The Promise of Low-Cost Interventions in Education

**Symposium Justification**

While many of the problems plaguing education are complex and expensive, many of the day-to-day issues impeding student success are considerably less cryptic and – potentially – less costly to intervene upon. Missing days of school, failing to complete tasks required for college matriculation, or low teacher motivation negatively impact student outcomes, yet these setbacks can often be attributed to maladaptive psychological processes rather than systemic breakdowns. To grapple with these daily setbacks, researchers, policymakers, and practitioners have started to consider how changing cognitive, social, or informational decision contexts can encourage desirable behaviors at scale (Benartzi et al., 2017; Rogers & Frey, 2015; Thaler, 2008). With a goal of being low-cost and scalable, these interventions can spark behavior change in various ways, such as targeting in-the-moment behaviors or enduringly changing consequential thoughts (Rogers & Frey, 2015). In the context of education, university-based researchers are increasingly partnering with schools, districts, and educational institutions to explore the impact of behavioral interventions on a wide range of educational outcomes (Dynarski, 2015).

The specific objectives of this session are to:

1. Describe how low-cost, scalable interventions can motivate behavior change in students and their supporters to increase student success across a wide range of outcomes;
2. Consider the advantages and disadvantages of behavioral interventions in education;
3. Engage the audience in a discussion of how researcher-practitioner partnerships can inform future behavioral interventions; and
4. Consider how these interventions can appropriately impact educational policy and be applied at scale.

The session consists of four presenters describing results from randomized field experiments. After the session organizer introduces the session, the first presenter introduces a novel solution using conversational artificial intelligence for the problem of “summer melt,” the phenomenon where college-intending high-school graduates fail to successfully matriculate. The second presents an experiment evaluating a mail-based intervention that mobilized parents to improve student attendance at scale by targeting commonly held misbeliefs about attendance. The third describes how instructor incentives can be structured to improve college student outcomes. The last presenter discusses how changes to the implementation of new technologies can dramatically increase adoption of strategies that improve student performance. The discussant will facilitate a discussion about the advantages and disadvantages of the new and growing trend of low-cost interventions in education, and how researcher-practitioner partnerships can inform scale-up efforts and educational policy more broadly.

**Organizer:** Carly D. Robinson, Harvard University

**Presenters:**
Lindsay C. Page, University of Pittsburg
Carly D. Robinson, Harvard University
Sally Sadoff, University of California San Diego
Todd Rogers, Harvard University

**Discussant:** Chris S. Hulleman, University of Virginia
BACKGROUND & OBJECTIVE: Deep reinforcement learning using convolutional neural networks is the artificial intelligence (AI) technology behind autonomous vehicles as well as algorithmic medical diagnostics, Facebook automated photo tagging, and programs that defeat world champions in Jeopardy, Chess, and Go. Although AI often performs technical tasks faster and better than people, it is less clear whether AI could substitute for human judgment in addressing individual needs. We investigate this possibility in the context students’ transition from high school to college and the many twists and turns where they can veer off course. Even after acceptance into college, students must navigate several well-defined but challenging tasks, such as completing the FAFSA, submitting transcripts, obtaining immunizations, accepting loans, and paying tuition, among others. Without support, students can stumble and succumb to “summer melt,” the phenomenon where college-intending students fail to matriculate (Castleman & Page, 2014).

Previous efforts to address summer melt have supported students with additional individual counselor outreach (Castleman, Owen, & Page, 2015) or through automated, customized text-message based outreach (Castleman & Page, 2015). Under both strategies, students could communicate with advisors one-on-one. Both approaches improved on-time college enrollment, however, scaling would require significant resources because of the need for a counselor to staff all follow-up communication. AI stands as a potential solution.

SETTING & POPULATION: We report on the use of this system at Georgia State University (GSU), a large, public postsecondary institution in Atlanta, GA.

INTERVENTION & RESEARCH DESIGN: We test whether a conversational AI system can efficiently support would-be college freshmen with the transition to college through personalized, text message-based outreach over the summer. To facilitate personalization, the system integrates with a university’s student information system and customizes outreach according to students’ progress on required tasks. During summer 2016, “Pounce,” the virtual assistant designed and implemented by AdmitHub (and named for the GSU mascot), sent text-based outreach to students admitted to the incoming class of 2016. To test the efficacy of the system to help students complete the required pre-enrollment tasks and matriculate at GSU by the fall, we implemented Pounce via a field experiment. At the outset of our study, some admitted students had already committed to GSU, while others were still choosing among their options (or had committed elsewhere). Consequently, we hypothesized that Pounce would function differently for these two groups. We stratified our sample and randomization accordingly.

