Using a Factorial Design to Maximize the Effectiveness of a Parental Text Messaging Intervention

Catherine Armstrong Asher
Ethan Scherer
James S. Kim
Harvard University

September 2020

Abstract:

Parental text messaging interventions are growing in popularity to encourage at-home reading, school-attendance, and other educational behaviors. These interventions, which often combine multiple components, frequently demonstrate varying amounts of effectiveness, and researchers often cannot determine how individual components work alone or in combination with one another. Using a 2x2x3 factorial experiment, we investigate the effects of individual and interacted components from three behavioral levers to support summer reading: providing updated, personalized information; emphasizing different reading views; and goal setting. We find that the personalized information condition scored 0.03 SD higher on fall reading assessments. Test score effects were enhanced by messages that emphasized reading being useful for both entertainment and building skills compared to skill building alone or entertainment alone.

Keywords: factorial design, literacy, parental text messaging, experimental design

Acknowledgments: The authors are thankful for feedback and advice from Nicole Pashley, Carly Robinson, Luke Miratrix, Monnica Chan, Eddie Kim, Thomas Kelley-Kemple, participants in the READS Lab, CARES Lab, and conference participants at AEFP. This work is supported by the Chan Zuckerberg Initiative and the Institute of Education Sciences, U.S. Department of Education through grant R305B150010. The opinions expressed are those of the authors and do not represent the views of the Institute or the U.S. Department of Education.

Background

Parental investment is a critical component of a child's development (Heckman, 2006; Greenman et al., 2011; Del Boca et al., 2014; Francesconi & Heckman 2016; Del Bono et al., 2017). Yet, there is wide variation in a parent's capacity to support learning at home (Greenman et al., 2011; Francesconi & Heckman 2016). One issue could be that parents are awash in an ocean of advice on websites, blogs, books, and from other parents and family members that form their beliefs and actions. Narrowing these options down and identifying trusted sources of information requires significant time investment itself. But with twenty-three percent of children under the age of 18 living with one parent (Kramer, 2019) and both parents working in sixty percent of married households (Bureau of Labor Statistics, 2017), time is a scare resource. These conditions make it harder for families to focus their attention and effort on consistently making informed home education choices, and dedicating time to regularly monitor their child's educational progress. Finding better ways to engage with parents has gained increasing importance in the context of the COVID-19 pandemic. Typical reasons why families might not be able to focus on their child's education, like the inability to commit sufficient attention when juggling multiple tasks or prioritizing the immediate over longer-term goals, has only been exacerbated by the stay-at-home orders and uncertainty around work and school security. Unemployment has skyrocketed and families have been forced to juggle work and children's education in unprecedented ways. Thus, better understanding optimal ways to engage families about their children's education will need to be identified to help now and in the future.

Both policymakers and researchers are increasingly using insights from behavioral sciences to enhance parental engagement. One subset of these interventions provides parents with younger children timely and actionable text messages, many of which contain reminders of

home literacy activities. Table 1 presents the collection of these studies. These interventions target several potential behavioral levers to increase parents' engagement. First, receiving a text message could refocus parents' attention to educational actions (Hurwitz et al., 2015; Kraft and Monti-Nussbaum, 2017; Mayer et al. 2019; Smythe-Leistico & Page, 2018 ; Cortes et al., 2019 ; Doss et al., 2019; Kim et al., 2019; York et al., 2018). Second, text messages can help parents navigate the challenge of short-term, immediate-cost behaviors (like reading to your child every day) whose benefits do not manifest until the future. Mayer et al. (2019) find setting goals and monitoring parents progress towards these goals mitigated these self-control concerns. Third, a subset of the studies attempts to reduce the complexity, or cognitive load, of parenting by breaking down parental education activities into several discrete components (Hurwitz et al., 2015; Kraft and Monti-Nussbaum, 2017; Cabell et al., 2019; Doss et al., 2019; Kim et al., 2019; York et al., 2018). They provide a combination of literacy facts or resources, a specific activity or practice, and extension activities that parents can implement without additional preparation. Finally, informational messages attempt to facilitate parent monitoring of their children's academic performance. For example, it is difficult to track and monitor a child's cumulative absences, and parents often underestimate how much school their child has missed (Smythe-Leistico & Page, 2018). One method to address these concerns is providing parents with up-to-date information on their child's performance to correct these biased beliefs.[1] Taken together, these studies show that parental text messages are of increasing importance as a tool to support educational outcomes for students.

While texting interventions in Table 1 demonstrated reduction in specific behavioral barriers, less is known about how messaging interventions inform parent's domain-specific beliefs. For example, all the literacy texting interventions in Table 1 focus the content of their

message on building literacy skills, a framework aligned with the instrumental view of reading that emphasizes the value of skill-building for future success (Baker, Scher, & Mackler, 1997). However, literacy scholars document that while some parents articulate this instrumental view, others emphasize an entertainment view, emphasizing reading for pleasure and enjoyment (Baker, Scher, Mackler, 1997). A series of small-sample studies finds that these parental beliefs differentially predict features of their children's reading experiences. Parents tend to promote types of reading activities for their children that align with their beliefs (Lynch, Anderson, Anderson, & Shapiro, 2006; Sonnenschein et. al, 1996; Sonnenschein et. al, 1997). Additionally, parents who endorse the entertainment view have children who score higher on measures of reading enjoyment and motivation (Baker, Cher, & Mackler, 1997; Baker & Scher, 2002). Correlational research indicates that students with exposure to the entertainment view tend to be better readers than similar students from homes who emphasize the instrumental value (Baker, Mackler, Sonnenschein, & Serpell, 2001; Sonnenschein et. al, 1997; Sonnenschein, Baker, Serpell, & Schmidt, 2000). Less is known, however, about whether these differences in views and activities are causally related to student outcomes, and how underlying student motivation maybe be influenced by the introduction of different views. Causally investigating how literacy parental texting intervention effect could vary based upon promoting particular beliefs remains a critically unexplored area of literacy parental texting research.

Furthermore, while the existing research indicates the focus of the message content is important, interventions often combine message types or other components to maximize the intervention's effectiveness. In Doss et al. (2019), for example, the messages combine facts and information about reading with specific activities for parents to complete with their children in both arms of the intervention. Similarly, Mayer et al. (2019) combine their text messaging with

an educational app and regular check-in meetings with the program provider. Thus, while it is clear that the characteristics of the intervention matter, it is hard to assess whether individual components could work as well as a cocktail of messages or if the synergies between the components drive the effects. Unfortunately, typical randomized controlled trials are ill-suited to efficiently unpack the effectiveness of individual intervention components and their combination to answer the question of what works, for whom, and why.

Additionally, this body of literature offers theoretical and empirical motivation to study potential heterogeneity in the effects of text messaging interventions across subpopulations of families. Several studies demonstrate larger effects for students with lower baseline performance (Cortes et al., 2018; Cabell et al., 2019; Cortes et al., 2019; Doss et al., 2019; York et al., 2019), arguing that tips for parents are most relevant for lower performing students. Moreover, many of these interventions are studied in populations of predominantly low-socioeconomic status families (Cabell, 2020; Cortes et al. 2018, Cortes et al., 2019; Hurwitz et al., 2015). Theory suggests that these populations may benefit more from these types of behavioral interventions due to the higher cognitive loads associated with poverty (Mani, Mullainathan, Shafir & Zhao, 2013), but as yet there is no evidence testing this hypothesis. Taken with the research on home literacy environments, we anticipate that the effects for parent texting interventions may vary based on multiple dimensions of student characteristics, including baseline reading motivation, baseline reading performance, and socioeconomic status.

The prior literature raises important unanswered questions for parental text messaging interventions. Our research seeks to understand what types of parental text messages are most effective, by investigating two primary research questions:

1. What components and combination of components of parental text messages affect parent behavior and child reading outcomes?
2. What student characteristics moderate these effects?

We address these questions by implementing a parental text message intervention to understand the effects of individual message components separately and in combination on family reading behaviors, student reading comprehension, and child reading behaviors. Specifically, we examine the effectiveness of three distinct components: personalizing messages with up-to-date information; the message's framing of reading as building skills, for entertainment or both; and setting reading goals at the beginning of the summer. We test these components of the text messaging campaign using a 2x2x3 factorial experiment with approximately 5000 elementary school students in a single school district in the southeast US. A factorial design is particularly well-suited to study an intervention with multiple distinct components that could be individually included or excluded, because it allows for the estimation of each component's individual effects as well as interaction effects of multiple components (Collins, Dziak, Kugler, & Trail, 2014; Somers, Collins, & Maier, 2014).

A unique feature of our intervention is our ability to both directly and indirectly measure parental reading behaviors in response to our text messages. All participants in our study had access to a summer reading program and accompanying educational reading app. Over sixty percent of our text messages encourage parents to have their child login and use the reading app. Given that we provide login passcodes only to the parents, we can observe parental engagement with summer reading based on app usage data. Similar to prior studies, we also directly survey a subsample of parents.

We find that in comparison to generic messages, messages containing personalized information increase parental reading supports of reading and student reading performance in the fall. Students of parents who received text messages were three percentage points more likely to use the app. These students also performed 0.03 standard deviations better on the formative reading scores as well as .02 standard deviations better on high stakes beginning of grade exams, though the latter was not statistically significant. Levering the interaction effects within our factorial design, we also find that the view of reading emphasized in the messages could enhance or detract from these information effects, with a combination of reading values magnifying the positive effects of personalized information when compared to personalized messages that focused on instrumental values alone.

We also explore potential mechanisms for the increase in reading comprehension test scores using our rich set of student-level outcomes. An important hypothesis from the view-of-reading literature posits that emphasizing entertainment values will increase a child's motivation, and ultimately improve reading comprehension. We directly test changes in reading motivation using student self-reports from both a paper survey, and among the students who logged into the app, responses to reading motivation questions directly embedded into the app, but find no evidence to support this claim. Additionally, we assess how the components of our messages differentially affect typically disadvantaged groups who might be most likely to benefit from the text messages: students from higher poverty neighborhoods, those that perform lower at baseline, and those with lower initial baseline motivation. While few of these results are statistically significant, broadly we find that our main results on message containing personalized information benefits both advantages and disadvantages groups. In addition, messages that

contained both instrumental and entertainment framing tended to benefit the disadvantaged groups.

We structure the remainder of the paper in the following way: first, we begin by providing a brief description of our intervention, data, measures, and methods used in our analysis; next, we present our main results and an investigation of potential mechanisms; we conclude with a discussion of our findings and their implications for texting interventions for parents.

The Intervention

Context and Study Eligibility

Students were recruited to participate in this study through their participation in a large, multi-school randomized control trial (RCT). For the larger RCT, all first and second grade students in the participating schools were invited to participate in both the curriculum and texting studies through an active consent process. For the curriculum study, either first or second grade students at each school were assigned to receive the MORE (Model of Reading Engagement) intervention curriculum, with the other grade serving as a control group who received the business-as-usual curriculum. This intervention consisted of a series of science- and social studies-themed lessons in the spring semester and the students' choices of 10 hardcopy books related to the same science and social studies topics presented in the classroom lessons.

