83	270.83
Control SD	418.67	418.67	418.67
N	351	351	351
School FE	N	Y	Y
Child Age Control	N	N	Y

Notes: The omitted group is the tablet-only group. Robust standard errors are in parentheses. *p<0.1; **p<0.05; ***p<0.01. Reading Time is measured as the total number of minutes read on the CAPER app during the entire 11-month (45 week) period.

Table 3: Treatment Effect of Digital Library on Literacy Skills

	(1)	(2)	(3)	(4)
	CELF	CELF	CELF Percentile	CELF Percentile
Tablet Only	4.98***	5.05***	8.92**	9.73***
	(1.87)	(1.83)	(3.80)	(3.72)
Reminders	2.10	2.18	2.72	3.88
	(1.91)	(1.88)	(3.76)	(3.73)
Goal Setting	4.53**	3.79*	7.58*	6.70*
	(1.99)	(1.95)	(3.89)	(3.81)
Constant	23.58***	24.75***	21.66***	4.20
	(1.85)	(6.42)	(3.20)	(12.86)
Control Mean	39.53	39.53	40.58	40.58
Control SD	17.43	17.43	32.02	32.02
N	300	300	300	300
Baseline CELF	Y	Y	Y	Y
School FE & Age	N	Y	N	Y

Notes: The omitted group is the control group. Robust standard errors are in parentheses. *p<0.1; **p<0.05; ***p<0.01.

Table 4: Treatment Effect of Reminder and Goal Setting Texts on Literacy Skills

	(1)	(2)	(3)	(4)
	CELF	CELF	CELF Percentile	CELF Percentile
Reminders	-2.93*	-2.96*	-6.27*	-6.15*
	(1.61)	(1.59)	(3.38)	(3.35)
Goal Setting	-0.46	-1.31	-1.31	-3.21
	(1.69	(1.66)	(3.48)	(3.40)
Constant	29.24***	30.76***	31.23***	16.72
	(1.72)	(7.06)	(2.98)	(14.08)
Control Mean	45.60	45.60	49.92	49.92
Control SD	15.79	15.79	31.91	31.91
N	257	257	257	257
Baseline CELF	Y	Y	Y	Y
School FE & Age	N	Y	N	Y

Notes: The omitted group is the tablet only group. Robust standard errors are in parentheses.
*p<0.1; **p<0.05; ***p<0.01.

