ANNENBERG
BROWN UNIVERSITY

EdWorkingPaper No. 19-15

Too little or too much? Actionable Advice in an Early-Childhood Text Messaging Experiment

Kalena E. Cortes
Texas A&M University

Hans Fricke
Stanford University

Susanna Loeb
Brown University

David Song
Stanford University

Text-message based parenting programs have proven successful in improving parental engagement and preschoolers' literacy development. The tested programs have provided a combination of (a) general information about important literacy skills, (b) actionable advice (i.e., specific examples of such activities), and (c) encouragement. The regularity of the texts – each week throughout the school year – also provided nudges to focus parents' attention on their children. This study seeks to identify mechanisms of the overall effect of such programs. It investigates whether the actionable advice alone drives previous study's results and whether additional texts of actionable advice improve program effectiveness. The findings provide evidence that text messaging programs can supply too little or too much information. A single text per week is not as effective at improving parenting practices as a set of three texts that also include information and encouragement, but a set of five texts with additional actionable advice is also not as effective as the three-text approach. The results on children's literacy development depend strongly on the child's pre-intervention literacy skills. For children in the lowest quarter of the pre-treatment literacy assessments, only providing one example of an activity decreases literacy scores by 0.15 standard deviations relative to the original intervention. Literacy scores of children in higher quarters are marginally higher with only one tip per week. We find no positive effects of increasing to five texts per week.

VERSION: January 2019

Cortes, Kalena, Hans Fricke, Susanna Loeb, and David Song. (2019). Too little or too much? Actionable Advice in an Early-Childhood Text Messaging Experiment. (EdWorkingPaper: 19-15). Retrieved from Annenberg Institute at Brown University: https://edworkingpapers.com/ai19-15

NBER WORKING PAPER SERIES

TOO LITTLE OR TOO MUCH? ACTIONABLE ADVICE IN AN EARLY-CHILDHOOD TEXT MESSAGING EXPERIMENT

Kalena E. Cortes
Hans Fricke
Susanna Loeb
David S. Song

Working Paper 24827
http://www.nber.org/papers/w24827

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2018

We thank Greg Duncan, Carolyn Heinrich, Tatiana Melquizo, Emily Penner, Lori Taylor, Sherri Widen, and seminar and conference participants at Stanford University, Texas A&M University, the University of California at Irvine, the University of Southern California, Vanderbilt University, and the Association for Education Finance and Policy for helpful feedback. Erika Byun, Stephanie Gullo, and J.B. Horsley provided outstanding research assistance. We are grateful for the financial and other support provided by the Laura and John Arnold Foundation, Knight Foundation, Smith Richardson Foundation, and Spencer Foundation. Any errors are attributable to the authors. Institutional support from Stanford University, Stanford's Center for Education Policy Analysis (CEPA) Labs, and Texas A&M University are also gratefully acknowledged. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Too little or too much? Actionable Advice in an Early-Childhood Text Messaging Experiment
Kalena E. Cortes, Hans Fricke, Susanna Loeb, and David S. Song
NBER Working Paper No. 24827
July 2018
JEL No. I21,I24,J18

ABSTRACT

Text-message based parenting programs have proven successful in improving parental engagement and preschoolers' literacy development. The tested programs have provided a combination of (a) general information about important literacy skills, (b) actionable advice (i.e., specific examples of such activities), and (c) encouragement. The regularity of the texts – each week throughout the school year – also provided nudges to focus parents' attention on their children. This study seeks to identify mechanisms of the overall effect of such programs. It investigates whether the actionable advice alone drives previous study's results and whether additional texts of actionable advice improve program effectiveness. The findings provide evidence that text messaging programs can supply too little or too much information. A single text per week is not as effective at improving parenting practices as a set of three texts that also include information and encouragement, but a set of five texts with additional actionable advice is also not as effective as the three-text approach. The results on children's literacy development depend strongly on the child's pre-intervention literacy skills. For children in the lowest quarter of the pre-treatment literacy assessments, only providing one example of an activity decreases literacy scores by 0.15 standard deviations relative to the original intervention. Literacy scores of children in higher quarters are marginally higher with only one tip per week. We find no positive effects of increasing to five texts per week.

Kalena E. Cortes
The Bush School
Texas A&M University
4220 TAMU
1049 Allen Building
College Station, TX 77843
and IZA
and also NBER
kcortes@tamu.edu

Hans Fricke
Graduate School of Education
Center for Education Policy Analysis
Stanford University
520 Galvez Mall
Stanford, CA 94305
and IZA
hfricke@stanford.edu

Susanna Loeb
Annenberg Institute for School Reform
Brown University
48 Lloyd Ave
Providence, RI 02906
and NBER
susanna_loeb@brown.edu

David S. Song
Graduate School of Education
Center for Education Policy Analysis
Stanford University
520 Galvez Mall
Stanford, CA 94305
dssong09@stanford.edu

I. Introduction

Parents almost invariably aim for their children to succeed in school and beyond, and often are their children's first teacher (Stevenson, Chen, and Uttal, 1990). Yet, many parents struggle to provide the necessary support due to limited resources, lack of information, and behavioral challenges. As a result, early home learning environments of children differ substantially (Bradley, Corwyn, Burchinal, McAdoo, and Coll, 2001).[1] These differences perpetuate discrepancies in educational attainment and professional success later in life (Heckman, 2006). To close learning gaps, a variety of programs has aimed at improving parenting practices. However, many of these parenting programs have shown only limited success, at least in part, due to high demands on parents' time, infrequency, and information delivery that is difficult for parents to operationalize. Some of the more successful programs are costly and difficult to scale (Aos, Lieb, Mayfield, Miller, and Pennucci, 2004; Duncan, Ludwig, and Magnusson, 2010; Karoly, Greenwood, Everingham, Houbé, Rydell et al., 1998).

Text-messaging interventions have emerged as a promising alternative or supplement due to their low cost, the widespread use of mobile phones, and their ease of scalability. These interventions have been shown to positively influence both student and parent outcomes in a wide array of educational settings.[2] In particular, a text-messaging program developed at Stanford University improved parental engagement and children's literacy development by overcoming

[1] For example, Hart and Risley (1995) estimate that children from low-income families hear about 30 million fewer words at the age of four than children from high-income families.

[2] Such interventions have been demonstrated to positively affect school and class attendance of students (Bergman and Chan, 2017; Groot, Sander, Rogers, and Bloomenthal, 2017), the number of course credits earned in high school (Kraft and Rogers, 2015), FASFA completion (Page, Castleman, and Meyer, 2016), and college enrollment rates (Castleman and Page, 2015, 2016). Moreover, these interventions have been particularly effective for children and parents from low-income backgrounds (Bergman, 2015; Castleman and Page, 2015, 2016; Bergman and Chan, 2017).

behavioral barriers to good parenting (Doss, Fahle, Loeb, and York, in press; York, Loeb, and Doss, in press). This texting intervention breaks down the complexities of parenting by providing a combination of *general information* about important literacy skills and parent-child activities (i.e., "FACT" text messages), *actionable advice* with specific examples of parent-child literacy activities (i.e., "TIP" text messages), and *encouragement/reinforcement* (i.e., "GROWTH" text messages).

In this study, we analyze the importance of content and frequency of the text messages in the context of the Stanford parenting program. Specifically, we focus on the provision of actionable advice in the form of concrete examples of early literacy activities. We extend prior findings by answering two questions. First, does the actionable advice (i.e., examples of activities) in the "TIP" message drive previous results or is the addition of general information and encouragement/reinforcement through the "FACT" and "GROWTH" texts more important for increasing parent-child interactions and child development? Second, does the provision of more activities through two additional "TIP" messages further improve parent-child interactions and child development?

We study these two interrelated questions in a randomized experiment. We assign parents of pre-kindergarten children into three experimental groups. The first group of parents only receives one "TIP" message on Wednesdays, henceforth the ***Tip*** program. The second group receives the original program (i.e., "FACT" message on Mondays, "TIP" message on Wednesdays, and "GROWTH" message on Fridays), which we call the ***FTG*** program (or original program). The third group receives the "FACT" message on Mondays, "TIP" messages on Tuesdays, Wednesdays, and Thursdays, and the "GROWTH" message on Fridays (i.e., the original program and two additional "TIP" messages), henceforth the ***FTTTG*** program.

Two prior studies have assessed the effectiveness of the ***FTG*** program. The first of these, York et al. (in press), estimated the average treatment effect compared to a control group and found substantial positive effects of the program on children's early literacy skills. These positive effects were largely driven by children who started the year in the lower half of literacy development. The second paper, Doss et al. (in press), tested whether the content of texts mattered or whether the benefits were driven solely by the reminder of getting a text about parenting. To investigate whether content mattered, the study compared the original ***FTG*** program to one that provided tips that better matched the skill level of the children. The personalized and differentiated program provided more difficult activities to children who demonstrated greater literacy skills on formative assessments in the beginning and middle of the school year. This second paper found that targeting texts based on skills improved results even further, as children who started the year in the bottom or top third of the literacy distribution benefited more from the differentiated program while those who started in the middle third did not.

Thus, to date, we have evidence that the *content* of the tips matters for the effectiveness of the ***FTG*** program. We do not, however, know whether the other elements of the ***FTG*** program are beneficial or whether one "TIP" text per week design provides enough actionable advice. Filling this gap in knowledge can guide a more efficient and effective program design and shed light on the process of parent behavior change that can inform a much broader array of programs. For instance, if the ***Tip*** program has the same effect as the ***FTG*** program on children's early literacy skills, sending one text-message per week instead of three text-messages per week is more efficient and perhaps less burdensome to already busy parents. It also shows that parents do not need information as much implementation help. Alternatively, if the ***FTTTG*** program has a more

positive effect than the ***FTG*** program, program effectiveness can be increased by providing more activities and that parents benefit from a greater quantity of support each week.

When designing the ***FTG*** program, we hypothesized that at least four factors may hinder good parenting practices. The first is *imperfect information* about what skills are important for children to develop.[3] The "FACT" texts provide this information. While parents can still do the suggested activities in the "TIP" texts and benefit from them even without this information, having the information from the "FACT" texts may allow parents to better extrapolate from the suggested activity in the "TIP" text and provide other similar opportunities to their children. Thus, including a "FACT" text could add further benefits. Both the ***FTG*** and the ***FTTTG*** programs include the "FACT", but the ***Tip*** program does not.

The second factor that potentially hinders parenting practices is the *cognitive load* of parenting, specifically, the burden that having to decide what activities to do with a child can place on a parent. Even if parents are knowledgeable about the "right" parenting practices, some parents might be overwhelmed by the complexity of raising a child. The cognitive demand of the multitude of choices to support their child's development involved in every parent-child interaction might lead them to engage sub-optimally with their child.[4] Moreover, the cognitive demand of parenting might be particularly challenging to parents who have financial concerns and other demands on

[3] Recent literature on the provision of information on educational choices provides mixed results. While Avery and Kane (2004) and Grodsky and Jones (2007) find little evidence that lack of information about the costs and benefits associated with higher education can explain differences in college attendance, other studies find that information can influence school choices, student outcomes, and major choices (Fricke, Grogger, and Steinmayr, 2018; Hastings and Weinstein, 2008; Valant and Loeb, 2014). Moreover, studies show that text messages to parents about their child's school absences increase parental involvement in supporting academic work, reduce absences (Rogers and Feller, 2016), and increase credit accumulation in high school (Kraft and Rogers, 2015).

