Motivation and Incentives in Education: Evidence from a Summer Reading Experiment

Jonathan Guryan
Associate Professor of Human Development and Social
Policy
Faculty Fellow, Institute for Policy Research
Northwestern University

James S. Kim
Associate Professor of Education
Graduate School of Education
Harvard University

Kyung Park
Assistant Professor of Economics
Wellesley College

Version: January 2015

DRAFT
Please do not quote or distribute without permission.

ABSTRACT

For whom and under what conditions do incentives work in education? In the context of a summer reading program called Project READS, the researchers test whether responsiveness to incentives is positively or negatively related to the student's baseline level of motivation to read. Elementary school students were mailed books weekly during the summer, mailed books and also offered an incentive to read, or assigned to a control group. They find that students who were more motivated to read at baseline were more responsive to incentives, suggesting that incentives may not effectively target the students whose behavior they are intended to change.

I. Incentives in Education

It has become increasingly common for economists and policymakers to suggest the use of incentives in educational settings. The draw to incentives for economists is clear. When faced with investment decisions by students that appear sub-optimally low – low high school completion rates, low effort and low attendance at early grades, at a time when the returns to education are at historically high levels (see e.g. Goldin & Katz, 2008) – economists turn to a central idea of economics, that individuals respond to incentives.

Economic studies of the effectiveness of incentives in education have yielded mixed results (Gneezy, Meier & Rey-Biel 2011). While some show promising results (e.g. Jackson 2010, Kremer, Miguel & Thornton 2009), many show positive results for some groups and no effects for others (Angrist & Lavy 2009, Fryer 2011), or positive results for tests in some subject areas and not for others (Bettinger 2008, Levitt, List, Necerkmann, & Sadoff, 2012).[1] More recent evidence from field trials in school districts suggest that incentives can induce short-term behavioral changes that improve math performance (Levitt et al., 2012) but incentives alone are not sufficient to improve or sustain gains in reading performance (Fryer, 2011; Fryer & Holden, 2012). Each of these previous studies focuses primarily on the question of whether incentives work. In this paper, we focus on a different question: for whom and under what conditions do incentives work? This question is particularly important because of the underlying reasons economists have been drawn to incentives, and because of the problems incentives are intended to address. As we explain, incentives are a promising tool

---

[1] One prominent recent example is the large-scale field experiment reported in Fryer (2010). Fryer implemented field experiments testing the effectiveness of incentives in 200 schools in the Chicago, New York City, and Dallas public schools districts. Students were paid to read books in Dallas, for mid-year test score performance in New York City, and for end of year grades in Chicago. Fryer cannot statistically rule out a zero effect in each city, though the precision of the estimates also allow for the possibility that the incentives would pass a benefit-cost test. Based on the pattern of results across cities, Fryer suggests that incentives based on inputs into the human capital production function (e.g. attendance, book reading) may be more effective than incentives based on outputs (e.g. test scores), possibly because students paid based on outputs may not know how to translate the inputs they can control into the outcomes that are incentivized.

because they might take the place of some cognitive failing or set of preferences that otherwise would have led students to make choices with large long-term benefits. A tool like this, if it is effective, should be effective for students with those cognitive deficits or preferences.

It is well documented that investments in education have high returns (Card 1999), and that those returns have risen dramatically over the past 30 years (Goldin and Katz 2008). Given the high returns, however, why is it that students are not choosing to invest in building human capital through schooling? And, if students are not choosing to respond to the incentives embodied in the labor market returns to education, why should we expect them to respond to nominally smaller incentives of the sort that are typically offered in educational settings?

One possibility is that students are myopic. Suppose that the ability to consider the future consequences of one's actions is a skill that develops as people age. In this case, the discount rate of a young child is not the discount rate his 40-year-old self would want him to use to make decisions. Since the lion's share of the benefits of education occur far into the future, it might make the 40-year-old self better off to provide short-term incentives to encourage behavior that someone with a low discount rate might choose.

Another possibility is that students lack the self-control necessary to engage in schoolwork when other immediate distractions draw their attention. Immediate and salient rewards may have a better chance at competing with other activities in their lives than investing in building human capital does on its own.

There are other similar possibilities, all of which have two things in common: 1) incentives take the place of some skill or trait of the student (e.g. forward-lookingness, self-control) that the policy maker deems to be in deficit relative to some optimal level; incentives act as substitutes for these drivers of behavior that some students have and others lack, and 2) these deficiencies are heterogeneous across students. Therefore, to generate more efficient allocation of effort and human capital investments, incentives must target the students who would not choose to engage in human capital building behavior on their own.

In this paper, we test whether incentives change behavior and learning for these students. Specifically, we test whether responsiveness to incentives is positively or negatively related to the student's level of reading motivation. We do so in the context of a summer reading program called Project READS. As a part of the program, which we describe in more detail in the next section, elementary school students are mailed books weekly during the summer. We implemented this book mailing program as a randomized experiment with three treatment arms. Students in the first treatment arm were mailed books as a part of the standard Project READS program. Students in the second treatment arm were mailed books as a part of Project READS, and were also offered an incentive to read the books they were mailed. Students in the third experimental group served as a control and were given books after post-testing occurred in the fall.

In the spring, as a part of pre-testing, we also collected baseline measures of students' reading motivation level. These measures allow us to distinguish students who enjoy reading and who are generally motivated to read from students who dislike reading and who are unmotivated. The former group of students is more likely to engage in behaviors, like reading and schoolwork, which have long-run returns. The latter group of students expresses an aversion to these behaviors and is precisely the group for whom incentives must be effective if they are to generate more efficient long-run allocation of student effort.

We find that, if anything, more motivated readers are more responsive to incentives to read. Students with greater motivation to read at baseline read more books in response to the incentives. In other words, incentives worked for the students who were more motivated to read before the incentives were introduced than they did for the students who were not already motivated to read.

While incentives induced students to read more over the summer, we did not find that the additional reading generated by the offer of incentives caused reading comprehension test scores to increase for the average student. However, when the most motivated students were incentivized to read books that were well-matched to their reading skill level, we found significant increases in reading comprehension scores. Moreover, these reading comprehension improvements sustained through the school year; we found similar effects on reading comprehension measured on the ITBS administered

in the fall after the students returned from summer break, and on the state end-of-grade standardized test (MCAS) administered in the spring after most of a school year had passed. The spring post-test results should be interpreted with caution since the control group were given books between fall and spring post-testing, but the fact that we find effects for spring and fall post tests for largely the same sets of students may suggest that the spring effects are the result of the intervention during the summer.

Together, these results suggest that the type of reading students do is important, and that incentives may induce productive behavior among students who are motivated to read and learn, but not for students who do not already possess that self-motivation. Furthermore, the results provide suggestive evidence that incentivizing intrinsically motivated children to read well-matched books may improve reading skills in a way that lasts throughout a school year.

II. Project READS

Project READS (Reading Enhances Achievement During Summer) is a voluntary summer reading program with three primary components. The core of the program is that books are mailed to participating students during the summer, typically one book a week for a total of 8-10 weeks. Effort is made to match the books to the students based on the student's reading ability and interests. Prior to the end of the school year, in the spring, teachers implement six lessons aimed at teaching reading strategies.

To assess the effectiveness of the program, reading pre-test assessments are given in the spring, and post-test assessments are given in the fall after the students return from summer vacation. Project READS has been implemented multiple times in multiple school districts (see e.g. Kim 2006; Kim & White 2008; Kim & Guryan 2010). When Project READS has been evaluated experimentally, a randomly selected control group of students has received books during the fall subsequent to post-testing.

Prior evaluations of Project READS have found significant positive effects of the program on reading test scores. Kim (2006) found that students randomly assigned to the program had significantly larger gains on the Iowa Test of Basic Skills (ITBS) between the spring and the subsequent fall than control group students who were not mailed books during the summer. Overall, students gained 0.08 standard deviations, which corresponds

to approximately 1.3 additional months of learning for the average student. Treatment effects were larger for black and Hispanic students, effect sizes of 0.22 and 0.14, respectively, and negative (-0.08) for Asian students.

In a subsequent experiment, Kim and White (2008) tested whether it was sufficient to provide access to books to generate reading score gains of this magnitude. The experiment had three treatment arms and a control group. Control students received books in the fall, after post-testing. All students participated in the series of reading skills lessons in the spring. Students in three separate READS treatment conditions were mailed eight books matched to their reading ability and interest, one per week, during the summer. Each week, all students in Project READS also received a postcard, which they were asked to return after reading the book. The contents of the postcards varied across the three experimental conditions. All postcards asked the students to indicate the title of the book and whether they read it. In a baseline treatment arm, this was the only content on the postcard. In a second treatment arm, the postcards asked questions of parents that were intended to encourage students to read aloud to their parents. In a third treatment arm, the postcards asked questions that were intended to encourage both oral reading and use of the comprehension strategies taught in the spring reading lessons.

Kim and White (2008) found the largest treatment effects for students primed to read aloud to their parents and to use the reading comprehension strategies while they read. Kim and White concluded that access to books is not sufficient to build reading skills; that additional supports or encouragement are necessary to induce children to read during the summer in a way that builds skills. They called these supports "scaffolding." These results are suggestive that to build reading skills, it is important that children not just have access to books and read them, but that they read constructively. When parents pointed students back to the spring reading lessons, the students used the strategies taught during those lessons and read in a way that built skills.

III. The importance of well-matched books

Providing access to books and encouraging children to read in constructive ways are key components of Project READS. The program also stresses the importance of matching books to kids along two dimensions. First, the books should match the interests

of the child, and they should be at a level of difficulty that matches the child's reading skills. Matching the child's interest is hypothesized to make it more likely that the child will read the book and read it intently. Motivation theorists have suggested that providing children with books that are personally interesting can increase the amount of time children spend reading for leisure (Guthrie & Humenick, 2004; McLloyd, 1979; Reynolds & Symons, 2001; Wigfield & Guthrie, 1997). Children who have opportunities to read interesting texts enjoy high levels of involvement during reading, have knowledge about the domain of interests, and recall information from interesting text (Renninger, 2000; Schiefele, 1999). Second, matching books to the child's reading ability has been found to be important for building reading skills. Prior experimental research indicates that children do not enjoy comprehension gains if they read books that are either too easy or too hard (Carver & Leibert, 1995; O'Connor, Bell, Harty, Larkin, Sackor, & Zigmond, 2002).

