www.sciencemag.org/content/350/6257/196/suppl/DC1

Supplementary Materials for

Math at home adds up to achievement in school

Talia Berkowitz, Marjorie W. Schaeffer, Erin A. Maloney,
Lori Peterson, Courtney Gregor, Susan C. Levine,* Sian L. Beilock*

*Corresponding author. E-mail: beilock@uchicago.edu (S.L.B.);
s-levine@uchicago.edu (S.C.L.)

Published 9 October 2015, *Science* **350**, 196 (2015)
DOI: 10.1126/science.aac7427

This PDF file includes:

Materials and Methods
Supplementary Text
Fig. S1
Table S1
Models S1 to S8

Other Supplementary Material for this manuscript includes the following:
(available at www.sciencemag.org/cgi/content/full/350/6257/196/DC1)

Accessory Data File S1 (as Excel file)

Materials and Methods

Participants

Children (N = 587; 287 males, 300 females) and their families were recruited from 22 schools across the Chicagoland area. 69.4% of the sample was from middle- to upper-middle class families, living in households making over $50,000 per year. Children ranged in age from 53 to 92 months, with an average age of 78.5 months. 83.9% of parents who filled out the primary caregiver survey were the child's mother (children's fathers made up 14.8% of survey respondents; the remaining 1.3% of respondents were grandparents or aunts/uncles).

Procedure

Families were recruited through schools where principals and 1st grade teachers agreed to participate in the study. We recruited schools throughout the Chicagoland area, including public, private and parochial schools. Research assistants attended back-to-school nights to inform parents about the study and to distribute consent forms. Schools were very interested in students participating in this study, and many more participants consented than we expected. We thus stopped accepting new participants once we surpassed our budget for iPad Minis. Information about family income, race, education level, and size were collected with the original consent form.

Children were assigned to group at the classroom level because parents often talk to each other within classrooms, and we wanted to minimize the chances that (a) families would become aware of the multiple treatment groups and (b) that families would ask to change conditions. In schools where there was only one 1st grade classroom, all students were assigned to the math group. In schools with two or more 1st grade classrooms, one classroom was randomly assigned to the reading group and the remaining classrooms were assigned to the math group. In all, students were spread among 40 classrooms and 22 schools.

Children were tested during two half-hour sessions in both the fall and the spring of the 2013-14 school year, as the first year of a larger, 5-year longitudinal study. Children were tested in a one-on-one session by one of eight trained researchers (two post-docs, two graduate students and four full-time research assistants made up the team of testers). At testing, researchers were blind to the group the classroom was assigned to and had no knowledge of parents' math anxiety levels.

The first session at both the beginning and end of the school year consisted of several different achievement measures. In the current study, we focus on the Woodcock-Johnson-III Applied Problems Subtest and the Woodcock-Johnson-III Letter-Word Identification Subtest, nationally normed measures of math and reading achievement, respectively *(14)*. The second session consisted of several different academic attitude measures.

Once student testing was completed at the beginning of the school year, parents were given the iPad Mini along with a survey to fill out and mail back.

Tasks

Child Measures

Woodcock-Johnson III Applied Problems subtest (14): As a measure of math performance, we administered the WJ III Applied Problems subtest, a subtest on a nationally normed, comprehensive test battery used to assess achievement skills of individuals between the ages of 2 through 90 years. The test consists of orally presented word problems that involve arithmetic calculations of increasing difficulty. For example, early problems involve single-digit arithmetic, while later problems involve money calculations and simple fractions. Form A of the test was used in the fall testing session, and Form B was used in the spring. Testing continued until basal (six items correct in a row) and ceiling (six items incorrect in a row) were established. 34 children were excluded from math analyses due to experimenter error leading to failure to achieve either a basal or ceiling on this subtest (24 in the fall: 16 in the math app group, 8 in the reading app group; 10 in the spring: 9 math app group, 1 reading app group).

Woodcock-Johnson III Letter-Word Identification (14): As a measure of reading we administered the WJ III Letter-Word Identification subtest, which measures the ability to identify letters and decode words at increasing levels of difficulty. It is administered using the same basal and ceiling criteria as the Applied Problems subtest, and again Form A was used in the fall and Form B in the spring. 25 children were excluded from reading analyses due to experimenter error leading to failure to achieve either a basal or ceiling on this subtest (7 in the fall: 4 math app group, 3 reading app group; 18 in the spring: 12 math app group, 6 reading app group).