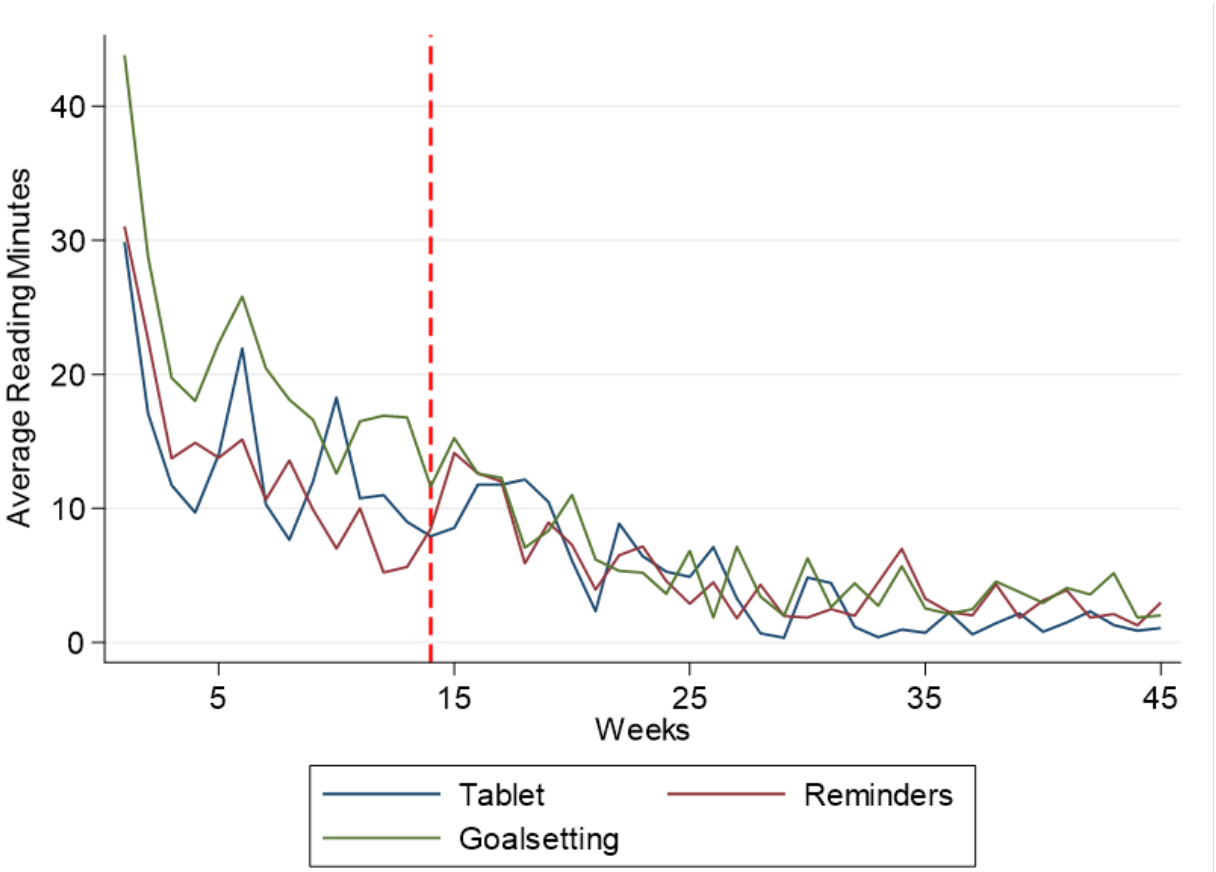
Table 5: Treatment Effect on Late-Night Reading

	(1)	(2)	(3)
	Reading minutes between 20pm to 6am	Reading minutes between 21pm to 6am	Reading minutes between 22pm to 6am
Reminders	23.94	16.79	11.93
	(18.27)	(14.85)	(11.57)
Goal Setting	87.58***	63.80***	33.31**
	(25.24)	(20.42)	(14.13)
Constant	82.24***	55.40***	33.85***
	(11.51)	(8.66)	(6.37)
Control mean	82.24	55.40	33.85
Control SD	122.91	92.42	68.03
N	351	351	351

Notes: The omitted group is the tablet-only group. Regressions include age and school fixed effects. Robust standard errors are in parentheses. *p<0.1; **p<0.05; ***p<0.01. Reading Time is measured as the total number of minutes read on the CAPER app during the entire 11-month (45 week) period.

FIGURES

Figure 1: Average Weekly Reading Time on the CAPER App over 45 Weeks



Notes: The data are from 351 retrieved tablets with reading time records. Among those, 42 children (11.97% of tablets with reading time records) never used tablets. The red dashed line at week 14 indicates the school closure because of COVID-19.