[4] Iyengar and Lepper (2000) show that substantial choice one can lead to inaction. By reducing the selection of jams offered to consumers, they increased purchases significantly. In education, Bettinger, Long, Oreopoulos, and Sanbonmatsu (2012) show that assistance in filling out the complex Free Application for Federal Student Aid (FAFSA) can increase the likelihood of submitting the application and of enrolling in college.

their time.[5] The "TIP" texts aim to lighten the cognitive load by providing a suitable, easy, and fun activity linked to the "FACT" text that the parent can have as a default.[6] The "GROWTH" texts include an activity that supplements the "TIPS" in providing default activities and reducing the cognitive load of finding and choosing and activity. The *optimal number* of tips is unclear. On the one hand, additional "TIP" texts may reduce the cognitive load further by giving parents default activities multiple times. On the other hand, the additional "TIP" texts may exacerbate cognitive demand as they provide more activities to read about and undertake. The ***FTG*** program has one "TIP" plus the "GROWTH" activity, while the ***Tip*** program reduces the number of activities to just one "TIP" and the ***FTTTG*** program increases the number of activities by adding two extra "TIPs."

The third factor is the *delayed gratification* of parenting, which in combination with time-inconsistent preferences, might lead parents to make suboptimal choices (DellaVigna, 2009; Thaler and Sunstein, 2008). Parents might be tempted to engage in activities that make their child happy at the present moment rather than engage in time consuming skill building and school readiness activities, which may be rewarded only in the long run. The "GROWTH" text aims to provide immediate gratification to parents with the encouraging words such as, "*you are doing a good job preparing your child for kindergarten*." Both the original ***FTG*** program and the ***FTTTG*** program contain the "GROWTH" texts and they may provide some benefit over the ***Tip*** program if delayed gratification is a salient factor.

[5] Researchers have shown the existence of the scarcity phenomenon both in the laboratory and in contexts such as farming (Shah, Mullainathan, and Shafir, 2012; Mani, Mullainathan, Shafir, and Zhao, 2013).
[6] Bandura and Schunk (1981) show that such an approach can lead to increases and test scores and student confidence.

Finally, the fourth factor underlying the design of the original ***FTG*** program was that *limited attention* could contribute to suboptimal parenting behavior.[7] While parents want to support the long-term development of their child, parenting requires sustained attention over many years while the immediate demands of life could distract parents from their commitment. Texting programs have proven effective at holding attention in a number of contexts such as learning, smoking cessation, weight loss, taking medicine, and exercise.[8] All three of the programs – ***Tip***, ***FTG***, and ***FTTTG*** – provide reminders to parents each week about parenting. However, there is a risk that more texts become bothersome to parents, leading them to stop paying attention or opt out of receiving the text messages altogether.

The question of whether the ***Tip*** and ***FTTTG*** programs improve upon the ***FTG*** program is an empirical one. On the one hand, the ***FTG*** and ***FTTTG*** programs provide general information and gratification, and may reduce the cognitive demand further with additional activities. On the other hand, additional burden of more texts may reduce the effectiveness of the programs. We find that the original ***FTG*** program has benefits in comparison to the ***Tip*** program, suggesting that the other elements of ***FTG*** are helping parents. Results from a parent survey suggest that providing only one "TIP" decreases self-reported parental engagement by 0.22 to 0.36 standard deviations (henceforth SDs) relative to the ***FTG*** program. However, the additional "TIPs" in the ***FTTTG*** program are also not beneficial for parents and, in some cases, may be detrimental. The negative effects of additional texts are seen in the opt out rates of parents who received the ***FTTTG*** program.

---

[7] For example, Karlan, McConnell, Mullainathan, and Zinman (2010) show that sending regular reminders can improve saving behavior.

[8] See for example, Evans, Wallace, and Snider (2012); Head, Noar, Innarino, and Harrington (2013); Miltello, Kelly, and Melnyk (2011); Pop-Eleches et al. (2011); Siopis, Chey, and Allman-Farinelli (2014); Traxler and Dearden (2005); Vodopivec-Jamsek, de Jongh, Gurol-Urganci, Atun, and Car (2012); and Whittaker et al. (2012).

That is, parents in the ***FTTTG*** program are *more likely* to opt out of the intervention by 2.9 percentage points compared to those receiving the ***FTG*** program (a 58 percent increase over the ***FTG*** program mean opt out of 5 percent), while parents who received the ***Tip*** program were 2.0 percentage points *less likely* to opt out (a 40 percent decline over the ***FTG*** program mean opt out of 5 percent). Furthermore, the effects on child literacy development strongly depend on the child's pre-intervention literacy skills. Specifically, for children in the lowest quarter of the pre-treatment literacy assessments, the ***Tip*** program decreases literacy scores by 0.15 SDs compared to also providing general information as well as encouragement and reinforcement in the ***FTG*** program. The literacy scores of children in the middle two quarters are higher in the ***Tip*** program than in the original program, but these effects are only marginally significant. We find no effects for providing additional examples of activities in the ***FTTTG*** program on children's literacy test scores.

II. Experimental Design

A. The Intervention

The ***FTG*** program is an eight-month-long text messaging pre-kindergarten program for parents of four year olds designed to help them support their children's academic development. The program was first introduced in the San Francisco Unified School District during the 2013-14 school year. The ***Tip*** program and the ***FTTTG*** program are similar to the original program in all aspects except in the number and content of texts per week. We designed them specifically for this study.

The three programs draw on research on literacy development (e.g., Lonigan and Shanahan, 2009), parenting practices (e.g., Reese, Sparks, and Leyva, 2010), and behavior change strategies (e.g., Abraham and Michie, 2008). In developing the scope and sequence of the literacy

curriculum, we consulted a number of state standards for early literacy skills (e.g., Abbot, Lundin, and Ong, 2008; Texas Education Agency, 2015). The texting curriculum is structured as a spiral curriculum – it starts simple and becomes progressively more advanced over the eight months of the intervention, and topics are reintroduced throughout the year for reinforcement. For example, the first week of the program focuses on parent-child conversations, while the last few weeks concentrate on developing high quality parent-child book reading routines.[9] The texts cover a wide range of literacy skills and related parenting practices, including: upper- and lower-case letter recognition, letter sound awareness, beginning sound awareness, rhyme, name writing, concepts of print, story comprehension, vocabulary, singing and listening to songs, self-narration, parent-child conversations, and parent-child book reading routines.

All text messages are couched in positive parenting practices (Parent Management Training: Patterson, Reid, and Dishion, 1992; Incredible Years: Webster-Stratton, 1992; Triple P-Positive Parenting Program: Sanders, Markie-Dadds, Tully, and Bor, 2000; VIPP-SD intervention: Van Zeijl et al., 2006; and Family Check-Up: Gardner, Burton, and Klimes, 2006) with the goal of making the activities fun and engaging for both parent and child. Our goal is for the activities to both increase children's school readiness skills and improve the parent-child relationship through continuous positive interactions. Most of the texts relate to existing family routines and activities. By building on these routines, we offer parents easy to implement choices and minimize costs of adopting beneficial behavior. Parents do not have to create entirely new routines; they merely add an additional step in their established parent-child interactions. For instance, some texts leverage bath time, commuting and travel, or family meals.

[9] See York et al. (in press) for a description of the original text development process.

We assign participating parents into three experimental groups, which differ in content and frequency of texts: ***Tip***, ***FTG*** and ***FTTTG***. The first group of parents receives one "TIP" message on Wednesdays. The second group receives a "FACT" message on Mondays, a "TIP" message on Wednesdays, and a "GROWTH" message on Fridays. The third group receives a "FACT" message on Mondays, a "TIP" message on Tuesdays, Wednesdays, and Thursdays, and a "GROWTH" message on Fridays.

To ensure comparability of treatments, the text message content is similar across treatment groups. The Monday "FACT" texts are the same in both the original and the ***FTTTG*** programs. The ***Tip*** program and the original program send the same "TIP" messages on Wednesdays. However, rearrangement and adaptation of a few texts are necessary in the ***FTTTG*** program in order to achieve a sensible progression of activities. Specifically, the ***FTTTG*** program sends Wednesday's "TIP" messages on Tuesdays, and the example in the "GROWTH" message is used as Wednesday's "TIP" message. Below we provide a literacy example of each treatment group that focuses on letter recognition:[10, 11]

Week Day	***Treatment: Tip***	***Treatment: FTG (original program)***	***Treatment: FTTTG***
Monday		**FACT:** Children need to know letters to learn how to read & write. Research shows that kids with good letter knowledge become good readers.	**FACT:** Children need to know letters to learn how to read & write. Research shows that kids with good letter knowledge become good readers.
Tuesday			**TIP:** Point out the first letter in your child's name in magazines, on signs & at the store. Have your child try.

[10] More examples can be found in York et al. (in press) and in Doss et al. (in press).
[11] On November 16, 2015, all parents in the study received the same welcome text, saying: "*Welcome to Ready4K! Every week, we will send you fun facts & easy tips to help you prepare your child 4K*!", and on June 27, 2016, all parents received the same farewell text, saying: "*READY4K: We hope you enjoyed Ready4K texts! Have a great summer & we'll begin texting you again in the fall*."

			Make it a game. Who can find the most?
Wednesday	**TIP:** Point out the first letter in your child's name in magazines, on signs & at the store. Have your child try. Make it a game. Who can find the most?	**TIP:** Point out the first letter in your child's name in magazines, on signs & at the store. Have your child try. Make it a game. Who can find the most?	**TIP:** Now point out each of the letters in your child's name. After you point to a letter ask: What sound does it make?
Thursday			**TIP:** See if your child can name the letters on common objects like a stop sign. Can s/he name all of the letters on the sign with your street's name?
Friday		**GROWTH:** Keep pointing out letters. You're preparing your child 4K! Point out each of the letters in your child's name. Ask: What sound does it make?	**GROWTH:** Keep pointing out letters to prepare your child 4K! Now have him/her make the letter sounds (ss, tt, oo & pp). What other words have those sounds?

B. Study Participants

We ran this study in partnership with the Dallas Independent School District (ISD). Parents of four-year-old preschoolers in the district received the texting intervention during the 2015-16 school year. The intervention launched in mid-November 2015, and delivered text messages through June 24, 2016.

The Dallas ISD is the second-largest public school district in Texas, and the 14th-largest district in the nation.[12] The district serves approximately 160,000 students in pre-kindergarten through the 12th grade in 224 schools. Of these, approximately 10,000 are pre-kindergarten students in 132 preschools. The Dallas ISD pre-kindergarten serves a diverse and economically disadvantaged student population. The main eligibility criteria for pre-kindergarten enrollment are

[12] The Dallas ISD encompasses the cities of Dallas, Cockrell Hill, Seagoville, Addison, and Wilmer, and parts of Carrollton, Cedar Hill, DeSoto, Duncanville, Farmers Branch, Garland, Grand Prairie, Highland Park, Hutchins, Lancaster and Mesquite.

that children are unable to speak and comprehend the English language or that children are eligible to participate in the National School Lunch Program. The preschool student body consists of approximately 67 percent Hispanic, 28 percent black, and 2 percent white students, with the remaining 3 percent including Asian, American Indian, and mixed-race students. Eighty percent of pre-kindergarten students are economically disadvantaged.[13]

To recruit parents for the study, we built on the district's existing school registration process for pre-kindergarten enrollment. When parents registered their children for preschool – a process that all parents must go through – they were invited to receive text messages and participate in the study. A study participation form, which included active consent and was vetted by the district, was available in both English and Spanish and was included in the district's preschool registration packet. After enrolling in the study, parents were able to receive texts in English or Spanish. Parents could choose to opt out of their program at any time during the school year.[14]

III. Data and Descriptive Statistics

A. Data Sources

In this study, we use information about the children, their parents, and their teachers. Parental information comes from three main sources. First, we obtained preferred texting language, age, and highest educational attainment from our enrollment forms. Second, we gathered opt out information from our texting platform, EZtexting.com. Parents were able to opt out by replying "Stop" or similar words to any text message. We use this opt out information as a measure of

[13] Numbers are based on official DISD enrollment statistics as of 12/15/2017. See https://mydata.dallasisd.org/SL/SD/ENROLLMENT/Enrollment.jsp for more information.