In other words, children are most likely to benefit from reading texts that are well-matched to their independent reading level (Wright & Stone, 2004). One tool for improving the match between readers and texts is the Lexile system. The Lexile framework is based on the Rasch model, which places measures of reader ability and the readability level of texts on a common scale (Rasch, 1980). The Lexile measure for text is based on measures of sentence length and word frequency, and is often used to determine whether a child's comprehension ability is matched to the readability level of text (U. S. Department of Education, 2001). Currently, 21 US states convert end-of-year reading scores into Lexiles. In addition, several widely used nationally-norm referenced tests yield Lexiles (e.g., Stanford 10, Iowa Test of Basic Skills, Gates-MacGinitie Reading Test). The Lexile system is a useful tool for measuring whether the readability of a book is within a child's independent reading level because it places measures of children's reading comprehension ability and the readability of text on a common developmental scale that increases from the early elementary grades to college.[2]

---

[2] Analyses based on the 1992 National Adult Literacy Study also suggest that mean Lexile scores vary by occupational status and by income levels (Campbell, Kirsch, Kolstad, 1992).

IV. The current study

The goal of the current experiment was to learn how the effectiveness of incentives varied with the level of motivation to read by students in the context of a reading program. We ask: For which students and under what conditions do incentives work? To answer this question, we designed and implemented a randomized field experiment in an urban school district in the northeastern United States. The median household income (average 2006-2010) in the district was about 25 percent higher than the national median. The district is very well educated: more than 90 percent of adults have high school degrees and close to three-quarters have college degrees. The district has fairly equal shares of non-Hispanic white and black students, and a smaller but significant proportion of Hispanic students.

The experiment took place in nine elementary schools, and included all rising 4th- and 5th-grade students in those schools.[3] Students were randomly assigned to one of three conditions, blocked by classroom. One-third were assigned to a control condition, one-third were assigned to participate in standard Project READS, and one-third were assigned to participate in Project READS with incentives.

A. The first treatment arm: Standard Project READS

The standard implementation of Project READS includes several components and is designed to tap children's motivation to read. The focus of the intervention is providing access to books during the summer. Each student who participated in standard Project READS was mailed one book at home each week, for ten weeks during the summer. In the spring towards the end of the school year and just prior to the summer when the books were mailed home, each student in the study filled out a survey, which we describe below, and took a reading pre-test. The reading test used was the Gates-MacGinitie reading test (GMRT), a commonly-used reading assessment that measures both vocabulary and reading comprehension.

---

[3] One classroom was a mixed classroom that included both rising 3rd and 4th graders, so a small number of rising 3rd graders were included in the study.

Students in Project READS also participated in six lessons focused on reading strategies that were designed for students reading on their own, away from the guidance of a teacher. These strategies replicated the lessons used in previous studies and focused on helping children use multiple strategies for independently reading books and reading aloud from texts at home during the summer months. During these lessons, the teacher also explained Project READS, told the students that they would be receiving books during the summer, and encouraged them to read the books.

Along with the books each week, the students also received a post-card. The students were told that they should mail back the post-card after they read each book. On the post-card were several questions. The post-card asked them to list the title of the book they read, what their favorite part was, and asked them to check off each of the reading strategies they used when reading the book. The former questions were intended to gather information about whether the student had in fact read the book. The latter questions were intended to prime students to think about the reading strategies from the spring lessons during their summer reading. These questions were similar to the ones that were found to generate positive impacts on reading scores in Kim & White (2008). These questions are thought to provide what is called in the reading and literacy literature, "scaffolding" (Meichenbaum, & Biemiller, 1998). They act in conjunction with the lessons to provide students with a guide for how to read effectively, and how to improve their skills as readers. In doing so, they do not perform the task for the child, rather they help to support the child's own effort, hence the term scaffolding.

In the fall, when they returned from summer vacation, all students in the study took the GMRT as a reading post-test. After post-testing, students in the control group received ten matched books. The school district preferred to have all students receive the same number of books, and only vary the timing. Because books were provided to control students after the post-test was administered in the fall, the experimental comparisons of fall GMRT scores measures the effect of the intervention(s) relative to a control group who had not yet received any books.

In the spring towards the end of the school year, students also took the Massachusetts Comprehensive Assessment System (MCAS) tests, which were the statewide end-of-grade standardized tests administered in all public schools in the state.

Since both treatment and control students received books before the MCAS was administered, differences in MCAS measure the effect of being mailed books during the summer relative to receiving the same number of books during the school year. While it is possible that the mailed books had a larger effect on reading than the books delivered during the school year, we are not able to test whether this is the case because we do not have measures of book reading during the school year. Nonetheless, we present estimates of the effect on the spring MCAS score along with this caveat because we think it is valuable to have a longer-term measure of test score effects to compare with the estimated effects directly after the summer ended.

B. The second treatment arm: Project READS plus incentives

Students randomly assigned to the second treatment arm received all of the components of standard Project READS and an incentive to read. This incentive to read was provided on top of the motivation to read that might have been provided by participation in Project READS. In addition to participation in the basic form of Project READS, these students were told that they had the opportunity to earn prizes based on the number of Project READS books they read during the summer. Students were told this after they had already chosen their books from the book fair, and after the spring lessons occurred. Students in the incentives treatment arm received a letter explaining that each Project READS book they read during the summer was worth 5 points. The letter read: "Congratulations! You were randomly selected to win prizes for reading books this summer!" Along with the letter, they received a catalog that showed various prizes that could be purchased with those points at the end of the summer. In the catalog were prizes with prices ranging from 5 to 50 points. The prices of prizes were set so that one point was approximately worth one dollar. The prizes included Captain Underpants books, art sets, board games, sports equipment, t-shirts and hats with local sports teams' logos, magazine subscriptions, and science kits. Students were told that they would earn points for reading a book if they returned a completed post-card, and that points would be awarded and prizes could be ordered at the end of the summer. With the maximum 50 points, students could buy a Razor scooter, an interactive robot toy, a watercolor paint and easel set, or combinations of prizes with lower-point prices. Students in the incentive

treatment arm also received the catalog showing the possible prizes every week along with their books. Thus, students were reminded each week about the incentives to read.

Of the 151 students in the incentives group, 100 read at least one READS book and received some points. Of the 100 students with points, 75 claimed a prize. Students were significantly more likely to choose non-book prizes than book prizes, at a ratio of approximately 5 to 1.

C. Book matching

The books each child received were matched to the student based on the reading skills of the student and the preferences expressed by the student. In the spring, soon before the summer when the books were to be sent, students in the study participated in a book fair, where they were asked to look through a set of 115 books and choose the 14 books they liked. From this list of 14 books, we chose the 10 books that were closest in reading difficulty to the student's reading skill level as measured by the match between the student's and the book's Lexile score.

We administered a standardized reading test at baseline (i.e. in the spring immediately preceding the summer when books were mailed home) to measure the students' reading skill level. The students' scores on this test were transformed into Lexiles, a measurement system that categorizes many children's books based on their difficulty (Rasch, 1980; Wright & Stone, 2004). The Lexile system, a proprietary product of MetaMetrics, is widely used in schools across the United States (Schnick, & Knickelbine, 2000). We then measured the match between books and a student's reading level based on the difference between his Lexile reading score and the Lexile rating of the book. Based on this match quality measure, books were categorized as being too easy, about right, or too hard for the student. The Lexile scale ranges from 200 to 1600, and the scale is calibrated so that a reader with a Lexile score equal to a book's Lexile rating is forecasted to comprehend 75 percent of the book's material without outside help or support (Schnick, & Knickelbine, 2000; U. S. Department of Education, 2001). This level is intended to allow the student to comprehend enough to follow the material but to be difficult enough to push the student to develop new comprehension skills. The developers of the Lexile framework consider a well-matched book to be from 100 lexiles below to 50 lexiles above the student's skill level. Research suggests that for developing

readers it may be most effective to read books that are slightly easier to comprehend than the student's skill level would allow, and that this might be particularly true during the summer when adult supports are less prevalent (Hiebert & Sailors, 2009; Kim & Guryan, 2010; Swanborn & de Glopper, 1999).

While our hope was to provide students with ten books that were well-matched to their reading level, in practice some students received books that were matched on reading difficulty better than others. The reason stemmed from our goal of providing books that were well-matched on two dimensions: reading difficulty and the child's interest. In retrospect, the set of books available at the book fair, from which students elicited their preferences, was too narrow in reading difficulty. As a result, students tended to chose books that were too close to the middle of the reading difficulty distribution. Even after 10 books that matched best were selected from among the 14 the students chose, some students received books that were too hard or too easy for them.

The pattern of text difficulty matching was closely related to the skill level of the reader. Because of the compressed nature of the book fair book difficulty distribution, high-ability readers tended to be matched to books that were too easy, while low-ability readers tended to be matched to books that were too hard. This pattern can be seen clearly in figure 1. The top panel of the figure shows the share of students in each tercile of the book difficulty distribution, separately for low-ability, middle-ability and high-ability readers. Book difficulty is measured relative to the reading ability of each student; it is equal to the book's Lexile rating minus the student's Lexile rating.

As can be seen in the figure, 88 percent of students in the bottom third of the reading-ability distribution received books that were in the top third of the book difficulty distribution relative to their reading level. By contrast, 84 percent of the top readers received books that were in the bottom third of the book difficulty distribution relative to their reading level. The students in the middle of the reading ability distribution received a better mix of books. About two-thirds received books that were well matched to their reading level; the remaining third were split fairly evenly between books that were on the hard end and on the easy end of the reading difficulty distribution.

Because reading experts believe reading books that are well-matched to a child's ability is extremely important for skill-building (Mesmer, Cunningham, & Hiebert,

2012), and because providing access to well-matched books is a key component of Project READS, we show results separately for students who received books that were too easy, just right and too hard. While it was not intended, and certainly unfortunate, the provision of books that were outside the recommended range of text comprehensibility provides an opportunity for two interesting empirical tests. Children who are induced to read books that are too hard for them should not experience reading score gains. Thus, the provision of poorly matched books allows for a falsification test. Furthermore, incentives may induce children to engage in behavior that has no benefit beyond earning the incentive payment; or incentives may only amplify incentivized behaviors that children view as independently beneficial. We investigate both of these possibilities.

V. Motivation

A key question we are interested in answering is whether incentives in education are more effective for students who are less self-motivated to make investments in their education. As described in the introduction, we are drawn to this question by the logic that undergirds the provision of extrinsic incentives in education. It has been well established that investments in education, in the form of additional years of schooling, have high returns (see e.g. Card, 1999). Evidence also suggests that school inputs at early ages have long-term labor market returns (Card & Krueger, 1992; Chetty et al., 2011). The association between wage inequality and the returns to schooling (Goldin & Katz, 2008) suggest that all forms of human capital investments, including effort expended at early ages, have very high returns.

If the long-term returns to exerting effort in grade school are so high, then why would providing incentives of the magnitude typically considered lead to more efficient allocation of effort? For incentives to induce optimal effort choices, the students who are induced to act by the incentives must be the ones who, in the absence of the incentive, would not have made decisions that properly weighed costs and benefits.