To be eligible for the text messaging study, the research team had to identify a valid cell phone number for a parent or guardian of each student. Phone numbers were either provided directly on the consent form or by the school district administrative records. Once a cell phone number was validated, the families were enrolled in the text messaging study. Participants in the

texting study were randomized into different texting interventions, including a small group assigned to a pure control condition that received no text messages.

At the end of the school year, all students in both the curriculum study and texting study received access to an educational reading app called MORE@Home, which contained six digital books as well as a series of reading-related activities matched to each book and broadly leveled according to the child's end-of-year reading ability. For those students in the MORE curriculum arm of the curriculum study, their MORE@Home accounts also provided access to leveled reading activities for each of the 10 books they had selected. On the whole, use of the app was relatively low across the sample. Among our pure control group of families, only 16% of student accounts were ever activated by their parents.

Text messaging intervention

The purpose of the text messaging intervention was to increase parental engagement with their children in summer reading activities and improve student learning. In line with recent research on the effect frequency and timing of text message interventions (Cortes et al., 2018; Cortes et al., 2019), families received text messages twice weekly over 9 weeks of summer vacation, with one message occurring earlier in the week and one message closer to the weekend. Messages were sent in either English or Spanish, based on the student's home language in district administrative records.[2] Messages were all sent through the Twilio messaging platform which was accessed through our sample database. This approach had several advantages: 1) message times were scheduled in advance; 2) messages were linked to user profiles containing other research information from the curriculum study and educational app database; 3) families could easily opt-out of messages, preserving consent; and 4) parental responses to text messages were logged on the user profiles. Some text messages were designed specifically to promote

usage of the educational reading app, while others encouraged a wider variety of reading activities. Each message contained a single topic, which generally covered one of the three larger themes: 1) reminders to engage in summer reading activities; 2) providing information about summer reading resources (including the educational app); and 3) monitoring progress throughout the summer. A comprehensive list of message themes can be found in Table 2.

Differentiating messages

To explore how text messaging features differentially influence parental and student engagement with a summer reading intervention, we differentiated the specific wording of each text message topic according to three separate factors that have shown promise in prior research: updated personalized information to correct parent informational misbeliefs (Smythe-Leistico & Page, 2018), goal-setting to reinforce immediate action (Mayer et al., 2019), and framing different views of reading to promote specific parental beliefs (Baker, Mackler, Sonnenschein, & Serpell, 2001; Sonnenschein et. al, 1997; Sonnenschein, Baker, Serpell, & Schmidt, 2000).

In the personalized information factor, some families received text messages that include student-specific information within the message. Some examples of the information that could be included were: the specific books the student had access to in the app, whether or not a student had logged into the educational app yet, or which books' activities the students had accessed on the app. Because the text messaging was integrated with the app's backend database, each student's information was updated continuously, ensuring that the messages reflected the students' most recent status. Families not in the personalization condition received more generic messages, but they still referred to individual students by name.

For the goal-setting factor, some families were invited to set a summer reading goal at the beginning of the intervention, with later messages periodically checking-in on their progress

towards that goal. We designed the goal setting to be a light-touch, low-cost, scalable version of other effective goal-setting studies (e.g., Mayer et al., 2019; Oreopoulos et al., 2020), with parents being asked not only to identify a goal but to make a plan for reaching it in the face of obstacles. However, without an in-person goal setting session or subsequent individualized follow-up, most families (more than 95% of the goal-setting condition) failed to complete the goal-setting exercise. For the few families who did set a reading goal, check-in messages would explicitly refer to that goal. For the families who did not set a goal, messages simply referred to "your summer reading goal."

Finally, the view-of-reading value's condition created three distinct groups who received differently framed messages over the course of the campaign. The instrumental-only-view condition emphasized reading as a process by which students develop specific skills important for future success. The entertainment-only-view condition emphasized reading as an enjoyable and fun activity. The combination-view received a balanced combination of entertainment- and instrumental-framed messages over the course of the summer. Importantly, however, this combination-view received the same total number of messages as the entertainment-only and instrumental-only conditions.

An example of how a text message might differ can be found in Table 3, but not all messages included components based on the levels of all 3 factors. A single message could meaningfully differentiate between conditions for just goal-setting factor, just the personalization factor, just the view-of-reading factor, or across any combination of those factors; in our text messages, 40% were relevant with regards to families' goal-setting condition, 47% differed with regards to personalization, and 85% were framed in reference to a specific view of reading. To ensure that the intended conditions were salient to families, we recruited colleagues to review

example text message versions and provide feedback on whether the messages clearly contained information aligned to specific conditions. In cases where colleagues' perspectives differed from our intention, we revised the message versions to increase or decrease the salience of a specific condition.

Methods

Sample

This study includes 5,172 rising second and third grade students, from 4,993 families, who attend thirty elementary schools in one large school district in the southeast. Demographic characteristics are presented in the first column of Table 4. Close to forty percent of the students are African American and an additional thirty percent are Hispanic. Approximately twenty percent are White, and approximately ten percent are Asian. Almost a quarter of students were receiving English-learner services. Our sample contains socioeconomic diversity but contains a larger proportion of students in low-SES neighborhoods relative to the district as a whole.

Research Design

We use a factorial experiment to compare the differential effectiveness of text messaging components. Traditional randomized controlled trials are only able to compare two treatment arms at a time, so investigating three potential multiple mechanisms would require multiple experiments, or a multi-arm RCT with a prohibitively large sample size. A factorial design, however, is particularly well-suited to study an intervention with multiple distinct components. In a factorial design, each intervention component is treated as its own factor, with different levels representing the treatment assignment. Each unit is randomized to a level for each factor independently. This design has the benefit of allowing the researcher to test the main effect of each intervention components on its own, as well as interactions of intervention components

(Collins, Dziak, Kugler, & Trail, 2014; Somers, Collins, & Maier, 2014). It is thus an appropriate design to address how the multiple levers targeted in texting messaging interventions contribute to an intervention. We use a full-factorial design, in which every factor is fully interacted with the other factors. The goal-setting and personalized information factors each have two levels (on, off), and the view-of-reading factor has three levels (instrumental view only, entertainment view only, both views presented), resulting in 12 different treatment combinations. Additionally, we assigned a small portion of the sample to a pure control condition, not receiving any text messages, as shown in Table 5. Separating out this pure control provides a business-as-usual condition to use as a benchmark for the magnitude of the factorial differences, but our research questions focus exclusively on the relative effectiveness of the different text message components as opposed to the effects of text messaging compared to no messaging.

To account for the presence of siblings in the sample, which could result in spillover and confusion for parents receiving two types of messages, random assignment to conditions occurred at the family-level. To improve precision and reduce the minimum detectable effect size, the sample was blocked at the school-by-grade level, the unit of treatment from the larger RCT, because average student reading levels and implementation fidelity of the larger intervention vary across schools. Within these blocks, the 4,993 families (5,172 children) with valid cell phone numbers were assigned to one of the 13 conditions.

A series of balance checks are presented in columns (2) through (5) of Table 4, comparing our sample across each factor of the intervention. Overall, we find few differences between experimental groups on baseline demographic characteristics or academic performance, and none that are statistically significant after applying a multiple hypothesis correction to

account for the number of student characteristics compared (Benjamini & Hochberg, 1995). The similarities between groups reflects a successful randomization process.

Data Sources

Baseline data from district administrative records of the pre-intervention year (2018-2019) is available for our entire study sample. These measures include enrollment information, student demographics, and reading and math test scores. The school district where this intervention took place does not collect student-level measures of socio-economic status; however, student neighborhoods, as determined by their census block-group, are categorized as being low-, middle-, or high-SES communities. Student-level outcome data from Fall 2019 is also provided by the district. We use the Measure of Academic Progress RIT score in literacy (MAP, Northwest Evaluation Association, 2011) a primary outcome, which measures foundational literacy skills. A second academic outcome, available for the rising third-grade cohort, is the Beginning of Grade (BOG) assessment, a statewide measure designed to measure student's progress towards proficiency on the End of Grade (EOG) accountability assessment at the end of third grade. The BOG is also used for school accountability growth and for the identification of exceptional teachers. Both assessment scores are standardized within grade to provide outcomes in standard-deviation units.

We assess the effect of the intervention on parents' behaviors and beliefs with two sources. First, we track family use of the MORE@Home app. Because 60% of the text messages parents received were related to the educational app, usage statistics reflect whether parents changed their own and their children's behavior in response to messages. Specifically, we are interested in whether or not parents logged their students into the app, the total number of books they accessed on the app, and the total number of minutes they spent engaged with the app. For

families who never logged into the app, total books and total minutes were recorded as zeros. We also collected self-reported outcomes from a subsample of parents who were randomly invited to complete a parent survey. The survey included questions about their summer reading activities beyond the educational app and their perceptions of the text messages they received.

To explore student-level mechanisms and provide a more information about students' summer reading experiences, we collected two waves of student surveys as well as qualitative measures of the students' use of the educational app. In the spring prior to the summer texting intervention, we measured student motivation using the Me and My Reading Profile (MMRP, Marinak, Malloy, Gambrell, & Mazzoni, 2015). In the fall, we administered a student survey that included both the MMRP and additional questions about students reading behaviors over the summer outside of the educational reading app. The fall survey was administered in the thirty study schools and thus does not cover either the students who moved out of the district or those who moved to a non-study school within the district. Students whose parents logged them into the app also provided additional data on whether they enjoyed the app, felt like a good reader, and found the app activities challenging. We also measured the percent of activities that they answered correctly. Because our sample is slightly different for each set of outcomes, we test for differential attrition rates based on the different factors of our treatment. Though not presented here, we find no significant differences in retention rates by condition for any of our outcome samples except for the app users sample, which was over-represented by students in the personalization condition

Empirical strategy

We analyze the twelve conditions in our factorial experiment using standard regression techniques that allow us to concurrently model the effects of each individual factor as well as their two-way combinations:

$$Y_{ij} = \beta_1 Pers_{ij} + \beta_2 EntView_{ij} + \beta_3 BothView_{ij} + \beta_4 Goals_{ij} + \beta_5 PersxEntView_{ij} + \beta_6 PersxBothView_{ij} + \beta_7 PersxGoals_{ij} + \beta_8 GoalsxEntView_{ij} + \beta_9 GoalsxBothView_{ij} + \beta_{10} PersxEntViewxGoals_{ij} + \beta_{11} PersxBothViewxGoals_{ij} + \Gamma X_{ij} + \phi + \varepsilon_{ij} \quad (1)$$

where Y_{ij} represents the outcome for individual i in randomization block j. The model also includes a vector of covariates X_{ij}, including student demographics and pre-test math and reading scores, as well as a set of fixed effects, ϕ, representing the randomization blocks. Standard errors are clustered at the family-level to account for the unit of randomization and the correlation of residuals among siblings.

