Variation in Motivational Appeals to Survey Completion: Lessons from a Randomized Experiment with Teachers

Robert Nathenson* [a]
Jonathan Supovitz [a]

Robert A Nathenson, PhD MSE MSc
Research Specialist
Consortium for Policy Research in Education (CPRE)
Email: robertna@upenn.edu
Phone: (215) 573-0773
* Corresponding author

Jonathan Supovitz, Ed.D.
Professor and Chair, Education Policy Division
Director, Consortium for Policy Research in Education (CPRE)

[a] Graduate School of Education
University of Pennsylvania
3440 Market Street, Suite 500
Philadelphia PA, 19104

Running Head
Variation in Motivational Appeals to Survey Completion

Key Words
Non-response Bias, Survey Attrition, Psychological Appeal, Financial Incentive, Behavioral Science

Acknowledgments
The authors would like to thank the National Science Foundation for its generous funding of this research (Grant # DRK-12 1316527). We would also like to thank Eric Foster of MDF Research for administering the survey, as well as the countless team members at CPRE who facilitated this work. This work also benefited from the feedback of many colleagues at the annual meetings of the Society for Research on Educational Effectiveness in the spring of 2018.

Variation in Motivational Appeals to Completing a Teacher Survey: Lessons from a Randomized Experiment

Abstract

Sample attrition increases the risk of statistical bias and hinders the ability to plausibly estimate causal effects when patterns of nonresponse are correlated with key variables of interest. Drawing on Leverage-Salience Theory and other work in the behavioral psychology field, we empirically capture the impact of distinct motivational appeals on the survey response rates of elementary education teachers in a large urban school district in the northeastern United States. During Spring 2017, teachers were randomized to receive one of six motivational appeals and were re-randomized to receive a different appeal each subsequent week, conditional on not having completed the survey. We observe the results on four different margins, which range in their time intensity (open email and click, start, and complete survey). We find that extrinsic rewards improved teacher response across all four margins and the social norm of reciprocity substantially improved teacher response along margins of lower time-intensity. As researchers continue to conduct multi-treatment arm studies and large-scale evaluations that can suffer from serious issues of sample attrition, this work highlights the contribution of message framing in survey response.

Introduction

Survey nonresponse and sample attrition, defined as "the loss of sample during the course of a study," has dramatically increased in the United States over the last several decades (Basson 2011; Tourangeau & Plewes 2013; What Works Clearinghouse 2015). This issue has affected a

broad array of research disciplines, from health (National Health Interview Survey) to education (the National Center for Education Statistics' National Household Education Survey), as well as popular national datasets (Current Population Survey; Survey of Income and Program Participation; Tourangeau & Plewes 2013). As evidenced by these studies, even those with strong research designs are not immune to such issues.

The fundamental problem with sample attrition is that it threatens statistical validity, as greater attrition increases the risk that the analytic sample is distinct in important ways from the underlying population of interest from which the sample was drawn. The issue has the potential to increase statistical bias in the estimates of the key parameters, which can impact the ability to plausibly estimate causal effects (Heckman 1979). As an example, using Finland's national death registry and biennial survey sample data, Mattila, Parkkari, and Rimpela (2007) find that Finnish males who did not complete the study survey had higher hazard ratios of death overall and higher hazard ratios of death from intoxication, disease, violence, and unintentional injury, specifically, as recorded in the national death registry, than the set of male survey respondents. Their respondent sample therefore significantly under-reported the risk of death for males from each of these factors (Mattila, Parkkari, & Rimpela 2007). In short, "when patterns of nonresponse…are significantly correlated with variables of interest in a survey, then the nonresponse contributes to biased estimates of those variables and is considered nonignorable" (Basson 2011).

The What Works Clearinghouse (WWC) has developed threshold standards to assess the severity of sample attrition in research studies (WWC 2015). The standards cover both (overall) attrition and differential attrition, the disparity in attrition across different groups of interest (e.g. treatment and control; teachers with traditional and alternative teaching credentials) (Lavrakas 2011; WWC 2015). Yet, most education research studies only briefly acknowledge these issues.

Multi-treatment arm studies and large-scale evaluations, particularly multi-site cluster randomized control trials, are especially at-risk for differential attrition, as loss of sample may vary substantially across sites and arms for non-random (and unobserved) reasons. For instance, if teachers in one site are provided an organized time and place to complete their survey while teachers in other sites are not, or if teachers in the treatment arm are more likely to complete their survey due to a positive association with the intervention than teachers in the control arm.

Early research by Don Dillman and colleagues suggests that a comprehensive approach to survey design, what they call the 'tailored design method,' can help alleviate issues of survey attrition by improving survey response rates.[1] They argue for the joining of psychological concepts from Social Exchange Theory, a framework for understanding behavior of individuals during interactions and the development of social norms that shape these interactions, to such survey features as mode, timing, and phrasing to create a "holistic data collection protocol" that can be tailored to the specific context of the research survey (Dillman, Smith, & Christian 2014, p. 24). Yet, empirical research on understanding the underlying psychological motivation for, and potential improvements to survey response, remains an under-developed area of focus in education research. Researchers therefore lack an extant body of empirical literature from which to draw when designing studies in which survey response is critical to calculating causal effects.

Recent developments in the behavioral psychology field hold strong potential for understanding and grappling with these issues. Insights from behavioral psychology, such as gentle nudges and choice architecture (e.g. Richard Thaler's *libertarian paternalism*), have been found to have a significant impact in a wide variety of real-world applications, including retirement savings (Center for Advanced Hindsight 2017), prescription adherence (Kessler et al

[1] Originally titled *Mail and Telephone Surveys: The Total Design Method* (1978), the most recent edition was published in 2014 as *Internet, Phone, Mail and Mixed-Mode Surveys: The Tailored Design Method. 4th edition.*

2018), and, in higher education, issues of college enrollment and financial aid (FAFSA) completion (Bettinger et al. 2012; Castleman & Page 2014). While much work in this area has shown great promise, researchers rarely describe how they crafted their message content or the framing used during the behavioral intervention. In this paper we help fill this gap by testing the differential impact of six distinct theory-driven motivational appeal messaging campaigns on survey response.

Drawing from Leverage-Salience Theory as well as the theoretical work of Singer and Ye (2013), McClelland's *Needs Theory,* Cialdini's *Social Norms*, and the extensive literature on extrinsic (i.e. financial) rewards, we examine the impact of six different psychological motivational appeals (altruism, reciprocity, achievement, affiliation, commitment and consistency, and extrinsic reward) on the survey response rate of grades 3-5 elementary education teachers in a large urban school district in the northeastern United States in the Spring of 2017. Teachers were randomized to receive one of six motivational appeals during the first week of the study and were re-randomized to receive a different appeal each subsequent week, conditional on not having completed the survey. Our research design allows us to empirically estimate within each week and across weeks the causal effect of the different motivational appeals on teachers' survey response. We observe the effect of each appeal on four different margins, which range in their time intensity from negligible (open email) and minimal (click survey) to moderate (start survey), to substantial (completing the ~25-minute survey). We discuss which financial and non-financial appeals perform best along each margin and provide empirical estimates of such classes of respondents as always-takers and never-takers. This research adds to the survey design and behavioral insights literature by providing researchers with empirical results of the differential effectiveness of distinct theory-driven psychological

appeals in the context of urban elementary school teachers. It has direct implications for researchers concerned about sample attrition, as the findings highlight potential opportunities to reduce sample attrition when designing the survey campaign component of research evaluations.

Theoretical Motivation

Leverage-Salience Theory, a form of Social Exchange Theory, conceptualizes survey response as a threshold in which an individual's propensity to respond is influenced by both the importance ('leverage') and saliency of each attribute in the survey request (Groves, Singer, & Corning 2000). Each attribute either increases (positive leverage) or decreases (negative leverage) the likelihood of an individual's survey response. For instance, a cash incentive may have positive leverage, but the subject matter of the survey may have negative leverage. The influence of these attributes on the "cooperation decision" (Groves, Singer, & Corning 2000) varies across individuals. At the same time, the saliency of each attribute indicates the importance placed on each attribute when making the request. For instance, the extent to which an interviewer emphasizes one attribute over another when conducting an in-person or phone interview, or the framing and messaging of an internet or mail survey. Threshold points should also vary by the magnitude of the request, such as the time commitment necessary to start or to complete a survey.