It is difficult to measure how far individual students are from their optimal level of effort in school. Instead, we measure students' motivation to read. We used a validated measure of children's reading motivation called the Motivation Reading Questionnaire (MRQ; Wigfield & Guthrie, 1999), which was based on a 27-item scale

administered at pretest. We normed each question to have mean zero and standard deviation one, and then took the average of the normed items. The MRQ yielded good reliability (alpha = .93) and was used as a baseline measure of reading motivation.[4] Because reading motivation is a multi-dimensional construct, the MRQ includes items that tap multiple dimensions of children's motivation to read. Other incentives studies have focused on both reading and mathematics achievement, and have therefore used either a general measure of academic self-regulation (Bettinger 2012; Ryan & Connell, 1989) or intrinsic motivation (Fryer 2011, Ryan 1982).

An important feature of our study is that motivation to read was measured at baseline, providing pre-intervention measures of children's motivation to read, whereas previous studies measured academic self-regulation and motivation at posttest. Our primary goal, as stated earlier, is to understand for whom incentives work (or do not work). Thus, we present two types of analyses. First, we estimate the main effects of the basic treatment and incentives on measures of reading behavior and reading skills. Second, we use data on children's reading motivation and the readability level of their books to understand how these two factors influence the relationship between incentives and reading behavior and reading scores.

VI. Research design and descriptive statistics

A. Empirical specification

Students in the study were randomly assigned to one of three conditions. Randomization was carried out at the individual level stratified by spring classroom, so that one-third of each spring classroom was assigned to each of three conditions. We estimate intent to treat (ITT) effects of each experimental condition relative to the control, though consent and assent were obtained prior to randomization. As a result, all students who were randomly assigned to Project READS, with or without incentives, were in fact mailed books.

---

[4] The reliability measure we report is commonly referred to as Cronbach's Alpha. It measures how highly correlated pairs of items are with each other, and can be thought of a measure of internal consistency.

In the analysis that follows we present two types of estimates: estimates of the effect of the treatments (1) on the level of book reading and test scores, and (2) on the gradient of each outcome with respect to baseline motivation to read. The former produces treatment effects on the level of the outcome of interest, while the latter produces estimates of how treatment effects vary with baseline reading motivation.

To estimate effects of the two interventions on the level of book reading and test scores, we estimate the following equations

(1) $\quad\quad\quad B_i = \alpha_0 + \alpha_1 R_i + \alpha_2 I_i + X_i'\beta^{LB} + \delta_c^{LB} + \varepsilon_i$

(2) $\quad\quad\quad T_i = \gamma_0 + \gamma_1 R_i + \gamma_2 I_i + X_i'\beta^{LT} + \delta_c^{LT} + u_i$

where B is a measure of book reading, T is a reading test score, R is an indicator for being assigned to the basic Project READS condition, I is an indicator for being assigned to the READS plus incentives condition, X is a vector of control variables including baseline test scores, gender, and race, the δ's are a full set of classroom dummy variables, ε and u are random error terms, and α, β, and γ are parameters to be estimated.[5] α_1 and α_2 are the causal effects of R and I, respectively, on book reading relative to the control group. γ_1 and γ_2 are the causal effects of R and I, respectively, on test scores relative to the control group.

To estimate the interaction of treatment effects with baseline motivation to read, which is also the treatment effect on the gradient of the outcome with respect to baseline motivation to read, we estimate the following equations

(3) $\quad B_i = \pi_0 + \pi_1 R_i + \pi_2 I_i + \pi_3 \left(R_i \times M_i \right) + \pi_4 \left(I_i \times M_i \right) + \pi_5 M_i + X_i'\beta^{MB} + \delta_c^{MB} + \eta_i$

(4) $\quad T_i = \lambda_0 + \lambda_1 R_i + \lambda_2 I_i + \lambda_3 \left(R_i \times M_i \right) + \lambda_4 \left(I_i \times M_i \right) + \lambda_5 M_i + X_i'\beta^{MT} + \delta_c^{MT} + \xi_i$

where M is our baseline measure of motivation to read, η and ξ are random error terms and the π's and λ's are parameters to be estimated.[6] π_3 and π_4 measure how much larger the treatment effects of R and I on book reading are for students who were more

---

[5] The LB superscript refers to *level* and *book reading*. The LT superscript refers to *level* and *test score*.
[6] The MB superscript refers to *motivation* and *book reading*. The MT superscript refers to *motivation* and *test score*.

motivated to read at baseline. λ_3 and λ_4 measure the same thing for treatment effects on test scores.

Randomization should ensure that conditional on the classroom fixed effects ε, u, η and ξ are uncorrelated with R and I. Furthermore, because the probability of assignment to each of the three conditions was the same in all classrooms (i.e. a $1/3^{rd}$ chance of each), ε, u, η and ξ should be uncorrelated with R and I even unconditional on the classroom fixed effects. If ε, u, η and ξ in fact uncorrelated with R and I, ordinary least squares will yield unbiased estimates of the parameters of interest, α_1, α_2, γ_1, γ_2, π_3, π_4, λ_3 and λ_4 .

B. Descriptive statistics and tests of balance of baseline characteristics

Table 1 presents means of variables measured prior to random assignment. The three columns on the left-hand-side show the means and standard deviations of each variable separately by experimental condition. The three columns on the right-hand-side show the differences in means between READS and control ("R vs. C"), READS plus incentive and control ("I vs. C") and READS plus incentive and READS ("I vs. R"), along with standard errors of the estimated difference.

The top panel of table 1 shows baseline (i.e. pre-random assignment) GMRT scores. The GMRT is a commonly used reading assessment that measures both vocabulary and reading comprehension. There were no statistically significant differences in reading scores at baseline across any pair-wise comparison of experimental groups. The control group had somewhat higher pre-treatment scores than the incentives group, who had somewhat higher baseline scores than the basic READS group, though none of the differences were statistically significant.

The middle panel of table 1 shows means of demographic characteristics. While overall, pre-assignment characteristics were largely balanced across the three experimental conditions, there were some differences. The incentives group had slightly more girls than the basic READS condition or the controls, though these differences were not statistically significant. The control group had fewer black students and more Hispanic students than the two treatment conditions, though the fraction non-white was balanced. The incentives group also had fewer students on free or reduced lunch than the

basic READS or control group. While some of these differences were statistically significant, as might be expected when making multiple comparisons with sample sizes of this magnitude, we cannot reject the joint hypothesis that the three experimental conditions were balanced in all of the demographic characteristics.

The bottom panel of table 1 shows means of reading motivation. There were no significant differences in motivation across the experimental conditions. The incentives group had somewhat higher motivation levels prior to treatment than the basic READS group, though the difference within the range expected due to sampling variation.

Though there did not appear to be more variation across experimental condition in baseline characteristics or pre-tests than would be expected by sampling variation, we have estimated effects from regressions that both do and do not control for pre-treatment observables and pre-tests. The results are not substantively different when these controls are excluded from the models.

VII. Effects on reading behavior

The first question we address is whether the provision of books and explicit incentives to read during the summer led to an increase in reading for the average student. Table 2 presents comparisons across the three experimental groups in three measures of book reading: the number of Project READS books read, the total number of books read during the summer, and the number of books the student checked out of the library over the summer. The former two are based on responses to a question on the fall survey that asked students to list the titles of books they had read during the summer. The number of Project READS books read is the number of titles listed by the student that match any of the books sent to that student as a part of Project READS. For students in the control group, titles are counted as Project READS books if they match one of the Project READS books that student was subsequently sent in the fall after post testing. Control students did not know during the summer or at the time the fall survey was conducted which books they would be sent, but some students read some of their assigned books

just by chance. The number of books read during the summer is the count of titles listed by the student, regardless of whether the titles were sent by Project READS. Total books are inclusive of Project READS books.[7] The library books measure is based on the student's response to questions on the fall survey that was administered after students returned from summer vacation.

A couple things should be noted about the first two measures of book reading. First, they are based on students' recall of the titles of the books they read. By requiring that students remember the titles of books, this measure of book reading is stricter than a pure self-report of number of books read. They are not, however, a perfect measure of book reading. Second, the measure of READS book reading by the control group is not zero because some by chance reported reading books that they were matched to, but would subsequently be sent after post-testing. Books were assigned to both treatment and control students in the same way, so treatment-control differences in READS books read is an estimate of how many more of the matched books treatment students read as compared with the number of those books they would have read in the absence of Project READS.

A. *Effect of book access and incentives on the number of books students read*

Columns 1-3 of table 2 show the average number of books reported read by students in each of the three experimental conditions. Columns 4-6 show the difference in means between each pair of experimental groups, along with the standard error of the difference. On average, students in the basic Project READS group reported reading 3.1 of the 10 books mailed to them. In comparison, students who were given explicit incentives to read the Project READS books read 4.4 of the 10 books mailed to them. The only difference between the READS and READS plus incentives conditions, of course, was the explicit incentive offered to students for reading books. In response to that incentive, students read a statistically significant additional 1.3 Project READS. This

_____

[7] On the fall survey, students self-reported the number of books they bought or received over the summer, the number of books they checked out of the library, and the total number of books they read. On average, the number of books bought/received plus the number of books checked out of the library equaled the total number of books students reported reading, suggesting that the self-reports were internally consistent.

increase represents a 41 percent increase in the number of READS books read compared with the students in the non-incentivized READS program.

Perhaps not surprisingly, students in both READS conditions, with and without incentives, read more READS books than the control students who were not mailed these books, increases of 2.7 and 4.0 books respectively. This result raises the natural question whether providing access to particular books, and incentivizing those books, causes an increase in reading, or a change in which books kids read. The second row of the table presents means of total books read, inclusive of READS books.

Students in the control group read on average 4.3 books during the summer. This represents more reading among the control group than was the case in most other sites where experimental tests of Project READS were implemented. The district where the current experiment was run appears to be one where reading was relatively prevalent in the absence of Project READS. Despite the high levels of baseline reading, Project READS both with and without incentives led to an increase in total reading. Students in Project READS without incentives read 1.2 more books than students in the control group. Students in Project READS who received incentives read 2.0 more books than the control group. Both of these differences are statistically significant. That the effect on total book reading is both positive and smaller than the effect on Project READS book reading suggests Project READS causes both an increase in book reading and a shift in the types of books students read.

About half of the effect of the incentives on reading of Project READS books appears to have come from shifting students away from reading non-incentivized books they otherwise would have read. The incentives caused students to read an additional 1.3 of the incentivized books, but only 0.8 more total books than the control group. The former difference is statistically significantly different from zero, while the latter is not.