Parent Measures

Parents were given a survey to fill out and mail back when they received the iPad Mini at the beginning of the school year. Parents were provided with a stamped envelope to return their survey. Families who did not return the survey were reminded multiple times via email to complete the survey. In total, 328 parents in the math group (83%) and 119 parents in the reading group (85.6%) returned the surveys. Response rates were approximately equal and high across groups, thus we assume that the missing surveys are random (i.e. it is not the case that high-math anxious parents in one group would be less likely to complete the survey)

Parent math anxiety was assessed using the short-Mathematics Anxiety Rating Scale (sMARS; *22*), which is a 25-item version of the widely used 98-item MARS. Parents responded to questions about how anxious different situations would make them feel (e.g. "studying for a math test," "calculating a tip at a restaurant," etc.). Responses were recorded on a Likert scale from 1 ("not at all") to 5 ("very much"). All analyses were performed on the average of the 25 items. Parent math anxiety in the math group ranged from 1 to 4.8, with a mean of 2.20 (SD=0.81), and parent math anxiety in the reading group ranged from 1-4.8 with a mean of 2.12 (SD=0.77). Parent math anxiety was not correlated with average app use in either the math group (r=-0.07, p=0.12) or the reading group (r=-0.01, p=0.91).

The survey also asked questions about the primary caregivers' previous math experience (i.e. the total number of math classes taken in high-school and college). On the returned surveys, 17 parents did not fill out information about their math experience (13 math; 4 reading). Parents' reported total math classes in the math group ranged from 0 to 15, with 4.80 (SD=2.17) as the mean; reported total math classes in the reading group ranged from 0-10 with 5.00 (SD=2.20) as the mean. To corroborate that total number of math courses taken tracked with parents' math knowledge, experimenters traveled to a subset of family homes in the math group (n=43) in the spring and administered a brief test of math ability to the primary care giver (Woodcock-Johnson III Fluency subtest; *13*). Test performance was significantly correlated with total number of math classes taken ($r=0.33$, $p=0.03$). Neither test performance ($r=0.16$, $p=0.31$) nor total number of math classes taken ($r=0.09$, $p=0.12$) was significantly correlated with math app usage.

Ipad App

After their children were tested in the fall, parents were given an iPad mini with the Bedtime Learning Together ('BLT') app preloaded onto it. The math version of BLT is based on "Bedtime Math," an app created and distributed by the Overdeck Family Foundation, and is available for free on iTunes and on the Android Market. The BLT app differs from the Bedtime Math app in the following ways. First, passage language was adapted to reflect the current year (e.g. a question might read, "If X was discovered in 1956, and it is now 2013, how many years ago was X discovered?" and the question and answer would be updated to say 2014 to reflect that BLT participants received the passage in 2014). Second, the questions used for the "Wee Ones" level were adapted so that all numbers used fell in the 0-10 range, so as to ensure that the lowest level question was developmentally appropriate for a preschool or kindergarten-aged child. Third, the passages were occasionally self-referential to Bedtime Math, and so all such references were removed from the BLT version of the app. Finally, the BLT version of the app was personalized to the specific family using it, so that each time the app was opened, parents could select which children were present at the time of usage. This ensured that we would know when the target child was present during app use, but also allowed parents to use the app with other children in the household.

Families were asked to use the app before bed with their children (or whenever worked best for them during their daily routine), preferably 4 times per week. Daily math (or reading) passages and questions were delivered to the families via the app in the afternoon each day. Additionally, a bank of extra problems was always available for use. Importantly, delivering passages via an iPad allowed us to track how often parents used the app with their children.

For both the math and reading groups, each passage had five questions that differed in difficulty from the kindergarten/first grade level to the fifth grade level. The levels were called “Wee Ones,” “Little Kids,” “Little Kids Bonus,” “Big Kids,” and “Big Kids Bonus.” Questions in the math version of the app covered topics that included counting skills, basic arithmetic (addition, subtraction, multiplication and division), fractions, geometry, and patterns. Questions in the reading version of the app dealt with recalling facts from the story, vocabulary skills, spelling, and making inferences from information in the passages. In both groups, the questions at each level attempted to align with the specific goals set forth in the Common Core Curriculum for the corresponding grade level. Below we provide sample passages and questions for the math and reading groups:

*Sample Math Passage and Questions***:**

Whipped cream was invented about 500 years ago, and is credited to a bunch of guys with long unpronounceable Italian and French names. But what made them think to whip up cream in the first place? Did they know what would happen? Never mind that there was no electricity back then - they had to whip it by hand. Luckily, it was worth the effort.