Figure 2: Literacy Skills Production with a Nag Factor

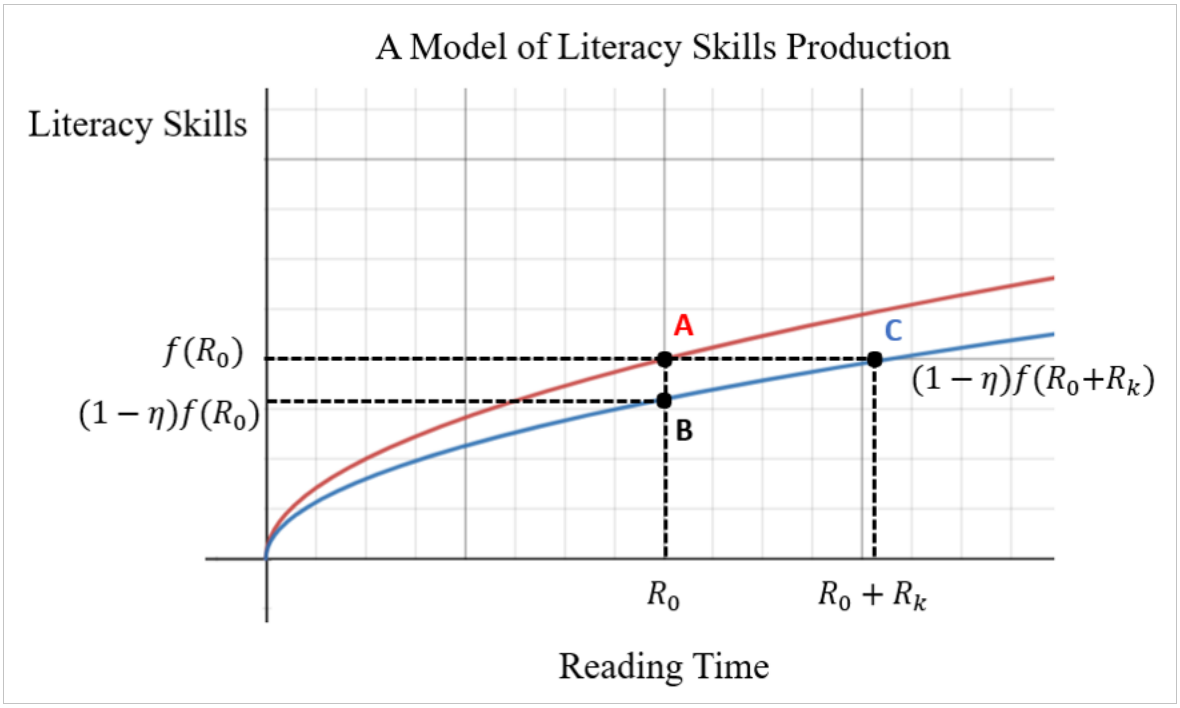
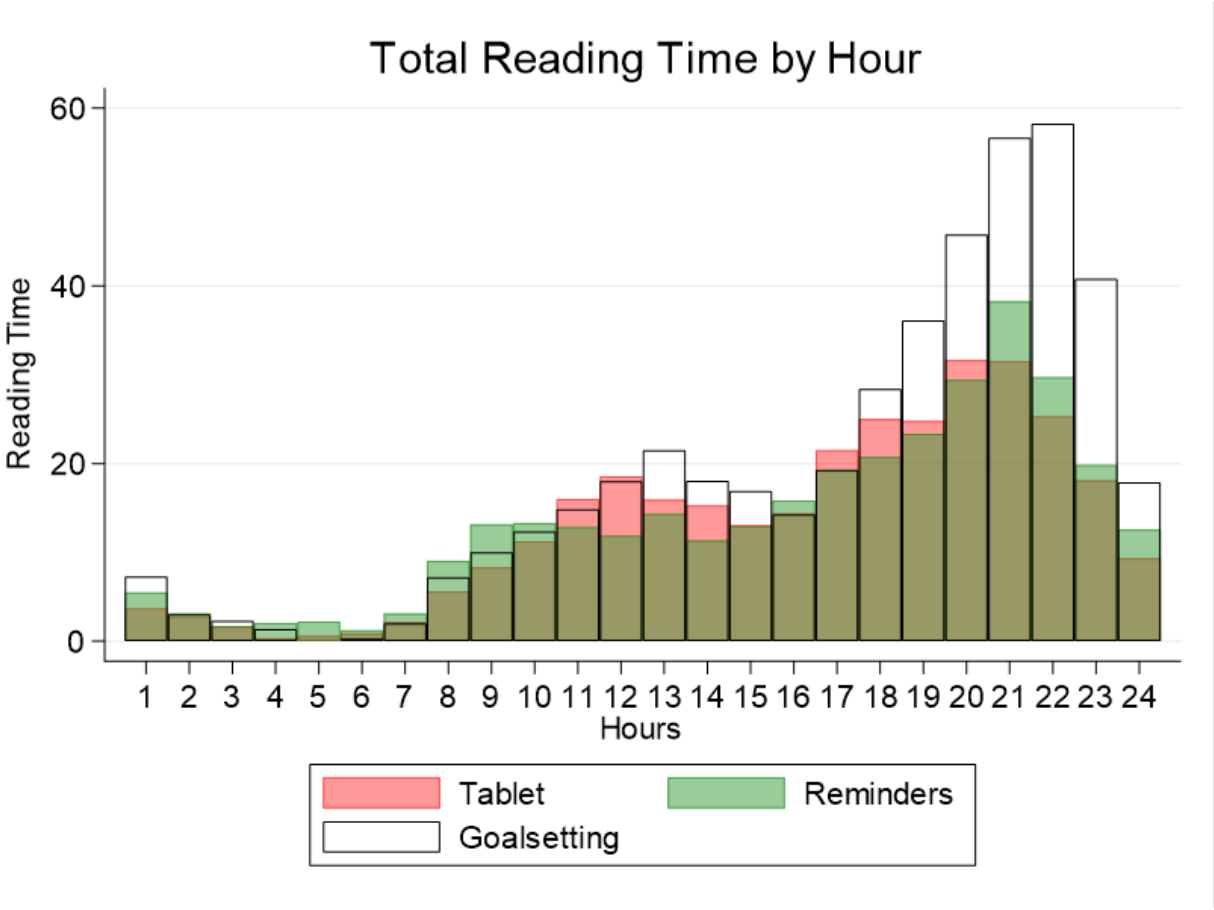