There were no significant differences in the number of books students borrowed from the library across the three experimental conditions. On average students in each experimental group borrowed approximately 2.7 library books, and receiving books in the mail did not appear to crowd out the amount of books students borrowed from the library.

In short, providing access to books during the summer caused an increase in book reading. Furthermore, on average students responded to the incentives provided. Adding an incentive to the book-mailing program caused a significant increase in reading of the incentivized books. About half of this response came in the form of reading more books, the rest by inducing students to read the incentivized books instead of other books. On net, the incentives caused a statistically insignificant increase in total reading.

B. Who responds to incentives?

As we explained in the introduction, the effectiveness of incentive-based policies in education depends not only on the effects of incentives on behavior on average. Given the aims of these policies and on their theoretical impetus, effectiveness depends importantly on which students respond to the incentives. A common argument for explicit immediate incentives in education is that students engage in too little behavior that has large long-term returns (e.g. attending school, reading books during the summer, expending effort in schoolwork, etc.). A related argument is that many students are not forward looking enough, and that the returns to working hard in school are inherently realized far in the future. There is, of course, heterogeneity across students in the need for incentives to augment the decision making process. For students who would make optimal decisions on their own, adding short-term incentives will either distort their behavior or have second-order effects. For incentives to address the problem for which they were designed, they must affect the behavior of students likely to make suboptimal decisions on their own, at least from the standpoint of their future selves.

In the case of reading books over the summer, we identify students likely to engage in less than optimal reading by measuring what we call their motivation to read and ask whether students with low motivation to read respond more to incentives to read.[8] Table 3 presents the slope of the relationship between books read and reading motivation, separately for the three experimental conditions. Each entry in columns 1-3 represents the

_____

[8] The reading motivation measure we use, the MRQ, includes items that measure intrinsic, extrinsic and social motivation to read. We have estimated models that investigate whether these types of motivation are independently related to responsiveness to reading incentives. Based on those estimates we are not able to reject the hypothesis that all three types of motivation are related to incentive responsiveness in the same way. For this reason we proceed by reporting results only for the overall measure of motivation to read.

relationship between a one standard deviation increase in reading motivation and the number of books read. For example, column 3 of the second row, which is an estimate of π_5, indicates that among those in the control group, students who scored one standard deviation higher on reading motivation read on average 0.428 more Project READS books. While this correlation is not significantly different from zero, the point estimate suggests that more motivated students indeed read slightly more during the summer even in the absence of any intervention.

A comparison of the gradient among control students and students enrolled in Project READS is informative of whether Project READS has a larger effect on summer reading behavior for more or less motivated students. Focusing first on the books that were targeted in Project READS, both versions of the summer reading program appear to have had a larger effect on reading for more motivated students. The estimate in column 4 of table 3, which corresponds to π_3, indicates that students who were one-standard deviation higher in measured motivation at baseline read 0.910 more of the Project READS books that were mailed to them, as compared with the control group. The interaction between motivation and the treatment effect was even more positive for the students offered incentives. A standard deviation increase in baseline reading motivation was associated with an additional 1.300 book increase in reading of the incentivized books (column 5 of table 3, which corresponds to π_4). The difference in motivation gradient between the students in basic Project READS and Project READS with incentives was 0.390 suggesting that if anything the incentives had a larger impact on reading for the more motivated students. The standard error on this difference is large, making it difficult to draw strong conclusions based on this estimate, but there does not appear to be evidence that the incentives affected behavior more strongly for the least motivated students.

C. Matching books to students

An important feature of this implementation of Project READS was the way that books were matched to students. In the spring before the summer book mailing occurred, students attended a book fair during the school day. There were a total of 115 books in the book fair. Students were encouraged to browse the bookshelves and to choose 14 books that they might want to read. From those 14 books, the 10 books that most closely

matched the student's Lexile reading level were chosen. Students in Project READS, with our without incentives, received those 10 books during the summer. Students in the control group received those 10 books in the fall following post-testing.

This process differed from prior implementations of Project READS in an important way. Prior implementations allowed students to express preferences of books they might like to read via a survey that allowed expression of preferences over a broader range of books. In retrospect, it is clear that the set of books included in the book fair encouraged students to express preferences over books with too narrow a range of reading difficulty. As a result, low-reading-ability students tended to choose books that were above their reading level, while high-reading-ability students tended to choose books that were below their reading level.

We expect that the effect of reading on reading skill varies with how well the book is matched to the child's reading ability (Mesmer, Cunningham, & Hiebert, 2012). For this reason, we examine experimental comparisons of post-tests separately by how well the books mailed to the student were matched by reading skill and text difficulty. As is clear from figure 1, this comparison is largely also a comparison of students in the three terciles of the baseline reading skill distribution. In one sense, the fact that some students received books that were better matched to their skill level than others allows us to investigate whether the quality of the match between a reader's skill and the text difficulty is important. In another sense, if we follow the literacy literature and make the assumption that reading poorly matched books has no effect on reading comprehension skills, we can treat students who received poorly matched books as a falsification test.

Table 4 shows comparisons by experimental condition of the characteristics of books that were matched to students, both the books' difficulty and how well the books' text complexity matched the students' baseline reading skills. The table confirms that, on average, book matching was comparable across experimental conditions. The first row shows the students' average Lexile score, an alternative measure of baseline reading skills, and the one that was used to match students to books. As was apparent in

comparisons of baseline GMRT scores, students in the basic READS condition had somewhat lower baseline reading skills than students in the other two experimental conditions.[9] On average, however, the students in each of the three experimental conditions chose books of approximately the same difficulty, and were assigned books of approximately the same difficulty. There were also no significant differences across experimental conditions in the match of books to student reading skills. While there was variation *within* experimental group in the degree to which their books were matched well to their reading skill level (as measured by the difference between the average Lexile of the books they were mailed and their baseline reading score), there was balance in the degree to which books were matched to student's reading skill level *across* experimental groups. This balance can be seen in the pairwise comparisons across experimental groups in table 4; a chi-squared test also fails to reject that the distribution of students across the terciles of book match quality was the same in each of the three experimental conditions. We take advantage of the within-experimental-group variation in book match quality in analyses that follow.

D. Effects on book reading by quality of the book match

Matching books properly to students' reading skill level was a goal of Project READS because it was believed to be important both for encouraging reading and for development of reading skills. For this reason, we examine the effect of providing book access, both with and without incentives, separately for students who were provided books properly and improperly matched to their reading level.

Table 5 presents estimated differences between experimental conditions in numbers of books read. We broke the sample into thirds based on the difference between the average text difficulty of the books assigned to the student and the student's baseline reading skill level, both measured in Lexiles. We call this difference, the book match quality. Columns 1-2 show results for students who were matched to books that were hard relative to their reading skill level. As shown in figure 1, these students were mostly

---

[9] For this reason, we have run all test score analyses both conditional and unconditional on baseline reading scores. We report results controlling for baseline scores and note in the text when there are differences in the estimates.

drawn from the less skilled readers. The regressions reported in the table control linearly, however, for baseline reading scores.[10] Columns 3-4 show results for students in the middle tercile of book match quality (i.e. those who were matched relatively well to books), and columns 5-6 show results for students in the bottom tercile of book match quality (i.e. those who were matched to books that were easy relative to their reading ability).

The top row of the table shows the control group mean of the same two measures of book reading that were shown in table 3. Control students in the groups that were matched to hard and medium difficulty books relative to their skill level read approximately 5.5 and 4.1 books during the summer, respectively. Control students matched to easy books read 3.2 books.

The next three rows of the table show estimated differences in book reading between pairs of experimental groups. Focusing on the second row, among students in the basic READS condition, there was a monotonic relationship between match quality and the number of Project READS books they read. Students assigned easier books read more. For students assigned books that were too hard and about right, basic READS had an insignificant positive effect on total reading, the point estimate being less than a full book increase. For students assigned books that were too easy, basic READS led to a larger increase of 2.4 more books read. Interestingly, there was not the same monotonic relationship for the students who received the incentives. The effect on reading READS books and total books was largest for the students with the best matched books. Students matched to easier books, who tended to be the most skilled readers responded to the incentives by reading both more Project READS books and more total books. Students with well-matched books, who tended to be in the middle of the reading skill distribution, responded most strongly to the incentivized version of Project READS.

A comparison of the basic and incentivized version of Project READS, shown in the bottom row, reveals that the incentives themselves only had large effects on reading behavior for the students whose books were matched properly to their reading level.

---

[10] Estimates are comparable in models that do not control for baseline reading scores.

Since these were the students in the middle of the reading skill distribution, it seems likely that it was the match quality of the books that drove this effect rather than an interaction of the effect of incentives with the skill level of the students. It appears that the students who responded to the incentives were the ones for whom the program worked as it was supposed, those who were provided access to books that were well-matched to their reading abilities.

Since the reading behavior of students appears to have been affected by how well the books the received were matched to their reading ability, we also examine the interaction of incentives and book access with baseline reading motivation separately for students with well-matched books. Table 6 shows the gradient of summer book reading with respect to baseline reading motivation separately for the terciles of book match quality, following the grouping from table 5. All slopes are estimated in a regression that controls for baseline reading scores. The top row shows the slope of books read with respect to reading motivation for the control group. For total book reading, none of the motivation gradients were statistically different from zero, though the point estimates are larger for students matched to easy books, who were the most skilled readers. The second row shows the estimated experimental difference in reading-motivation gradients between students in basic READS and the control group. For students matched to hard books and those with well-matched books, assignment to READS induced more reading of the mailed books by the students who were more motivated at baseline.

The most interesting results in the table are in the bottom two rows, which compare the incentives group to the control group and the basic READS groups, respectively. Recall that the results in table 3 suggested that the incentives had a larger effect on book reading for students who were more motivated to read at baseline. Breaking the sample up by how well the books were matched shows that this result was strongest for the students who received well-matched books. Among those students, book access with incentives had a significantly larger effect on reading behavior for students who were more motivated to read at baseline. The point estimate implies that students who are one standard deviation more motivated read 3.3 more Project READS books and 2.2 more total books relative to the control group, though the latter estimate is not significant at traditional levels (p=0.17). The final row reports comparisons of motivation

gradients between students in the basic READS and incentivized READS experimental conditions. Among students with well-matched books, incentives had a significantly stronger effect on reading of the incentivized books for more motivated students. The point estimate implies that for students who received well-matched books, a one standard deviation increase in baseline reading motivation was associated with a 2.0 book larger response to the incentives. Among students with poorly matched books, the motivation gradient was not significantly different from zero, and point estimates were negative.