The main effects (captured in β_1 through β_4) provide information about the average effect of each text messaging factor, and the two-way interactions (captured in β_5 through β_9) tell us whether the effects of one factor depend on the levels of the other factors. The three-level view-of-reading factor has been separated into two variables, EntView (for entertainment only) and BothView (for a combination of values). Families in the instrumental view-of-reading condition serve as the reference category for both of those variables. The three-way interaction terms (β_{10} & β_{11}) are included purely to allow us to estimate the main effects and two-way interactions concurrently; given their difficulty of interpretation, they are not parameters of interest for this paper.

To facilitate interpretation of effects, treatment variables are coded using effect coding with variables taking on values of -1 or 1 (Kugler, Trail, Dziak, & Collins, 2012). This

parameterization of the treatments highlights the benefits from using effect coding. Because our sample is evenly divided across conditions, this parameterization allows both the main effects (the difference between levels of each factor or marginal effects) and the interaction effects (the additional effect of receiving a particular combination of factors) to estimated concurrently (Hardy, 1993). However, because level differences require the treatment indicators moving from -1 to 1 instead of from 0 to 1 as in the standard dummy coding, we must multiply all coefficients (and their standard errors) by 2. Thus, we can interpret the main effect β parameters as follows:

- The main effect of receiving personalized information (vs. not, as if in a traditional RCT) is $2\beta_1$.
- The main effect of framing messages with the entertainment views of reading (compared to only instrumental-only view, as if in a traditional RCT) is $2\beta_2$.
- The main effect of framing both views of reading (compared to only instrumental-only view, as if in a traditional RCT) is $2\beta_3$.
- The main effect of setting goals (vs. not, as if in a traditional RCT) is $2\beta_4$.

Interaction terms are interpreted as normal once they have been scaled up: the additional effect of one factor in the presence of another factor. In all tables in this paper, we have already adjusted the point estimates and standard errors for ease of interpretation. Because we present multiple outcomes in each domain, we also test our confirmatory results to the sensitivity of false discoveries, using the Benjamini-Hochberg procedure with a false discovery rate (FDR) set to 0.05 (Benjamini & Hochberg, 1995) by outcome domain.

Results

Effects on parental behaviors and beliefs

We first consider how the different types of text messages affect parental behaviors and beliefs (see Table 6). Panel A presents the main effects of each component, and Panel B presents the interaction effects of how combining two factors can enhance or detract from their main effects, but for each column, both panels come from the same fitted model. Our primary measures of behavioral change in parents are captured in their use of the educational reading app with their children. For both the probability of ever logging into the app, the total minutes spent on the app and the number of books completed, the personalized messages were significantly more effective than non-personalized messages. Families receiving personalized messages were three percentage points more likely to use the app (ES=0.08), spent about an extra 1.6 minutes using the app (ES = 0.11) and completed an additional 0.7 books worth of activities (ES=0.12). These effects remain statistically significant after multiple-hypothesis corrections (Benjamini-Hochberg, 1995).

To facilitate this interpretation, Figure 1 presents model-based predicted outcomes for our sample, grouped first by personalization, and aggregated across the goal setting condition. From these, we can more easily identify common trends that are not immediately apparent from the model output. This figure highlights that personalization improved app use for all three view-of-reading conditions, but that compared to our control condition, both generic and personalized messages were effective at increasing parental engagement with reading activities. These results are consistent with prior work on absenteeism using postcards that providing personalized information to parents debiases their belief when information is hard to obtain (Robinson et al., 2018; Rogers & Feller, 2018). We also consider self-reported behaviors and beliefs from the subsample of parents who responded to our parent survey. Because the survey sample is relatively small, we only examine the main effects of each factor. Among this group, parents

receiving personalized activities also report engaging in more frequent reading activities with their students (ES = 0.046) but this difference is not significant. These results demonstrate that personalized text messages caused parents to login and use the reading application more.

For the other text message types, while the point estimates are generally not significant, it is worth noting that unlike the personalization condition, the reading behaviors in the app and the self-reported behaviors trend in the opposite direction. For example, there are no significant main effects on directly-observed app usage or self-reported frequency of reading activities from changing the view-of-reading emphasized in text messages or providing parents with the opportunity to set a summer reading goal. However, the magnitudes of the reported differences are quite large for self-reported reading activities and are often the opposite sign between the two sources of data. The goal-setting and view-of-reading components do have an effect on parental beliefs about the intervention. Anchoring messages around goal setting increased the likelihood that parents found the text messages useful by 32 percentage points, as did receiving messages framed around both views of reading compared to only the instrumental view (by 36 percentage points). These features also made parents slightly more likely to recommend text messages to other parents, but these differences are not significant. We speculate that the divergence of app and self-reported results could be because the content of these messages change parental activities beyond the limited behaviors that we can observe in the app.

We see that the goal-setting and view-of-reading effects are significant in the interaction effects in Panel B. The effectiveness of a combination of values does seem to increase app usage when combined with goal setting, even though the individual components were not significant on their own. The combined effect of goal-setting and a combination of view-of-reading was significantly more than either of the individual effects of app usage: 4 percentage points higher

login rate, 1.8 additional minutes on the app, and an addition 0.7 books completed. These effects are similar in magnitude to the main effects of personalization described above but need to be considered in the context of the non-significant main effects of the goal-setting and view-of-reading factors. Thus, goal setting alone does not seem to have an effect on directly observed app usage, but the combination of goal setting and changing the view of reading does affect these outcomes.

Effects on student reading performance

To understand whether these changes in parental behaviors and beliefs translate into effects for student outcomes, we present effects on reading scores in Table 7, following the same panel structure to investigate both main effects and interaction effects.

We see evidence of transfer from the behavioral effects of personalized information to student test score outcomes. Personalized text messages significantly improve Fall MAP scores by 0.03 standard deviations. While not significant, the point estimate for the effect of personalization on BOG scores is also positive, though slightly smaller than the effect on MAP (ES = 0.02). The MAP effect remains marginally significant after correcting for multiple hypotheses. Despite evidence that goal setting and view-of-reading differences affected parental experiences with the intervention, we find no significant main effects for changing either of these components on student test scores. We interpret the positive effects of personalization as transfer of the increased parental reading behaviors with the app because the app activities were designed to support the specific foundational literacy skills measured by the MAP assessment. The BOG effects represent farther transfer, which may be why they are smaller and less precise. The differences are not due to differences in sample; in sensitivity analyses we find similar effects on the MAP when we limit the sample to rising third graders who also took the BOG.

Panel B shows the interaction effects when we consider whether the observed main effects of each factor depend on the levels of the other components. Though the point estimates are only significant for Fall BOG scores, we see a similar pattern across both outcomes: personalization with the entertainment-only view is slightly less effective than personalization with instrumental only view (for the BOG, ES = -0.077, $p<0.05$), but personalization with both reading views presented is more effective than personalization with instrumental only values emphasized (for the BOG, ES = 0.104, $p<0.001$). Taken together with the positive interaction effects of the combined view and goal setting on parental reading behaviors, these effects on student test scores are further evidence that the view-of-reading framing of messages has the potential to enhance or detract from the effects of other factors.

Figure 2 shows these results visually, where we plot the average predicted test score outcomes for each of the six groups determined by their levels of personalization and the view-of-reading emphasized (collapsing across levels of goal setting). The personalized text results are in green and the generic texts are in blue. The graph also moves from no entertainment on the left to the most entertainment on the right. From the figure it is easy to see across almost all the interaction conditions that personalization outperforms no personalization because the green markers are almost always above the blue markers across MAP and BOG. In addition, focusing on the blue markers, absent personalization of text messages, students perform about 0.05 standard deviations better on fall reading assessments when they receive text messages emphasizing the entertainment view as opposed the instrumental view of reading. And while we saw from the output that personalization improves test scores on average, Figure 2 highlights that this effect is two to three times larger for students receiving a combination of reading values than for those receiving a single type of reading value. Figure 2 also includes the standardized

outcomes for our pure control group as a reference. The observed average control outcome highlights that our effects are not simply due to the mode of our intervention -- the specific content and framing of the message matters. While the most effective types of text messages are helpful for student reading, less-effective messages may be worse than no messages at all.

Mechanisms and student-level moderators

In this section, we investigate mechanisms for the transfer of effects from parents to student. Specifically, we explore why the effects of view-of-reading content that parents received in text messages appear to amplify only the effect of message personalization for student reading.

Student self-reports

Table 8 columns (1) and (2) show the self-reported reading behaviors of the students. Column (1) shows that students whose parents received personalized messages reported reading approximately 2.5 percentage points more of their available books on the MORE@Home reading app,[3] but in column (2) there is no effect on visiting the library. There are no significant main effects for either goal-setting or view-of-reading in column (1) or (2), but the point estimates are positive for goal-setting and the combination view-of-reading conditions, and negative for the entertainment-only condition. Looking at the interaction effects for the proportion of books read in column (1), we continue to see that emphasizing the entertainment value tended to mitigate or counteract other effects, particularly any non-significant effects of goal-setting. Students in both the goal-setting and the entertainment values conditions read 3.1 percentage points fewer of the available books than would have been predicted by the main effects of goal setting and the entertainment-view alone.

We see a different pattern when we consider the differential effects on students reading motivation as measured by the MMRP in column (3). There are very few main effects, though students receiving a combination of reading values emphasized had fall reading motivations that were lower (ES = 0.078) than students who only received instrumental messages. This negative effect of the combination value is amplified among students who were also in the goal-setting condition (ES = -0.122). On the other hand, we see that the combination of goal-setting and the entertainment value increased student reading motivation (ES = 0.102). We see a similar pattern, though smaller and not significant, when we look at the interaction effects between personalization and the view-of-reading, where entertainment enhances the personalization condition.

Although the strength of these relationships is not statistically significant, these results are particularly interesting because they point in the opposite direction of our main test score results. Students whose parents received the entertainment-only-view messages reported reading less, and if they were also in the personalization or goal-setting conditions tended to score lower on reading assessments. However, they reported higher levels of reading motivation. This suggests that the entertainment-themed messages did encourage students to see reading as more enjoyable, but that this did not translate into the types of behaviors associated with improving reading skills. On the other hand, students who received a combination of reading values in their message in addition to either goal setting or personalization used the educational app more and scored higher on reading assessments. But these same students reported lower reading motivations when they returned to school in the fall. One hypothesis is that if students were engaged in activities framed as instrumental, they found them harder, and thus less enjoyable.

However, overall this challenge improved their reading scores more than easier activities that may have seemed fun to a young student.