Whipping air bubbles into cream makes it take up a lot more "volume," or space. In the Bedtime Learning Together test kitchen, 1 cup of heavy cream generated 3 cups of whipped cream. With something as important as dessert, that's a key fact.

Questions. Wee ones (counting on fingers): If you can whip 2 cups of heavy cream into 6 cups of whipped cream, how many cups of air did you whip into it? Little kids: If you're making whipped cream for a party, and 1 cup of heavy cream makes 3 cups of whipped cream, how much whipped cream does 6 cups make? Bonus: If when no one's looking you slurp up 9 cups of the whipped cream, how much heavy cream did it take to make that? Big kids: If a can of whipped cream holds 6 cups, and when you open it, it kind of explodes and squirts 1 1/2 cups on you, how much is left in the can? Bonus: If you then try to squirt half of what's left into your mouth, how much is left after that?

Answers: Wee ones: 4 cups of air. Little kids: 18 cups of whipped cream. Bonus: 3 cups of heavy cream (by the way, at 800 calories per cup, that's 2400 calories - about as much food as a grown-up eats in a day. We don't recommend eating that all at once). Big kids: 4 1/2 cups. Bonus: 2 1/4 cups left.

Sample Reading Passage and Questions:

Whipped cream was invented a long time ago, and is credited to a bunch of guys with long unpronounceable Italian and French names. But what made them think to whip up heavy cream in the first place? Did they know what would happen? Never mind that there was no electricity back then - they had to whip it by hand. Luckily, it was worth the effort.

Whipping air bubbles into heavy cream makes it take up a lot more “volume,” or space. In the Bedtime Learning Together test kitchen, just a little heavy cream generated a lot of fluffy whipped cream! With something as important as dessert, that’s a key fact!

Questions. Wee ones: How did people make whipped cream before we had electricity? Little Kids: What’s the difference between whipped cream and cream? Bonus: Which countries were the inventors of whipped cream from? Big Kids: How does whipping heavy cream turn it into whipped cream? Bonus: In the sentence, “In the Bedtime Learning Together test kitchen, just a little heavy cream generated a lot of fluffy whipped cream,” what does the word “generated” mean?

Answers: Wee Ones: It was whipped by hand. Little Kids: Whipped cream has air whipped into it so it is fluffier than cream. Bonus: France and Italy. Big Kids: The whipping process creates air bubbles. Bonus: To "generate" means to make or create.

App Use

App use was calculated by dividing the total number of times families used the BLT app between fall and spring testing by the number of weeks families had access to the app. Weeks of access ranged from 21 to 30 (Mean=25.43, SD=1.97) for the math group and 21 to 27 (Mean=25.26, SD=1.69) for the reading group. For example, if a family had access to the app for 22 weeks between fall and spring testing, and used the app 88 times during this period, their average app usage would be 4 times/week. In this way, we were able to obtain a measure of how often, on average, children were exposed to the app. App use in the math group ranged from 0 to 4.30 times/week [Mean=1.19, SD=0.97 times/week]. App use in the reading group ranged from 0 to 6.28 times/week [average app use: M=1.61, SD=1.35 times/week].

Supplementary Text

Results

In addition to the children excluded from the relevant math or reading analyses because they failed to hit basal or ceiling on the measures as detailed above, 32 twins were excluded since they were often split between different classrooms, and therefore would likely have been assigned to different groups and thus would have had access to both apps. An additional 21 families dropped out by calling to tell us that they were moving to another district, or that they no longer wished to be in the study: 16 in the math group (3.8%); 5 in the reading group (3.0%). Moreover, for an additional 31 children, app data suggested they never opened the app. We left these children in as 0 users. However, excluding them from the analyses does not change the significance level of any of the reported outcomes. Finally, it is important to note that 26 families consented to be in the study, but failed to pick up an iPad Mini following fall testing. Given that these families did not start the study, we did not include them in our original 587 children number. All other children were included in the original 587 tally.