[14] In prior work, we found that nearly 90 percent of economically disadvantaged families had unlimited text messaging plans (York et al. (2017)).

parents' overall experience of the texting program. Third, we surveyed parents after the texting intervention ended at the end of the school year. We collected measures of parent-child engagement, such as reading and literacy activities, and overall satisfaction with the texting program. Parents were invited to participate in the survey by text, email, and regular mail during the months of August through September 2016. We offered parents 20 dollars for completing the survey. Ultimately 664 parents did so. For our analysis, we only consider the 648 parents who answered all questions, a response rate of only 18.6 percent. Though the survey response rate is low, the treatment status did not affect survey participation (see attrition analysis in Section IV). Parents who did not answer the survey are on average less educated and older, and they are less likely to be black and more likely to be Hispanic.[15]

The parent survey included four series of questions. The first addressed the parents' experience with the texting program. For instance, the survey asked if parents received and read the texts, and if parents used the information and found it helpful. It also asked if parents would have liked to have received more or less information. The second series of questions asked about parents' confidence in supporting their child's school readiness skills in literacy and math skills and in improving their child's behavior.[16] The third series of questions assessed the frequency of activities when reading a book to their child such as letting the child turn the pages, talking about pictures, asking questions, and underlining words with the finger. The fourth series of questions assessed the frequency of reading related activities, including among others helping the child to write her name, practicing word sounds, and helping to learn more words.

[15] Results are available upon request.

[16] Studies have suggested that literacy-only curriculum may both increase children's literacy skills and enable them to more readily acquire math skills (e.g., Purpura, Hume, Sims, and Lonigan, 2011). Thus, we also asked parents about supporting their child's math skills.

The child information comes from the Dallas ISD administrative student records. These data include demographic information, such as age, gender, race and ethnicity, and an indicator for low socio-economic status.[17] The data also include our main child literacy outcome measure, the Circle Assessment System (hereafter referred to as Circle). Circle is a one-on-one literacy assessment that takes less than ten minutes per child to complete. All children in this study were assessed with either the English- or Spanish-language version of Circle.[18] Circle assesses language and literacy skills along three distinct dimensions: rapid letter naming, rapid vocabulary naming, phonological awareness. Specifically, the rapid letter naming task measures a child's alphabet knowledge (a one-minute timed assessment task); the rapid vocabulary naming task evaluates a child's ability to name common objects (a one-minute timed assessment task); and the phonological awareness task assesses a child's understanding of sound (approximately five minutes).[19] The phonological awareness is a sum of the following four subtasks:[20] *rhyming* (i.e., the ability to distinguish if two words rhyme when spoken), *alliteration* (i.e., the ability to indicate same beginning sound(s) between two or more words), *syllabication* (i.e., the ability to separate a word into parts), and *onset-rime* (i.e., the ability to blend two parts of a word together when segmented between the beginning consonant(s) and the rest of the word). For Spanish speakers, the phonological awareness assessment only includes rhyming, alliteration, and syllabication.

---

[17] Students are classified as economically disadvantaged if they qualify for free or reduced price lunches or if other district specific criteria apply.

[18] In the Circle-3 test sample 31.23 and 66.54 percent of students were assessed in the English- and Spanish-language version of Circle, respectively. Small percentage of students (i.e., 2.23 percent; 65 students) in the Circle-3 test sample was assessed in both languages. For students who had both English and Spanish test scores, the higher score was used.

[19] In Appendix 1 we provide further details on each assessment, the administration of the assessments, and examples of each assessment.

[20] Only students who took the English-version of Circle were given the onset-rime task of the phonological awareness test. For this reason, we only analyzed three of the four subtasks. We standardized the phonological awareness composite score within language to take into account the fact that the English composite score includes onset-rime task and the Spanish composite score does not.

The Circle assessment is administered three times during the school year: Circle-1 is carried out in beginning of the year (September/October 2015), Circle-2 is carried out in the middle of the year (January/February 2016), and Circle-3 is carried out at the end of the year (April/May 2016). Since the first assessment of Circle occurred before the intervention started, we use Circle-1 results as covariates in all regression specifications. Our main set of child outcomes comes from the third assessment (Circle-3), as parents and children had the most exposure to the program.[21]

The teacher information also comes from the Dallas ISD administrative data. These data include teachers' gender, race/ethnicity, experience in years, and the number of hours they were absent in the school year. For each child, we use mean characteristics of all of their teachers during the school year.

B. Descriptive Statistics

Table 1 presents descriptive statistics on the sample of parents represented in the randomization sample (N=3,473), Circle-3 test sample (N=2,920), and parent survey sample (N=648) by children, parental, and teacher characteristics. As shown in the first column of Table 1, about 11 percent of children in the randomization sample are black, 85 percent are Hispanic, and 2 percent are Asian and white, respectively, and the majority of the sample is of low socioeconomic status (95 percent). The average fall age of children in this sample is 4.7 years. The demographic composition of population of four year olds in the Dallas ISD is similar the randomization sample.

[21] We launched in the Dallas ISD on November 16, 2015 and the intervention ended on June 24, 2016.

Also shown in Table 1 are the child's pre-intervention Circle-1 raw test scores. Children, on average, name 4.8 letters in one minute, identify 10.3 objects in one minute, and get 10.1 items correct in the phonetic awareness assessment. To put these tests into context: the maximum score a child can achieve in rapid letter naming is 52 letters (i.e., 26 lower-case and 26 upper-case letters in the alphabet);[22] the maximum score a child can achieve in rapid vocabulary naming is 55 words; and the maximum composite score a child can achieve in phonological awareness is 28 in total across all four subtasks (i.e., nine in rhyming, seven in alliteration, seven in syllabication, and five in onset-rime). Also, according to the CIRCLE's technical manual a child from the ages of 4 and less than 4.5 as of September 1st should be able to name eight letters in one minute, identify 16 objects in one minute, and get 11 items correct in the phonetic awareness assessment at the beginning of the school year. At the beginning of the intervention, our analytical sample of preschoolers (four year olds) had literacy skills equivalent to average three year old children.[23]

Regarding parental characteristics (shown in panel B), 30 percent have less than a high school degree, 26 percent have a high school degree, and 22 percent have some college or higher. The average fall age of parents in this sample is 31.2 years. About two-thirds of parents, 64 percent, chose to receive texts in Spanish, while 36 percent chose English.

As for average teacher characteristics (shown in panel C), most teachers are female (76 percent), and the average experience is 8.4 years in the district. The racial and ethnic teacher composition in the district differs that of the student population: 19 percent are black and 46 percent are Hispanic, but a higher percentage of teachers were white (32 percent) compared to the

[22] The Spanish alphabet has 27 letters, thus, the maximum score a child can achieve is 54 letters (i.e., 27 lower-case and 27 upper-case letters).
[23] A child older than 3.5 but younger than 4 as of September 1st should be able to name 7 letters in one minute, identify 10 objects in one minute, and get 10 items correct in the phonetic awareness assessment.

student population. A side-by-side comparison of all three samples by child, parent, and teacher characteristics shows similar characteristics by most of these covariates.

IV. Empirical Strategy

A. Estimating Treatment Effects

We estimate the treatment effects of one specific activity in the ***Tip*** program and of two additional activities in the ***FTTTG*** program in comparison to the original texting program, ***FTG***, with the following model specification:

$$y_{is} = \alpha + \beta_1 \cdot Tip_{is} + \beta_2 \cdot FTTTG_{is} + \delta \cdot X_{is} + \gamma_s + \varepsilon_{is}$$

where y_{is} is the outcome of interest of parent (or child) i in pre-school site s. The main parent outcomes are survey measures of parent engagement and program experience, and opt out of the program. The child outcomes are the Circle-3 literacy test scores, standardized within testing language (English or Spanish) to have standard deviation one and mean zero. The variables Tip_{is} and $FTTTG_{is}$ are binary indicators of whether a parent received only one "TIP" message with one specific activity per week or the "FACT", "TIP" , "TIP" , "TIP" , and "GROWTH" messages per week, respectively, in comparison to receiving the original program (omitted category in all regression specifications) with a "FACT", "TIP", and "GROWTH" texts message. X_{is} is a vector of covariates that includes child characteristics (i.e., age, gender, race/ethnicity, low-SES status, and pre-intervention Circle-1 test scores), parental characteristics (i.e., age and highest educational attainment), and lastly, teacher characteristics (i.e., gender, race/ethnicity, years of experience, and hours absent in school year). γ_s are pre-school site fixed effects and ε_{is} is a parent-level (or child-level) error-term. Standard errors are clustered at the pre-school site level. The coefficients of

interest are β_1 and β_2 as they measure the causal effects of receiving ***Tip*** or ***FTTTG*** programs in comparison to original texting intervention, the ***FTG*** program. To explore treatment effect heterogeneity, we also estimate the above model specification in quarters of the child's pre-intervention Circle-1 test.

B. Randomization Checks

In any randomized experiment the only difference that should exist among the experimental groups is, in expectation, the treatment status itself. Any systematic difference of observed and unobserved characteristics between the treatment groups could produce biased estimates of the treatment effects.

To check whether the randomization was successful based on observed characteristics, we estimate a series of pre-school site fixed effects models in order to evaluate covariate balance. These fixed effects models take the following functional form:

$$X_{is} = \alpha + \beta_1 \cdot Tip_{is} + \beta_2 \cdot FTTTG_{is} + \gamma_s + \varepsilon_{is}.$$

We regress the child, parent, and teacher characteristics, X_{is}, on the Tip_{is} and $FTTTG_{is}$ treatment indicators. If the randomization was successful, then the coefficients β_1 and β_2 should be statistically insignificant.

Table 2 presents the randomization checks for all three samples previously shown in Table 1, and illustrates little covariate imbalance. Specifically, of the 62 point estimates in each sample, only six estimates of the randomization sample are statistically significant at the ten percent level or less; only six estimates of the Circle-3 test sample are statistically significant at the ten percent level or less; and only four estimates of the parent survey sample are statistically significant at the ten percent level or less. These numbers are largely what we would expect due to chance.

Furthermore, as shown at the bottom of Table 2 (panel C, last row), we cannot reject the null hypothesis that the coefficients are jointly equal to zero in each of the three samples. Nonetheless, in all regression analyses, we include the full set of covariates.

C. Attrition Analysis

We analyze attrition in both the child outcome data (i.e., Circle-3 test sample) and parent outcome data (i.e., parent survey sample) by testing whether attrition differs by treatment status. If the attrition rate of parents who received the ***Tip*** or ***FTTTG*** programs systematically differ to that of the comparison group (i.e., ***FTG*** program) in a way that is related to our study outcomes, then our treatment effects would be biased. For instance, if lower performing children whose parents received the ***Tip*** program leave the school district at a higher rate than children whose parents received the ***FTG*** program, the treatment effects are likely biased upward.