In the simple two arm randomized trial case, in which an individual is randomly assigned to either the control (0) or treatment (1) group, any underlying variation in the leverage individuals' place on the set of survey attributes should be equal across the groups. The same assumption should hold in randomized studies with multiple treatment arms. Therefore, if all individuals receive the same request, including the same message framing, then the nature of the

request should introduce no new source of differential nonresponse bias across treatment arms. The same holds for requests across multiple timepoints (e.g. follow-up requests to complete a survey) – as long as additional requests are uniform across treatment arms, the additional requests should not introduce a source of differential non-response bias.

On the other hand, making the framing of a request (or follow-up requests) conditional on an evaluation site or treatment arm does introduce a potential source of bias. If the propensity to respond and the cooperation decision are affected by variation in the saliency of the requests, then the set of respondents across treatment arms, for instance, may no longer be assumed to be alike. Even if the proportion of respondents is the same across arms, the underlying motivation of those who responded may be different as they were presented with different requests. The extent to which this is an issue depends on whether differential response (and nonresponse) are correlated with the outcomes of interest and the magnitude of the nonresponse (WWC 2015).

One way to implement multiple request framings while minimizing the threat to statistical bias is to randomize message framing across respondents. If initial and follow-up requests are allocated independent of treatment assignment, then variation in saliency of the requests should introduce no new source of differential non-response bias. It is under these conditions that we seek to assess the differential impact of six underlying psychological appeals (i.e. psychological attributes) on elementary education teachers' survey response rates in the context of a large urban school district in the northeastern United States.

Selecting Motivational Appeals

Singer and Ye (2013) describe three primary motivators for survey response: altruism, egoistic (e.g. extrinsic incentives), and reasons associated with the survey (e.g. survey length and topic).

Our decision to use an extrinsic rewards appeal as one of the six motivational appeals draws on their work and the extensive literature suggesting that extrinsic motivators, such as goods and monetary (financial) rewards, positively influence response rates.[2] Incentive prepayments have been shown to improve response rates over promise of payment upon completion (Collins et al. 2000), while the size of incentives also enhances response rates when paid in advance (Church 1993). More recently, Göritz's (2006) meta-analysis suggests a significant impact of material incentives on starting and completing web-based surveys.

Despite the strong evidence of the contribution of extrinsic motivations on response rates, there has been relatively little empirical work conducted on the psychology underlying the decision of a potential survey respondent to participate or not. Singer (2002) went so far as to argue that most research on survey responses has been non-theoretical. We derive our additional motivational theories of survey uptake from the theoretical and empirical work of Robert Cialdini (*Social Norms Theory*; Cialdini & Trost 1998) and aspects of David McClelland's *Needs Theory* (McClelland 1985).

The theory of social norms was developed by Cialdini and colleagues to understand the ways in which people think and behave within social contexts. Social norms include what people believe others usually do and what people believe most others approve or disapprove of (Cialdini, Kallgren, & Reno 1991). Three important social norms identified by Cialdini are reciprocity, commitment and consistency, and altruism. Reciprocity is the feeling of obligation to reciprocate another person's concession with a concession of one's own, which can manifest as acting in kind or moving one's position in accordance with another's adjustment (Cialdini et al. 1975). These types of moderation strategies have been found to produce a significant increase in compliance with the request (Cialdini & Trost 1998; O'Keefe & Hale 1998). For commitment

[2] For early work in this area see research by Huck & Gleason 1974, Armstrong 1075, and Church 1993.

and consistency-based compliance, a central assumption is that people act consistently with their self-views and prior commitments in order to serve the ultimate motivation of maintaining or enhancing the self-concept (Cialdini & Trost 1998). Lastly, and paralleling Singer and Ye's (2013) work, Cialdini and colleagues argue selfless behaviors create bonds of connection to others through altruistic acts (Cialdini & Kenrick 1976). Cialdini, Brown, Lewis, Luce & Neuberg (1997) expanded on this idea with a theory of altruism-empathy in which acting altruistically leads to a greater sense of self and thus, in the overlap, selflessness can be seen as an act of helping the self.

The second overriding psychological framework used in this work focuses on McClelland's needs theory, a motivational model that seeks to explain peoples' motivation for thought and action. It posits that people are driven by and interpret the world and its functions through the lens of their motivating drive - a need for *power*, a need for *affiliation*, or an underlying desire for *achievement* (McClelland 1985). The importance of each of these three needs varies among individuals and the level of people's motivation in different situations depends upon their personal need orientation. In this study we focus on the need for affiliation and the need for achievement. Achievement motivation theory (Atkinson 1957; McClelland 1961) states that some peoples' motivations and activities are driven by an expectation that their performance will lead to excellence and success and, alternatively, that lack of action or fear of other action will arouse a fear of failure. Thus, the theory of achievement motivation focuses primarily upon subjects' internal grappling and resolution of the conflict between the opposing tendencies of achieving success and avoiding failure. In the drive for affiliation, McClelland and colleagues theorized that highly affiliative individuals wish to maintain others' affections and are motivated to seek out interpersonal relationships and friendships. In work environments, the

drive for affiliation leads highly affiliative individuals to feel rewarded by being involved in work that involves a high level of interaction with others (McClelland 1982; McClelland 1985). In a series of studies using an assessment of subjects' interpretation of pictures of human activities, McClelland and his colleagues and adherents built a strong empirical base for his theory (Winter & McClelland 1978; Winter 1987; Pennebaker 2011).

Research Questions

This study includes two sets of research questions. (1) Our primary goal is to examine the impact of distinct motivational appeals on inducing survey response. With the large literature on extrinsic incentives, we hypothesized that the extrinsic rewards appeal would yield the highest response rate of our six appeals. Conversely, we hypothesized that the altruism appeal would yield the lowest response. We expected the impact of the other four appeals to fall between these two, but did not have an ex ante belief of which would have the largest effect on survey response.

(2) Our second research question contributes to and expands upon the literature on respondent types in the potential outcomes framework - always-takers, compliers, defiers, and never-takers (Angrist, Imbens, & Rubin 1996). Our goal is to obtain empirical estimates of the proportion of a sample that should reasonably be expected to be nudged to completing a survey through targeted psychological appeals and savvy messaging strategies (the compliers), as well as the proportion of always-takers and never-takers. Our survey design presents a unique opportunity to add empirical estimates to these theoretical types of respondents. By asking what proportion of teachers in our sample were likely to respond to any appeal, were persuaded to respond through a motivational appeal, and did not respond at all, we are able to place empirical

estimates around these classes of respondent types.

Study Setting

This study was part of the Spring 2017 follow-up survey assessment of an evidence-based multi-school study of the Ongoing Assessment Project (OGAP), a mathematics professional development and support program, conducted by the Consortium for Policy Research in Education (CPRE) in the Philadelphia School District. The research efforts were funded by a math science partnership (MSP) grant from the National Science Foundation. OGAP seeks to develop teachers' ability to analyze student work products and to assess students' level of mathematical understanding along domain specific learning progressions (i.e. additive, multiplication, fractions, etc.). Teachers are taught to utilize this information to inform their instructional responses to improve students' understanding of the content. The 2016-2017 study consisted of a second cohort of schools receiving OGAP. The first cohort was part of a two-year (2014-2016) multi-cluster randomized control study in Philadelphia and an adjoining district. The 2016-2017 study included mathematics teachers in schools that received training at a summer institute and ongoing support throughout the school year ('treatment' schools), and mathematics teachers in schools that did not ('control' schools). Control schools consisted of a demographically similar set of schools that attended the Philadelphia School District's summer institute in literacy. Teachers were surveyed in the summer/fall of 2016 (pre) and spring of 2017 (post) to examine any differential gains in the ability to assess students' learning progression between teachers who did and did not receive ongoing OGAP support.