Children's Math Achievement

All analyses were performed on students' W scores, a transformation of the students' raw score into a Rasch-scaled score with equal intervals (a score of 500 is the approximate average performance of a 10-year-old). Because of its properties as an interval scale with a constant metric, the W score is recommended for use in studies of individual growth. However, for ease of interpretation and illustrative purposes in the figures in the manuscript and SOM, we used grade equivalent growth scores, a measure of how many months of knowledge children gained across the school year.

Relation Between Parents' Math Anxiety and Children's Math Achievement:

Below (Table-S1) we provide fall and spring math W score data for the entire sample as a function of group and app usage, split by parent math anxiety, to give the reader a descriptive view of the data. We first conducted an "Intent-to-Treat" analysis on fall (or beginning-of-year) achievement scores, to ensure that there were no differences between groups at the beginning of the school year. For children of high math anxious parents, we found no significant difference in fall math achievement between children in the math and reading groups $[\hat{\beta}_{11} = 2.27, t = 0.78, p = 0.44]$. Similarly, for children with low math anxious parents, we found no difference in fall math achievement for children in the math and reading groups $[\hat{\beta}_{21} = 2.39, t = 0.60, p = 0.03]$(Model-S3). See the main text for the "Intent-to-Treat" analysis on end-of-year math achievement scores.

We also conducted an instrumental variable analysis on end-of-year math achievement (controlling for beginning-of-year math achievement), which estimates the effect of dosage on those whose dosage was induced by randomization. We found a significant effect of math app use on children of high-math anxious parents $[\hat{\beta}_{21} = 4.12, t = 2.16, p = 0.03]$, but no significant effect of math app use on children of low-math anxious parents $[\hat{\beta}_{31} = -0.65, t = -0.42, p = 0.67]$ (Model-S6). We note that this result is nearly identical to the "as treated" result in the main paper, suggesting that the "as treated" result is negligibly influenced by selection bias.

Accounting for Parent's Math Knowledge in Children's Math Achievement:

Given that people with higher levels of math anxiety typically know less math, one might wonder if the relation between parents' math anxiety and children's gains in math achievement are due to parents' math knowledge instead. We used parents' reported number of math classes taken in high-school and college as a proxy for math knowledge.

Controlling for parent's math knowledge, we still see that for children in the math group, gains in math achievement across the school year as a function of app use are different for children of low- and high-math-anxious parents. Specifically, controlling for parent's math knowledge, for children with high math anxious parents, we found a significant effect of usage $[\hat{\beta}_{30} = 2.60, t = 2.99, p = 0.003]$(Model-S7), such that higher math app usage results in higher end-of-year math achievement. Even when controlling for parent math knowledge, children of high-math-anxious parents who used the math app about once a week (*Bin-1*) grew significantly more in math achievement than children of high-math-anxious parents who used the app the least (*Bin 0*) $[\hat{\beta}_{30} = 7.83, t = 2.97, p = 0.004]$ (Model-S4). However, children of high-math-anxious parents who used the app two or more times a week (*Bin-2+*) did not show significant growth over children who used the app once a week (*Bin-1*) $[\hat{\beta}_{30} = -0.03, t = -0.02, p = 0.99]$ - a finding similar to what was reported in the main paper.

Similarly, controlling for parents' math knowledge, for children of low-math-anxious parents, the parallel model to high-math-anxious parents also yielded a significant effect of usage $[\hat{\beta}_{30} = 2.73, t = 2.39, p = 0.02]$. Similar to that reported in the main paper, when controlling for parents' math knowledge, for children of low-math-anxious parents, the only significant effect was that, at higher doses of math app use (*Bin-*

2+), these children grow significantly more in math achievement than those who interacted with their parents around the app less often (*Bin-1*) $[\hat{\beta}_{30} = 7.43, t = 2.80, p = 0.006]$. However, there was no significant difference between those who used the app the least (*Bin-0*) and those who used it once a week (*Bin-1*) $[\hat{\beta}_{30} = -4.31, t = -1.55, p = 0.13]$, nor between those who used the app the least (*Bin-0*) and those who used it the most (*Bin-2+*) $[\hat{\beta}_{30} = 1.70, t = 1.07, p = 0.29]$.