Thus, to check for selective attrition from the Circle-3 test sample and parent survey sample, we estimate the following pre-school site fixed effect models:

$$A_{is} = \alpha + \beta_1 \cdot Tip_{is} + \beta_2 \cdot FTTTG_{is} + \delta \cdot X_{is} + \gamma_s + \varepsilon_{is}$$

where A_{is} is a binary indicator that equals one (and zero otherwise) if a child (or parent) does not appear in the Circle-3 test sample or parent survey sample.[24] We also control for child, parent, and teacher characteristics, and pre-school site fixed effects.

Table 3 reports the results for the attrition analysis. The first row shows that attrition in the Circle-3 test sample overall does not differ by treatment status (panel A). Both coefficients are close to zero and are statistically insignificant. However, when examining attrition by quarters of

[24] Children are included in the Circle-3 test sample if they have test scores in both Circle-1 and Circle-3 assessments.

Circle-1 literacy scores, the coefficients for the ***Tip*** treatment are marginally statistically significant for the two middle quarters. In the second quarter, the likelihood of missing Circle-3 literacy scores is 4.4 percentage points higher than for those who received the original program. In the third quarter, this probability is 4.7 percentage points lower. The main concern here is whether the treatment status affects attrition. Taking a closer look at these two middle quarters, the effect on the attrition pattern does not appear systematic and may simply be due to noise. That said, given this differential attrition in these two quarters, we assess the robustness of our main results with a bounding analysis in Section V. Our results remain largely unchanged. Panel B of Table 3 shows that neither treatment group is affected by selective attrition into the parent survey. All reported estimates are statistically insignificant.

V. Main Results

A. Results on Parental Program Experience and Engagement

We find evidence that the experience of some parents was tempered by increasing the number of text messages that they received per week. As shown in Table 4 (panel A), for the full sample, parents who received the ***Tip*** program were 2 percentage points less likely to opt out of the program compared to those receiving the original program, ***FTG***. In other words, a 40 percent decline in opting out of the intervention over the base mean opt out of 5 percent in the ***FTG*** program. In line with these results, parents who received the ***FTTTG*** program were about 2.9 percentage points more likely to opt out compared to those receiving the original program (a 58 percent increase in opting out over the base mean opt out of 5 percent in the ***FTG*** program). Furthermore, these results are driven mostly by parents whose children are in the lowest quarter of their pre-intervention literacy assessment. These parents were 4.4 percentage points less likely to

opt out if they were in the ***Tip*** program, and 4.9 percentage points more likely to opt out of the ***FTTTG*** program compared to the ***FTP*** program.

Panel B shows supporting evidence for the opt out results based on the parent survey data. In the survey, we asked parents about their overall experience with the text messages (for instance, how helpful the texts were to them, if they thought they received too few or too many texts, etc.).[25] For the most part, parent's responses to these questions align with the parental opt out data. For instance, the ***FTTTG*** program decreased the frequency of parents reading the texts by 0.25 SDs compared to the original texting program. Interestingly, parents who received only the ***Tip*** program also decreased the frequency of reading the texts by 0.19 SDs compared to the original texting program. These parents also reported that the "ideal" number of texts should be lower than parents in the original program. Overall, while some parents opted out more with three texts than with one texts, parents reported that they like three texts per week the best. Five texts per week increased opt out and was not preferred to three texts per week in the survey reports.

We further find evidence of the benefits of three texts relative to one or five when looking at parent-reported interactions with their child. Table 5 shows the treatment effects on parent's reading and literacy activities, and parental understanding of child development based on the parent survey.[26] As shown across all panels of Table 5, parents in the ***Tip*** program decreased the

[25] Parents could answer the questions: "When you received Ready4K texts, did you READ them?" and "Did you USE the information in Ready4K texts?" these two questions were on a four point Likert scale (never, sometimes, most of the time, always), the question "How HELPFUL was the information in Ready4K texts?" was on a four point Likert scale (not helpful, a little helpful, helpful, very helpful), the question "Was the NUMBER of Ready4K texts that you received each week not enough, too many or just right?" was on a three point Likert scale (not enough, too many, just right), and the question "To what extent would you DISCOURAGE or RECOMMEND Ready4K texts to other parents?" was on a three point Likert scale (discourage, neither discourage nor recommend, recommend). Answers are standardized to have mean zero and standard deviation one.

[26] In Panel A, parents could answer the question: "When you READ to your child, HOW OFTEN do you do the following things?" this question was on a four point Likert scale (never, sometimes, often, always). In Panel B, parents were asked to answer: "Last week, HOW MANY TIMES did you do each of the following READING RELATED

frequency by which they did literacy activities with their child. For example, parents were significantly less likely to report: talking about the pictures in a book; asking their child questions about what is happening in the story; underlining words with their finger as they read to their child; practicing word sounds with their child (e.g., milk starts with "mmm"); and helping their child learn new words. The size of these effects ranges from approximately 0.21 to 0.36 SDs. Parents in the ***Tip*** program also indicated a significantly lower confidence in how to develop their child's literacy skills compared to the original program. The ***FTTTG*** program greatly decreased self-reported frequency by which parents underline words as they read (0.30 SDs) and read to their child (0.19 SDs). All together, these survey results suggest that providing only one actionable activity decreases self-reported parental reading and literacy activities, and providing additional actionable activities does so to a lesser extent.

B. Results on Child's Reading and Literacy Development

Although parents who received the ***Tip*** program were less likely to opt out of the texting intervention, they were less likely to engage in parent-child reading and literacy activities and to understand their child's literacy development. These two effects on parent behavior could have an offsetting impact on their child's literacy development. On the one hand, parents opting to stay in the texting intervention because they find the one tip per week helpful could lead to learning gains for their child. On the other hand, parents' reduced engagement with their child relative to the ***FTG*** program because of the missing buy-in (i.e., FACT text) and reinforcement (i.e., GROWTH text)

---

activities with your child?" this questions was on a four point Likert scale (not at all, once or twice, 3 or 4 times, more than 4 times). In Panel C, parents could answer the question: "How much do you AGREE with each of the following statements? I know what I can do to help my child develop/improve … ?" this questions was on a four point Likert scale (strongly disagree, disagree, agree, strongly agree). All answers were standardized to have mean zero and standard deviation one.

could negatively affect their child's development. Parents in the ***FTTTG*** program were more likely to opt out of the intervention and less likely to engage in reading and literacy activities. In this case, we would expect the changes in these behaviors if anything to reduce their child's literacy development compared to the original program.

Table 6 reports the treatment effects on children's literacy skill development. While we find, on average, no treatment effects of either program compared to the ***FTG*** program (all coefficients are close to zero and statistically insignificant) on children's Circle-3 test, the effect of the texting programs is clearly dependent on the pre-intervention literacy skill distribution of the child. We find that the ***Tip*** program substantially decreased literacy skills for low performing children, while children whose parents received the ***FTTTG*** program did not significantly differ from those in the original program. In particular, for children in the lowest quarter of the pre-treatment Circle-1 literacy assessment, providing parents the ***Tip*** program decreased children's literacy scores by 0.15 SDs compared to the original program. Point estimates suggest that the literacy scores of children in next higher two quarters increased by approximately 0.11 to 0.12 SDs compared to the original texting program, but these effects are only marginally significant. Pooling these two middle quarters, we find a statistically significant effect of 0.11 SDs ($p < 0.001$). The effect of the ***FTTTG*** program are also close to zero and statistically insignificant across all quarters of Circle-1 literacy scores.

The effects on the overall literacy scores reported in Table 6 are most strongly driven by effects on children's rapid vocabulary naming and phonological awareness. Table 7 shows treatment effects of the ***Tip*** and ***FTTTG*** programs for the various sub-tasks of Circle-3 language and literacy skills assessment test: (1) rapid letter naming, (2) rapid vocabulary, and (3) phonological awareness. Phonological awareness in turn can further be disaggregated into: (4)

rhyming, (5) alliteration, and (6) syllabication. We report estimates across quarters of the Circle-1 literacy scores. Estimates for the full sample are close to zero and statistically insignificant.[27] The negative treatment effects of the ***Tip*** program for children in the lowest quarter are driven by a significant decrease in phonological awareness (-0.27 SDs). In particular, rhyming and syllabication decreased by 0.32 and 0.20 SDs, respectively, compared to children's test scores whose parents received the original texting program, ***FTG*** program. The remainder of sub-tasks (i.e., rapid letter naming, rapid vocabulary naming, and alliteration) in the lowest quarter have negative, but small and statistically insignificant point estimates. The literacy gains of children in the higher two middle quarters are driven by rapid vocabulary naming (0.14 SDs), alliteration (0.20 SDs), and syllabication (0.19 SDs). The effects of the ***FTTTG*** program on the subtasks are mostly statistically insignificant with the exception of phonological awareness, which is lower for children in the lowest quarter compared to the original program (-0.20 SDs), and this effect appears to be mostly driven by the effect on rhyming and syllabication (-0.22 to -0.23 SDs).

C. Robustness Checks

In order to further probe our main results, in this section we conduct two robustness checks. First, we provide evidence that different model specifications do not change our results. If randomization is successful, observable and unobservable parent and child characteristics should not differ in expectations between treatment groups. As a result, the inclusion of different covariates in the models should not substantially change the estimated coefficients. All results reported in Tables 4 through 7 are controlling for child, parent, and teacher covariates, as well as

[27] These results are available on request.

pre-school site fixed effects. Table 8 reports again the estimates for the overall child literacy development from Table 6 (shown in model 1), estimates excluding the Circle-1 literacy scores from the set of covariates (shown in model 2), and estimates only conditional on pre-school site fixed effects (shown in model 3). Across all model specifications, the results are very similar.

Second, we assess the robustness of our results to the systematic attrition differences across programs within quarters of the baseline literacy skill distribution. If literacy development of children without Circle-3 test scores differs on average from those children with test scores, our results would be biased. To understand the scope of this problem, we estimate the effects of the programs on the overall literacy score including children with missing scores. We estimate nine different regressions for the full sample and within the second and third quarters – those that exhibited differential attrition rates. In each regression, we impute missing scores with a different percentile (1^{st} to 9^{th} percentile) of the observed distribution in the respective sample. All models include site fixed effects and the full set of covariates.

Figure 1 shows these results. The dotted lines correspond to the baseline effects of the ***Tip*** (light grey) and ***FTTTG*** (dark grey) programs, respectively, not including children with missing scores. The solid blue and dashed red lines correspond to the effects with missing scores imputed at a given percentile of the observed distribution in the respective sample. As shown in the top panel, the results for the ***Tip*** and ***FTTTG*** programs in the full sample are robust, but this is not surprising given that we did not find a significant impact of the treatments on attrition for these samples. However, the estimated effects of the ***Tip*** program on the literacy score in the second and third quarter of the baseline skill distribution are not as robust. Recall, that the ***Tip*** program appeared to have increased the likelihood of missing any Cirlce-3 scores by 4.4 percentage points in the second quarter and decreased the likelihood by 4.7 percentage points in the third quarter. In

the second quarter, the estimated effect of the ***Tip*** program increases assuming higher Circle-3 literacy scores for those children with missing values. When we assign all missing values to have scores at the 5th percentile, the estimates are significantly different from zero. Conversely, in the third quarter, the estimates decrease assuming higher Circle-3 literacy scores. The estimates are only significant at the first three percentiles.