Research Design

Motivational appeals

The survey randomization utilized six distinct motivational appeals. Each appeal included targeted messaging in the subject line of the email, in the body of the email, and a tailored image. For instance, under the altruism appeal for teachers who received OGAP professional development throughout the year, the subject line of the email read "Show us your giving nature by completing your OGAP survey today." For teachers in the control group who did not receive OGAP professional development, the subject line of the email read "Show us your giving nature by completing your NSF survey today." The only language in the appeals that varied between teachers who did and did not participate in OGAP was the substitution of National Science Foundation and NSF for Ongoing Assessment Project and OGAP, respectively. The extrinsic reward appeal included the financial reward in the subject line, in the first line of the email text, and in the associated image. Under all other appeals the financial reward was mentioned solely in the last line of the appeal (all survey participants received $25 for completing the survey). The six subject line messages are detailed below (for full versions of the email subject, body, and image, see Appendix figures A1-A8):

Achievement - "Cross the finish line by completing your OGAP survey today!"

Affiliation - "Join your peers and complete your OGAP survey today!"

Reciprocity - "Help us help you – complete your OGAP survey today!"

Commitment and consistency – "Demonstrate your commitment to education and complete your OGAP survey today!"

Extrinsic reward - "Receive your $25 reward for completing your OGAP survey today!"

Altruism – "Show us your giving nature by completing your OGAP survey today!"

Survey Randomization

During the spring of 2017, 405 teachers were randomly assigned to receive one of the six motivational emails.[3] The email contained a link to their OGAP/NSF assessment as well as a unique survey code to access the assessment. The random design ensured that at survey onset there should not have been any observed or unobserved differences in a teacher's innate proclivity to respond to the survey across theme.

The survey randomization occurred over a six-week span with a maximum of twelve emails per teacher. We count the beginning of each week as Thursday mornings, as the first email was sent on a Thursday at approximately 6:05am (May 4, 2017). If a teacher did not complete the survey by the following Monday afternoon they were sent a 2nd email under the same motivation appeal at approximately 6:05 am that Tuesday. If a teacher had not completed their survey by the end of the first week (i.e. mid-afternoon that Wednesday), they were randomly assigned to receive a different motivational email for that Thursday.[4] Surveys were administered through Constant Contact, an online marketing company that can be used to distribute messages to listservs.

This process was repeated for each week of the study.[5] No teacher ever received the same motivational appeal in different weeks. Although a slight time gap existed between when the research team determined to send a subsequent email to a teacher (during the afternoon) and

[3] The randomization was conducted using Stata's *runiformint* command. For reproducibility, a seed number was randomly selected from the serial number of a U.S. treasury bill.
[4] The survey campaign sent all emails at approximately 6:05am according to the following schedule: For Week 1: Th 5/4 and Tu 5/9; Week 2: Th 5/11 and Tu 5/16; Week 3: Th 5/18 and Tu 5/23: Week 4: Th 5/25 and Tu 5/30; Week 5: Th 6/1 and Tu 6/6; Week 6: Th 6/8 and Tu 6/13
[5] A different unique U.S. Treasury bill serial number was used each week to set the seed number for the rerandomization of teachers to new appeals.

when the email was sent (approximately 6:05am the following day), in practice no teacher completed the survey during any of these gaps across the entire survey time-period. Across the six weeks, ten teachers were unable to be reached (for part or all of the study) because their emails bounced and five more teachers unsubscribed to constant contact, yielding an analytic sample of 390 unique teachers and 3,347 teacher observations.[6]

Outcome Measures

The study includes four sequential binary outcome measures – whether the teacher (1) opened the email; (2) clicked on the link in the email to the survey; (3) started the survey; and (4) completed the survey (0 no, 1 yes for each outcome). Email opens and survey clicks were observed through Constant Contact while survey starts and completions were observed from survey logs provided by the third party survey administrator.

When deciding to open an email, teachers only observed the subject line and potentially the first few words of the email body as a preview (depending on the browser). We interpret variation on this first outcome measure across motivational appeals as capturing the effect of a subject line 'teaser' on teachers' decision to open the email. It does not speak to changing teacher behavior on other margins, including starting or finishing the survey.

The effect of the full motivational appeal (subject, body, and image) is observed for the other three outcomes. Variation in clicking the survey link and starting the survey capture the differential impact an appeal has on eliciting a quick behavioral response (i.e. response on the extensive margin). It takes little effort to click on a survey link and there is little time commitment involved in inputting a code to start a survey. On the other hand, the final outcome of interest, completing the survey, can be interpreted as representing the differential ability of an

[6] Mean imputation was used to impute the years of experience for one teacher.

appeal to motivate a teacher to spend upwards of 25-30 minutes completing the survey (i.e. response on the intensive margin).

Empirical Analysis

We first descriptively examine overall response rates by motivational appeal for each outcome of interest, as well as the weekly response rates and cumulative response rates. We then examine how random assignment to a motivational appeal influenced survey participation in a panel data logistical regression framework. Lastly, we discuss how the research design can be used to calculate empirical estimates of respondent types in the potential outcomes framework.

Under the logistic function, g, let p represent the probability of a teacher completing each outcome of interest (click email, open, start, or complete survey) and $ln\left(\frac{p}{1-p}\right)$ represent the log odds of doing so. The probability of each outcome of interest can be specified as a linear combination of the log odds of the logistic function, such that

$$g(p) = logit(p) = log\left(\frac{p}{1-p}\right) = \beta_0 + \delta_1 M_{it} + \beta_1 Tr_i + \beta_2 W_{it} + \beta_3 W_{it}^2 + \beta_4 D_{it} + \beta_5 X_i + \varepsilon_{it} \quad (1)$$

Under the above specification, the δ parameter is the key coefficient of interest as it represents the effect of the motivational appeal on the response of teacher *i* at time *t*. *M* represents which of the six motivational appeals each teacher received at each time point (a series of dummy variables). A priori we hypothesized that the altruism appeal would be the weakest of the six appeals and therefore should be considered as the reference appeal. As this spring 2017 randomized teacher survey design was part of a broader research agenda with teachers that did (treatment) and did not (control) receive mathematics professional development, *Tr* is an

indicator that accounts for the professional development treatment assignment. Week, *W,* and quadratic week, W^2, account for the potential for survey exhaustion (i.e. declining survey response) across weeks. *D* is a binary indicator for day of the week (Thursday-0 or Tuesday-1), which accounts for whether responses were more likely during the initial or follow-up email of an appeal. X_i represents a vector of teacher-level control variables (years teaching, measured as continuous, indicator for math teacher, indicator for special education teacher, and indicator for English language learner/English as a second language teacher).

Although the weekly randomization of the appeals is a key methodological strength of this study, assuming the probability of each outcome of interest is independent across weeks is a strong claim, as it assumes no carryover effect from a previous week's appeal to the current week's response. Stated formally, let $M \in$ appeal set A (altruism, achievement, affiliation, reciprocity, commitment and consistency, extrinsic reward) and A_{it} a permutation of ordering of appeals up to and including time *t*, e.g. $A_{i3} = M_{it}, M_{it-1}, M_{it-2}$. Independence of response across weeks assumes $p_{it} \perp A_{it-1} \; \forall \; A_{it-1}$. For instance, $p_{it}|altruism_{it}, reciprocity_{it-1} = p_{it}|altruism_{it}, extrinsic\ reward_{it-1}$. To account for the potential of a carryover effect, we include a one-week lag of the previous week's appeal as a covariate in an additional specification (M2). We assume that any lagged effect operates strictly through the most recent previous appeal, and that the influence of further previous appeals (i.e. *t-2)* is negligible.

$$g(p) = logit(p) = log\left(\frac{p}{1-p}\right) = \beta_0 + \delta_1 M_{it} + \delta_2 M_{it-1} + \beta_1 Tr_i + \beta_2 W_{it} + \beta_3 W_{it}^2 + \beta_4 D_{it} + \beta_5 X_i + \varepsilon_{it} \quad (2)$$

Results

Descriptive Results

Our ex ante assumption that altruism would be the weakest of the six appeals is supported in the empirical results. The response rate for altruism was almost always lower than for the other appeals across each of the four outcomes. For instance, the proportion of teachers who opened an altruism appeal over the full time-period was the lowest of all appeals (48.1%, Table 1, Panel B). We therefore utilize the altruism appeal as our reference category for all results.[7]

First Week

During the first week of the experiment, nearly 60% of teachers opened either the first survey email on Thursday or the follow-up email (under the same theme) on Tuesday (232 of 390; see Table 1, Panel A).[8] The conversion rate of moving from opening an email to clicking on the survey link was nearly 40% (89 of 232). Of teachers who clicked on the survey link, 86.5% went on to start the survey (77 of 89), and nearly 80% of teachers who started the survey completed it (61 of 77). Overall, nearly 20% (77 of 390) of teachers started and 15.6% (61 of 390) completed the survey during the first week.