ANALYSIS: The data to which we have access for assessing outcomes represents an important contribution of this paper, as it provides an unusually rich look into how this type of intervention can impact students’ success in navigating the college transition and matriculation process. In

---

1 For more information on AdmitHub, see http://www.admithub.com/.
prior studies, researchers examined students’ postsecondary intentions at the time of high school completion and whether students successfully matriculated to their intended (or any) postsecondary institution the following fall, with little understanding of whether additional support led students to complete required tasks at a greater rate. Given the partnership with GSU, we observe not only students’ intentions and postsecondary enrollment outcomes but also their success or failure in completing required pre-matriculation tasks as well as the intervention’s impact on each of these process measures. In short, these data provide a window into the processes that can derail planned college matriculation. The student-level pre-matriculation milestones that we observe include: filing commitment to enroll in GSU; submitting FAFSA; having verification hold on financial aid; submitting final high school transcript; submitting housing deposit; RSVPing for orientation; attending orientation; having immunization hold on registration; accepting any student loan; accepting a Stafford loan; completing loan counseling; setting up a tuition payment plan; and enrolling in the fall of 2016.

To assess the impact of Pounce on student-level outcomes, we use a linear probability model specification. We focus on intent-to-treat estimates, given that the Pounce outreach may have led to improvements in student outcomes, even for those who never responded to the text-based outreach. The models that we fit are of the following general form:

$$Y_{ij} = \alpha_j + \beta_1TREATMENT_{ij} + \gamma X_{ij} + \varepsilon_{ij}, \quad (1)$$

where $\alpha_j$ represents a fixed effect for GSU commitment status at the time of randomization (including in the analyses that pool data across commitment status groups); $TREATMENT_{ij}$ is an indicator for assignment to the treatment group, $X_{ij}$ represents a vector of student-level covariates; and $\varepsilon_{ij}$ is a residual error term. Our estimates of the $\beta_1$ coefficient indicate whether targeting students for Pounce outreach serves to improve student success on the outcome measure considered.

**RESULTS:** The intervention had a significant, positive impact on GSU-committed students. GSU-committed treatment students were 3.3 percentage points more likely to enroll, a 21 percent reduction in summer melt. Consistent with the theory of action underlying summer melt interventions more broadly, the outreach improved students’ success with accessing financial aid, submitting required paperwork and attending orientation, among other requirements.

**CONCLUSION:** Of students who completed high school in 2014, 68 percent—some two million individuals—transitioned directly to postsecondary education. The matriculation process and its corresponding challenges remain reasonably consistent over time. Thus, AI-enabled systems hold promise to provide transitioning students with personalized support for staying on track while not burdening universities with excessive staff time demands. The system can alleviate the need for staff to respond to common questions and free their time for issues that only a human can solve. Just as self-driving cars perform poorly in bad weather, AI-enabled advising technology cannot handle all of the challenges students face in the complex terrain of accessing postsecondary education.
BACKGROUND & OBJECTIVE: Despite the well-documented association between attendance in kindergarten and elementary school and positive student outcomes (Balfanz & Byrnes, 2013), there is little experimental research on how to reduce student absenteeism. What’s more, many parents believe attendance in early grades is not as important as attendance in later grades (Ehrlich et al., 2014). We conducted a large-scale randomized field experiment evaluating the impact of a low-cost, parent-focused intervention on student attendance in kindergarten and elementary school. The light-touch intervention motivated parents to improve student attendance at scale in grades K-5 by targeting commonly held parental misbeliefs undervaluing the importance of regular K-5 attendance as well as the number of school days their child has missed.

Parents’ beliefs about the value of schooling and attendance may influence their motivation to engage in their child’s education (Hoover-Dempsey & Sandler, 1997). A useful theoretical framework for understanding the role of perceived value in education is the expectancy-value model (Atkinson, 1957; Eccles et al., 1983). In the present context, parents’ beliefs about the utility value of attending school regularly in the early grades may affect their behaviors and involvement in their child’s early education. Parents of high-absence students also tend to underestimate the number of days of school their child has missed, shining a light on another barrier to improving student attendance: even if parents value daily attendance in the early grades, they will not be motivated to help their child attend school more if they do not perceive that their child’s attendance is substandard.

SETTING & POPULATION: The sample consisted of 10,967 households across ten school districts in a diverse county in California. Our sample was limited to all kindergarten students (who had no prior school year data) and all first through fifth grade students who were in the bottom 60th percentile of attendance of participating districts countywide during the prior school year. In households with two or more qualifying K-5 students attending the same district (18.3%), we randomly selected one student to receive treatment.