Qualitative app experience among app users

We continue to explore the motivation hypothesis by looking at what students reported about their experiences with the MORE@Home reading app. We know that the app users are only a subset of our total sample, and that they are not necessarily representative of all students. Despite having a meaningfully smaller sample, students who received personalized messages were more likely to enjoy the app activities and to find them challenging, though these differences are not statistically significant. There is suggestive evidence that receiving a combination of values increased students' correct answers on the app, but there are no other significant main effects for the goal setting or reading values conditions.

The interaction effects align with the self-reported reading motivation results. Students in entertainment value and personalization conditions in column (5) felt like much better readers (ES = 0.191) than students who were in instrumental value and personalization condition. On the other hand, students in both the combination of values and personalization conditions reported feeling like worse readers than students who were in one or the other (ES = -0.154). These amplification patterns for the values conditions play out in their interaction with goal setting as well, but those differences are not statistically significant. This may indicate that students who were encouraged to use the reading app for fun tended to feel good about themselves afterwards, whereas students who received targeted "mixed messages" about enjoying and learning from the app tended to feel slightly worse.

Patterns across theory-driven subgroups

Overall, we find that the transfer of effects from parental behaviors to student outcomes may be shaped by how parents communicated the value of reading to students, and how these expectations shaped students' affective experience with their summer reading activities. Existing literature suggests that these effects may be concentrated among certain groups of students, so we compare our full-sample effects for three sets of theory-driven subgroups: better versus worse baseline readers, students from mid- or high-SES neighborhoods versus students from low-SES neighborhoods, and more motivated versus less motivated readers. Based on existing theory and empirical results, we expect to see larger effects for students who might be categorized as being disadvantaged (worse-than-average readers at baseline, from a low-SES neighborhood, and less motivated to read) compared with the subgroups representing a form of advantage (better-than-average readers at baseline, from a mid-high SES neighborhood, and already motivated to read).

Figure 3 shows the impacts of each independent component on student test scores. In these charts, the advantaged groups are shown in green, the disadvantaged groups are show in blue, and we also report the full sample effects in black. We further breakdown each subgroup by shape using circles for baseline reading, squares for neighborhood SES, and triangles for baseline reading motivation. The categories at the bottom of the chart represent each of the four main-effects comparisons from our empirical model. The biggest takeaway is that the main effects of each component are fairly consistent across the different subsamples. As indicated by the overlapping 95% confidence intervals, there are very few significant differences, particularly with regards to socio-economic status and baseline reading motivation. Overall, the personalization of the text messages benefits almost all of the subgroups. However, the effects on the BOG assessment are better for good readers and worse for poor readers, though we do not see this for the MAP assessment. These results are congruent with prior work on literacy texting

interventions. While York et al. (2019) found that lower performing students benefited more from the general texting strategy, Doss et al. (2019) found that by leveling the content of the message, the lower and upper ends of the distribution benefited the most because the messages were no longer targeted at the mean respondent. Our results align with that pattern. The results for lower-SES family neighborhoods are more surprising since prior theoretical work on cognitive demand suggests there should be larger effects for this group. However, it is plausible that the lack of educational structures during summer means that all families, regardless of SES, find it more difficult to monitor activities and behaviors, leading to the formation of inaccurate beliefs that are corrected by the personalized messages.

Additionally, we find that good readers benefit significantly more from the entertainment-framed condition than poor readers, and poor readers benefit significantly more from the combination value condition than better readers on the MAP. We find a similar pattern on the BOG assessment, but they are not significant. Furthermore, while the effects are not statistically significant, another pattern is that texts conveying a combination of reading views, benefits most disadvantaged subgroups on both the MAP and BOG. These results point to the possibility that exposure to more entertainment-based messaging alone will not benefit lower performing readers or low-SES families.

Next we consider the interaction effects in Figure 4. The layout of Figure 4 is similar to Figure 3, but now each group of point estimates now refers to one of the two-way interaction effects from our empirical model. Similar to the main effects, the overlapping 95% confidence intervals indicate there are very few significant differences in the effects across subgroups, indicating that our findings are broadly consistent for most students. Of particular importance, the interaction of personalization and entertainment is generally small and often negative,

particularly for the BOG. These negative interactions indicate that while entertainment-themed messages and personalized messages may have positive effects independently, the combined effect is not additive. Unlike the main effects, there are no effects that differ significantly between subgroups. Taken together, these main effect and interaction effect findings do not support theory that messages are differentially effective based upon student's socio-economic status or baseline reading motivation.

Discussion

In this study, we demonstrate how a factorial design applied to a parent literacy texting intervention can improve both parental engagement with summer reading resources and student academic outcomes. The main contribution of our work is that it clearly demonstrates that variation in the content and tone of messages can lead to important differences in both parental behaviors and student academic outcomes. Moreover, our use of a factorial design allows us to estimate not only the impact of each individual component, but also how those components' effects interact.

We find that personalized messages improve parental use of an educational reading app across a variety of measures and transfer to student reading performance in the fall. Furthermore, though the parent outcomes were not shown in the subgroup figures, there was little difference between the subgroups effects for logging on and using the app. This clear effectiveness of correcting these parents' misbeliefs through information aligns with similar interventions that have addressed student absenteeism in early grades (Robinson et al., 2018; Rogers & Feller, 2018; Smythe-Leistico & Page, 2018). These prior studies demonstrate that communication with parents corrected the parents' misbeliefs on the number school days their child missed. In our

context, this suggests that parents may be overestimating how much time they are spending with their child on reading over the summer.

Our findings that more personalized messages outperform generic messages on reading comprehension outcomes is also consistent with prior work on literacy reminders (Cabell et al., 2020; Doss et al., 2018), but also differs from their experiments in important ways. First, our texting intervention was only 9-weeks during the summer compared with a 10-month period encompassing the school year and the summer. In addition, the differentiation in the prior studies' messages provided different recommended behaviors based on student reading levels, whereas our personalized messages provide new information to families about their own behaviors. Finally, the content of our messages was not focused exclusively on reading activities, but rather included messages related to reading resources, reminders to engage in reading activities, and checking-in on progress. Thus, our work demonstrates that shorter sustained periods of engagement with more personalized information (as opposed to personalized requests) can also be effective and may suggest that if the texting continued throughout the year that the effects would have been even larger.

Home literacy theory suggests that emphasizing an entertainment-view of reading could lead to improved student motivation and literacy outcomes (Baker, Mackler, Sonnenschein, & Serpell, 2001). We do not see broad evidence of transfer to self-reported reading motivation or positive feelings about app-based reading activities. Additionally, there are no average effects on student test scores. However, we do find that emphasizing the entertainment view significantly improves test scores for students who are already good readers and hurts outcomes for students who are not.

At the same time, we do find that presenting a combination of the entertainment-view and the instrumental-view amplifies the positive effects of personalization, whereas only presenting the entertainment view detracts from the effects of personalization. Yet, we find suggestive evidence that for our sample these effects do not operate through student motivation. Students in the personalization and entertainment condition saw statistically significant and positive effects on whether they felt like a good reader, as well as positive effects on a survey of reading motivation, with the converse being true for when both the entertainment and instrumental views-of-reading were present. Thus, these entertainment themed messages did encourage students to feel more motivated, but this did not translate into the types of behaviors associated with improving reading skills. One potential mechanism could be that adding instrumental messages encouraged parents to practice skill-building with their children, which can be less enjoyable, but would improve their reading scores. This type of messages would be particularly useful for parents if they are also receiving specific feedback about the amount of reading they were doing on the app.

We were surprised by the limited effectiveness of the goal-setting component of our intervention, which tended to have point estimates close to zero across multiple subgroups and outcomes. Mayer and colleagues (2019) found large positive effects on the use of an educational app in their comprehensive behavioral intervention (PACT) that leveraged goal setting, reminders, and social pressure. However, there are two key differences that may explain these discrepant findings. First, the goal-setting component of PACT was intensive, involving weekly face-to-face meetings with program staff. This was possible because they were only working with 80 students. Out of necessity, our goal-setting intervention was much lighter touch, because it had to be scaled up to reach almost 5,000 families. With so few parents electing to set goals

through our online form, the goal-setting factor of our intervention should best be described as a goal-oriented framing of summer reading as opposed to the commitment device used in Mayer. In addition, goal setting was just one of three major components of the PACT. We have seen from this study that our own positive effects were largely driven by information-sharing, which aligns more closely with the reminder and social-pressure components of the PACT intervention. It is possible that the large positive effects Mayer et al. find are primarily due to the other components of the intervention.

Policy Implications

One implication of our work is that not only that the content of the message matters, but it is insufficient to consider how elements of message content like personalization or value-framing operate independently. First, Figure 2 shows that non-personalized messages perform worse than sending no text at all, whereas the personalized messages are clearly better. Recent research examining 126 nudge randomized controlled trials "nudge units" in government covering over 23 million individuals found statistically significant, but considerably smaller effects on nudging interventions when compared to the published literature, which they attribute to publication bias and power (DellaVigna & Linos, 2020). However, the authors do not test for content differences, which our work shows could be a crucial feature. Second, for distal outcomes, like beginning of third grade exam scores, neither personalization nor an emphasis of particular view-of-reading is statistically significant on its own. However, combining these types of messages, personalization with a combination of reading views is more effective than personalization with instrumental only values emphasized or than a combination of values presented in generic messages. These nuances seem particularly important for our more distal outcomes, such as test scores, highlighting that the values framing could be particularly

important for reading behaviors that involve parental engagement but are not related to the educational app. As text messaging campaigns become increasingly popular in arenas including education, civic engagement, and health, these findings highlight that not all messages are equally effective, for more distal outcomes in particular and that factorial designs are valuable tools to unpack these nuances.

Relatedly, in education there has been little causal research on whether informing individuals to expand the frame of choices could influence their behavior. For example, all the prior intervention literature on parental texting for younger children focuses exclusively on building skills. Other international research in education nudges has found that reframing a cash transfer for educational had remarkable effects on parent behavior (Benhassine et al., 2015), but we're aware of no other research in the areas. Our work suggests that messages that potentially broaden and inform the parent's choice set by exposing them to additional views could be important for future research not only for parental texting literacy interventions, but in educational nudges more broadly.

A second implication is emerging evidence that varying types of content that parents receive could be beneficial, particularly for more sustained parent engagement through text message. York et al. (2019) find that their effects of the intervention were much weaker in the first year of implementation. While they are unable to empirically test the hypothesis, one idea was that in the second-year parents received text messages that combined literacy, mathematics, and socioemotional skills, rather than literacy alone. Particularly because their intervention was over 8 months, the additional variety of messages could have helped, even though it potentially reduced the number of literacy activities by parents. Similarly, we find that when paired with student information, a combination of reading-views dominated either type of view. In an age

where groups like advertisers and political campaign bombard parents with text messages, a single source that can vary the types of actions and activities could be beneficial. At the same time, especially if the text can be sustained over longer periods of time, there might be opportunities to apply the lessons learned about message content and framing from text messaging interventions that move beyond "nudges" into sustained assistance and support for parents.