Descriptively, there was no significant difference in the likelihood of opening an email by theme over the course of the first week. Teachers receiving the extrinsic reward appeal in the first week were, however, more likely to click on the survey (27 percentage points), start the

[7] To assess the differential impact of each appeal on teachers' response rates, we compare the impact of the other appeals as compared to the weakest appeal, which is typically altruism. Even when the response rate under altruism isn't the weakest for a particular outcome, there is less than one-half a percentage point difference between the response rate under altruism and the lowest appeal (see Table 1 Panels B and C). We therefore compare every other appeal to altruism for each outcome of interest. We acknowledge that using the appeal with the lowest response has the potential to inflate the odds ratio of other appeals. As a sensitivity to our findings we also test all empirical models using alternative motivational appeals as the reference category (achievement, affiliation, reciprocity, or commitment and consistency), the results of which are described later in the text.

[8] In Table 1 Panels A and B the denominator is the total number of teachers.

survey (21 percentage points), and complete the survey (20 percentage points) than teachers who received the altruism appeal ($p<.01$ for each).

Full Time-Period

Across the six weeks of the survey experiment (see Table 1, Panel B), 87.4% of teachers opened at least one email and over sixty percent of teachers clicked on at least one link to the survey. 57.4% (224) started the survey and 48.4% of teachers completed the survey (185). This indicates that 12.6% of teachers were never-respondents, receiving two emails for each of the six motivational appeals but not responding to any of them at any point during the survey experiment.

Of teachers who started the survey, nearly 42% did so after receiving an extrinsic reward appeal (Table 2). The next highest percentage was for the reciprocity appeal, at just under 15%. Similarly, 43% of teachers who completed the survey did so after receiving an extrinsic reward appeal (and before receiving a different appeal). The next highest completion percentage occurred under the affiliation appeal (13.5%). Reciprocity was 3rd, at 12.4%.

Figures 1a-1d graphically display the open, click, start, and complete response rates by theme by week. The results are largely consistent across weeks. They detail the sizeable gap in response rate by week between teachers that received the extrinsic reward appeal and other appeals. For instance, the percentage of teachers clicking on the survey under the extrinsic appeal never reached below 29%. It never reached above 20% for any of the other appeals other than week one.

Figures 2a-2d display the cumulative effect across the six-week time-period for the same

outcome measures as in Figures 1a-1d, respectively.[9] While the extrinsic reward clearly outperforms the other motivational appeals on each outcome, the reciprocity appeal is seen to perform second best for opening, clicking, and starting the survey across the full study time period.

Descriptive t-tests mirror the results seen in the figures. Two themes outperformed the reference altruism appeal across the full experimental time frame. Receiving an extrinsic reward appeal was positively associated with each outcome ($p<.001$). Under this appeal teachers' responses were 18, 27, 25, and 18 percentage points higher for opening the email and clicking, starting, and completing the survey, respectively, than teachers who received the altruistic appeal (Table 1, Panel B). Receiving the reciprocity appeal ('*Help us help you'*) was also positively associated with a teacher ever clicking and ever starting the survey (both $p<.1$) as compared to teachers who received the altruism appeal. There were no differences between reciprocity and altruism for completing the survey.

Because teachers were able (and observed) to start a survey at one point but finish the survey days later, we interpret additional emails as providing additional appeals for a teacher to complete their survey. We therefore respecify each outcome in these descriptive findings as a percentage of emails sent.[10] Results are reported in Table 1, Panel C. The impact of the reciprocity appeal on the likelihood of clicking the survey remains significant ($p<.05$). Additionally, the reciprocity appeal is observed to have induced more teachers to open the email than the altruism appeal ($p<.1$), at a rate of 41%, which was 2nd only to the extrinsic appeal. The extrinsic findings are similar under both specifications.

A decline in response rate is expected with longer exposure to the survey campaign. This

[9] The denominator is the total number of teachers that received the appeal.
[10] Rather than as a proportion of teachers sent each theme.

may be attributable to such factors as survey exhaustion and the potential increase of never-takers in the remaining sample. However, the decline in the response rate across motivational appeals (as measured by the decrease in slope between weeks) was not uniform (see Figures 2a-2d). The rate of decline was noticeably lower under the extrinsic reward appeal than the other appeals. By the end of week six, teachers still were noticeably responding to the extrinsic rewards appeal. In total, over 20% (80) of the 390 teachers in the analytic sample completed the survey after receiving this appeal. The next closest was affiliation, at 6.4% (25). In contrast, with nearly flat slopes, the week-to-week improvement in average response across four of the five other themes was minimal by the end of the survey (other than for opening the email). The sole exception is the reciprocity appeal, which had the largest response rate of all non-extrinsic appeals (excluding survey completion).

While this survey's randomization occurred independently of whether a teacher worked in an OGAP treatment or control school during the 2016-2017 school year, it is possible that survey response rates varied between teachers working in schools that received ongoing OGAP professional development and teachers in schools that did not. There were no differences in the likelihood of opening an email across teachers in treatment and control schools. Teachers in control schools had a 2.5-3 percentage point higher response rate for clicking, starting, and completing the survey ($p<.01$). Subanalyses by theme indicate that teachers in control schools were more likely to respond to the achievement, commitment and consistency, and extrinsic appeals (not reported).

We also examined how survey completion varied across themes, conditional on starting the survey. Only 20 teachers completed the altruism appeal (the fewest of any appeal), and yet altruism had the highest conversion rate from starting to completing (87%). The second best

conversion rate was affiliation at 83%. Extrinsic reward was third (79%). Of all teachers who completed the survey, 61% did so in one session (only logging in once). Fifteen teachers started the survey under one appeal and finished under a different appeal. Of these, eight completed under the extrinsic rewards appeal, three under altruism, three under affiliation, and one under reciprocity. Six teachers took one additional week to complete the survey, three took two additional weeks, five took three additional weeks, and one took five weeks. There was no interesting variation across the psychological appeals in the number of sessions a teacher took to complete the survey, either in the first week or across the full study period. Nor did the amount of time teachers took to complete the survey significantly vary across weeks.

Multivariate Results

We examine results during the first week of the survey and then across the whole study period within a longitudinal logistic framework. We additionally examine teachers' responses conditional on not responding during the first week and include a one-week lag effect to account for any carryover influence of the previous week's appeal.

Consistent with the descriptive results, there was no observed difference in the odds ratio of a teacher opening an email by theme during the first week of the survey, after controlling for day of the week, whether the teacher was in an OGAP treatment or control school, and teacher characteristics. Mirroring the descriptive findings, teachers that received the extrinsic rewards appeal were over five times as likely to click on the survey than teachers who received the altruism appeal (odds ratio of 5.49) and over three times as likely to start and to complete the survey (Table 3, panel A). Higher responses were seen on Thursday, during the first email appeal, than during the follow-up email on Tuesday mornings ($p<.1$ or $p<.05$ for each outcome).

Across the entire survey time frame, both the extrinsic reward and reciprocity appeals induced greater teacher survey response. Teachers that received the reciprocity appeal were more likely to open the email, click on the survey, and start the survey, than teachers that received an altruism appeal (odds ratio of 1.45, 1.81, and 1.93, respectively; see Table 3, Panel B). Teachers that received the extrinsic reward appeal had odds ratios of 2.86, 6.19, 7.95, and 5.11 (all $p<.001$) for each of the four outcomes, respectively.