INTERVENTION & RESEARCH DESIGN: The intervention consisted of delivering personalized information to parents of high-absence students through a series of mail-based communications. We sent six rounds of treatment over the course of the school year to treatment households, sending on average 5.15 mailings to each household (after accounting for opt-outs, bounce backs, etc.). Specifically, this study explored whether sending parents mailers that: 1) emphasize the utility value of regular school attendance in the early grades, and 2) accurately report how many days their child has been absent has an impact on student absences (compared to a control group).

At the end of the school year, the research team conducted a 15-minute phone survey of eligible households (both treatment and control) to learn whether the intervention impacted parental beliefs.
ANALYSIS: We registered an analysis plan on Open Science Framework before receiving outcome data from the school districts and pre-specified our hypotheses. Our major hypothesis was: Students who received treatment mailings will have improved attendance as compared to students in the control group.

To assess these hypotheses, we first employed Fisher Randomization Tests (FRT) to obtain exact \( p \)-values to determine whether there was a statistically significant treatment impact on student absences (Athey & Imbens, 2016). Second, we fit linear regression models to estimate the Average Treatment Effect (ATE) of random assignment to the treatment condition on student absences. To examine the ATE on chronic absenteeism, we used logit regression models. Our final models adjusted for student-level demographic indicators, student’s previous year absences (when available), and the student’s school and grade level.

We also explored the extent to which the treatment impacted parental beliefs about the importance of schooling in the early grades and whether the treatment corrected parents’ (possibly incorrect) beliefs about how many days their child was absent. We conducted a factor analysis to create latent variables that summarize parental beliefs toward education and attendance, and then evaluated the ATE on parental beliefs.

RESULTS: We find that students in households assigned to receive attendance mailings were absent for 0.53 fewer days over the course of the school year, on average, than students in households that did not receive attendance mailings. This translates to a 7.7% reduction in absences compared to students in the control condition, and resulted in students attending 3,486 more days of school over the course of the year (0.53 days * 6,579 students in the treatment condition). This also corresponds with a 14.9% reduction in chronic absenteeism.

A follow up survey showed that the intervention partly corrected parents’ incorrect beliefs regarding the number of days their child had been absent, increasing parental accuracy by approximately one day. The mailings also made parents more likely to agree with statements about the value of schooling in the early grades and the importance of regular attendance.

CONCLUSION: The intervention can be economically implemented by schools, costing about $10 per incremental school day generated. Furthermore, the intervention mobilizes the efforts of a costless resource for schools and students: parents. Schools need to empower and inform parents if they can be expected to effectively intervene upon their child’s education. This study demonstrates that targeting parental beliefs about the importance of regular K-5 attendance provides a cost-effective solution for reducing student absenteeism.
Presentation 3: Improving College Instruction through Incentives

Sally Sadoff

Background / Context: Recent work has demonstrated that instructor quality matters at the post-secondary level, though the effect sizes vary widely. Student performance is estimated to improve between 0.05 and 0.3 standard deviations (SD, hereafter) with a 1 SD increase in instructor quality. These effects are generally smaller for instructors at research universities and larger for instructors at for-profit institutions (Braga, Paccagnella, & Pellizzari, 2016; Brodaty & Gurgand, 2016; Carrell & West, 2010; De Vlieger, Jacob, & Stange, 2017; Hoffmann & Oreopoulos, 2009). Research has also examined the correlation between instructor type—adjunct or tenure-track faculty—and performance. Some studies find little relationship and others find that adjunct instructors are more effective (Bettinger & Long, 2010; Carrell & West, 2010; Figlio, Schapiro, & Soter, 2015; Hoffmann & Oreopoulos, 2009). However, little is known about how to improve instructor quality at the post-secondary level. While a growing literature examines the effect of teacher incentives at the K-12 level, to our knowledge no previous study has tested the impact of instructor incentives at the post-secondary level.

Purpose / Objective / Research Question: In a randomized experiment, we test the effect of performance-based incentives for community college instructors. We examine the effects of instructor incentives alone and also test for evidence of complementarities with student incentives. Our incentives are framed as losses given in the form upfront bonuses that instructors pay back at the end of the semester if they do not meet performance targets. The design is based on Fryer et al. (2012) who find large effects of loss-framed incentives among elementary and middle school teachers. We also elicit instructor's preferences to work under loss-framed contracts compared to gain-framed contracts (rewards given at the end of the semester).

Setting & Population: We conducted the experiment in the 2016-2017 school year at Ivy Tech Community College of Indiana (Ivy Tech). Our study includes 136 instructors and 5,912 student-course observations (3826 unique students).