Finally, our work continues to build on the nudging literature that messaging intervention can successfully address attention limitations that parents and others face at a much lower cost than other interventions with similar effect sizes. We estimated a back-of-the-envelope cost of this intervention of less than $4 per student, which included the cost to build the connection and integration between a student database and the texting software, the staffing associated with composing each text message, and responding to questions and messages received from families, as well as the cost of sending the text messages through Twilio's platform. While our test score effect sizes are modest, the cost-effectiveness would compare favorably with the Tennessee STAR class size experiment (Schanzenbach, 2006), and other literacy interventions, including Reading Partners (Jacob, Armstrong, Bowden, & Pan, 2016), Project READS (Kim et al., 2016), and others (Hollands, Kieffer, Shand, Pan, Cheng, & Levin, 2013).

Future directions

Our work also offers several lessons for the design of future evaluation of parental texting interventions. First, our text messages were direct to parents, yet only a small proportion of the families responded to our parent survey, which limited our ability to further unpack some of the mechanisms that drove our results. Particularly because factorial design allows researchers to more efficiently explore heterogenous effects, ensuring sufficient effort is dedicated to this data

source will be important for future work. Similarly, we solicited our initial goal-setting survey via text message and received a poor response rate on the parental goals, despite being based in goal-setting theory about the importance of plans (Oettingen & Reininger, 2016). Perhaps an alternative in-between our work and Mayer et al., is to call parents for the survey. In other fields, like political science, phone calls encouraging voters to "make a plan" about voting significantly increased turnout (Nickerson & Rogers, 2010; Rogers et al., 2015) and now is common practice in voter outreach. Finally, as noted in other work (York et al., 2019; Kraft and Monti-Nussbaum, 2017) there is learning curve to sending these parental text messages at scale, thus we suggest that district and policymakers commit to these interventions in order to reap the benefits after the setup cost.

Notes

1. Other studies have used informational messaging campaigns administered through a non-texting medium to reduce absenteeism among students (Rogers & Feller, 2018; Rogers et al., 2017; Robinson, Lee, Dearing, & Rogers, 2018).

2. Our students had several other home languages designated, in addition to English and Spanish, though none represented more than (1%) of students. Families whose home language was neither English nor Spanish received English messages due to resource constraints.

3. As described previously, all students received access to 6 electronic books and activities via the MORE@Home app, but half also received a set of 10 hardcopy books and additional app activities. Thus, for one half of the sample, the percent of available books read is a proportion of the 16 total books they had access to via the program.

References

Baker, L., Mackler, K., Sonnenschein, S., & Serpell, R. (2001). Parents' interactions with their first-grade children during storybook reading and relations with subsequent home reading activity and reading achievement. *Journal of School Psychology*, 39(5), 415–438. https://doi.org/10.1016/S0022-4405(01)00082-6.

Baker, L., & Scher, D. (2002). Beginning readers' motivation for reading in relation to parental beliefs and home reading experiences. *Reading Psychology*, 23(4), 239–269. https://doi.org/10.1080/713775283

Baker, L., Scher, D., & Mackler, K. (1997). Home and family influences on motivations for reading. *Educational Psychologist*, 32(2), 69–82. https://doi.org/10.1207/s15326985ep3202_2

Benhassine, N., Devoto, F., Duflo, E., Dupas, P., & Pouliquen, V. (2015). Turning a shove into a nudge? A "labeled cash transfer" for education. *American Economic Journal: Economic Policy*, 7(3), 86–125. https://doi.org/10.1257/pol.20130225

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B* (Methodological), 57(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Cabell, S. Q., Zucker, T. A., DeCoster, J., Copp, S. B., & Landry, S. (2019). Impact of a parent text messaging program on pre-kindergarteners' literacy development. *AERA Open*, 5(1), 233285841983333. https://doi.org/10.1177/2332858419833339

Collins, L. M., Dziak, J. J., Kugler, K. C., & Trail, J. B. (2014). Factorial experiments: Efficient tools for evaluation of intervention components. *American Journal of Preventive Medicine*, 47(4), 498–504. https://doi.org/10.1016/j.amepre.2014.06.021

Cortes, K. E., Fricke, H., Loeb, S., Song, D. S., & York, B. N. (2019). Too little or too much? Actionable advice in an early-childhood text messaging experiment. *Education Finance and Policy*, 1–44. https://doi.org/10.1162/edfp_a_00304

Cortes, K., Fricke, H. D. U., Loeb, S., Song, D., & York, B. (2019). When behavioral barriers are too high or low – how timing matters for parenting interventions (No. w25964; p. w25964). National Bureau of Economic Research. https://doi.org/10.3386/w25964

Damgaard, M. T., & Nielsen, H. S. (2018). Nudging in education. *Economics of Education Review*, 64, 313–342. https://doi.org/10.1016/j.econedurev.2018.03.008

Del Boca, D., Flinn, C., & Wiswall, M. (2014). Household choices and child development. *The Review of Economic Studies*, 81(1), 137–185. https://doi.org/10.1093/restud/rdt026

DellaVigna, S., & Linos, E. (2020). RCTs to scale: Comprehensive evidence from two nudge units (No. w27594; p. w27594). National Bureau of Economic Research. https://doi.org/10.3386/w27594

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

Bureau of Labor Statistics. (2019). Employment characteristics of families—2019 [Press release]. U.S. Department of Labor, 14.

Francesconi, M., & Heckman, J. J. (2016). Child development and parental investment: Introduction. *The Economic Journal*, 126(596), F1–F27. https://doi.org/10.1111/ecoj.12388

Greenman, E., Bodovski, K., & Reed, K. (2011). Neighborhood characteristics, parental practices and children's math achievement in elementary school. *Social Science*

Research, 40(5), 1434–1444. https://doi.org/10.1016/j.ssresearch.2011.04.007

Guryan, Hurst, Kearney, G., Hurst, Kearney, & Guryan, J. (2008). Parental education and parental time with children. The Journal of Economic Perspectives, 22(3), 23–26. http://doi.org/10.1257/jep.22.3.23.

Hardy, M. (1993). Regression with dummy variables. SAGE Publications, Inc. https://doi.org/10.4135/9781412985628

Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. Science, 312(5782), 1900–1902. https://doi.org/10.1126/science.1128898

Hollands, F. M., Kieffer, M. J., Shand, R., Pan, Y., Cheng, H., & Levin, H. M. (2016). Cost-effectiveness analysis of early reading programs: A demonstration with recommendations for future research. Journal of Research on Educational Effectiveness, 9(1), 30–53. https://doi.org/10.1080/19345747.2015.1055639

Hurwitz, L. B., Lauricella, A. R., Hanson, A., Raden, A., & Wartella, E. (2015). Supporting Head Start parents: Impact of a text message intervention on parent–child activity engagement. Early Child Development and Care, 185(9), 1373–1389. https://doi.org/10.1080/03004430.2014.996217

Jacob, R. T., Armstrong, C., Bowden, A. B., & Pan, Y. (2016). Leveraging volunteers: An experimental evaluation of a tutoring program for struggling readers. Journal of Research on Educational Effectiveness, 9(S1), 67–92. https://doi.org/10.1080/19345747.2016.1138560

Kim, J. S., Asher, C. A., Burkhauser, M., Mesite, L., & Leyva, D. (2019). Using a sequential multiple assignment randomized trial (SMART) to develop an adaptive K–2 literacy intervention with personalized print texts and app-based digital activities. AERA Open.

https://doi.org/10.1177/2332858419872701

Kim, J. S., Guryan, J., White, T. G., Quinn, D. M., Capotosto, L., & Kingston, H. C. (2016). Delayed effects of a low-cost and large-scale summer reading intervention on elementary school children's reading comprehension. Journal of Research on Educational Effectiveness, 9(sup1), 1–22. https://doi.org/10.1080/19345747.2016.1164780

Kraft, M. A., & Monti-Nussbaum, M. (2017). Can schools enable parents to prevent summer learning loss? A text-messaging field experiment to promote literacy skills. The ANNALS of the American Academy of Political and Social Science, 674(1), 85–112. https://doi.org/10.1177/0002716217732009

Kugler, K. C., Trail, J. B., Dziak, J. J., & Collins, L. M. (2012). Effect coding versus dummy coding in analysis of data from factorial experiments. The Methodology Center, Pennsylvania State University. Retrieved from https://www.methodology.psu.edu/

Lynch, J., Anderson, J., Anderson, A., & Shapiro, J. (2006). Parents' beliefs about young children's literacy development and parents' literacy behaviors. Reading Psychology, 27(1), 1–20. https://doi.org/10.1080/02702710500468708

Mani, A., Mullainathan, S., Shafir, E., & Zhao, J. (2013). Poverty impedes cognitive function. Science, 341(6149), 976–980. https://doi.org/10.1126/science.1238041

Marinak, B. A., Malloy, J. B., Gambrell, L. B., & Mazzoni, S. A. (2015). Me and My Reading Profile: A tool for assessing early reading motivation. The Reading Teacher, 69(1), 51–62. https://doi.org/10.1002/trtr.1362

Mayer, S. E., Kalil, A., Oreopoulos, P., & Gallegos, S. (2018). Using behavioral insights to increase parental engagement: The Parents and Children Together intervention. Journal of Human Resources, 0617-8835R. https://doi.org/10.3368/jhr.54.4.0617.8835R

Measure of Academic Progress. (2011). NWEA.

Nickerson, D. W., & Rogers, T. (2010). Do you have a voting plan?: Implementation intentions, voter turnout, and organic plan making. Psychological Science, 21(2), 194–199. https://doi.org/10.1177/0956797609359326

Kramer, S. (2019). U.S. has world's highest rate of children living in single-parent households. Pew Research Center: FactTank News in the Numbers. Retrieved September 15, 2020, from https://www.pewresearch.org/fact-tank/2019/12/12/u-s-children-more-likely-than-children-in-other-countries-to-live-with-just-one-parent/

Oettingen, G., & Reininger, K. M. (2016). The power of prospection: Mental contrasting and behavior change: Mental contrasting and behavior change. Social and Personality Psychology Compass, 10(11), 591–604. https://doi.org/10.1111/spc3.12271

Oreopoulos, P., Petronijevic, U., Logel, C., & Beattie, G. (2020). Improving non-academic student outcomes using online and text-message coaching. Journal of Economic Behavior & Organization, 171, 342–360. https://doi.org/10.1016/j.jebo.2020.01.009

Robinson, C. D., Lee, M. G., Dearing, E., & Rogers, T. (2018). Reducing student absenteeism in the early grades by targeting parental beliefs. American Educational Research Journal, 55(6), 1163–1192. https://doi.org/10.3102/0002831218772274

Rogers, T., & Feller, A. (2018). Reducing student absences at scale by targeting parents' misbeliefs. Nature Human Behaviour, 2(5), 335–342. https://doi.org/10.1038/s41562-018-0328-1

Rogers, T., Duncan, T., Wolford, T., Ternovski, J., Subramanyam, S., & Reitano, A. (2017). A randomized experiment using absenteeism information to "nudge" attendance (REL 2017-252; p. 26). U.S. Department of Education, Institute of Education Sciences,

National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Mid-Atlantic. http://ies.ed.gov/ncee/edlabs.