VI. Conclusion

This study analyzes content and frequency of a text messaging program aimed at supporting parental engagement in the literacy development of preschoolers. Our analysis provides three main takeaways. First, parents are more likely to opt out of the intervention as the frequencies of texts increases. This pattern is particularly strong for parents of the lowest performing children. Second, the original program results in greater parent satisfaction and parental engagement in reading and literacy activities than either the ***Tip*** or the ***FTTTG*** programs. Parents who received only one text message per week were the least likely to engage in literacy building activities or to report knowing what to do to build literacy skills. Third, we find no difference in child literacy assessment results by treatment status for the full sample, but for children in the lowest quarter of baseline skills, the original program resulted in greater learning than the ***Tip*** program. For children in the middle half of the distribution, the ***Tip*** program may have resulted in the greater learning.

Overall our findings suggest that text messaging programs can include too little and too much information. The original program of three texts per week was more effective at changing parent-reported behaviors and increasing learning for lower achieving children than the ***Tip*** program. York et al. (in press) estimated that the original program in comparison to a placebo program increased literacy development of children below the median of the base line skill

distribution by 0.31 SDs. We find that a one "TIP" per week in comparison to the original program increases the literacy development of children in the lowest quarter of the base line skill distribution by 0.15 SDs. Taken at face value, these two results suggest that one activity alone may still increase child literacy development by about 0.16 SDs for lower performing children compared to no treatment. No group benefited from five texts per week in the ***FTTTG*** program relative to fewer texts. Moreover, an increased number of texts led to greater program attrition and lower parental engagement.

For parenting programs, as well as for other interventions aimed at changing adult behavior, it is easy to assume that more is better. Recent programs – some but not all using text-messaging to remind and provide information – have shown the light-touch interventions can have large effects, effects that are, in many cases, quite a bit larger than more intensive traditional program. These light-touch programs have often provided easy-to-implement suggestions over extended periods of time. Yet even in these light-touch programs, the balance between too much and too little can be quite salient. This study is the first that we know of to directly test quantity, particularly with a population of low-income adults. The results point clearly to the possibility of too much, even for information that recipients welcome and for aimed goals that they prioritize. For this population, five contacts per week was simply too much for parents and showed no benefit for children. At the lower levels, the differences between one and three were less clear, with parents favoring three but only the lowest performing children benefiting from the additional contacts.

References

Abbott, D., Lundin, J., and Ong, F. (Eds.) (2008). *California Preschool Learning Foundations (Vol. 1)*. Sacramento, CA: California Department of Education.

Abraham, C., and Michie, S. (2008). A taxonomy of behavior change techniques used in interventions. *Health Psychology, 27*(3), pp. 379-387.

Aos, S., Lieb, R., Mayfield, J., Miller, M., and Pennucci, A. (2004). *Benefits and costs of prevention and early intervention programs for youth*. Olympia, WA: Washington State Institute for Public Policy.

Avery, C., and Kane, T.J. (2004). Student perceptions of college opportunities. The Boston COACH Program. In Caroline Hoxby (ed.) *College choices: The economics of where to go, when to go, and how to pay for it*, pp. 355-394, Chicago: The University of Chicago Press.

Bandura, A., and Schunk, D.H. (1981). Cultivating competence, self-efficacy, and intrinsic interest through proximal self-motivation. *Journal of Personality and Social Psychology, 41*(3), pp. 586-598.

Bergman, P. (2015). Parent-Child Information Frictions and Human Capital Investment: Evidence from a Field Experiment. *CESifo Working Paper Series, 5391*. Retrieved from https://ssrn.com/abstract=2622034

Bergman, P., and Chan E.W. (2017). Leveraging Parents through Technology: The Impact of High-Frequency Information on Student Achievement. *CESifo*. Retrieved from https://www.cesifo-group.de/dms/ifodoc/docs/Akad_Conf/CFP_CONF/CFP_CONF_2017/ee17-Hanushek/Papers/ee17_Bergman.pdf

Bettinger, E.P., Long, B.T., Oreopoulos, P., and Sanbonmatsu, L. (2012). The role of application assistance and information in college decisions: Results from the H&R Block FAFSA experiment. *Quarterly Journal of Economics, 127*(3), pp. 1205-1242.

Bradley, R.H., Corwyn, R.F., Burchinal, M., McAdoo, H.P., and Coll, C.G. (2001). The home environments of children in the United States part II: Relations with behavioral development through age thirteen. *Child Development, 72*(6), pp. 1868-1886.

Castleman, B.L., and Page L.C. (2015). Summer nudging: Can personalized text messages and peer mentor outreach increase college going among low-income high school graduates? *Journal of Economic Behavior & Organization, 115*, pp. 144-160.

Castleman, B.L., and Page L.C. (2016). Freshman Year Financial Aid Nudges. *Journal of Human Resources, 51*(2), pp. 389-415.

DellaVigna, S. (2009). Psychology and economics: Evidence from the field. *Journal of Economic Literature, 47*(2), pp. 315-372.

Doss, C., Fahle, E., Loeb, S., and York, B. (in press). Supporting parenting through differentiated and personalized text messaging: Testing effects on learning during kindergarten. *Journal of Human Resources*.

Duncan, G.J., Ludwig, J., and Magnuson, K.A. (2010). Child development. In Phillip B. Levine and David J. Zimmerman (eds.), *Targeting investments in children: Fighting poverty when resources are limited*, pp. 27-58. Chicago: The University of Chicago Press.

Evans, W.D., Wallace, J.L., and Snider, J. (2012). Pilot evaluation of the text4baby mobile health program. *BMC Public Health*, *12*, pp. 1031-1041.

Fricke, H., Grogger, J., and Steinmayr, A. (2018). Exposure to academic fields and major choice. *Economics of Education Review, 64*, pp. 199-213

Gardner, F., Burton, J., and Klimes, I. (2006). Randomized controlled trial of parenting intervention in the voluntary sector for reducing child conduct problems: outcomes and mechanisms of change. *Journal of Child Psychology and Psychiatry*, *47*(11), pp. 1123-1132.

Grodsky, E., and Jones, M.T. (2007). Real and imagined barriers to college entry: *Perceptions of cost. Social Science Research*, *36*(2), pp. 745-766.

Groot, B., Sanders, M., Rogers, T., and Bloomenthal, E. (2017). I get by with a little help from my friends: Two field experiments on social support and attendance in further education colleges in the UK. *The Behavioral Insights Team*. Retrieved from: http://38r8om2xjhhl25mw24492dir.wpengine.netdna-cdn.com/wp-content/uploads/2017/04/Study-Supporter-WP_April-2017.pdf

Hart, B., and Risley, T.R. (1995). *Meaningful differences in the everyday experience of young American children, 1st ed*. Baltimore, MD: Paul H. Brookes Publishing.

Hastings, J.S., and Weinstein, J.M. (2008). Information, school choice, and academic achievement: Evidence from two experiments. *Quarterly Journal of Economics*, *123*(4), pp. 1373-1414.

Head, K.J., Noar, S.M., Iannarino, N.T., and Harrington, N.G., (2013). Efficacy of text messaging-based interventions for health promotion: A meta-analysis. *Social Science and Medicine, 97*, pp. 41-48.

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

Iyengar, S., and Lepper, M. (2000). When choice is demotivating: Can one desire too much of a good thing? *Personality & Social Psychology*, *79*(6), pp. 995-1006.

Karlan, D., McConnell, M., Mullainathan, S., and Zinman, J. (2016). Getting to the top of mind: How reminders increase saving. *Management Science*, *62*(12), pp. 3393-3411.

Karoly, L.A., Greenwood, P.W., Everingham, S.S., Houbé, J., Rydell, C.P., and Chiesa, J. (1998). Investing in our children: What we know and don't know about the costs and benefits of early childhood interventions. Santa Monica, CA: RAND.

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

Lonigan, C.J., and Shanahan, T. (2009). Developing early literacy: Report of the National Early Literacy Panel. *National Institute for Literacy*. Retrieved, December 15, 2005 from https://eric.ed.gov/?id=ED508381

Mani, A., Mullainathan, S., Shafir, E., and Zhao, J. (2013). Poverty Impedes Cognitive Function. *Science*, *341*(6149), pp. 976-980.

Miltello, L.K., Kelly, S.A., and Melnyk, B.M. (2011). Systematic review of text-messaging interventions to promote healthy behaviors in pediatric and adolescent populations: Implications for clinical practice and research. *Worldviews on Evidence-Based Nursing*, *2012*(2), pp. 66-77.

Page, L.C., Castleman, B., and Meyer K. (2016). Customized nudging to improve FAFSA completion and income verification. Working Paper. Retrieved from https://ssrn.com/abstract=2854345

Patterson, G.R., Reid, J.B., and Dishion, T.J. (1992) *Antisocial Boys: A Social Interactional Approach*. Eugene, OR: Castalia.

Pop-Eleches, C., Thirumurthy, H. Habyarimana, J.P., Zivin, J.G., Goldstein, M.P., de Walque, D., and Bangsberg, D.R. (2011). Mobile phone technologies improve adherence to antiretroviral treatment in a resource-limited setting: a randomized controlled trial of text message reminders. *AIDS*, *25*(6), pp. 825-834.

Purpura, D.J., Hume, L.E., Sims, D M., and Lonigan, C.J. (2011). Early literacy and early numeracy: The value of including early literacy skills in the prediction of numeracy development. *Journal of experimental child psychology*, *110*(4), pp. 647-658.

Reese, E., Sparks, A., and Leyva, L. (2010). A review of parent interventions for preschool children's language and emergent literacy. *Journal of Early Childhood Literacy*, *10*(1), pp. 97-117.

Rogers, T., and Feller, A. (2016). Reducing Student Absences at Scale. Working Paper. Retrieved from http://scholar.harvard.edu/files/todd_rogers/files/reducing.pdf

Sanders, M.R., Markie-Dadds, C., Tully, L.A., and Bor, W. (2000). The triple P-positive parenting program: A comparison of enhanced, standard, and self-directed behavioral family intervention for parents of children with early onset conduct problems. *Journal of Consulting and Clinical Psychology*, *68*(4), pp. 624-640.

Shah, A.K., Mullainathan S., and Shafir, E. (2012). The consequences of having too little. *Science*, *338*(6107), pp. 682-685.

Siopis, G., Chey, T., and Allman-Farinelli, M. (2014). A systematic review and meta-analysis of interventions for weight management using text messaging. *Journal of Human Nutrition and Dietetics*, *28*(suppl. 2), pp. 1-15.

Stevenson, H.W., Chen, C., and Uttal, D.H. (1990). Beliefs and achievement: A study of black, white, and Hispanic children. *Child Development*, *61*(2), pp. 508-523.

Texas Education Agency. (2015). *Texas Prekindergarten Guidelines*. Retrieved from https://tea.texas.gov/index2.aspx?id=2147495508

Thaler, R.H., and Sunstein, C.R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. New Haven, CT: Yale University Press.

Traxler, J., and Dearden, P. (2005) The Potential for Using SMS to Support Learning and Organisation in Sub-Saharan Africa. Proceedings of Development Studies Association Conference, Milton Keynes, September 2005.

Valant, J., and Loeb, S. (2014). Information, choice, and decision-making: Field experiments with adult and student school choosers. Working Paper.

Van Zeijl, J., Mesman, J., Van IJzendoorn, M.H., Bakermans-Kranenburg, M.J., Juffer, F., Stolk, and Alink, L.R.A. (2006). Attachment-based intervention for enhancing sensitive discipline in mothers of 1- to 3-year-old children at risk for externalizing behavior problems: A randomized controlled trial. *Journal of Consulting and Clinical Psychology*, *74*(6), pp. 994-1005.

Vodopivec-Jamsek, V., de Jongh, T., Gurol-Urganci, I., Atun, R., and Car, J. (2012). Mobile phone messaging for preventive health care (review). *Cochrane Database of Systematic Reviews*, *2012*(12).