VIII. Effects on test scores

We now turn to the effect of book access and incentives on reading test scores. We present results separately for vocabulary and comprehension reading test scores, and for the two combined. Access to books and summer reading on the scale manipulated by Project READS is designed to affect reading comprehension skills. Interventions like READS and incentives experiments are not designed to induce reading of enough text to significantly improve a child's vocabulary (Fryer, 2011; National Reading Panel, 2000; Swanborn, & de Glopper, 1999). For this reason, we focus on comprehension scores, but present both. We also show effects on the GMRT, which was administered in the fall soon after the students returned from summer vacation, and on the MCAS, which was administered in the spring after most of a school year had elapsed. As discussed above, the MCAS results should be interpreted with caution since the control students were given 10 books after the fall GMRT was administered, but before the MCAS was administered.

The comparison of reading test scores by experimental condition is shown in table 7. The first three columns show the unconditional mean test scores separately for the three experimental conditions. Columns 4-6 show the regression-adjusted average differences between each pair of experimental groups. Regressions control for baseline test scores. As can be seen in the table, there was no measurable effect of Project READS, with or without incentives, on reading comprehension scores. All three point estimates for reading comprehension are small and statistically indistinguishable from zero. The students in Project READS plus incentives scored significantly higher in vocabulary than those in Project READS, though the low score by the Project READS

students, not a high score by the incentivized students, drives this difference. There were also no significant differences in spring MCAS scores across the three experimental groups.

It is perhaps more interesting to examine the effect of book access and reading incentives separately by how well the books were matched to students' reading skills. For early readers, reading books that are too easy or too difficult may not effectively build reading skills. To this end, we present results for test score outcomes separately by book match quality in tables 8 and 9. Table 8 shows differences in average reading scores by experimental condition. None of the 36 coefficients reported in the table are statistically significant. Although students read more during the summer, the null effect on test scores appears to have been consistent regardless of book match quality.

Table 9 shows the gradient of test scores with respect to baseline reading motivation. Recall that the largest interactions between baseline motivation and reading behavior were for students with well-matched books. For the most part, the test score-motivation gradients were not significantly different among the experimental groups for students with poorly matched books. There was one significant positive estimate comparing students in the incentivized and basic READS condition matched to books that were too easy. This difference appears to have been driven as much by a negative gradients among the basic READS students, compared with controls, as by a positive gradient among the incentives group.

In contrast, for students with well-matched books there appears to have been a stronger effect on reading comprehension scores for students who were more motivated to read at baseline. For students one standard deviation above the mean of motivation, the incentives plus READS intervention increased reading comprehension scores by 0.295 standard deviations. This effect was insignificantly larger for the incentivized students than for those in the basic Project READS intervention. Notably, the incentives had a larger effect on spring MCAS scores for students who were more motivated at baseline. The effects of incentives on spring MCAS scores for highly motivated students were similar in magnitude to the effects on fall GMRT scores.

A balance test shows that the correlation between baseline motivation and baseline reading comprehension was stronger in the incentives group than the READS or

control group, but only among students matched to easy or hard books. The regressions reported in table 9 control for baseline reading comprehension, and the test score treatment effects for high motivation students are seen for the one group where this imbalance was not present, those with well-matched books, suggesting that the difference in correlation between baseline motivation and baseline reading comprehension across experimental groups likely does not account for the treatment effects shown in table 9.

We interpret these results with caution because the mismatching of books was not planned. It is interesting, however, to consider one possible interpretation of the test score results in conjunction with both the reading behavior responses and the results of previous studies. In this study, the response of reading behavior to incentives was stronger for more motivated students, only among those who received well-matched books. And, when we look at how treatment effects on reading comprehension varied with baseline motivation we also find that more motivated students experienced reading skill gains relative to less motivated ones, most strongly among those who received well-matched books. It is also interesting that this pattern does not show up for vocabulary scores. It would seem unlikely that the amount of reading induced by this intervention could increase a student's vocabulary. Prior research suggests that children would need massive exposure to books and print over several years to enjoy gains in their reading vocabulary (Nagy, Anderson, & Herman, 1987; Swanborn, & de Glopper, 1999). That we see the same result for in the spring MCAS test that was administered after almost a full school year had passed lends credence to the possibility that this test score difference was driven by the program rather than a statistical accident.

One possible interpretation of the pattern of test score results is that multiple things are necessary for access to books during the summer to increase reading comprehension skills. Consistent with the findings from Kim & White (2008), perhaps more is necessary than just increased access to books. Maybe what is important is that children read books that are well matched to their reading level, and that they do so in a constructive and purposeful way. Simply reading may not be enough; reading deliberatively may be necessary. There are multiple ways to encourage children to read deliberatively and constructively. As Kim & White (2008) found, priming students to use

reading strategies and to read aloud to their parents appears to be one way. Perhaps, extrinsic incentives when provided to intrinsically motivated students is another.

It is interesting to consider the possibility that incentives may be effective for students who are already motivated to learn. When these motivated students encounter an incentive to engage in learning behavior, it seems possible that their internal motivation counteracts any desire to do only what is necessary to earn the contingent reward. For students who do not possess this internal motivation, perhaps reading induced by incentives is less likely to be done in a way that also builds lasting reading skills.

IX. Conclusion

This experimental study addresses both questions about the efficacy of incentives-based educational interventions and for whom and under what conditions such interventions work best. We extend prior research by examining whether incentives to read are more effective for students who were more or less motivated to read at baseline. Economic theory suggests that for incentives to help children make more optimal educational investments, they should change the behavior of those who are least likely to make investments with long term benefits on their own. To test whether this is the case, we randomly assigned students to a status quo control condition and two treatment conditions, one in which students were mailed books during the summer, and one in which students were mailed books and given rewards based on how many of those books they read. Because we measured motivation to read among all of the students in the study, treatment and control alike, we were able to show that, in fact, students who were the most motivated to read at baseline responded most strongly to the incentives (i.e. their summer reading increased most). In light of these findings, rewards may not be a well-targeted strategy for increasing educational investments by less-motivated students.

It is interesting to compare these findings with Leuven, Oosterbeek, and van der Klauww's (2010) finding that incentives had positive effects on academic achievement for the students who were highest achieving at baseline and negative effects for students who were lowest achieving at baseline; Bettinger's (2010) finding that incentives had the largest effects on math scores for students with the highest (and lowest) baseline math scores, but no effects on students in the middle of the distribution; and Angrist and

28

Lavy's (2009) finding that girls responded more than boys to an incentive to pass a matriculation exam. While there may be other explanations for these patterns, it is interesting to wonder whether the heterogeneity in responses to incentives in education that emerges in these studies is related.

While incentives induced students to read more over the summer in the current study, we did not find that the additional reading generated by the offer of incentives caused reading comprehension test scores to increase for the average student. However, when the most motivated students were incentivized to read books that were well-matched to their reading skill level, we found significant increases in reading comprehension scores. Moreover, these reading comprehension improvements sustained through the school year; we found similar effects on reading comprehension measured on the ITBS administered in the fall after the students returned from summer break, and on the state end-of-grade standardized test (MCAS) administered in the spring after most of a school year had passed. These results suggest that the type of reading students do is important, and that incentives may induce productive behavior among students who are motivated to read and learn, but not for students who do not already possess that self-motivation. These results provide suggestive evidence that incentivizing intrinsically motivated children to read well-matched books may improve and sustain performance on non-incentivized test score outcomes. The findings also suggest that – building on the important studies of whether incentives work – we should now be asking how incentives can be effectively incorporated into existing real-world education interventions to encourage children to invest in activities that best promote learning in both the short- and long-run.

References

Angrist, Joshua, and Victor Lavy. 2009. "The Effects of High Stakes High School Achievement Awards: Evidence from a Randomized Trial," *Review of Economics and Statistics* 99(4): 1384-1414.

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

Bettinger, Eric P. 2010. "Paying to Learn: The Effect of Financial Incentives on Elementary School Test Scores," NBER Working Paper No. 16333.

Campbell, A., Kirsch, I.S., & Kolstad, A. 1992. "Assessing Literacy: The Framework for the National Adult Literacy Survey." Washington, DC: National Center for Education Statistics, U.S . Department of Education.

Card, David, and Alan Krueger. "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States," *Journal of Political Economy* 100(1): 1-40.

Card, David. 1999. "The Causal Effect of Education on Earnings," in Orley C. Ashenfelter and David Card (eds.) *Handbook of Labor Economics* Elesevier, 1801-1863.

Carver, R. P., & Leibert, R. E. (1995). The effect of reading library books at different levels of difficulty upon gain in reading ability. *Reading Research Quarterly*, *30*, 26-48.

Chetty, Raj, John N. Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan. 2011. "How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project STAR," *Quarterly Journal of Economics* 126(4): 1593-1660.

Fryer, Roland G. Jr. 2011. "Financial Incentives and Student Achievement: Evidence from Randomized Trials," *Quarterly Journal of Economics* 126(4): 1755-1798.

Fryer, Roland G. Jr. and Richard T. Holden. 2012. "Multitasking, Learning, and Incentives: A Cautionary Tale," NBER Working Paper No. 17752.

Gneezy, Uri, Stephan Meier, and Pedro Rey-Biel. 2011. "When and Why Incentives (Don't) Work to Modify Behavior," *Journal of Economic Perspectives* 25(4): 191-210.

Goldin, Claudia and Lawrence F. Katz. 2008. *The Race Between Education and Technology* Cambridge, MA: Harvard University Press.

Guthrie, John T. and Nicole M. Humenick. 2004. "Motivating Students to Read: Evidence for Classroom Practices that Increase Reading Motivation and Achievement," in Peggy D. McCardle and Vinita Chhabra (eds.) *The Voice of Evidence in Reading Research*. Baltimore, MD: Paul H. Brookes, 329-354.

Guthrie, John. T., Allan Wigfield, Pedro Barbosa, Kathleen Perencevich, Ana Taboada, Marcia H. Davis, Nicole T. Scafiddi, and Stephen Tonks. 2004. "Increasing reading comprehension and engagement through concept-oriented reading instruction," *Journal of Educational Psychology, 96*(3): 403-423.

Hiebert, Elfrieda H. and Misty Sailors. 2009. *Finding the Right Texts: What Works for Beginning and Struggling Readers*. New York: Guilford Press.

Jackson, C. Kirabo. 2010. "A Little Now for a Lot Later: A Look at a Texas Advanced Placement Incentive Program," *Journal of Human Resources* 45(3): 591-639.

Kim, James S. 2006. "Effects of a Voluntary Summer Reading Intervention on Reading Achievement: Results from a Randomized Field Trial," *Educational Evaluation and Policy Analysis* 28(4): 335-355.

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

Kim, James S. and Thomas G. White. 2008. "Scaffolding Voluntary Summer Reading for Children in Grades 3 to 5: An Experimental Study," *Scientific Studies of Reading* 12(1): 1-23.