Rogers, T., Milkman, K. L., John, L. K., & Norton, M. I. (2015). Beyond good intentions: Prompting people to make plans improves follow-through on important tasks. Behavioral Science & Policy, 1(2), 33–41. https://doi.org/10.1353/bsp.2015.0011

Schanzenbach, D. Whitmore. (2006). What have researchers learned from project STAR? Brookings Papers on Education Policy, 2006(1), 205–228. https://doi.org/10.1353/pep.2007.0007

Sénéchal, M., & Young, L. (2008). The effect of family literacy interventions on children's acquisition of reading from kindergarten to grade 3: A meta-analytic review. Review of Educational Research, 78(4), 880–907. https://doi.org/10.3102/0034654308320319

Smythe-Leistico, K., & Page, L. C. (2018). Connect-text: Leveraging text-message communication to mitigate chronic absenteeism and improve parental engagement in the earliest years of schooling. Journal of Education for Students Placed at Risk (JESPAR), 23(1–2), 139–152. https://doi.org/10.1080/10824669.2018.1434658

Somers, M.-A., Collins, L., & Maier, M. (2014). Design options for an evaluation of head start coaching: Review of methods for evaluating components of social interventions (OPRE Report #2014-81; Produced by American Institutes for Research for Head Start Professional Development: Developing Evidence for Best Practices in Coaching.). U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research and Evaluation. https://www.mdrc.org/sites/default/files/hspd_task_3_2_review_final_11_19_2014new_title.pdf

Sonnenschein, S., Baker, L., Serpell, R., Scher, D., Fernandez-Fein, S., & Munsterman, K. A. (1996). Strands of emergent literacy and their antecedents in the home: Urban preschoolers early literacy development (No. 48; p. 49). National Reading Research Center.

Sonnenschein, S., Baker, L., Serpell, R., Scher, D., Truitt, V. G., & Munsterman, K. (1997). Parental beliefs about ways to help children learn to read: The impact of an entertainment or a skills perspective. Early Child Development and Care, 127(1), 111–118. https://doi.org/10.1080/0300443971270109

Sonnenschein, S., Baker, L., Serpell, R., & Schmidt, D. (2000). Reading is a source of entertainment: The importance of the home perspective for children's literacy development. In Play and literacy in early childhood: Research from multiple perspectives (pp. 107–124). Lawrence Erlbaum Associates Publishers.

Tanaka, Y., & Yamano, T. (2015). Risk and time preference on schooling: Experimental evidence from a low-income country. 37.

York, B. N., Loeb, S., & Doss, C. (2018). One step at a time: The effects of an early literacy text messaging program for parents of preschoolers. Journal of Human Resources, 0517-8756R. https://doi.org/10.3368/jhr.54.3.0517-8756R

Table 1

Parent Text Messaging Studies

Paper	Components of intervention	Average Effects	Significant Subgroup Effects
Reminders and Complexity Reduction of Home Literacy Activities			
Cabell et al. (2019)	Varied message focus on literacy activities or health and well-being: messages utilized student name and gender, importance of the topic, actionable acitivity, and encoragement	Test scores: no significant differences	High performers: better with the literacy texts Lower performers: better with the health texts
Cortes et al. (2018)	Varied frequency (one, three, or five messages per week) of home literacy activities: messages sequenced a fact, reading activity tip, and a growth activity	Test scores: no significant differences	Low performers: -0.15 SD lower for 1 message/week; no difference of 5 messages/week
Cortes et al. (2019)	Varied timing (weekday or weekend) of literacy activities: messages sequenced a fact, reading activity tip, and a growth activity	Test scores: no significant differences	Low Performers: +0.06 SD
Doss et al. (2019)	Varied personalized activities to reading level or not of home literacy activities: messages sequenced a fact, reading activity tip, and a growth activity	Amount of reading activities: 0.26 SD Reading level: 0.18 SD on reading level Ease of building reading skills (parents): 0.32 SD For general campaign: smaller, less significant effects	Bottom/top quartile: 2.5 time more likely to move up a reading level 21.5pp more likely to exceed expectations
Hurwitz et al. (2015)	Reminders of home literacy: message focused on activity and encoragement	Amount of reading activities: 0.3	Males: 1 additional activity
Kim et al. (2019)	Reminder of home literacy: messages focused on activity and encouragement. Used reading app.	Test scores: 0.10 SD, marginally significant	
Kraft and Monti-Nussbaum (2017)	Reminder of home literacy: message focused on facts, resources, and activities	Parent behavior: more likely to attend conferences and take their child to the library. Test scores: mixed/marginally significant effects	3rd graders: 0.21 SD 4th graders: 0.29 SD
York et al. (2019)	Reminder of home literacy: Messages sequenced a fact, reading activity tip, and a growth activity	Parental involvement at home: 0.16 SD Parental involvement at school by 0.14 SD Test scores: 0.11 SD	Low performers: 0.31 SD
Goal Setting for Literacy Activities			
Mayer et al. (2018)	Reminder of weekly reading goal: intervention involved goal setting, social rewards, reading app	Reading amount: 1 SD	Higher discount parents drove effects
Debiasing Parent Beliefs			
Smythe-Leistico & Page (2018)	Reminders on attendance and home literacy: the text encouraged two-way communication	Absenteeism: -11 percentage points	Spanish-speakers: qualitatively large, but not significant

Table 2

Text Message Topics, Organized by Theme

Summer Reading Resources	**Reminders to Engage in Reading Activities**	**Monitoring progress**
· App feature: availability of books	· Tip: talking about books	· Kickoff & goal setting
· App feature: personalized activities/goal reminder	· App is great	· Monitoring: check-in on goals/progress
· App feature: "catching words"	· Tip: planning time to read	· Monitoring: check-in on goals/progress
· Resource: chromebooks available at library	· Social pressure: number of users	· Closeout & survey preview
· App feature: new content	· Tip: talk about favorite book	
· App feature: 2 sets of activities		

Table 3

Comparing Message Variations Across Conditions for Two Example Messages

	Generic	Personalized information
Message 1		
Instrumental view	The MORE@Home app contains personalized activities for each book that will help [StudentFirstName] **develop reading skills**.	The MORE@Home app contains personalized activities for each book. [StudentFirstName] can **practice different reading skills** for both **[Book1Title] and [Book2Title]**.
Entertainment view	The MORE@Home app contains personalized activities for each book. We think [StudentFirstName] **will have fun doing them all!**	The MORE@Home app contains personalized activities for each book. [StudentFirstName] can **have fun exploring them** for both **[Book1Title] and [Book2Title]**.
Message 2		
No goals	We are already 6 weeks into summer vacation! Hopefully you and [StudentFirstName] are making progress on your summer reading list!	We are 6 weeks into summer vacation and [StudentFirstName] **has used [NumBooksAccessed] books on the MORE@Home app**. Keep reading to get through all of them!
Goals	We're already 6 weeks into summer vacation! At this point, you should be about **1/2 of the way to your summer reading goal**.	We're already 6 weeks into summer vacation! At this point, you should be about 1/2 of your way to your **summer reading goal of [BookGoal] books**.

Table 4

Baseline characteristics and balance checks

	Full Sample	Personalization vs. Not Difference	Entertainment vs. Instrumental Difference	Combination vs. Instrumental Difference	Goal Setting vs. Not Difference
White (%)	0.184	0.004	-0.000	-0.008	-0.001
	(0.388)	(0.010)	(0.012)	(0.012)	(0.010)
Black (%)	0.386	0.008	-0.023	-0.011	0.008
	(0.487)	(0.014)	(0.017)	(0.017)	(0.014)
Hispanic (%)	0.317	-0.008	0.022	0.001	-0.004
	(0.465)	(0.013)	(0.016)	(0.016)	(0.013)
Asian (%)	0.079	-0.001	0.008	0.023*	0.003
	(0.270)	(0.008)	(0.009)	(0.010)	(0.008)
Limted English Proficiency	0.226	-0.016	0.014	-0.000	0.013
	(0.418)	(0.012)	(0.015)	(0.015)	(0.012)
Low SES	0.405	0.017+	0.004	-0.004	-0.004
	(0.491)	(0.010)	(0.012)	(0.012)	(0.010)
Med SES	0.385	-0.000	0.002	0.011	0.004
	(0.487)	(0.011)	(0.014)	(0.014)	(0.011)
High SES	0.205	-0.018*	-0.009	-0.010	-0.001
	(0.404)	(0.007)	(0.009)	(0.009)	(0.007)
Standardized Spring MAP ELA RIT	-0.000	0.040	0.001	-0.014	-0.022
	(0.988)	(0.027)	(0.033)	(0.033)	(0.027)
Standardized Spring MAP Math RIT	-0.000	0.052+	0.001	0.005	-0.042
	(0.988)	(0.027)	(0.032)	(0.033)	(0.027)
N	5175	4678	3119	3128	4678

Source: district administrative records

Notes: Each row represents a separate regression. Point estimates reflect the condition-reference group differences derived from a model that includes indicators for the randomization block. Robust standard errors clustered at the family level (in parentheses).