Webster-Stratton, C. (1992). *The Incredible Years: Basic program manual*. Seattle, WA: The Incredible Years.

Whittaker, R., McRobbie, H., Bullen, C., Borland, R., Rodgers, A., and Gu, Y. (2012). Mobile phone-based interventions for smoking cessation (review). *Cochrane Database of Systematic Reviews*, 2012(11).

York, B.N., Loeb, S., and Doss, C. (in press). One Step at a Time: The Effects of an Early Literacy Text Messaging Program for Parents of Preschoolers. *Journal of Human Resources*.

Appendix 1.

CIRCLE Assessment System Subtest Descriptions: Language and Literacy Skills

A. Rapid Letter Naming

This assessment is given to evaluate a student's ability to identify letters of the alphabet.

Directions: Child is given a total of 60 seconds to identify letters that appear on the screen. The student must respond within 3 seconds. If 3 seconds elapse without a response the item is automatically scored as incorrect. A response should be recorded if the child correctly names the letter, if the child says: "I don't know", or if the child provides the incorrect response.

Time to Administer: 1 minute per child

B. Rapid Vocabulary Naming

The Rapid Vocabulary Naming subtest attempts to gain insight into a child's expressive vocabulary skills.

Directions: Child is given a total of 60 seconds to identify pictures as they appear on the screen. There are different pictures for each wave of the assessment. The Rapid Vocabulary Naming assessment includes 2 untimed warm-up items. The teacher conducts a practice session with the warmup items and give feedback for both practice items:

Correct response: "Good job."
Incorrect response: "That was a good try, but this is a ball. Let's try some more. You say 'ball.'

After a picture appears on the screen, the student must respond within 3 seconds. If 3 seconds elapse without a response the item is automatically scored as incorrect. A response should be recorded if the child correctly names the letter, if the child says: "I don't know", or if the child provides the incorrect response.

Time to Administer: 1 minute per child

C. Phonological Awareness

This measure is used to assess a child's understanding of sound in his/her language.

There are four subtests in the Phonological Awareness assessment:

1)	**Rhyming 1:**	Ability to distinguish if two words rhyme when spoken.
2)	**Alliteration:**	Ability to give two or more words that have the same sound(s) at the beginning of the words.
3)	**Syllabication:**	Ability to separate a word into its parts.
4)	**Onset-Rime:**	Ability to blend two parts of a word together when segmented between the beginning consonant(s) and the rest of the word.

General Instructions: The teacher allows a 5 second wait time for the student to respond. Any time longer than 5 seconds is considered a no response. All of the subtests contained within the

Phonological Awareness (PA) subtest include a sample item. This is an auditory assessment and students do not see the teachers screen.

Approximate Time to Administer: 5 minutes per child

1) Rhyming 1

The Rhyming 1 subtest of the PA subtest contains 9 test items that evaluate whether a child can identify whether or not two words rhyme. For the 2015-16 year, these 9 test items were: house-mouse, make-cake, girl-dog, pig-puppy, jump-pump, in-down, sun-sleep, night-light, and mop-top.

Directions: The child repeats each word pair prior to indicating if the words are the same or not. After the teacher records the child's response.

2) Alliteration

The Alliteration subtest is another task that asks children to provide a "yes" or "no" answer to whether or not a pair of words start with the same sound. This subtest contains a sample item and 7 test items. For the 2015-16 year, these 7 test items were: nut-nail, sock-sail, foot-tie, rain-mouse, boat-box, log-light, and kind-glue.

Directions: The child repeats each word pair prior to indicating if the words are the same or not. After the teacher records the child's response.

3) Syllabication

In the Syllabication subtest, children are asked to demonstrate knowledge of how words can be broken down into syllables. There are 7 test items, as well as a sample item. For the 2015-16 year, these 7 test items were: big, ball, wagon, hat, water, candy, and banana.

Directions: The teacher will say a word and clap the word parts and ask the child to say how many parts he/she hears in the word.

4) Onset-Rime

Onset-Rime subtest of the PA subtest includes a sample item and 5 test items. This subtest evaluates one of the key components of phonological processing (i.e., blending) within single syllable words. For the 2015-16 year, these 5 test items were: hat, man, dad, hot, and pig.

Directions: The teacher breaks up a word into sounds. The child repeats the parts and says the word. The teacher then records whether the child's response was correct or incorrect.

Figure 1: Attrition Sensitivity Analysis for Circle-3 Literacy Test Score

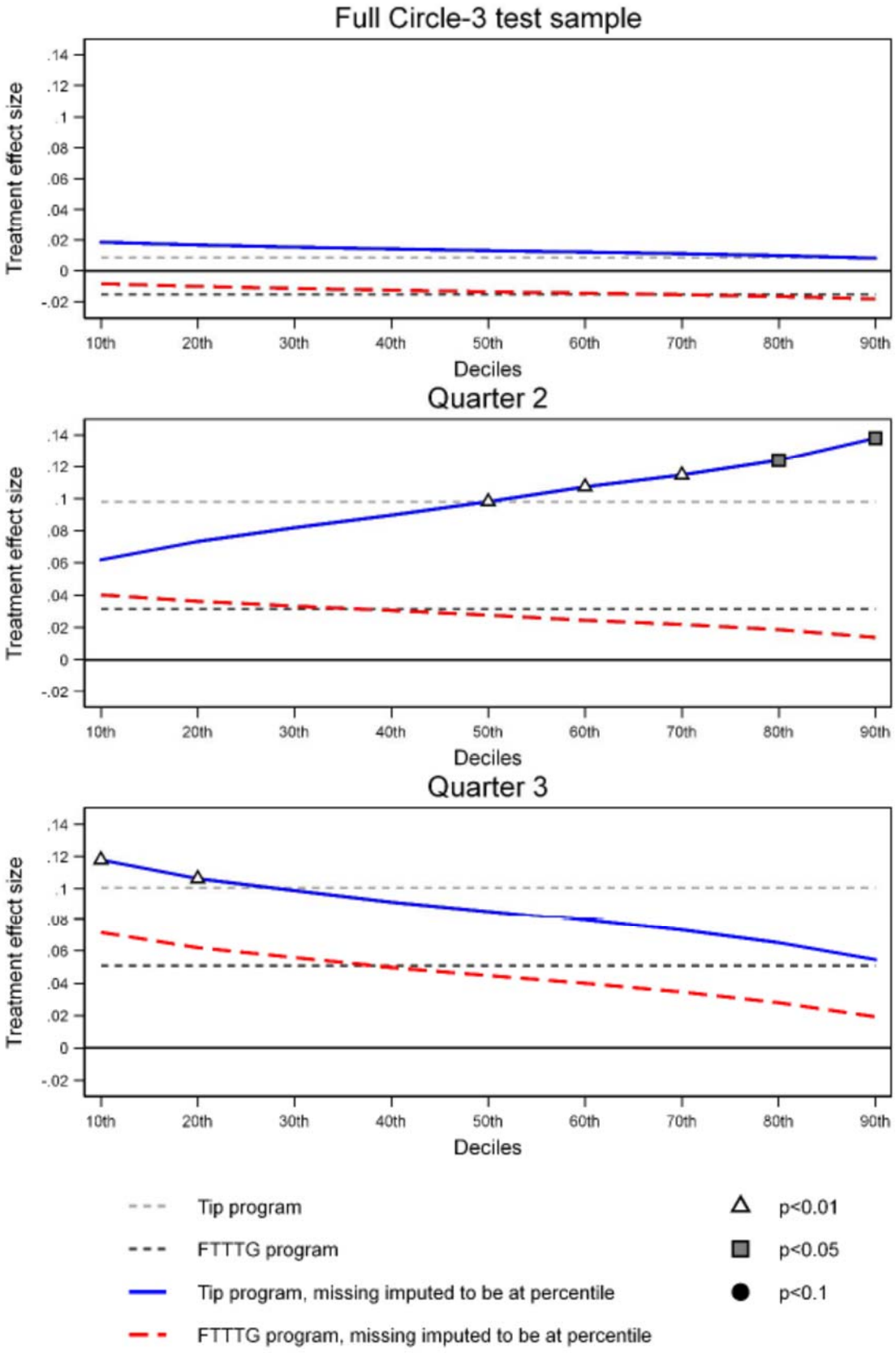

Table 1: Summary Statistics - Means and Standard Deviations

	Randomization Sample	Circle-3 Test Sample*	Parent Survey Sample
Panel A: Children Characteristics			
Age (in years)	4.68	4.71	4.68
	(0.33)	(0.29)	(0.34)
Female	0.50	0.50	0.48
Race and ethnicity			
Black	0.11	0.10	0.13
Hispanic	0.85	0.86	0.83
Asian	0.02	0.02	0.01
White	0.02	0.02	0.02
Other	0.01	0.00	0.01
Low socioeconomic status	0.95	0.95	0.95
Circle-1 assesment (raw) scores[a]			
Rapid letter naming (RLN)	4.77	4.73	4.97
	(7.33)	(7.50)	(7.49)
Missing RLN	0.07	0.00	0.07
Rapid vocabulary (RV)	10.33	10.27	10.82
	(6.74)	(7.00)	(6.98)
Missing RV	0.07	0.00	0.08
Phonetic awareness (PA)	10.13	10.22	10.27
	(4.86)	(5.08)	(4.96)
Missing PA	0.08	0.00	0.09
Panel B: Parental Characteristics			
Age (in years)[a]	31.16	31.27	30.69
	(5.89)	(5.83)	(5.96)
Missing age	0.23	0.24	0.18
Highest education level[b]			
Less than high school	0.30	0.31	0.28
High school	0.26	0.26	0.39
Some college	0.13	0.12	0.19
Associate degree	0.04	0.04	0.06
Bachlelor's degree	0.03	0.03	0.05
Graduate degree	0.02	0.02	0.02
Missing education	0.22	0.23	0.01
Texting language			
English	0.36	0.34	0.38
Spanish	0.64	0.66	0.62
Panel C: Teacher Characteristics			
Average female	0.76	0.76	0.78
Average race and ethnicity			
Black	0.19	0.19	0.20
Hispanic	0.46	0.46	0.44
White	0.32	0.32	0.33
Other	0.03	0.03	0.07
Average missing teacher demographics	0.01	0.01	0.01
Average experience (in years)	8.36	8.49	8.27
	(5.88)	(5.96)	(5.79)
Average hours of absence used[a]	49.64	49.86	48.90
	(33.27)	(34.48)	(33.37)
Average missing hours of absence	0.12	0.12	0.14
Observations	3,473	2,920	648

Notes: Numbers in parentheses are standard deviations for continuous variables. *Circle-3 test sample is conditioned on having Circle-1 test scores. [a]Missing data values imputed to be the mean. [b]Missing data values set to zero.