Moreover, in contrast to the different patterns of math achievement gains observed for children with parents who were high and low in math anxiety as a function of math app usage, children's math achievement gains across the school year do not depend on parents' math knowledge and math app usage $[\hat{\beta}_{40} = -0.14, t = -0.51, p = 0.61]$ (Model-S8). Finally, it is important to point out that the math parents did with their kids on the app was relatively simple (e.g., addition and subtraction). It seems likely that using the math app was giving parents more and better ways to talk to their children about math rather than enhancing their own math skills, as the math provided was geared to young children.

Reading Group

We also looked at whether gains in math achievement for children in the reading group varied as a function of app use for children of low- and high-math-anxious parents. Unlike in the math group, for children with high math anxious parents in the reading group, we found no significant effect of app usage $[\hat{\beta}_{20} = 0.01, t = 0.21, p = 0.84]$. For children of low math anxious parents in the reading group, the parallel model also yielded no overall significant effect of usage $[\hat{\beta}_{20} = -0.18, t = -0.13, p = 0.90]$ (Model-S4; also see Table-S1).

Children's Reading Achievement

In addition to looking at changes in performance on the WJ Math Applied Problems subtest as a function of app use, we also looked at how app usage in the math and reading groups influenced achievement on the WJ III Letter-Word Identification subtest. We first performed a second "Intent-to-Treat" analysis in which we looked at the impact of condition (math vs. reading app) on children's end-of-year reading achievement (controlling for beginning-of-year reading achievement) independent of actual app usage, separating parents on the basis of whether they were lower or higher in math anxiety (median split). We did not find a significant effect of condition for children of high-math anxious parents $[\hat{\beta}_{21} = -0.32, t = -0.12, p = 0.91]$, or for children of low-math anxious parents $[\hat{\beta}_{31} = 0.35, t = 0.14, p = 0.89]$(Model-S3). Furthermore, we did not find a significant group x average use interaction (controlling for fall reading achievement) $[\hat{\beta}_{21} = 0.46, t = 0.35, p = 0.73]$(Model-S1). Figure S1 graphs children's reading achievement gains as a function of app use.

It is perhaps not surprising that using the reading app did not affect children's gains in Letter-Word Identification across the school year given that many parents are already doing reading related activities with their children at home, and that this subtest measures decoding. Additionally, given that the math app is verbal in nature, and similar reading passage content was used in the reading and math apps, enhancements to reading comprehension (e.g., through opportunities to learn new vocabulary) was likely similar across the math and reading apps, although this was not tested in the current study.

References

22 L. Alexander, C. Martray. *Measure. and Eval. in Counseling and Dev.,22*(3),143-150 (1989).

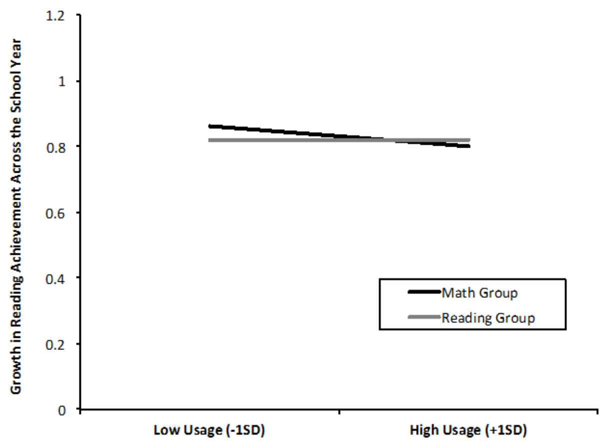

Fig. S1.

Estimated number of months of reading knowledge children gained across the school year (1 is equivalent to 9 months or 1 school year) as a function of average weekly reading app use. Lower usage and higher usage refers to values at -1SD and +1SD from the mean of usage.

Table S1.

Children's Fall and Spring WJ Math Applied Problems W-Scores by Average App Use Bins and Group for (a) children of high math anxious parents and (b) children of low math anxious parents. A 2 point change in W Score is approximately equal to 1 month of growth.