Intervention & Research Design: We conducted the experiment in the fall and spring semester of the 2016-2017 school year. In the fall semester, we randomly selected instructors to receive performance bonuses based on the number of their students who passed an objective course exam with a 70 percent or higher.

We offered performance bonuses worth $50 per student and framed these incentives as “losses”. We sent all treatment instructors upfront bonuses at the beginning of the semester equivalent to the amount they would receive if half of their students passed the exam. At the end of the semester, if fewer than half of the instructors' students passed the exam, they returned the difference between their final reward and the upfront bonus. If over half of the students passed the exam, the instructor received additional bonuses. In the spring semester, we also randomly selected course sections to receive student incentives. In the sections that received student incentives, students who passed the exam received free tuition for one summer course (worth approximately $400). Finally, we elicit instructors’ preferences for loss-framed contracts (upfront bonuses) versus gain-framed contracts (semester-end bonuses).
Data Collection and Analysis: We pre-registered our analysis plan. See [https://osf.io/fbxpw/Methods](https://osf.io/fbxpw/Methods).

We test two hypothesis -- first, that instructor incentives improve student outcomes, and second, that instructor incentives are more effective when combined with student incentives -- using the following equation estimated using a random effects model with standard errors clustered at the instructor level:

\[
Y_{ij} = \beta_0 + \beta_1 Z_{ij}^1 + \beta_2 Z_{ij}^2 + \beta_3 Z_{ij}^3 + \beta_4 X_i + \beta_5 X_s + \beta_6 X_j + U_j + \varepsilon_i + \varepsilon_s + \varepsilon_j,
\]

where \(Y_{ij}\) is the outcome for student \(i\) in section \(s\) taught by instructor \(j\); \(Z^t_{ij}\) is an indicator variable for whether section \(s\) taught by instructor \(j\) is assigned to treatment \(t\) = \{1,2,3\} (1=Instructor incentives, 2=Student incentives, 3=Combined Instructor-Student incentives); \(X_i\) denotes a vector of student covariates; \(X_s\) represents course-specific covariates (such as whether the course is a co-requisite course); \(X_j\) represents instructor-specific covariate (such as full-time or adjunct status); \(U_j\) represents instructor-specific random effects; due to the randomization, the error terms \(\varepsilon_i, \varepsilon_s, \varepsilon_j\) are mechanically uncorrelated to the \(Z^t_{ij}\) terms. \(\beta_1\) provides our test of the effect of instructor incentives. \(\beta_3\) is our estimate of the effect of combined instructor and student incentives.

Findings / Results: We find that instructor incentives alone increase exam performance by an estimated 0.19 - 0.2 SD \((p < 0.01)\). Instructor incentives decrease course withdrawals by 4.3 percentage points \((p < 0.01)\) and increase course grades by 0.17 grade points or 0.11 SD \((p = 0.016)\). The overall effects are driven by large impacts among adjunct instructors: the estimated effects on exam scores for adjunct instructors of 0.26-0.27 SD \((p < 0.01)\) is significantly larger than the estimated effects of 0.075-0.09 SD for full-time faculty (not significant at the 10% level).

We find no evidence that student incentives increase the impact of instructor incentives when offered in combination; and, in fact, find suggestive evidence that combined incentives are less effective than instructor incentives alone (we also find little evidence that student incentives improve performance when offered alone; we did not power the experiment to estimate the effects of student incentives alone so these estimates are noisy).

In our elicitation of instructor preferences for contract, we find that at baseline instructors prefer gain-framed contracts, but that experience with the loss-framed incentives significantly increases instructors' preference for these types of contacts, making them (close to) indifferent between the contract types.

Conclusions: Our results demonstrate that loss-framed incentives can improve post-secondary instruction at a low cost (about $22 per student). Our setting is also interesting because teaching assignments and employment contracts are far more flexible in higher education than in most K-12 settings, particularly for part-time adjunct instructors, who instruct courses at a far lower cost than full-time faculty. The dramatic impact our incentives have on adjunct instructors suggest that there could be substantial gains from reconsidering the contracts offered to part-time instructors.
Presentation 4: Is Technology Useless? Impact of Actionable Information Delivered to Parents and How to Increase Demand for It.

Todd Rogers & Peter Bergman

BACKGROUND & OBJECTIVE: Recent research has shown that technology that automates the sending out of frequent and low-cost text message alerts to parents with timely information about their children’s academic progress can be potent at increasing student success (Bergman, 2014; Bergman & Chan, 2017). This impact is only possible when parents elect to receive the communications. However, many families that might otherwise benefit do not elect to receive the communications. We study how several psychologically-informed implementation strategies affect parent take-up of these communications as well as their impact.