Response rates varied by whether teachers worked in a treatment or a control school. The response rates for clicking ($p<.05$), starting ($p<.05$), and completing the survey ($p<.01$) were higher for teachers in control schools than treatment schools. There was no difference in the odds of opening the email, however. Day of the week influenced completing the survey (but not the other outcomes), with lower odds of doing so on the follow-up survey on Tuesdays than the initial Thursday emails. Special education teachers were about half as likely to click the survey ($\beta=.56$, $p<.05$) and start the survey ($\beta=.51$, $p<.1$) as non-special education teachers. Having greater years of experience was associated with a very slight decrease in the odds of opening the email and completing the survey (not reported).

One-Week Lag Effect

Because teacher responses during the first week include those teachers that were likely to respond to any appeal (i.e. 'always-takers'), we also assess variation in teacher responses excluding the first week (weeks 2-6; Model 1, Table 4). We include a one-week lag effect (Model 2, Table 4) to account for any carryover influence of the previous week's appeal.

Excluding first week completers, on the lower time intensity margins of opening the email and clicking the survey, teachers were more likely to respond to both the reciprocity and

extrinsic reward appeals. These findings remain after the inclusion of the one-week lag.[11] Teachers were also more likely to respond to the extrinsic reward appeal on the higher time intensity outcomes of starting and completing the survey.[12] Most of the prior week appeals' coefficients are not significant, with a few exceptions. Having received the affiliation appeal in the previous week was negatively associated with clicking, starting, and completing the survey ($p<.05$) as were both achievement and reciprocity for completing the survey.

Model Sensitivities

Using the appeal with the lowest response rate as the reference category has the potential to inflate odds ratios. As a sensitivity to our findings we also test all empirical models using alternative motivational appeals as the reference category (achievement, affiliation, reciprocity, or commitment and consistency). Across all empirical models the significant effect of extrinsic rewards on each outcome remains. The positive effect of reciprocity on opening the email and clicking the survey also remains after accounting for the potential of a one-week carryover effect. In fact, teachers receiving the reciprocity appeal were more likely to open and/or click the email than every other appeal (except extrinsic rewards; see Appendix Table 1). Lastly, in an attempt to account for the testing of multiple hypotheses, we follow McDonald's (2014) implementation of the Benjamini-Hochberg (BH) procedure.[13] For the models in Table 4, we find that extrinsic reward remains significant for all four outcomes for both model specifications. Reciprocity remains significant for opening the email under both specifications and for clicking

[11] Teachers were also more likely to start the survey under the reciprocity appeal, but this effect is eliminated with the inclusion of the one-week lag.

[12] We also asked whether there is any discernible pattern to teachers' responses after receiving but declining to act on an extrinsic reward appeal. With the inclusion of the one-week lag, teachers who did not respond to an extrinsic rewards appeal were more likely to respond to a reciprocity appeal for opening the email ($p<.01$) and to an affiliation appeal for starting the survey ($p<.1$).

[13] We rank all independent variables by their p-value, multiply the rank by the number of independent variables, and divide by a false discovery rate. We select a false discovery rate of .25.

the survey without the one-week lag effect.

Respondent Types

Our survey design presents a unique opportunity to add empirical estimates to the theoretical types of respondents in the potential outcomes framework - always-takers, compliers, defiers, and never-takers – as relates to survey response (Angrist, Imbens, & Rubin 1996). In our design we observe up to six weeks of behavior, with teachers potentially receiving all six motivational appeals (and 12 total emails). In order to maximize our potential response rates to the survey (which is part of a broader ongoing research agenda), after the conclusion of the survey randomization we continued to email all teachers who had yet to complete the survey. Teachers received a maximum of eight additional extrinsic rewards emails between June 15 and June 26, 2017, potentially quadrupling their exposure to the extrinsic rewards appeal. Receiving additional extrinsic rewards appeals increased the proportion of teachers who started the survey 14.7 percentage points, from 57.4 to 72.1% (281), and the proportion who completed the survey nearly the same number of percentage points, from 47.4 to 61.0% (238).

Teachers who completed the survey within the first-week can be thought of as a combination of always-takers – teachers who would have completed the survey under any appeal – and compliers – teachers who complete the survey as a result of the specific appeal. In total, 15.6% (61) of teachers completed the survey in the first week and 19.7% began the survey (77). As previously discussed, completion rate varies by motivational appeal, from 8% for commitment and consistency to 32% for extrinsic reward in the first week, suggesting that a majority of first-week responders are compliers rather than always-takers. As a lower bound, the proportion of always-takers could be construed as the response rate for the worst-performing

appeal (8% under commitment and consistency), which corresponds to a scenario in which a real-world survey includes a particularly poor messaging strategy. For an average generic survey not specifically focused on messaging and not offering an extrinsic reward (a common scenario), the average first week response rate of the non-extrinsic appeals may present a better estimate of the proportion of always-takers. In our study the average first week response rate of the non-extrinsic appeals is 12.7%.

We classify never-takers as teachers who never responded to any of the motivational appeals during the study time-period, despite having received all six, as well as not responding to any of the eight additional extrinsic reward appeals sent between June 15 and June 26, 2017. In total, 8.7% of teachers never opened a single email, implying at no point did they engage with the survey campaign. 26.7% never clicked on the survey and 28.0% never started the survey. The difference in non-response between the clicking the survey margin and the opening the email margin identifies the portion of teachers who viewed the survey campaign and initial survey instructions but chose not to respond further. We believe the 28.0% of teachers who did not start the survey most accurately represents the set of never-takers, as it includes both teachers who did not engage at all with the survey campaign and those who did engage but actively decided not to begin th**e** survey. This margin most closely parallels the participatory aspect of most experimental treatments, such as engaging in a teacher professional development program or fidelity to an experimental drug trial.

Placed together, approximately 41% of teachers could be considered as either always-takers or never-takers.[14] Assuming monotonicity and therefore no defiers (i.e. no teachers chose not to complete the survey because they received it), then just under 60% of teachers may be

[14] The 12.7% 1st week non-extrinsic average for always-takers in combination with the 28.0% of teachers classified as never-takers.

construed as compliers. We therefore estimate that three-fifths of teachers have the potential to be nudged to completion through targeted psychological appeals and savvy messaging strategies (and an additional 13% would complete regardless of messaging). In our study, we calculate that approximately 27% of these compliers completed the survey during the randomization under the five non-extrinsic rewards appeals, 32% completed the survey across the six-week randomization because of the extrinsic reward, and an additional 23% were nudged to completion through additional exposure to extrinsic rewards subsequent to the six-week randomization.[15] The remainder of complier teachers (18%) are those that started but did not complete the survey.

Discussion

In this study we conducted a six-week randomized teacher survey campaign as part of the spring 2017 follow-up survey to a mathematics professional development program in a large urban school district. We were interested in assessing the impact of appealing to different underlying psychological motivations on teacher response rates, including reciprocity, extrinsic reward, and four other theory-driven appeals. We examined variation in teachers' responses to the appeals on four margins, which ranged from least to most time intensive: opening an email and clicking, starting, and completing the survey. We estimated that approximately 60% of the teachers could be construed as compliers and therefore had the potential to be nudged to completion through targeted psychological appeals and savvy messaging strategies.

The extrinsic rewards and reciprocity appeals outperformed all other appeals across a variety of time periods and outcomes. Both appeal types were effective under just the subject line

[15] Of the 390 teachers, we estimate 231 are compliers (59.3%). Of these, 63 completed the survey not in the first week and not under the extrinsic rewards appeal (42 completed in the first week). An additional 80 teachers completed the survey under the extrinsic reward appeal – 19 in the first week and 61 later. Of those in the first week, we estimate that 7 would have completed the survey irrespective of the appeal.

'teaser' and under the entire appeal (subject, body, and image). The empirical results for extrinsic reward are consistent with previous work on the power of financial incentives (e.g. Goritz 2006). Additionally, this work provides empirical results supporting the influence of Cialdini's social norm of reciprocity along margins of lower time-intensity, while there is no evidence to support the other social norms (commitment and consistency, altruism) or McClelland's motivational needs theory (achievement, affiliation), at least in this setting.