Kremer, Michael, Edward Miguel, and Rebecca Thornton. 2009. "Incentives to Learn," *Review of Economics and Statistics* 91(3): 437-456.

Leuven, Edwin, Hessel Oosterbeek, and Bas van der Klaauw. 2010. "The Effect of Financial Rewards on Students' Achievement: Evidence from a Randomized Experiment," *Journal of the European Economic Association* 8(6): 1243-1265.

Levitt, Steven D., John A. List, Susanne Neckermann, and Sally Sadoff. 2012. "The Behavioralist Goes to School: Leveraging Behavioral Economics to Improve Educational Performance," NBER Working Paper No. 18165.

McLloyd, Vonnie. 1979. "The effects on extrinsic rewards of differential value on high and low intrinsic interest," *Child Development, 50,* 1010–1019.

Meichenbaum, Donald and Andrew Biemiller. 1998. *Nurturing Independent Learners: Helping Students Take Control of their Education*. Cambridge, MA: Brookline Books.

Mesmer, Heidi. A., Cunningham, James. A., & Hiebert, Elfrieda. H. 2012. "Toward a theoretical model of text complexity for the early grades: Learning from the past, anticipating the future," *Reading Research Quarterly, 47*, 235-258.

Nagy, Wiliam E., Richard C. Anderson, and Patricia A. Herman. 1987. "Learning Word Meanings from Context during Normal Reading," *American Educational Research Journal* 24(2): 237-270.

O'Connor, R., Bell, K. M., Harty, K. R., Larkin, L. K., Sackor, S. M., & Zigmond, N. (2002). Teaching reading to poor readers in the intermediate grades: A comparison of text difficulty. *Journal of Educational Psychology*, *94*, 474-485.

Rasch, G. A. 1980. Probabilistic models for some intelligence and attainment tests. Chicago: University of Chicago Press.

Reynolds, P. L., & Symons, S. 2001. "Motivational variables and children's text search," *Journal of Educational Psychology*, 93, 14–23.

Schnick, Thomas and Mark Knickelbine. 2000. *The Lexile Framework, an Introduction for Educators*. Durham, NC: Metametrics, Inc.

Swanborn, M.S.L. and Kees de Glopper. 1999. "Incidental Word Learning While Reading: A Meta-Analysis," *Review of Educational Research* 69(3): 261-285.

U. S. Department of Education. 2001. "Assessing the lexile framework: Results of a panel meeting," NCES 2001-08. Washington, DC: U. S. Department of Education, National Center for Education Statistics.

Wigfield, A., & Guthrie, J. T. 1997. "Motivation for reading: individual, home, textual, and Classroom perspective," *Educational Psychologist, 32*, 57–135.

Wright, B. D., & Stone, M. H. 2004. *Making measures*. Chicago: The Phaneron Press, Inc.

Table 1. Means and Mean Differences of Baseline Characteristics by Treatment Status

Demographic Characteristics	Experimental Group Means			Mean Differences		
	READS	READS + Incentive	Control	R vs. C	I vs. C	I vs. R
Female	0.478	0.547	0.500	-0.022	0.047	0.069
	(0.043)	(0.043)	(0.042)	(0.060)	(0.060)	(0.060)
3rd Grade	0.058	0.073	0.043	0.015	0.030	0.015
	(0.020)	(0.020)	(0.020)	(0.028)	(0.028)	(0.028)
4th Grade	0.493	0.496	0.471	0.021	0.025	0.004
	(0.043)	(0.043)	(0.042)	(0.060)	(0.060)	(0.060)
5th Grade	0.449	0.431	0.486	-0.036	-0.055	-0.019
	(0.043)	(0.043)	(0.042)	(0.060)	(0.060)	(0.060)
White	0.297	0.299	0.321	-0.024	-0.022	0.002
	(0.039)	(0.040)	(0.039)	(0.055)	(0.056)	(0.056)
Black	0.478	0.431	0.379	0.100*	0.052	-0.048
	(0.042)	(0.042)	(0.042)	(0.059)	(0.059)	(0.060)
Hispanic	0.094	0.088	0.150	-0.056	-0.062*	-0.007
	(0.027)	(0.027)	(0.027)	(0.038)	(0.038)	(0.038)
Other	0.130	0.182	0.150	-0.020	0.032	0.052
	(0.031)	(0.031)	(0.031)	(0.043)	(0.043)	(0.044)
Free or Reduced Lunch	0.486	0.431	0.571	-0.086	-0.141**	-0.055
	(0.048)	(0.045)	(0.049)	(0.067)	(0.068)	(0.050)
Reading Ability at Baseline						
GMRT Vocabulary (NCE)	-0.098	0.013	0.084	-0.182*	-0.071	0.111
	(0.096)	(0.116)	(0.103)	(0.105)	(0.115)	(0.106)
GMRT Comprehension (NCE)	-0.041	0.012	0.029	-0.070	-0.017	0.053
	(0.104)	(0.103)	(0.085)	(0.118)	(0.126)	(0.118)
GMRT Total (NCE)	-0.071	0.026	0.045	-0.115	-0.019	0.096
	(0.102)	(0.108)	(0.092)	(0.109)	(0.120)	(0.108)
Motivation at Baseline						
Motivation	-0.046	0.066	-0.019	-0.027	0.086	0.113
	(0.070)	(0.097)	(0.097)	(0.112)	(0.127)	(0.120)
No. of obs.	138	137	140		415	

Note: The left-hand portion of the table shows means of selected variables by experimental group. The right-hand portion of the table shows mean differences between pairs of experimental groups. There were a total of 415 students in the study. "R" refers to the group that received the basic READS intervention without incentives, "I" refers to the group that received READS plus incentives, and "C" refers to the control group. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 2: Summer Book Reading by Experimental Condition

	READS (1)	READS + Incentive (2)	Control (3)	R vs. C (4)	I vs. C (5)	I vs. R (6)
Project Reads Books	3.127	4.385	0.410	2.717***	3.975***	1.258**
	(0.321)	(0.512)	(0.116)	(0.336)	(0.530)	(0.471)
Total Books	5.474	6.311	4.276	1.198***	2.034***	0.836*
	(0.363)	(0.453)	(0.291)	(0.408)	(0.535)	(0.453)
Books from library	2.597	2.850	2.722	-0.125	0.129	0.254
	(0.179)	(0.174)	(0.203)	(0.210)	(0.257)	(0.227)
Observations	138	137	140		415	

Note: The columns 1-3 show average number of books read by experimental condition. Columns 4-6 report regression adjusted differences in books read between pairs of experimental conditions. Regressions control for baseline reading comprehension scores to match the specification reported for test score outcomes. Project READS books refers to the number of self-reported books read among the Project READS titles matched to the student. Total books refers to the total number of titles that the student self-reported reading during the summer. Books from library refer to the number of books the student reported taking out of the library over the summer. "R" refers to the group that received the basic READS intervention without incentives, "I" refers to the group that received READS plus incentives, and "C" refers to the control group. Standard errors are clustered by classroom and in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 3: Motivation Gradient in Summer Books Read

	READS (1)	READS + Incentive (2)	Control (3)	R vs. C (4)	I vs. C (5)	I vs. R (6)
Project Reads Books	0.685*	1.075**	-0.225	0.910**	1.300**	0.390
	(0.354)	(0.434)	(0.193)	(0.386)	(0.478)	(0.584)
Total Books	1.279***	1.093**	0.428	0.850*	0.665	-0.185
	(0.294)	(0.469)	(0.302)	(0.473)	(0.581)	(0.599)
Books from library	0.367**	0.411**	0.478***	-0.111	-0.066	0.044
	(0.163)	(0.162)	(0.144)	(0.266)	(0.248)	(0.210)
Observations	138	137	140		415	

Note: The columns 1-3 show the slope of the number of books read against baseline reading motivation by experimental condition. Columns 4-6 report regression adjusted differences in that slope between pairs of experimental conditions. Regressions control for baseline reading comprehension scores to match the specification reported for test score outcomes. Motivation is scaled to have mean 0 and standard deviation 1 in the sample. Project READS books refers to the number of self-reported books read among the Project READS titles matched to the student. Total books refers to the total number of titles that the student self-reported reading during the summer. Books from library refer to the number of books the student reported taking out of the library over the summer. "R" refers to the group that received the basic READS intervention without incentives, "I" refers to the group that received READS plus incentives, and "C" refers to the control group. Standard errors are clustered by classroom and in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Experimental Group Average and Mean Differences in Student Lexile and Book Match

	Experimental Group Means			Mean Differences		
	READS	READS + Incentive	Control	R vs. C	I vs. C	I vs. R
Lexile	789.802	790.301	791.946	-2.144	-1.646	0.498
	(1.096)	(0.644)	(1.453)	(1.657)	(1.441)	(1.248)
Avg Lexile Matched Books	679.729	675.054	676.463	3.266	-1.409	-4.676
	(5.874)	(6.450)	(5.184)	(7.864)	(6.112)	(8.765)
Avg Lexile Book Fair Books	685.699	682.196	685.871	-0.172	-3.675	-3.503
	(6.122)	(5.593)	(5.818)	(7.184)	(6.412)	(6.611)
Match Quality	118.178	108.140	114.718	3.460	-6.578	-10.039
	(6.567)	(7.450)	(5.833)	(9.142)	(6.911)	(10.589)

Notes: Match quality is defined as the difference between the child's lexile and the average lexile of the child's 10 matched Project Reads books. Negative values of match quality indicate that the books are harder relative to the child's reading level and positive values imply that the books are easier. Regressions control for baseline reading comprehension scores to match the specification reported for test score outcomes. "R" refers to the group that received the basic READS intervention without incentives, "I" refers to the group that received READS plus incentives, and "C" refers to the control group. Standard errors are clustered by classroom and in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 5: Experimental Differences in Summer Books Read by Books' Match Quality

Book match quality	Hard		Well matched		Easy	
	READS Books	Total Books	READS Books	Total Books	READS Books	Total Books
	(1)	(2)	(3)	(4)	(5)	(6)
Control mean	1.251	5.493	0.643	4.176	-0.626	3.223
	(0.322)	(0.554)	(0.225)	(0.506)	(0.375)	(0.584)
READS-Control	1.776***	0.181	2.690***	0.962	3.654***	2.396***
	(0.441)	(0.718)	(0.596)	(0.814)	(0.538)	(0.676)
Incentives-Control	2.960***	1.082	4.694***	2.683**	4.381***	2.330***
	(0.591)	(0.753)	(1.230)	(1.004)	(0.819)	(0.822)
Incentives-READS	1.184	0.9	2.004*	1.721*	0.727	-0.066
	(0.702)	(0.848)	(0.996)	(0.932)	(0.996)	(0.932)
Observations	415	415	415	415	415	415