+p<0.10, *p<0.05, **p<0.01, ***p<0.001

Table 5

Treatment Conditions in Factorial Design with Effect Coding

Group	Received Text Messages	Value	Goals Condition	Value	Personalization Condition	Value	Reading Value Condition	Value (Entertain)	Value (Both)
1	No	0		.		.		.	.
2	Yes	1	No Goals	-1	No Personalization	-1	Instrumental Only	-1	-1
3	Yes	1	No Goals	-1	No Personalization	-1	Entertainment Only	1	0
4	Yes	1	No Goals	-1	No Personalization	-1	Both	0	1
5	Yes	1	No Goals	-1	Personalization	1	Instrumental Only	-1	-1
6	Yes	1	No Goals	-1	Personalization	1	Entertainment Only	1	0
7	No	1	No Goals	-1	Personalization	1	Both	0	1
8	Yes	1	Goals	1	No Personalization	-1	Instrumental Only	-1	-1
9	Yes	1	Goals	1	No Personalization	-1	Entertainment Only	1	0
10	Yes	1	Goals	1	No Personalization	-1	Both	0	1
11	Yes	1	Goals	1	Personalization	1	Instrumental Only	-1	-1
12	Yes	1	Goals	1	Personalization	1	Entertainment Only	1	0
13	Yes	1	Goals	1	Personalization	1	Both	0	1

Table 6

Differential Effects of Text Messaging Components on Parent Outcomes

	Observed reading behaviors in app			Self-reported behaviors & beliefs		
	Ever logged in (%)	Minutes on app	Books completed	Frequency of reading activities	Found texts helpful (%)	Would recommend texts to others (%)
Panel A - Main Effects						
Personalization vs. Not	0.028*	1.626**	0.681**	0.046	0.023	0.013
	(0.013)	(0.601)	(0.257)	(0.139)	(0.168)	(0.069)
Entertainment vs. Instrumental	0.004	-0.386	0.033	-0.273	-0.090	-0.112
	(0.018)	(0.847)	(0.364)	(0.169)	(0.212)	(0.081)
Combination vs. Instrumental	-0.020	-0.179	-0.345	0.249	0.360+	0.068
	(0.018)	(0.853)	(0.353)	(0.174)	(0.214)	(0.081)
Goals vs. Not	-0.013	-0.245	-0.187	0.131	0.318*	0.027
	(0.013)	(0.606)	(0.261)	(0.122)	(0.151)	(0.060)
Panel B - Two-Way Interaction Effects						
Personalization x Entertainment	-0.013	-0.282	-0.149	0.191	0.211	0.012
	(0.018)	(0.847)	(0.361)	(0.168)	(0.198)	(0.084)
Personalization x Combination	0.006	-0.116	-0.187	-0.255	-0.097	-0.111
	(0.018)	(0.858)	(0.355)	(0.171)	(0.210)	(0.087)
Personalization x Goals	0.014	0.365	0.138	-0.044	-0.243	-0.051
	(0.013)	(0.604)	(0.258)	(0.121)	(0.162)	(0.063)
Entertainment x Goals	-0.021	-0.947	-0.464	-0.089	0.058	0.021
	(0.018)	(0.849)	(0.362)	(0.186)	(0.226)	(0.090)
Combination x Goals	0.041*	1.858*	0.686+	0.129	0.280	-0.003
	(0.018)	(0.851)	(0.352)	(0.185)	(0.244)	(0.099)
N	4678	4678	4678	319	266	263

Source: app-use records, survey of parent subsample, district administrative records

Notes: Point estimates derived from effect-coded regressions that include all treatment factors and their interactions, as well as the following covariates: gender, race/ethnicity, participation in gifted program, participation in Special Education, English learner status, neighborhood SES, language of text messages, baseline reading and math scores, and indicators of the randomization block. Robust standard errors clustered at the family level (in parentheses).

+p<0.10, *p<0.05, **p<0.01, ***p<0.001

Table 7

Differential Effects of Text Messaging Components on Student Test Scores

	MAP	Beginning of Grade
Panel A - Main Effects		
Personalization vs. Not	0.035*	0.021
	(0.017)	(0.024)
Entertainment vs. Instrumental	0.028	0.021
	(0.023)	(0.034)
Combination vs. Instrumental	0.025	-0.009
	(0.024)	(0.034)
Goals vs. Not	0.007	0.008
	(0.017)	(0.024)
Panel B - Two-Way Interaction Effects		
Personalization x Entertainment	-0.012	-0.076*
	(0.023)	(0.034)
Personalization x Combination	0.027	0.105**
	(0.024)	(0.034)
Personalization x Goals	-0.005	-0.010
	(0.017)	(0.024)
Entertainment x Goals	-0.009	0.016
	(0.023)	(0.034)
Combination x Goals	-0.011	-0.028
	(0.024)	(0.034)
N	3959	2038

Source: district administrative records

Notes: Point estimates derived from effect-coded regressions that include all treatment factors and their interactions, as well as the following covariates: gender, race/ethnicity, participation in gifted program, participation in Special Education, English learner status, neighborhood SES, language of text messages, baseline reading and math scores, and indicators of the randomization block. Robust standard errors clustered at the family level (in parentheses).

+p<0.10, *p<0.05, **p<0.01, ***p<0.001

Table 8

Differential Effects of Text Messaging Components on Alternate Student Outcomes

	Student Self-Reports			Qualitative Experience Among App-Users			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Proportion of books read (%)	Visited library (%)	Reading motivation	Enjoyed activities	Felt like a good reader	Performance on app activities	Found activities challenging
Panel A - Main Effects							
Personalization vs. Not	0.025*	0.009	-0.000	0.023	-0.003	0.083	0.070
	(0.011)	(0.017)	(0.033)	(0.064)	(0.065)	(0.061)	(0.065)
Entertainment vs. Instrumental	-0.019	-0.015	0.011	-0.026	-0.125	-0.095	-0.059
	(0.016)	(0.024)	(0.047)	(0.093)	(0.096)	(0.084)	(0.089)
Combination vs. Instrumental	0.018	0.022	-0.078+	0.079	0.115	0.154+	0.068
	(0.016)	(0.024)	(0.048)	(0.089)	(0.092)	(0.086)	(0.091)
Goals vs. Not	0.002	0.008	0.003	-0.050	-0.031	-0.023	0.078
	(0.011)	(0.017)	(0.033)	(0.065)	(0.067)	(0.061)	(0.066)
Panel B - Two-Way Interaction Effects							
Personalization x Entertainment	-0.001	-0.008	0.035	0.114	0.191*	0.147+	-0.132
	(0.016)	(0.024)	(0.047)	(0.092)	(0.094)	(0.086)	(0.093)
Personalization x Combination	-0.003	0.002	0.003	-0.050	-0.154+	-0.093	0.090
	(0.016)	(0.024)	(0.047)	(0.091)	(0.091)	(0.086)	(0.093)
Personalization x Goals	-0.011	-0.005	0.048	0.058	0.052	-0.026	-0.068
	(0.011)	(0.017)	(0.033)	(0.066)	(0.066)	(0.060)	(0.064)
Entertainment x Goals	-0.031+	0.009	0.102*	0.074	0.054	0.087	-0.024
	(0.016)	(0.024)	(0.047)	(0.093)	(0.094)	(0.083)	(0.093)
Combination x Goals	0.010	-0.053*	-0.122**	-0.134	-0.073	-0.112	0.025
	(0.016)	(0.024)	(0.047)	(0.090)	(0.089)	(0.085)	(0.092)
N	3472	3418	3490	995	995	1155	996

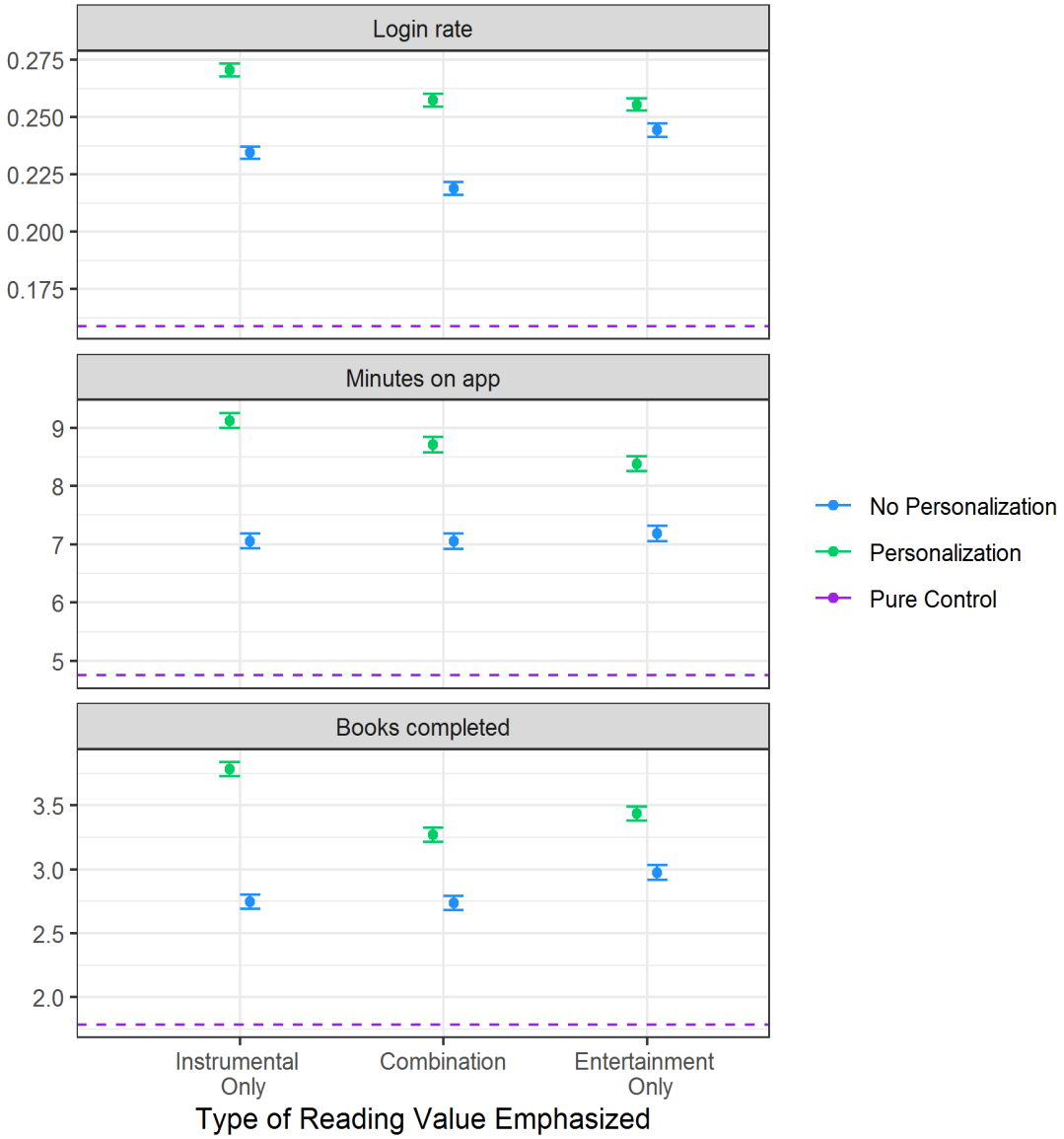
Source: student survey, district administrative records, app-use records

Notes: Point estimates derived from effect-coded regressions that include all treatment factors and their interactions, as well as the following covariates: gender, race/ethnicity, participation in gifted program, participation in Special Education, English learner status, neighborhood SES, language of text messages, baseline reading and math scores, and indicators of the randomization block. Robust standard errors clustered at the