Because teacher time is highly regimented throughout the working week and subject to tight constraints, a financial reward can change the cost-benefit tradeoff of how a teacher chooses to spend their limited free time – completing a survey for someone else or conducting another activity that more directly tracks to a teacher's individual utility function (e.g. working on a lesson plan, checking social media).

Interestingly, the reciprocity appeal also appears to change this tradeoff calculation. Steeped in language of assistance ('Help us help you'), teachers who received the reciprocity appeal were more likely to open the email and click the survey than the reference appeal of altruism. Steeping our campaign in terms of assistance also elicited a response beyond what is observed in the non-extrinsic appeals – affiliation, achievement, commitment and consistency. While ex ante we cannot distinguish between groups of teachers more likely to respond to one type of appeal, the empirical results suggest that there is a set of teachers particularly aligned with appeals of assistance.

We caution, however, that response rates vary depending on the margin of interest. Opening an email can be almost automatic and clicking a survey link is a negligible time investment. The upfront cost to starting a survey is also minimal but completing our survey took substantial time (approximately 25 minutes). Because the opportunity cost to completing the

survey was much higher than on the other margins, it is not surprising the extrinsic reward was the only appeal that significantly impacted survey completion. Our findings should therefore be interpreted with a note of caution – we find that reciprocity induces a higher response on outcomes which require lower time commitments. While promising, we cannot generalize to what response rates might have been had the survey been shorter (e.g. 5-10 minutes). We also conducted the survey amongst grades 3-5 elementary school teachers in a large urban school district in the Northeast and cannot generalize to educators in other settings or industries. Further research is necessary to examine the effect of non-financial motivations in additional settings and under other survey conditions.

As two final contributions to research on non-response and survey attrition, our work highlights the importance of piloting messaging campaigns at the onset of a study, while our research design of randomizing the initial and follow-up appeals provides an analytical approach that can minimize the threat of differential attrition while testing the efficacy of various motivational appeals. Greater response to reciprocity and extrinsic reward appeals emerged early within the study period, suggesting researchers could implement what we call an 'adaptive saliency' approach in which they adjust message framing and content after observing initial response rates, and continue to dynamically incorporate such feedback across a study's time frame. Doing so would enable researchers to more finely tune messaging to the specific context of their study, with a potential positive impact on survey response. While in this paper we have largely focused on optimal messaging in the context of survey non-response, this suggestion is also applicable to the framing and messaging of interventions (e.g. text campaigns to fill out a FAFSA; Castleman & Page 2016).

Conclusion

This study highlights the importance of message saliency for surveys in an educational setting, furthering both the survey design and behavioral insights literatures. In particular, financial incentives are especially powerful when viewed easily at the initial point of contact, i.e. within the subject line and at the beginning of an email. The study also provides empirical evidence that the social norm of reciprocity has the potential to improve teacher response, particularly along margins of lower time-intensity. In evaluative and other settings where researchers face tight budget constraints, researchers should consider highlighting the reciprocal aspects of the work. Researchers could also consider combining extrinsic and reciprocity appeals, such as a tiered system in which an extrinsic reward is offered only after initial non-completion. More broadly, as the research community continues to conduct large multi-site evaluations that can suffer from serious issues of sample attrition, it is important to recognize the contribution of message framing in survey response.

References

Albaum, G., Evangelista, F., & Medina, N. (1998). Role of response behavior theory in survey research: A cross-national study. *Journal of Business Research*, *42*(2), 115-125.

Angrist, J.D., Imbens, G.W., & Rubin, D.B. (1996). Identification of Causal Effects Using Instrumental Variables. *Journal of the American Statistical Association*, 91(434): 444-455.

Armstrong, J.S. (1975) Monetary incentives in mail surveys, *Public Opinion Quarterly,* 39, pp. 111–116.

Atkinson, J. W. (1957). Motivational determinants of risk-taking behavior. *Psychological review*, *64*(6p1), 359.

Basson, D. "Nonignorable Nonresponse." In Lavrakas, P.J., Ed. (2011). *Encyclopedia of Survey Research Methods*. Thousand Oaks, CA: Sage Publications.

Berger, J. M., Zelditch, M., Anderson, B. & Cohen, B. (1972). Structural Aspects of Distributive Justice: A Status Value Formulation. Pp. 119-46 in *Sociological Theories in Progress.* Vol. 2, Joseph Berger, Morris Zelditch, Jr., and Bo Anderson, eds.. Boston: Houghton Mifflin.

Bem, D. J. (1972). Self-Perception Theory, in Advances in Experimental Social Psychology, vol. 6, L. Berkowitz, ed., Academic, Press, New York.

Bettinger, E.P., Long, B.T., Oreopoulos, P., & Sanbonmatsu, L. The Role of Application Assistance and Information in College Decisions: Results from the H&R Block FAFSA Experiment. *The Quarterly Journal of Economics*, 127(3): 1205-1242.

Castleman, B.L. & Page, L.C. (2016). Freshman Year Financial Aid Nudges: An Experiment to Increase FAFSA Renewal and College Persistence. *Journal of Human Resources*, 51(2): 389-415.

Castleman, B.L. & Page, L.C. (2014). *Summer Melt: Supporting Low-Income Students Through the Transition to College*. Cambridge, MA: Harvard Education Press.

Center for Advanced Hindsight, Common Cents Lab. 2017. Annual Report.

Church, A.H. (1993) Estimating the effects of incentives on mail response rates: a meta-analysis, *Public Opinion Quarterly,* 57, pp. 62–79.

Cialdini, R.B., & Kenrick, D.T. (1976). Altruism as hedonism: a social development perspective on the relationship of negative mood state and helping. *Journal of personality and social psychology*, *34*(5), 907-914.

Cialdini R.B. & Trost M.R. 1998. Social influence: social norms, conformity, and compliance. ln *The Handbook of Social Psychology,* ed. DT Gilbert, ST Fiske, G Lindzey, 2: 151-92. Boston: McGraw-Hill. 4th ed.

Cialdini, R.B, Vincent, J.E., Lewis, S.K., Catalan, J., Wheeler, D., & Darby B.L. 1975. Reciprocal concessions procedure for inducing compliance: the door-in-the-face technique. *J. Personal. Soc. Psychol.* 31:206-15.

Cialdini, R.B., Brown, S.L., Lewis, B.P., Luce, C., & Neuberg, S.L. (1997). Reinterpreting the empathy–altruism relationship: When one into one equals oneness. *Journal of personality and social psychology*, *73*(3), 481-493.

Cialdini, R.B., Kallgren, C.A., & Reno, R.R. (1991). A focus theory of normative conduct: A theoretical refinement and reevaluation of the role of norms in human behavior. In *Advances in experimental social psychology*, 24, 201-234.91.

Collins, R.L., Ellickson, P.L., Hays, R.D. & McCaffrey, D.F. (2000) Effects of incentive size and timing on response rates to a follow-up wave of a longitudinal mailed survey, *Evaluation Review,* 24, 4, pp. 347–363.

Dillman, D.A. (1978). *Mail and telephone surveys: The total design method* (Vol. 19). New York: Wiley.

Dillman, D.A., Smyth, J.D., & Christian, L.M. (2014). *Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method.* New Jersey: John Wiley & Sons.

Festinger, L. (1962). *A theory of cognitive dissonance* (Vol. 2). Stanford university press.

Göritz, A.S. (2006). Incentives in web studies: Methodological issues and a review. *International Journal of Internet Science*, *1*(1), 58-70.

Groves, R.M. & Couper, M.P. (1998). *Nonresponse in Household Interview Surveys*. New York: John Wiley.

Groves, R.M., Singer, E., & Corning, A. (2000). "Leverage-Saliency Theory of Survey Participation." *Public Opinion Quarterly*, 64(3): 299-308.

Heckman, J.J. (1979). Sample Selection Bias as a Specification Error. *Econometrica*, 47(1): 153-161.

Homans, G. (1974). Social Behavior: Its Elementary Forms. New York: Harcourt, Brace & World.

Huck, S.W. & Gleason, E.M. (1974) Using monetary inducements to increase response rates from mail surveys, *Journal of Applied Psychology,* 59, pp. 222–225.