Table 2: Randomization Checks -
The Effect of Treatment Status on Pre-Treatment Covariates in the Randomization Sample, Circle-3 Test Sample, and Parent Survey Sample

Panel A: Regressions of Child Characteristics on Treatment Status (3 texts per week, FTG program, is the omitted category in all regressions)

	Randomization Sample (N=3,473)		Circle-3 Test Sample* (N=2,920)		Parent Survey Sample (N=648)	
	Treatment Effect Estimates:		Treatment Effect Estimates:		Treatment Effect Estimates:	
	Tip	*FTTTG*	*Tip*	*FTTTG*	*Tip*	*FTTTG*
Age (in years)	-0.011	-0.012	-0.014	-0.003	-0.012	-0.027
	(0.013)	(0.012)	(0.014)	(0.014)	(0.040)	(0.040)
Female	0.025	-0.004	0.036	-0.008	0.048	-0.025
	(0.021)	(0.019)	(0.023)	(0.020)	(0.063)	(0.060)
Black	-0.003	-0.01	0.005	-0.008	-0.005	-0.016
	(0.010)	(0.011)	(0.011)	(0.011)	(0.032)	(0.029)
Hispanic	-0.001	0.005	-0.005	0.007	-0.004	0.000
	(0.012)	(0.012)	(0.012)	(0.012)	(0.037)	(0.035)
Asian	0.000	0.002	0.000	0.000	-0.003	0.004
	(0.003)	(0.005)	(0.004)	(0.006)	(0.010)	(0.010)
White	0.001	0.003	0.000	0.004	-0.003	-0.003
	(0.007)	(0.005)	(0.007)	(0.005)	(0.018)	(0.013)
Other	0.003	0.000	0.001	-0.002	0.016	0.014
	(0.003)	(0.003)	(0.003)	(0.003)	(0.017)	(0.015)
Low socioeconomic status	0.008	0.011	0.013	0.018**	0.046*	0.046*
	(0.008)	(0.008)	(0.009)	(0.009)	(0.027)	(0.024)
Circle-1 rapid letter naming (RLN)[a,b]	0.042	0.061	0.045	0.041	-0.007	0.112
	(0.041)	(0.041)	(0.047)	(0.046)	(0.107)	(0.123)
Circle-1 RLN Missing[c]	-0.002	-0.010	---	---	-0.038	-0.004
	(0.010)	(0.011)			(0.029)	(0.019)
Circle-1 rapid vocabulary (RV)[a,b]	0.061	0.025	0.058	0.018	-0.048	0.199**
	(0.041)	(0.033)	(0.047)	(0.039)	(0.097)	(0.099)
Circle-1 RV Missing[c]	-0.001	-0.009	---	---	-0.03	-0.008
	(0.010)	(0.011)			(0.033)	(0.022)
Circle-1 phonological awareness (PA)[a,b]	0.058	0.015	0.054	-0.001	-0.094	0.047
	(0.039)	(0.034)	(0.045)	(0.039)	(0.090)	(0.100)
Circle-1 PA Missing[c]	0.000	-0.009	---	---	-0.02	-0.007
	(0.010)	(0.011)			(0.034)	(0.020)

Notes: Each row represents a separate regression model (only the coefficients of the treatments status are reported). All regressions include pre-school site fixed effects. Standard errors are clustered at the pre-school site level. The omitted reference group in all regressions is the original texting program of 3 texts per week (i.e., *FTG* program). *Circle-3 test sample is conditioned on having Circle-1 test scores. [a]Missing data values imputed to be the mean and all regressions include a dummy variable for categorical variables with missing values. [b]All Circle-1 test score variables are in standard deviation units. [c]Due to low outcome values in the dependent variable we are unable to run balance checks. Statistical significance levels: *p<0.10; **p<0.05; ***p<0.01.

Table 2 (Continued): Randomization Checks -
The Effect of Treatment Status on Pre-Treatment Covariates in the Randomization Sample, Circle-3 Test Sample, and Parent Survey Sample

Panel B: Regressions of Parental Characteristics on Treatment Status (3 texts per week, FTG program, is the omitted category in all regressions)

	Randomization Sample (N=3,473)		Circle-3 Test Sample* (N=2,920)		Parent Survey Sample (N=648)	
	Treatment Effect Estimates:		Treatment Effect Estimates:		Treatment Effect Estimates:	
	Tip	*FTTTG*	*Tip*	*FTTTG*	*Tip*	*FTTTG*
Age (in years)	-0.421*	-0.475*	-0.192	-0.517*	0.718	-0.221
	(0.243)	(0.255)	(0.275)	(0.273)	(0.621)	(0.640)
Missing age	-0.001	0.000	0.004	0.008	-0.052	-0.032
	(0.017)	(0.015)	(0.018)	(0.018)	(0.045)	(0.038)
Less than high school	0.038**	0.024	0.034*	0.021	0.045	0.013
	(0.019)	(0.018)	(0.020)	(0.020)	(0.044)	(0.046)
High school	-0.029*	-0.033*	-0.035*	-0.035*	-0.016	0.016
	(0.016)	(0.017)	(0.018)	(0.019)	(0.057)	(0.052)
Some college	-0.011	0.001	-0.011	-0.001	-0.043	0.005
	(0.012)	(0.015)	(0.014)	(0.017)	(0.046)	(0.047)
Associate degree	-0.002	0.002	0.001	0.002	-0.006	-0.045*
	(0.008)	(0.008)	(0.009)	(0.008)	(0.032)	(0.026)
Bachlelor's degree	-0.002	-0.006	0.002	-0.002	0.012	0.007
	(0.007)	(0.006)	(0.008)	(0.007)	(0.027)	(0.024)
Graduate degree	-0.005	-0.001	-0.005	-0.002	0.008	0.004
	(0.006)	(0.006)	(0.006)	(0.006)	(0.016)	(0.016)
Missing education	0.012	0.014	0.013	0.017	0.001	0.000
	(0.015)	(0.014)	(0.017)	(0.016)	(0.008)	(0.007)

Notes: Each row represents a separate regression model (only the coefficients of the treatments status are reported). All regressions include pre-school site fixed effects. Standard errors are clustered at the pre-school site level. The omitted reference group in all regressions is the original texting program of 3 texts per week (i.e., *FTG* program). *Circle-3 test sample is conditioned on having Circle-1 test scores. [a]Missing data values imputed to be the mean and all regressions include a dummy variable for categorical variables with missing values. [b]All Circle-1 test score variables are in standard deviation units. [c]Due to low outcome values in the dependent variable we are unable to run balance checks. Statistical significance levels: *p<0.10; **p<0.05; ***p<0.01.

Table 2 (Continued): Randomization Checks -
The Effect of Treatment Status on Pre-Treatment Covariates in the Randomization Sample, Circle-3 Test Sample, and Parent Survey Sample

Panel C: Regressions of Teacher Characteristics on Treatment Status (3 texts per week, FTG program, is the omitted category in all regressions)

	Randomization Sample (N=3,473)		Circle-3 Test Sample* (N=2,920)		Parent Survey Sample (N=648)	
	Treatment Effect Estimates:		Treatment Effect Estimates:		Treatment Effect Estimates:	
	Tip	*FTTTG*	*Tip*	*FTTTG*	*Tip*	*FTTTG*
Average female	-0.005	-0.002	-0.001	0	0.01	0.006
	(0.008)	(0.007)	(0.008)	(0.007)	(0.020)	(0.026)
Average black	0.020**	0.001	0.017**	-0.003	0.02	-0.023
	(0.008)	(0.007)	(0.008)	(0.008)	(0.023)	(0.021)
Average Hispanic	-0.009	0.011	0.000	0.018	-0.001	0.024
	(0.011)	(0.011)	(0.011)	(0.011)	(0.024)	(0.025)
Average white	-0.007	-0.01	-0.011	-0.011	-0.006	-0.005
	(0.010)	(0.008)	(0.010)	(0.008)	(0.023)	(0.021)
Average missing teacher demographic data	0.000	0.001	0.000	0.000	-0.001	0.002
	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)	(0.002)
Average experience (in years)	-0.021	-0.215	-0.061	-0.194	0.273	-0.509
	(0.206)	(0.167)	(0.214)	(0.172)	(0.582)	(0.543)
Average hours of absence used[a]	-0.252	0.438	0.62	1.004	1.532	-1.756
	(1.001)	(0.935)	(1.114)	(1.050)	(2.269)	(2.178)
Average missing hours of absence	0.003	0.002	0.003	0.004	0.006	0.008
	(0.004)	(0.005)	(0.005)	(0.005)	(0.010)	(0.014)
SUR test for joint significance (p-value)	0.170	0.687	0.238	0.436	0.144	0.225

Notes: Each row represents a separate regression model (only the coefficients of the treatments status are reported). All regressions include pre-school site fixed effects. Standard errors are clustered at the pre-school site level. The omitted reference group in all regressions is the original texting program of 3 texts per week (i.e., *FTG* program). *Circle-3 test sample is conditioned on having Circle-1 test scores. [a]Missing data values imputed to be the mean and all regressions include a dummy variable for categorical variables with missing values. [b]All Circle-1 test score variables are in standard deviation units. [c]Due to low outcome values in the dependent variable we are unable to run balance checks. Statistical significance levels: *p<0.10; **p<0.05; ***p<0.01.

Table 3: Treatment Effects on Attrition in Circle-3 Test Sample and Parent Survey Sample

Panel A: Circle-3 Test Sample - Whether any Circle-3 Test Outcomes are Missing

	Tip	*FTTTG*	Adj. R^2	N
Full sample	-0.004	-0.006	0.59	3,473
	(0.010)	(0.009)		
By quartiles:[a]				
Quarter 1 (lowest)	-0.028	-0.017	0.21	780
	(0.025)	(0.027)		
Quarter 2	0.044*	0.002	0.32	778
	(0.022)	(0.017)		
Quarter 3	-0.047*	-0.029	0.12	779
	(0.024)	(0.026)		
Quarter 4 (highest)	0.011	0.023	0.01	777
	(0.019)	(0.020)		

Panel B: Parent Survey Sample - Whether any Parent Survey Outcomes are Missing

	Tip	*FTTTG*	Adj. R^2	N
Entire sample	0.012	0.000	0.10	3,473
	(0.015)	(0.016)		

Notes: Each row represents a separate regression model (only the coefficients of the treatments status are reported). All regressions include controls for student characteristics (age, gender, race/ethnicity, low-SES status, and Circle-1 test scores), parental characteristics (age and highest educational attainment), teacher characteristics (gender, race/ethnicity, years of experience, and hours of absence used), and pre-school site fixed effects. Standard errors are clustered at the pre-school site level. The omitted reference group in all regressions is the original texting program of 3 texts per week (i.e., *FTG* program). [a]Quarters are based on student's Circle-1 tests prior to the intervention. Statistical significance levels: *p<0.10; **p<0.05; ***p<0.01.

Table 4: Treatment Effects on
Parents' Opting-out of the Intervention and Overall View of Receiving the Weekly Texts

Panel A: Parental Opt Out Rates

	Tip	*FTTTG*	Adj. R^2	N
Full sample	-0.020**	0.029***	0.03	2,920
	(0.008)	(0.010)		
By quartiles:[a]				
Quarter 1 (lowest)	-0.044**	0.049**	0.09	715
	(0.019)	(0.020)		
Quarter 2	-0.011	0.01	-0.04	728
	(0.020)	(0.024)		
Quarter 3	-0.038*	0.003	-0.01	735
	(0.022)	(0.026)		
Quarter 4 (highest)	-0.019	0.032	-0.01	742
	(0.024)	(0.026)		

Panel B: Parental Survey Responses to Receiving the Weekly Texts

	Tip	*FTTTG*	Adj. R^2	N
Received texts (0/1)	-0.022	-0.013	0.10	648
	(0.022)	(0.016)		
Read texts (std.)	-0.188*	-0.250**	-0.04	648
	(0.111)	(0.119)		
Uses texts (std.)	-0.134	-0.169	-0.02	648
	(0.127)	(0.128)		
Texts are helpful (std.)	-0.134	0.070	0.05	648
	(0.122)	(0.112)		
Receive too many texts (0/1)	-0.021	0.038	0.02	648
	(0.027)	(0.028)		
Not enough texts (0/1)	0.041	0.016	0.03	648
	(0.026)	(0.025)		
Ideal # of texts (std.)	-0.547***	-0.158	0.03	648
	(0.120)	(0.111)		
Recommend texts (std.)	-0.132	-0.043	-0.01	648
	(0.111)	(0.125)		

Notes: Each row represents a separate regression model (only the coefficients of the treatments status are reported). All regressions include controls for student characteristics (age, gender, race/ethnicity, low-SES status, and Circle-1 test scores), parental characteristics (age and highest educational attainment), teacher characteristics (gender, race/ethnicity, years of experience, and hours of absence used), and pre-school site fixed effects. Standard errors are clustered at the pre-school site level. The omitted reference group in all regressions is the original texting program of 3 texts per week (i.e., *FTG* program). [a]Quarters are based on student's Circle-1 tests prior to the intervention. Statistical significance levels: *p<0.10; **p<0.05; ***p<0.01.