+p<0.10, *p<0.05, **p<0.01, ***p<0.001

Figure 1

Predicted Reading Behaviors, by Personalization and View-of-Reading Conditions

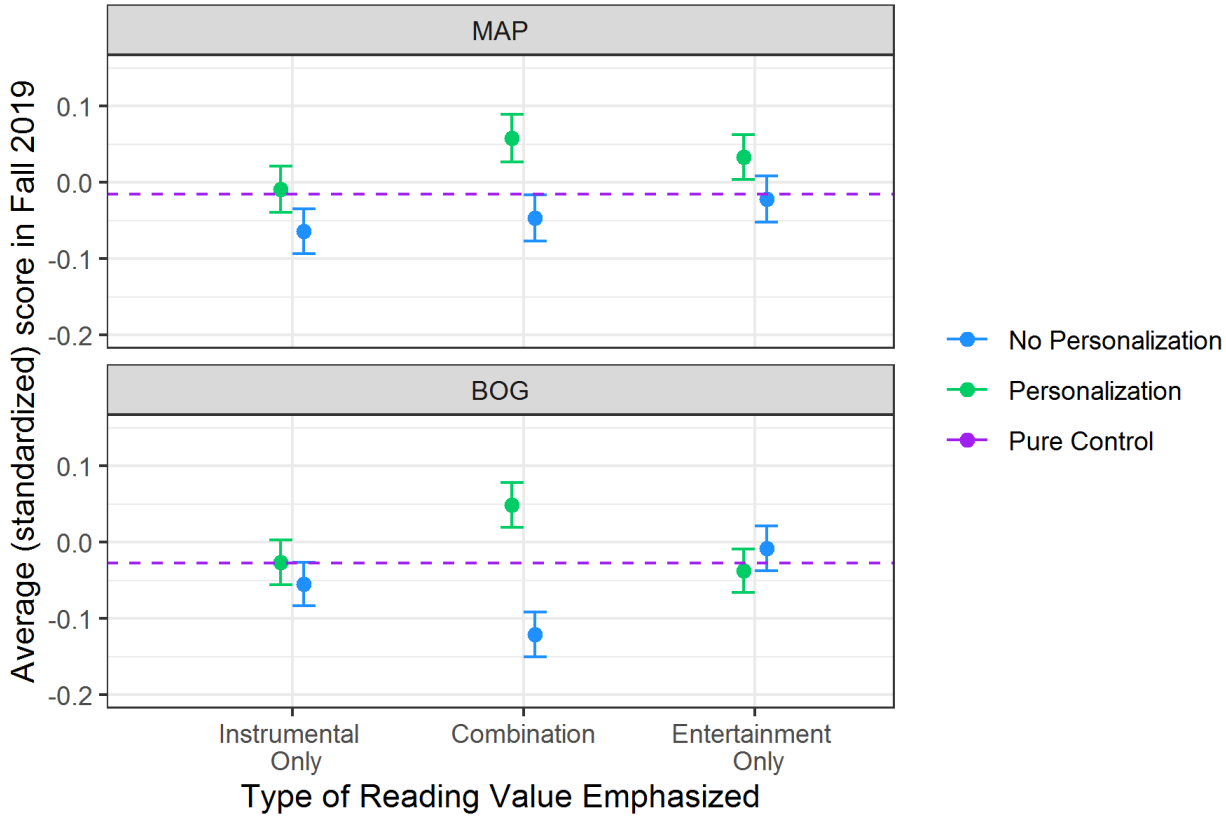


Source: app use data, district administrative records

Notes: Predicted values and standard errors for the six groups are calculated from model parameters and aggregated by the personalization and view-of-reading conditions, collapsing across the values condition. The pure control value is observed directly from the data.

Figure 2

Predicted Test Scores, by Personalization and View-of-Reading Conditions

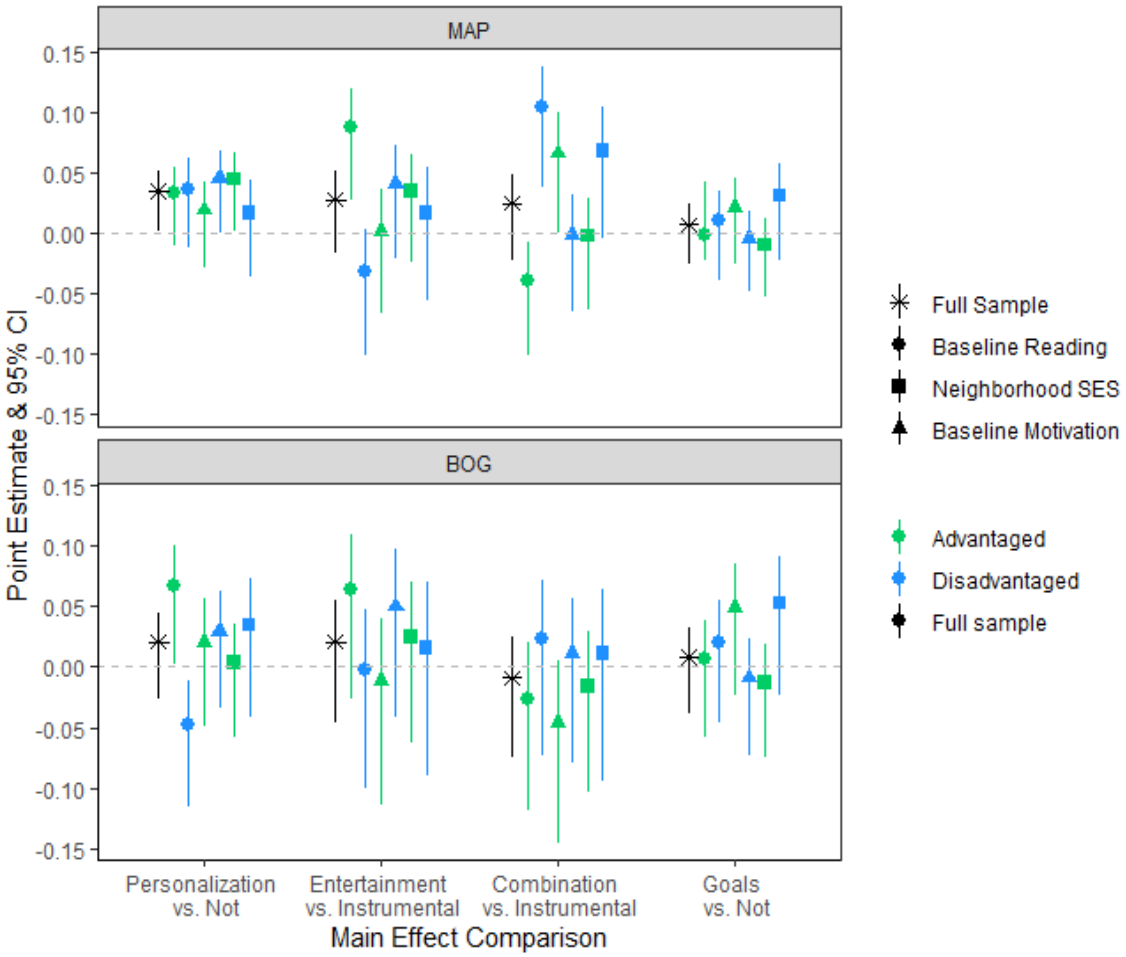


Source: app use data, district administrative records

Notes: Predicted values and standard errors for the six groups are calculated from model parameters and aggregated by the personalization and view-of-reading conditions, collapsing across the values condition. The pure control value is observed directly from the data.

Figure 3

Main Effects on Test Scores, Full Sample and Subgroups

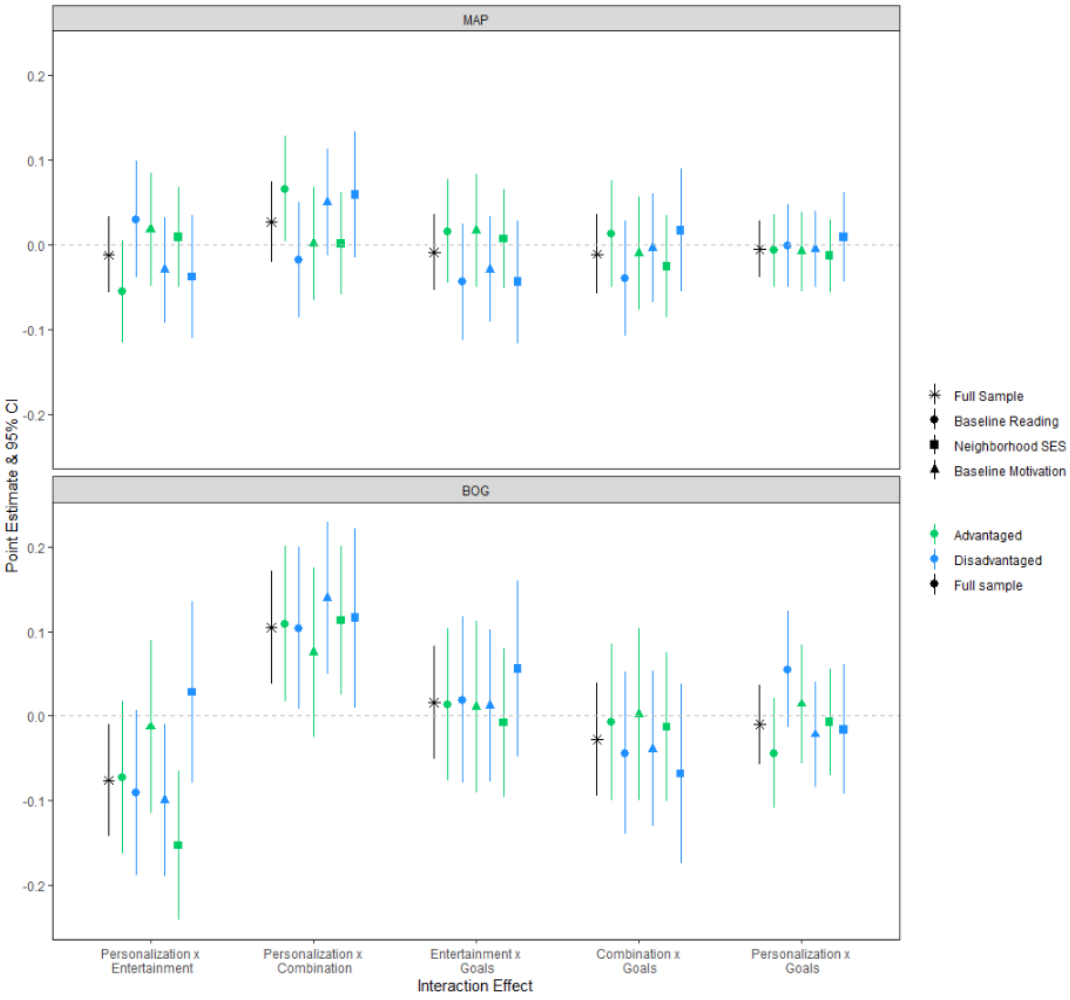


Source: app use data, district administrative records

Notes: Point estimates and confidence intervals presented represent the results from the empirical model fitted separately for each sample.

Figure 4

Interaction Effects on Test Scores, Full Sample and Subgroups



Source: app use data, district administrative records

Notes: Point estimates and confidence intervals presented represent the results from the empirical model fitted separately for each sample.