(a)	Entire Math Group		Reading Group	
	Fall W-Score (SD)	Spring W-Score (SD)	Fall W-Score (SD)	Spring W-Score (SD)
Overall	459.15 (16.11)	474.12 (18.73)	457.76 (13.56)	469.02 (19.00)
Bin 0	456.00 (16.78)	465.33 (18.51)	456.50 (10.17)	465.70 (14.77)
Bin 1	454.29 (15.68)	472.27 (18.30)	455.60 (15.38)	468.53 (18.36)
Bin 2+	465.36 (14.09)	481.63 (16.35)	459.63 (13.86)	470.71 (21.31)

(b)	Entire Math Group		Reading Group	
	Fall W-Score (SD)	Spring W-Score (SD)	Fall W-Score (SD)	Spring W-Score (SD)
Overall	462.88 (18.86)	478.54 (21.01)	458.86 (18.08)	475.40 (19.47)
Bin 0	459.73 (17.92)	476.70 (22.78)	447.43 (18.90)	466.00 (17.32)
Bin 1	462.27 (17.69)	474.71 (19.06)	463.63 (18.04)	479.95 (21.48)
Bin 2+	465.74 (21.04)	485.53 (21.42)	461.17 (15.92)	476.90 (18.20)

Model S1.

$$Y_{ij} = \beta_{01}C_j + \beta_{10}F_{ij} + \beta_{20}U_{ij} + \beta_{21}C_jU_{ij} + r_{0j} + e_{ij}$$

Sample	Matched Sample	
Outcome (Y)	WJ Applied Problems Spring '14	WJ Letter-Word ID Spring '14
N	329	335
Page in manuscript	6	S14
Fixed Effects		
Group (C)	474.25***	469.15***
Fall W Score (F)	0.79***	0.68***
Average App Use – Continuous (U)	0.23	0.33
Group (C) * App Use (U)	4.03**	0.46

***p<0.001, **p<0.05

Model S2.

$$Y_{ij} = \beta_{00} + \beta_{10}F_{ij} + \beta_{20}U_{ij} + r_{0j} + e_{ij}$$

Sample	Entire Math Group	Reading Group
Outcome (Y)	WJ Applied Problems Spring '14	
N	370	130
Page in manuscript	6	6
Fixed Effects		
Intercept	475.30***	471.12***
Fall W Score (F)	0.78***	0.85***
Average App Use - Continuous (U)	2.88***	0.22

$^{***}p<0.001$, $^{**}p<0.05$

Model S3.

$$Y_{ij} = \beta_{10}F_{ij} + A_{ij} * (\beta_{20} + \beta_{21}C_j) + (1 - A_{ij}) * (\beta_{30} + \beta_{31}C_j) + r_{0j} + e_{ij}$$

Sample	Matched Sample[*]		
Outcome (Y)	WJ Applied Problems Spring ‘14	WJ Applied Problems Fall ‘13	WJ Letter-Word ID Spring ‘14
N	278	278	279
Page in Manuscript	8	S10	S14
Fixed Effects			
Fall W Score (F)	0.87***	---	0.69***
High Math Anxiety (A)	470.01***	457.93***	468.38***
High Math Anxiety (A) * Group (C)	5.25**	2.27	-0.32
Low Math Anxiety (1-A)	475.10***	458.02***	470.62***
Low Math Anxiety (1-A) * Group (C)	-0.61	2.39	0.35

***$p<0.001$, **$p<0.05$, *Only includes those participants for whom we have parent math anxiety measures

Model S4.

$$Y_{ij} = \beta_{00} + \beta_{10} F_{ij} + \beta_{20} U_{ij} + r_{0j} + e_{ij}$$

Sample	Entire Math Group*		Reading Group*	
	High Math Anxious Parents	Low Math Anxious Parents	High Math Anxious Parents	Low Math Anxious Parents
Outcome (Y)	WJ Applied Problems Spring '14			
N	156	154	49	63
Page in Manuscript	8	8	S13	S13
Fixed Effects				
Intercept	474.28***	478.55***	2.53***	475.39***
Fall W Score (F)	0.81***	0.81***	1.15***	0.82***
Average App Use – Continuous (U)	2.83**	2.76**	0.01	-0.18

Sample	Entire Math Group*					
	High Math Anxious Parents – Bin 0 v Bin 1	High Math Anxious Parents – Bin 1 v Bin 2+	High Math Anxious Parents – Bin 0 v Bin 2+	Low Math Anxious Parents – Bin 0 v Bin 1	Low Math Anxious Parents – Bin 1 v Bin 2+	Low Math Anxious Parents – Bin 0 v Bin 2+
Outcome (Y)	WJ Applied Problems Spring '14					
N	94	114	104	105	123	79
Page in Manuscript	8	8	Not in text	9	9	9
Fixed Effects						
Intercept	469.17***	477.01***	470.13***	478.66***	468.58***	479.98***
Fall W Score (F)	0.81***	0.82***	0.84***	0.89***	0.78***	0.80***
Average App Use – Binned (U)	8.08**	0.52	4.31***	-4.35	7.31**	1.60