Table 5: Treatment Effects on Parental Reading Activities, Literacy Activities, and Understanding of Child Development

Panel A: Frequency of How Often Parents Did Reading Activities with their Child (standardized)

	Tip	*FTTTG*	Adj. R^2	N
Mean Reading Activity	-0.202	-0.156	0.06	648
	(0.127)	(0.112)		
Let child hold the book	-0.052	-0.160	0.03	648
	(0.107)	(0.117)		
Show child book parts	-0.073	-0.142	0.04	648
	(0.113)	(0.119)		
Talk about the pictures in a book	-0.218*	-0.046	0.01	648
	(0.126)	(0.108)		
Ask child questions about the story	-0.230*	-0.084	0.02	648
	(0.119)	(0.105)		
Underline words as you read	-0.269**	-0.299**	0.08	648
	(0.106)	(0.118)		
Show child that we read from left to right	-0.112	-0.115	0.00	648
	(0.127)	(0.118)		
Ask child questions about the story after you read it	-0.074	0.066	0.04	648
	(0.129)	(0.103)		

Panel B: Frequency of How Often Parents Did Literacy Activities with their Child (standardized)

	Tip	*FTTTG*	Adj. R^2	N
Mean Literacy Activity	-0.281**	-0.144	0.06	648
	(0.130)	(0.122)		
Helped child write his/her name	-0.021	-0.041	-0.04	648
	(0.126)	(0.131)		
Pointed out letters on objects	-0.195	-0.110	0.01	648
	(0.126)	(0.128)		
Sang a song or nursery rhyme with child	-0.107	-0.099	-0.02	648
	(0.126)	(0.130)		
Practiced word sounds	-0.309**	-0.071	0.03	648
	(0.119)	(0.125)		
Said rhyming words with child	-0.188	-0.101	0.04	648
	(0.120)	(0.123)		
Asked child about his or her day	-0.357***	-0.203	0.03	648
	(0.122)	(0.123)		
Played a literacy learning game with child	-0.148	-0.041	0.04	648
	(0.118)	(0.108)		
Helped child learn new words	-0.245*	-0.023	0.07	648
	(0.126)	(0.119)		
Worked on reading skills during family activities	-0.293**	-0.190	0.10	648
	(0.118)	(0.120)		
Helped child sound out a word	-0.147	-0.074	0.05	648
	(0.130)	(0.122)		
Read to child	-0.211*	-0.188*	0.04	648
	(0.119)	(0.110)		

Panel C: Parents' Understanding of How to Develop Skills for their Child (standardized)

	Tip	*FTTTG*	Adj. R^2	N
Mean Skill Development	-0.219*	0.034	0.11	648
	(0.115)	(0.101)		
Literacy development	-0.210*	0.009	0.06	648
	(0.112)	(0.115)		
Math development	-0.184	0.042	0.10	648
	(0.126)	(0.107)		
Behavioral development	-0.185	0.038	0.11	648
	(0.113)	(0.097)		

Notes: Each row represents a separate regression model (only the coefficients of the treatments status are reported). All regressions include controls for student characteristics (age, gender, race/ethnicity, low-SES status, and Circle-1 test scores), parental characteristics (age and highest educational attainment), teacher characteristics (gender, race/ethnicity, years of experience, and hours of absence used), and pre-school site fixed effects. Standard errors are clustered at the pre-school site level. The omitted reference group in all regressions is the original texting program of 3 texts per week (i.e., *FTG* program). Statistical significance levels: *p<0.10; **p<0.05; ***p<0.01.

Table 6: Treatment Effects on
Overall Circle-3 Language and Literacy Skills Assesment Test (standardized)

	Tip	*FTTTG*	Adj. R^2	N
Full sample	0.016	-0.016	0.41	2,920
	(0.032)	(0.033)		
By quartiles:[a]				
Quarter 1 (lowest)	-0.152**	-0.054	0.44	715
	(0.073)	(0.072)		
Quarter 2	0.118*	0.049	0.40	728
	(0.060)	(0.071)		
Quarter 3	0.111*	0.050	0.35	735
	(0.063)	(0.060)		
Quarter 4 (highest)	0.002	-0.062	0.43	742
	(0.052)	(0.048)		

Notes: Each row represents a separate regression model (only the coefficients of the treatments status are reported). All regressions include controls for student characteristics (age, gender, race/ethnicity, low-SES status, and Circle-1 test scores), parental characteristics (age and highest educational attainment), teacher characteristics (gender, race/ethnicity, years of experience, and hours of absence used), and pre-school site fixed effects. Standard errors are clustered at the pre-school site level. The omitted reference group in all regressions is the original texting program of 3 texts per week (i.e., *FTG* program). [a]Quarters are based on student's Circle-1 tests prior to the intervention. Statistical significance levels: *p<0.10; **p<0.05; ***p<0.01.

Table 7: Treatment Effects on Circle-3 Language and Literacy Skills Assesment Test by Specific Sub-tests (standardized)

1) Rapid Letter Naming		*Tip*	*FTTTG*	Adj. R^2	N
	Quarter 1 (lowest)	-0.077	0.060	0.35	715
		(0.092)	(0.092)		
	Quarter 2	0.087	-0.044	0.37	728
		(0.084)	(0.075)		
	Quarter 3	0.086	0.049	0.32	735
		(0.085)	(0.085)		
	Quarter 4 (highest)	0.016	-0.065	0.33	742
		(0.074)	(0.073)		
2) Rapid Vocabulary Naming		*Tip*	*FTTTG*	Adj. R^2	N
	Quarter 1 (lowest)	-0.107	-0.018	0.47	715
		(0.086)	(0.078)		
	Quarter 2	0.142**	0.154*	0.44	728
		(0.065)	(0.081)		
	Quarter 3	0.068	-0.004	0.38	735
		(0.072)	(0.069)		
	Quarter 4 (highest)	0.029	-0.019	0.41	742
		(0.085)	(0.072)		
3) Phonological Awareness [a]		*Tip*	*FTTTG*	Adj. R^2	N
	Quarter 1 (lowest)	-0.270***	-0.204*	0.30	715
		(0.093)	(0.108)		
	Quarter 2	0.124	0.036	0.27	728
		(0.084)	(0.100)		
	Quarter 3	0.179**	0.106	0.23	735
		(0.090)	(0.080)		
	Quarter 4 (highest)	-0.038	-0.103	0.25	742
		(0.068)	(0.067)		
4) Rhyming		*Tip*	*FTTTG*	Adj. R^2	N
	Quarter 1 (lowest)	-0.322***	-0.222*	0.16	715
		(0.108)	(0.127)		
	Quarter 2	0.093	-0.034	0.15	728
		(0.105)	(0.102)		
	Quarter 3	0.041	0.044	0.15	735
		(0.108)	(0.086)		
	Quarter 4 (highest)	-0.060	-0.151*	0.13	742
		(0.083)	(0.086)		
5) Alliteration		*Tip*	*FTTTG*	Adj. R^2	N
	Quarter 1 (lowest)	-0.103	-0.028	0.23	715
		(0.095)	(0.107)		
	Quarter 2	0.196**	0.101	0.20	728
		(0.088)	(0.104)		
	Quarter 3	0.162	0.132	0.17	735
		(0.101)	(0.085)		
	Quarter 4 (highest)	0.003	-0.045	0.20	742
		(0.070)	(0.073)		
6) Syllabication		*Tip*	*FTTTG*	Adj. R^2	N
	Quarter 1 (lowest)	-0.200*	-0.229*	0.23	715
		(0.105)	(0.126)		
	Quarter 2	-0.014	-0.033	0.18	728
		(0.090)	(0.108)		
	Quarter 3	0.185**	0.057	0.12	735
		(0.083)	(0.089)		
	Quarter 4 (highest)	-0.018	-0.030	0.14	742
		(0.073)	(0.081)		

Notes: All regressions include controls for student characteristics (age, gender, race/ethnicity, and low-SES status), parental characteristics (age and highest educational attainment), teacher characteristics (gender, race/ethnicity, years of experience, and hours of absence used), and pre-school site fixed effects. Standard errors are clustered at the pre-school site level. The omitted reference group in all regressions is the original texting program of 3 texts per week (i.e., *FTG* program). Quarters are based on student's Circle-1 tests prior to the intervention. aPhonological awareness is a composite score of the following assessments: rhyming, alliteration, and syllabication. Statistical significance levels: *p<0.10; **p<0.05; ***p<0.01.

Table 8: Robustness Check for Table 6 Results -
Comparison of Regression Models for Treatment Effects on Overall Circle-3 Test (standardized)

	Model 1: Table 6 Results	Model 2:	Model 3:
	Pre-Kindergarten site fixed effects with full set of controls[1], including initial Circle-1 assesment scores	Pre-Kindergarten site fixed effects with only full set of controls[1]	Pre-Kindergarten site fixed effects and no controls[1]
Panel A: Full sample			
Tip	0.016	0.045	0.04
	(0.032)	(0.035)	(0.037)
FTTTG	-0.016	0.003	-0.009
	(0.033)	(0.038)	(0.039)
N	2,920	2,920	2,920
Adj. R^2	0.41	0.25	0.21
Panel B: By quartiles [a]			
		Quarter 1 (lowest)	
Tip	-0.152**	-0.178**	-0.183**
	(0.073)	(0.072)	(0.085)
FTTTG	-0.054	-0.076	-0.101
	(0.072)	(0.071)	(0.079)
N	715	715	715
Adj. R^2	0.44	0.39	0.35
		Quarter 2	
Tip	0.118*	0.136**	0.147**
	(0.060)	(0.062)	(0.068)
FTTTG	0.049	0.027	0.018
	(0.071)	(0.078)	(0.076)
N	728	728	728
Adj. R^2	0.40	0.31	0.27
		Quarter 3	
Tip	0.111*	0.092	0.098
	(0.063)	(0.067)	(0.065)
FTTTG	0.05	0.061	0.045
	(0.060)	(0.064)	(0.065)
N	735	735	735
Adj. R^2	0.35	0.29	0.28
		Quarter 4 (highest)	
Tip	0.002	0.018	0.028
	(0.052)	(0.055)	(0.053)
FTTTG	-0.062	-0.027	-0.045
	(0.048)	(0.052)	(0.049)
N	742	742	742
Adj. R^2	0.43	0.32	0.30

Notes: [1]Full set of controls include: student characteristics (age, gender, race/ethnicity, and low-SES status); parental characteristics (age and highest educational attainment); and teacher characteristics (gender, race/ethnicity, years of experience, and hours of absence used). Standard errors are clustered at the pre-school site level. The omitted reference group in all regressions is the original texting program of 3 texts per week (i.e., *FTG* program). [a]Quarters are based on student's Circle-1 tests prior to the intervention. Statistical significance levels: *p<0.10; **p<0.05; ***p<0.01.