Kessler, J.B., Troxel, A.B., Asch, D.A., Mehta, S.J., Marcus, N., Lim, R., Zhu, J., Shrank, W., Brennan, T., & Volpp, K.G. 2018. Partners and Alerts in Medication Adherence: A Randomized Control Trial. *Journal of General Internal Medicine*, forthcoming.

Lavrakas, P.J. "Differential Nonresponse." In Lavrakas, P.J., Ed. (2011). *Encyclopedia of Survey Research Methods*. Thousand Oaks, CA: Sage Publications.

Mattila, V.M., Parkkari, J., & Rimpela, A. (2007). "Adolescent Survey Non-Response and Later Risk of Death. A Prospective Cohort Study of 78,609 Persons with 11-Year Follow-Up." *BMC Public Health*, 7(87).

McClelland, D.C. (1961). *The achieving society*. Princeton, NJ: Van Nostrand.

McClelland, D.C. (1985). Human motivation. Glenview, IL: Scott, Foresman.

McDonald, J.J. (2014). *Handbook of Biological Statistics (3rd Ed.)*. Baltimore, MD: Sparky House Publishing. http://www.biostathandbook.com/

O'Keefe, D.J., & Hale, S.L. (1998). The door-in-the-face influence strategy: A random-effects meta-analytic review. *Annals of the International Communication Association*, 21(1): 1-33.

Pennebaker, J.W. (2011). *The secret life of pronouns: What our words say about us*. New York: Bloomsbury Press.

Singer, E. (2002). The use of incentives to reduce nonresponse in household surveys. In Groves, R.M., Dillman, D.A., Eltinge, J.L., & Little, R.J.A. (Eds.), *Survey Nonresponse* (pp. 163–178). New York: Wiley-Interscience.

Singer, E. & Ye, C. (2013). "The Use and Effects of Incentives in Surveys." *The Annals of the American Academy of Political & Social Science*, 645(1): 112-141.

Thaler, R.H. & Sunstein, C.R. Libertarian Paternalism. *American Economic Review*, 93(2): 175-179.

Tourangeau, R. & Plews, T.J., Eds. (2013). *Nonresponse in Social Science Surveys: A Research Agenda*. National Research Council of the National Academies. Washington DC: The National Academies Press.

What Works Clearinghouse. (2015). *WWC Standards Brief for Attrition*. Washington, D.C.: Institute of Education Sciences. https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_brief_attrition_080715.pdf

Winter, D. G. & McClelland, D. C. (1978). Thematic analysis: An empirically derived measure of the effects of liberal arts education. *Journal of Educational Psychology, 70*(1), 8-16.

Winter, D. G. (1987). Leader appeal, leader performance, and the motive profiles of leaders and followers: A study of American presidents and elections. *Journal of Personality and Social Psychology, 52*(1), 196-202.

Main Figures

Figure 1a.

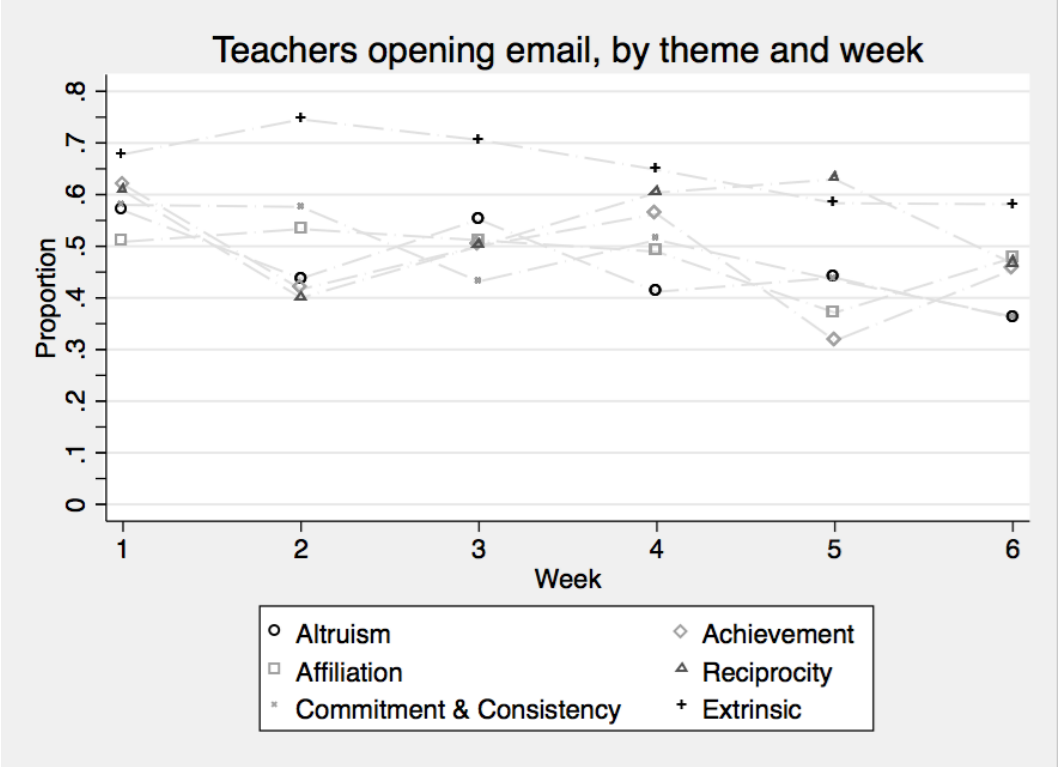


Figure 1b.

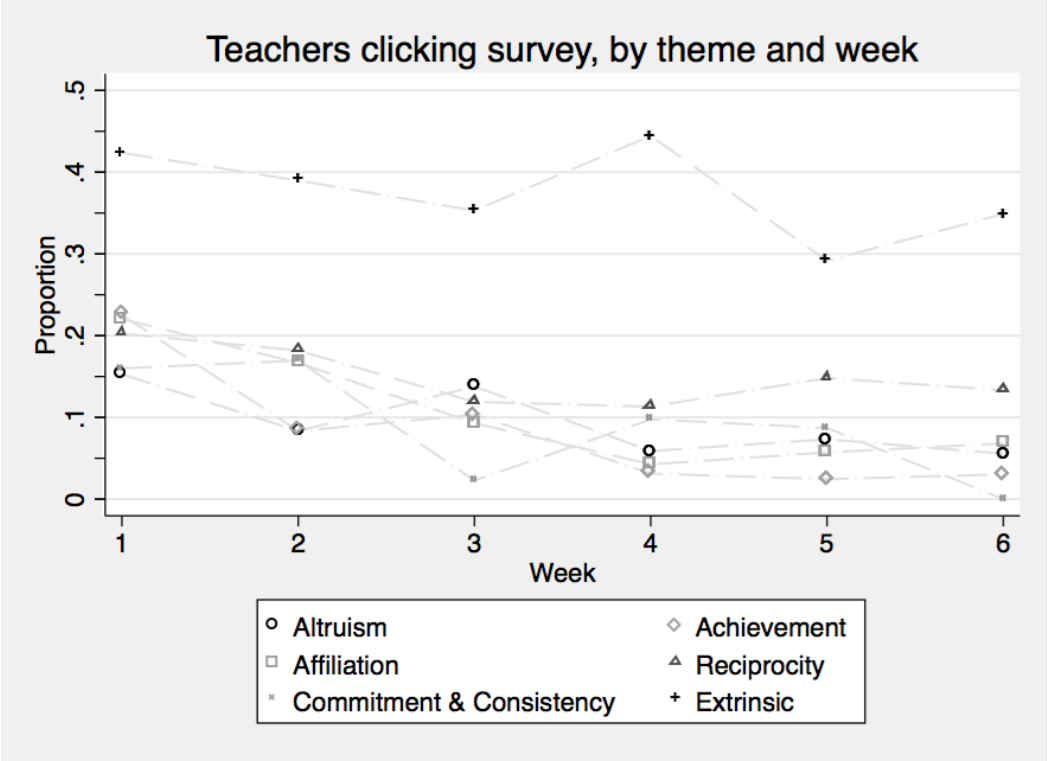


Figure 1c.