****p*<0.001, ***p*<0.05, *Only includes those participants for whom we have parent math anxiety measures

Model S5.

$$Y_{ij} = \beta_{00} + \beta_{10}F_{ij} + \beta_{20}A_{ij} + r_{0j} + e_{ij}$$

Sample	Entire Math Group* – High v Low Math Anxious Parents		
	Bin 0	Bin 1	Bin 2+
Outcome (Y)	WJ Applied Problems Spring '14		
N	72	127	111
Page in Manuscript	9	9	9
Fixed Effects			
Intercept	474.73***	472.13***	485.36***
Fall W Score (F)	0.94***	0.79***	0.78***
Parent Math Anxiety (A)	-7.94**	3.44	-3.60

***p<0.001, **p<0.05, *Only includes those participants for whom we have parent math anxiety measures

Model S6.

$$Y_{ij} = \beta_{00} + \beta_{10}F_{ij} + A_{ij} * (\beta_{20} + \beta_{21}\hat{d}_j) + (1 - A_{ij}) * (\beta_{30} + \beta_{31}\hat{d}_j) + r_{0j} + e_{ij}$$

Sample	Matched Sample*
Outcome (Y)	WJ Applied Problems Spring '14
N	278
Page in manuscript	S10
Fixed Effects	
Fall W Score (F)	0.86***
High Math Anxiety (A)	471.34***
Low Math Anxiety (1-A)	476.86***
High Math Anxious (A) * estimated dosage ($\hat{d}$)	4.12**
Low Math Anxious (1-A) * estimated dosage ($\hat{d}$)	-0.65

***p<0.001, **p<0.05, *Only includes those participants for whom we have parent math anxiety measures

Model S7.

$$Y_{ij} = \beta_{00} + \beta_{10}F_{ij} + \beta_{20}K_{ij} + \beta_{30}U_{ij} + r_{0j} + e_{ij}$$

Sample	Math Group	
	High Math Anxious Parents	Low Math Anxious Parents
Outcome (Y)	WJ Applied Problems Spring '14	
N	149	148
Page in Manuscript	S12	S12
Fixed Effects		
Intercept	474.33***	478.74***
Fall W Score (F)	0.82***	0.82***
Parent Total Math Classes (K)	0.10	0.11
Average App Use – Continuous (U)	2.60**	2.73**

Sample	Entire Math Group*					
	High Math Anxious Parents – Bin 0 v Bin 1	High Math Anxious Parents – Bin 1 v Bin 2+	High Math Anxious Parents – Bin 0 v Bin 2+	Low Math Anxious Parents – Bin 0 v Bin 1	Low Math Anxious Parents – Bin 1 v Bin 2+	Low Math Anxious Parents – Bin 0 v Bin 2+
Outcome (Y)	WJ Applied Problems Spring '14					
N	89	109	100	100	118	78
Page in Manuscript	S12	S12	Not in text	S12	S12	S12
Fixed Effects						
Intercept	464.79***	477.40***	470.52***	478.73***	468.58***	479.79***
Fall W Score (F)	0.82***	0.82***	0.86***	0.90***	0.79***	0.81***
Parent Total Math Classes (K)	0.45	0.11	-0.31	-0.07	-0.01	-0.10
Average App Use – Binned (U)	7.83**	-0.03	4.05***	-4.31	7.43**	1.70

$^{***}p<0.001$, $^{**}p<0.05$, *Only includes those participants for whom we have parent math classes

Model S8.

$$Y_{ij} = \beta_{00} + \beta_{10}F_{ij} + \beta_{20}U_{ij} + \beta_{30}K_{ij} + \beta_{40}U_{ij}K_{ij} + r_{0j} + e_{ij}$$

Sample	Entire Math Group*
Outcome (Y)	WJ Applied Problems Spring '14
N	297
Page in manuscript	S12
Fixed Effects	
Intercept	476.36***
Fall W Score (F)	0.82***
Average App Use – Continuous (U)	3.52**
Parent total math classes (K)	0.42
Parent math classes (K) * App Use (U)	-0.14

***p<0.001, **p<0.05, *Only includes those participants for whom we have parent math anxiety measures