SETTING & POPULATION: We conduct a field experiment across 12 middle and high schools in the District of Columbia Public Schools (N=6,291). In these 12 schools in 2015, 81% of students were Black, 16% Hispanic, and just under 2% white. Across the entire school district, 67% of all enrolled students in 2015 were Black, 17% Hispanic, and 12% white. The 2015 graduation rate for DCPS as a whole was 64%, and the graduation rate for the four high schools in our sample was 68%. Overall, 25% of all DCPS students met ELA proficiency on the PARCC assessment, and 21% met math proficiency. In our 12 school sample, 9% of students met ELA proficiency, and 5% met math proficiency on the PARCC assessment in 2015.

INTERVENTION & RESEARCH DESIGN: Teachers in these schools use a gradebook technology in classrooms for tracking of attendance, assignment grades, and quiz and exam grades. We studied an intervention that sent automated text-messages directly to parents’ mobile phones conveying three types of message: (1) student absences, (2) missing assignments, and (3) if their child’s grade average was below 70%. Families electing to receive the messages received an average of 19 over the semester. Parents were randomized into one of four conditions, as described below. Outcome measures reflect district administrative data.

Those in the Standard condition were told by text message that they could adopt the technology by enrolling on the district website, which is standard practice. Those in the Simplified condition were told by text message that they could adopt the technology by replying “start” in response to a text message (Bettinger, Long, Oreopoulous, & Sanbonmatsu, 2012). Those in the Automatically Enrolled condition were told by text message that they could adopt the technology passively by not opting out of being enrolled by default, which parents could do by responding “stop” to any text-message alert (Madrian & Shea, 2001).

Finally, a survey of large district leaders sheds light on why they tend to not implement using Automatic Enrollment (N = 100).

ANALYSIS: We are interested in two primary outcomes. First, we are interested in how implementation strategy (as reflected in condition assignment) affects adoption of the text message parent alert service. Second, we are interested in how implementation strategy (as reflected in condition assignment) affects student academic performance. For the latter, we use two measures of academic performance: number of courses a student fails, and average semester grade point average (GPA).
We estimate the causal effect of condition assignment by the OLS regression

\begin{equation}
  y_i = \alpha_0 + \beta_1 T_i + X_i + \epsilon_i
\end{equation}

where \( y_i \) is the outcome variable for student \( i \); \( T_i \) is a vector of indicators for assignment to one of the three treatment groups or the Control condition; \( X_i \) is a vector of pre-intervention student-level covariates; and \( \epsilon_i \) is an error term. The standard errors are estimated using the Huber-White robust estimator for the variance-covariance matrix. Student-level covariates included in \( X_i \) are baseline GPA, a continuous measure of the number of times the parent had ever logged into the parent portal prior to the intervention, the number of student absences prior to the start of the intervention, and an indicator for students who are Black or African-American.

The first outcome variable we test is average second semester grade point average. Students receive numeric grades on a 100-point scale in each course, as well as letter grades ranging from A+ to F. Letter grades of a D- or below are considered failing. We calculated an average term GPA for each student from individual course grades received in language, math, science, history, and arts courses. We then calculated each student’s second semester GPA by averaging her third and fourth term GPAs.

We also use equation (1) to test the effect of treatment on the number of courses failed in the second semester. To pass a course, students must have a final grade of 64 or above on a 100-point scale, which is equivalent to a “D” letter grade.

**RESULTS:** The Standard implementation strategy—proactive online signup—induces negligible adoption (<1%). The Simplified implementation strategy modestly increases adoption (8%)—especially among parents of higher-performing students. The Automatic Enrollment implementation strategy dramatically increases adoption (96%). The Standard and Simplified implementations generate no detectable increases in student performance. However, Automatic Enrollment meaningfully increases GPA and reduces student course failures. District leaders fail to anticipate the massive difference in adoption between Automatic Enrollment (62%), Simplified (42%), and Standard (37%).

**CONCLUSION:** How new educational technologies are implemented can radically affect their impact on district objectives. This impact is underappreciated by district leaders, which can help explain why the typical district implementation strategy for new technologies like this parent text message alert requires active parent opt-in (rather than opt-out). This results in lost opportunities for student achievement, and inefficient district resource expenditure. More broadly, simple changes to the implementation of new technologies can lead to radically different conclusions about whether new technologies are valuable and their ability to close achievement gaps.
References


Bergman, P. (2014). *Evidence from a field experiment on parent-child information frictions and human capital investment.* Retrieved from