Figure 3: Total Reading minutes on the CAPER App by Hour over 45 Weeks



Notes: The data are from 351 retrieved tablets with reading time records.

APPENDIX 1: SAMPLE TEXT MESSAGES

Table A1.1 Text Message for Parents in the Control Group and Tablet Only Group

Intro Text	Survey Text
Welcome to Children and Parents Engaged in Reading (CAPER)!	Hi [first name]. We will be at [school name] on [date] for the survey! It takes 15 minutes, and you will receive $10.

Table A1.2 Text Messages for Parents in the Reminders Group

Introductory texts		
Intro Text 1	Intro Text 2	Intro Text 3
Welcome to Children and Parents Engaged in Reading (CAPER). The CAPER app is made for you and your child to read together.	The tablet is a tool for YOU to read to your child. Every week you will get text reminders to use CAPER with your child.	The CAPER app has many books for you and your child to choose from and read together! Remember to use it often.

Treatment texts			
Monday	Wednesday	Friday	Saturday
Don't forget to start the new week by finding time to read CAPER books to your child.	Does [child_name] have a favorite CAPER book? Read it to [him/her] tonight!	Remember that reading with your child is important for [his/her] future. Read a CAPER book tonight!	The weekend is a great chance to spend time reading CAPER books to [child_name]!

Table A1.3 Text Messages for Parents in the Goal Setting group

Introductory texts		
Intro Text 1	Intro Text 2	Intro Text 3
Welcome to Children and Parents Engaged in Reading (CAPER). The CAPER app is made for you and your child to read together.	The tablet is a tool for YOU to read to your child. Every two weeks CAPER will suggest a goal for how much time to spend reading to [child_name]'s. Try to reach each goal!	Each week you will get a text telling you how much time you read to your child using CAPER last week. Only the time spent reading to [child_name] using CAPER counts towards your goal. Use CAPER to reach your goal – and more!

Treatment texts			
Monday	Wednesday	Friday	Saturday
Help your child be a reader by staying on track with your reading goal!	Are you on track to meet your goal? Reading different CAPER books can help your child learn and help you reach your goal.	In the past two weeks you read an average of [feedback] minutes per week to [child_name] using CAPER. This week try reading [goal] minutes.	How are you doing on your goal? The weekend is a great chance to spend time reading CAPER books to [child_name]!

APPENDIX 2: SUPPLEMENTAL TABLES

Table A2.1: CAPER Tablets Distributed and Collected by Treatment Group

Treatment	Number of tablets distributed	Number of tablets returned	% of tablets returned	Number of tablets not returned
Control	0	0	0	0
Tablet only	158	114	72%	44
Reminders	163	125	77%	37
Goal Setting	158	112	71%	47
Total	479	351	73%	128