Note: The table reports book reading results separately by tercile of the book match quality distribution. The first row reports the number of books read by students in the control group. The next three rows report regression adjusted differences in books read between pairs of experimental conditions. Regressions control for baseline reading comprehension scores to match the specification reported for test score outcomes. Project READS books refers to the number of self-reported books read among the Project READS titles matched to the student. Total books refers to the total number of titles that the student self-reported reading during the summer. "READS" refers to the group that received the basic READS intervention without incentives, "Incentives" refers to the group that received READS plus incentives, and "Control" refers to the control group. Standard errors are clustered by classroom and in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 6: Motivation Gradient in Books Read by Books' Match Quality

Book match quality	Hard		Well matched		Easy	
	READS Books	Total Books	READS Books	Total Books	READS Books	Total Books
	(1)	(2)	(3)	(4)	(5)	(6)
Control gradient	-0.012	0.231	-0.571	0.237	0.155	0.957
	(0.085)	(0.377)	(0.366)	(0.553)	(0.136)	(0.546)
READS-Control	0.832***	1.062**	1.267*	0.934	0.018	0.561
	(0.194)	(0.498)	(0.721)	(0.817)	(1.446)	(1.309)
Incentives-Control	0.702**	0.343	3.311**	2.207	-0.066	-0.361
	(0.323)	(0.600)	(1.289)	(1.498)	(0.599)	(0.819)
Incentives-READS	-0.130	-0.719*	2.044	1.273	-0.084	-0.922
	(0.369)	(0.421)	(0.369)	(1.518)	(1.599)	(1.354)
Observations	415	415	415	415	415	415

Note: The table reports the relationship between book reading and baseline reading motivation, separately by tercile of the book match quality distribution. The first row reports the slop of book reading with respect to baseline reading motivation, for the control group. The next three rows report regression adjusted differences in that slope between pairs of experimental conditions. Regressions control for baseline reading comprehension scores to match the specification reported for test score outcomes. Motivation is scaled to have mean 0 and standard deviation 1 in the sample. Project READS books refers to the number of self-reported books read among the Project READS titles matched to the student. Total books refers to the total number of titles that the student self-reported reading during the summer. "READS" refers to the group that received the basic READS intervention without incentives, "Incentives" refers to the group that received READS plus incentives, and "Control" refers to the control group. Standard errors are clustered by classroom and in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 7: Reading post-test scores by experimental condition

	READS (1)	READS + Incentive (2)	Control (3)	R vs. C (4)	I vs. C (5)	I vs. R (6)
GMRT Fall 2008 post-test:						
Comprehension GMRT	-0.027	-0.015	0.044	-0.072	-0.059	0.012
	(0.047)	(0.056)	(0.049)	(0.060)	(0.054)	(0.064)
Vocabulary GMRT	-0.065	0.079	0.083	-0.148**	-0.004	0.144*
	(0.060)	(0.064)	(0.058)	(0.072)	(0.073)	(0.080)
Total GMRT	-0.062	0.000	0.064	-0.126*	-0.064	0.062
	(0.053)	(0.057)	(0.051)	(0.069)	(0.056)	(0.069)
MCAS Spring 2009 post-test:						
MCAS Total English Language Arts	0.026	-0.006	0.022	0.004	-0.028	-0.032
	(0.068)	(0.053)	(0.051)	(0.096)	(0.067)	(0.077)

Note: The table reports reading test score results. Columns 1-3 show average reading test scores by experimental condition. Columns 4-6 report regression adjusted differences in reading test scores between pairs of experimental conditions. Regressions control for baseline reading comprehension scores. The top 3 rows report results for the vocabulary and reading comprehension portions of the Gates-MacGinitie Reading Test and for the total. The bottom row reports results for the English Language Arts portion of the Massachusetts Comprehensive Assessment System (MCAS) test, the state-wide end of grade standardized test in Massachusetts at the time. Sample sizes were 400 for GMRT comprehension, 405 for GMRT vocabulary, 397 for GMRT total, and 409 for MCAS. Test scores are standardized using normal curve equivalent (NCE) scores to have mean 0 and standard deviation 1 in the sample. "R" refers to the group that received the basic READS intervention without incentives, "I" refers to the group that received READS plus incentives, and "C" refers to the control group. Standard errors are clustered by classroom and in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 8: Experimental Difference in Reading Post-Tests by Books' Match Quality

Book match quality	Hard				Well Matched				Easy			
	Fall - GMRT			Spring	Fall - GMRT			Spring	Fall - GMRT			Spring
	Comp	Vocab	Total	MCAS ELA	Comp	Vocab	Total	MCAS ELA	Comp	Vocab	Total	MCAS ELA
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
READS-Control	-0.09	-0.167	-0.177	-0.155	0.001	-0.075	-0.049	0.191	-0.134	-0.189	-0.153	-0.015
	(0.122)	(0.157)	(0.152)	(0.243)	(0.083)	(0.102)	(0.081)	(0.125)	(0.110)	(0.111)	(0.102)	(0.094)
Incentives-Control	-0.067	0.092	-0.046	-0.095	-0.04	0.018	-0.035	0.17	-0.08	-0.115	-0.11	-0.117
	(0.096)	(0.140)	(0.108)	(0.150)	(0.121)	(0.132)	(0.119)	(0.154)	(0.104)	(0.118)	(0.099)	(0.128)
Incentives-READS	0.023	0.259	0.131	0.06	-0.041	0.092	0.014	-0.021	0.054	0.074	0.043	-0.102
	(0.120)	(0.160)	(0.137)	(0.192)	(0.106)	(0.098)	(0.088)	(0.144)	(0.106)	(0.098)	(0.088)	(0.144)
Observations	400	405	397	409	400	405	397	409	400	405	397	409

Note: The table reports the differences in reading post-tests by experimental condition, separately by tercile of the book match quality distribution. Each row reports regressions adjusted differences in that slope between pairs of experimental conditions. Regressions control for baseline reading comprehension scores. Columns 1-3, 5-7, and 9-11 report results for the vocabulary and reading comprehension portions of the Gates-MacGinitie Reading Test and for the total. Columns 4, 8 and 12 report results for the English Language Arts portion of the Massachusetts Comprehensive Assessment System (MCAS) test, the state-wide end of grade standardized test in Massachusetts at the time. Test scores are standardized using normal curve equivalent (NCE) scores to have mean 0 and standard deviation 1 in the sample. "READS" refers to the group that received the basic READS intervention without incentives, "Incentives" refers to the group that received READS plus incentives, and "Control" refers to the control group. Standard errors are clustered by classroom and in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 9: Motivation Gradients in Reading Post-Tests by Books' Match Quality

Book match quality	Hard				Well Matched				Easy			
	Fall - GMRT			Spring	Fall - GMRT			Spring	Fall - GMRT			Spring
	Comp	Vocab	Total	MCAS ELA	Comp	Vocab	Total	MCAS ELA	Comp	Vocab	Total	MCAS ELA
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Control gradient	-0.101	-0.030	-0.056	0.019	-0.010	0.066	0.029	0.122	0.055	0.039	0.037	0.114
	(0.067)	(0.098)	(0.085)	(0.128)	(0.064)	(0.082)	(0.068)	(0.069)	(0.079)	(0.082)	(0.073)	(0.097)
READS-Control	0.141**	-0.007	0.047	0.105	0.024	-0.101	-0.030	-0.182	-0.252**	-0.115	-0.184*	-0.269*
	(0.069)	(0.083)	(0.085)	(0.136)	(0.082)	(0.113)	(0.086)	(0.158)	(0.112)	(0.140)	(0.102)	(0.154)
Incentives-Control	0.055	-0.031	0.013	-0.092	0.295**	0.104	0.204*	0.202*	0.040	-0.055	-0.028	-0.209
	(0.131)	(0.150)	(0.140)	(0.173)	(0.114)	(0.129)	(0.109)	(0.112)	(0.133)	(0.160)	(0.141)	(0.196)
Incentives-READS	-0.086	-0.024	-0.033	-0.198*	0.271**	0.204*	0.235**	0.384**	0.292*	0.059	0.155	0.06
	(0.127)	(0.118)	(0.113)	(0.098)	(0.107)	(0.106)	(0.092)	(0.156)	(0.145)	(0.192)	(0.161)	(0.163)
Observations	400	405	397	409	400	405	397	409	400	405	397	409

Note: The table reports the relationship between reading post-tests and baseline reading motivation, separately by tercile of the book match quality distribution. Each row reports regression adjusted differences in that slope between pairs of experimental conditions. Regressions control for baseline reading comprehension scores. Columns 1-3, 5-7, and 9-11 report results for the vocabulary and reading comprehension portions of the Gates-MacGinitie Reading Test and for the total. Columns 4, 8 and 12 report results for the English Language Arts portion of the Massachusetts Comprehensive Assessment System (MCAS) test, the state-wide end of grade standardized test in Massachusett at the time. Test scores are standardized using normal curve equivalent (NCE) scores to have mean 0 and standard deviation 1 in the sample. Motivation is scaled to have mean 0 and standard deviation 1 in the sample. "READS" refers to the group that received the basic READS intervention without incentives, "Incentives" refers to the group that received READS plus incentives, and "Control" refers to the control group. Standard errors are clustered by classroom and in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Figure 1: Joint distribution of book match quality and baseline reading skill
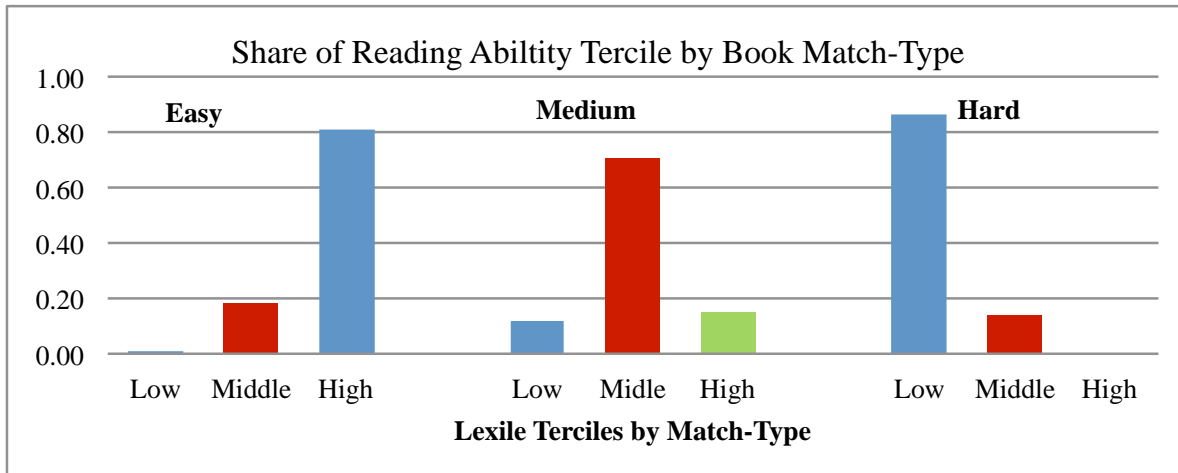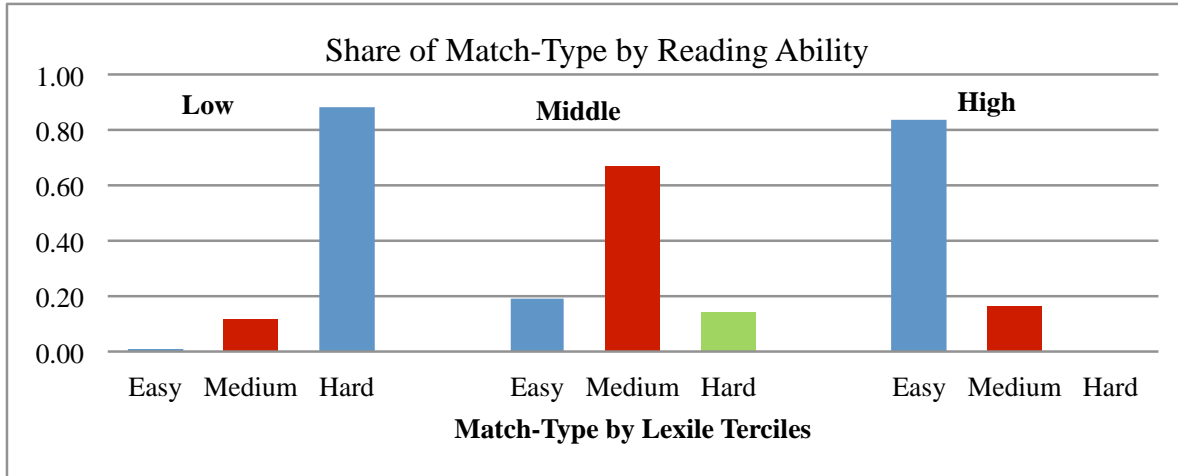
Note: The top panel of the figure shows the distribution of book match quality separately by tercile of the baseline Lexile reading score distribution. The bottom panel shows the distribution of baseline Lexile reading score separately by tercile of the book match quality distribution. Book match quality is defined as the difference between the student's baseline Lexile reading score and the average Lexile of the books that were matched to the student.

Figure 2: Incentives prize catalog sample pages
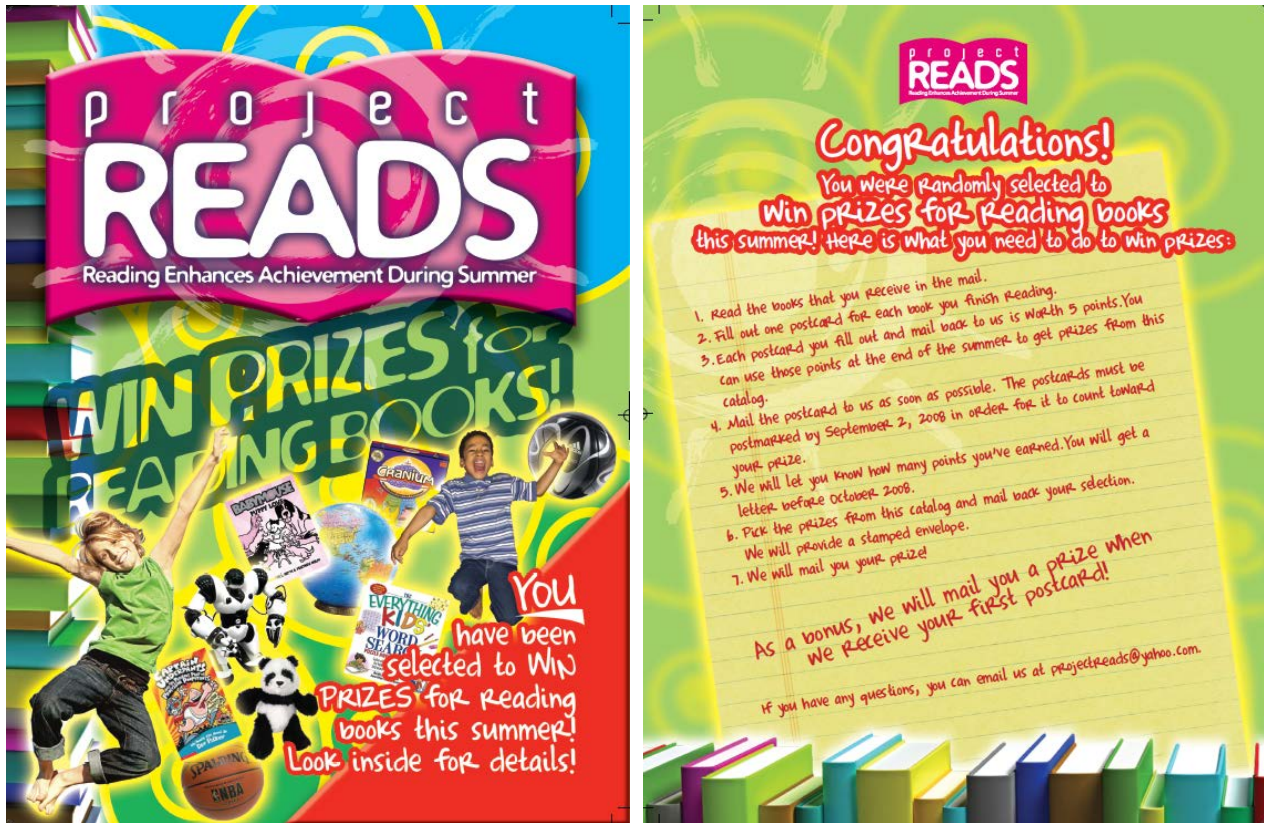
Appendix Figure: Reading Motivation Questionnaire items

★ **Directions for PART 1**: Each question asks how you <u>feel</u> about reading. Circle only <u>one</u> Garfield cat. If you're not sure about the answer to a question, try to make your best guess. **There is no right or wrong answer.** Please follow along with me while I read each of the statements, and then circle your answer.

1. How do you feel when you read a book on a rainy Saturday?

2. How do you feel when you read a book in school during free time?

3. How do you feel about reading for fun at home?

4. How do you feel about getting a book for a present?

5. How do you feel about spending free time reading?

6. How do you feel about starting a new book?

7. How do you feel about reading during summer vacation?

8. How do you feel about reading instead of playing?

9. How do you feel about going to a bookstore?

10. How do you feel about reading different kinds of books?

★ **Directions for PART 2**: We are interested in your reading.

The sentences tell how some students feel about reading. Listen to each sentence and decide whether it talks about a person who is like you or different from you. There are no right or wrong answers. We only want to know how you feel about read. <u>For many of the statements, you should think about the kinds of things you read in your class.</u> Here are some ones to try before we start on the ones about reading:

1. I like ice cream.	Very Different From Me	A Little Different From Me	A Little Like Me	A Lot Like Me
	1	2	3	4

If the statement is **very different from you**, circle a 1.

If the statement is **a little different from you**, circle a 2.

If the statement is **a little like you**, circle a 3.

If the statement is **a lot like you**, circle a 4.

2. I like spinach.	Very Different From Me	A Little Different From Me	A Little Like Me	A Lot Like Me
	1	2	3	4

If the statement is **very different from you**, circle a 1.

If the statement is **a little different from you**, circle a 2.

If the statement is **a little like you**, circle a 3.

If the statement is **a lot like you**, circle a 4.

Okay, we are ready to start on the ones about reading. Remember, when you give your answers you should think about the things you are reading in your class. <u>There are no right or wrong answers: we are just interested in YOUR ideas about reading.</u> <u>To give your answer, circle ONE number on each line.</u> The numbers are right under each statement.

Let's turn the page and start. Please follow along with me while I read each of the statements, and then circle your answer.

	Very Different From Me	A Little Different From Me	A Little Like Me	A Lot Like Me
1. I visit the library often with my family.	1	2	3	4
2. I like having the teacher say I read well.	1	2	3	4
3. I try to get more answers right than my friends.	1	2	3	4
4. I like hard, challenging books.	1	2	3	4
5. I make pictures in my mind when I read.	1	2	3	4
6. My friends sometimes tell me I am a good reader.	1	2	3	4
7. If the teacher discusses something interesting I might read more about it.	1	2	3	4
8. I like being the best at reading.	1	2	3	4
9. I often read to my brother or my sister.	1	2	3	4
10. I like to get compliments for my reading.	1	2	3	4

	Very Different From Me	A Little Different From Me	A Little Like Me	A Lot Like Me
11. I like it when the questions in books make me think.	1	2	3	4
12. I like to finish my reading before other students.	1	2	3	4
13. I am happy when someone recognizes my reading.	1	2	3	4
14. I have favorite subjects that I like to read about.	1	2	3	4
15. I like being the only one who knows an answer about something we read.	1	2	3	4
16. My parents often tell me what a good job I am doing in reading.	1	2	3	4
17. I sometimes read to my parent(s).	1	2	3	4
18. It is important for me to see my name on a list of good readers.	1	2	3	4
19. My friends and I like to trade things to read.	1	2	3	4

20. I usually learn difficult things by reading.	Very Different From Me 1	A Little Different From Me 2	A Little Like Me 3	A Lot Like Me 4
21. I am willing to work hard to read better than my friends.	Very Different From Me 1	A Little Different From Me 2	A Little Like Me 3	A Lot Like Me 4
22. I read to learn new information about topics that interest me.	Very Different From Me 1	A Little Different From Me 2	A Little Like Me 3	A Lot Like Me 4
23. I enjoy a long, involved story or fiction book.	Very Different From Me 1	A Little Different From Me 2	A Little Like Me 3	A Lot Like Me 4
24. I like to read about new things.	Very Different From Me 1	A Little Different From Me 2	A Little Like Me 3	A Lot Like Me 4
25. I talk to my friends about what I am reading.	Very Different From Me 1	A Little Different From Me 2	A Little Like Me 3	A Lot Like Me 4
26. I like to help my friends with their schoolwork in reading.	Very Different From Me 1	A Little Different From Me 2	A Little Like Me 3	A Lot Like Me 4
27. I like to tell my family about what I am reading.	Very Different From Me 1	A Little Different From Me 2	A Little Like Me 3	A Lot Like Me 4

11