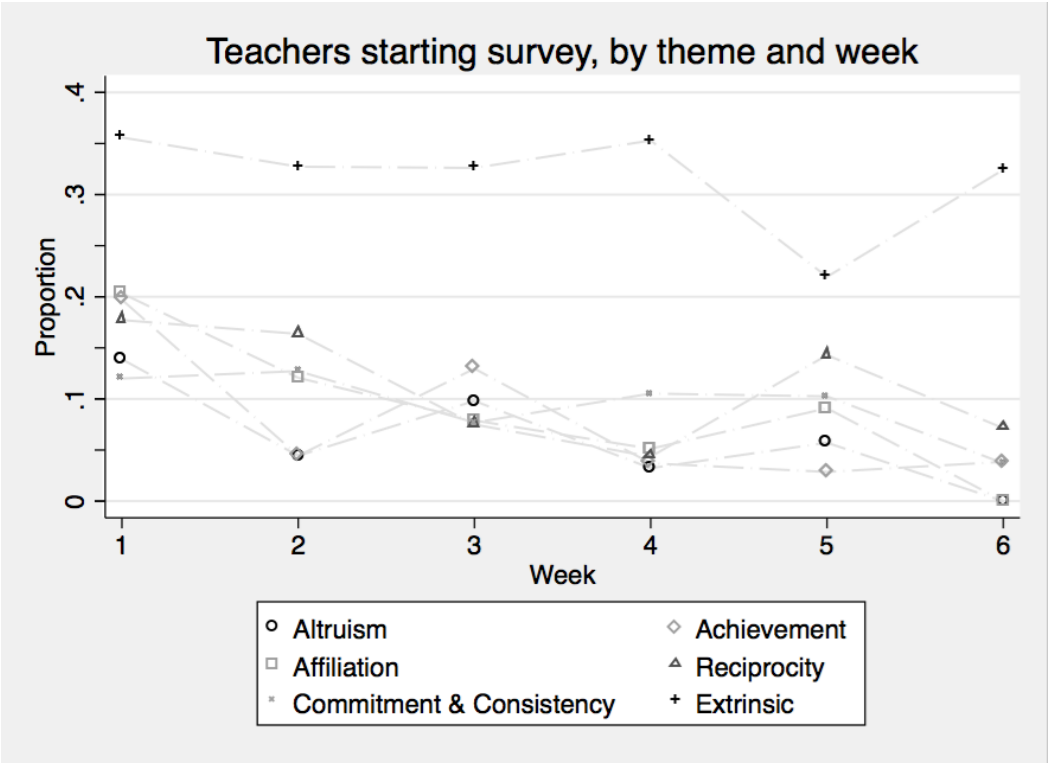


Figure 1d.

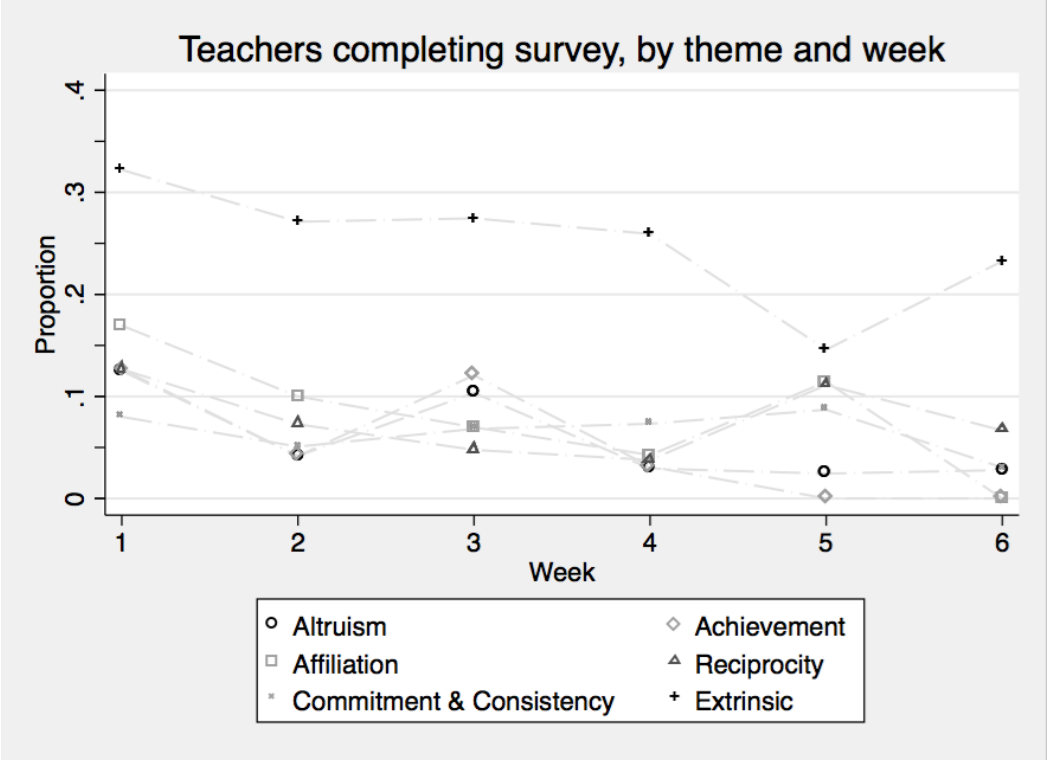


Figure 2a.

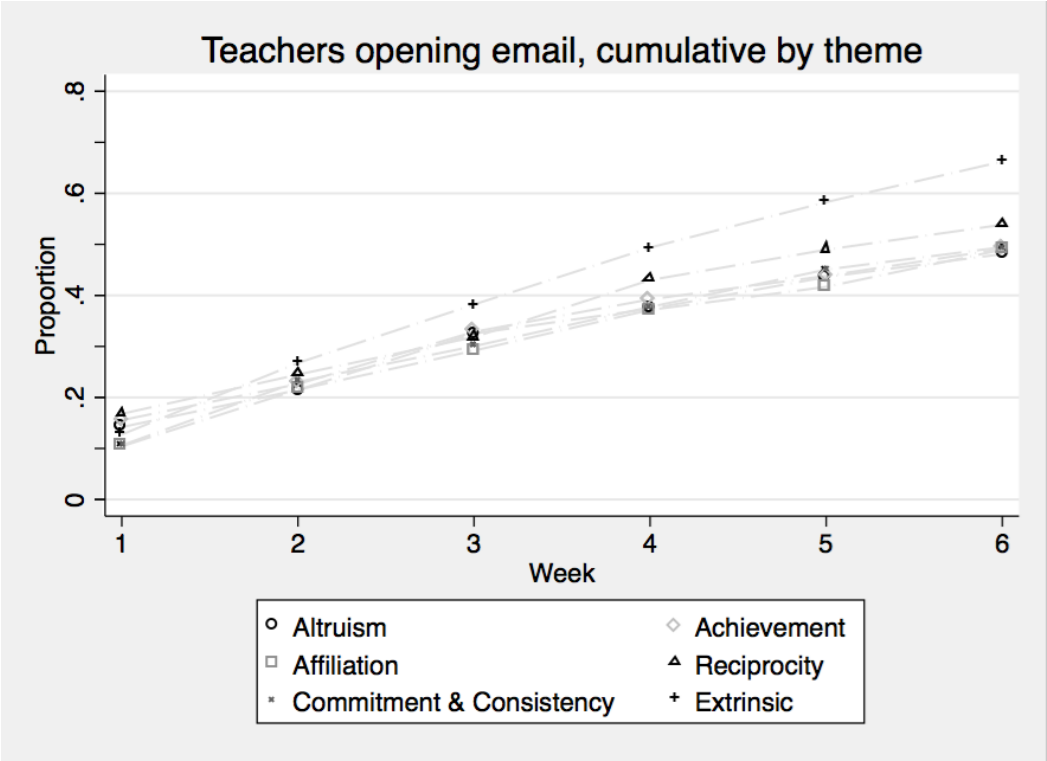


Figure 2b.

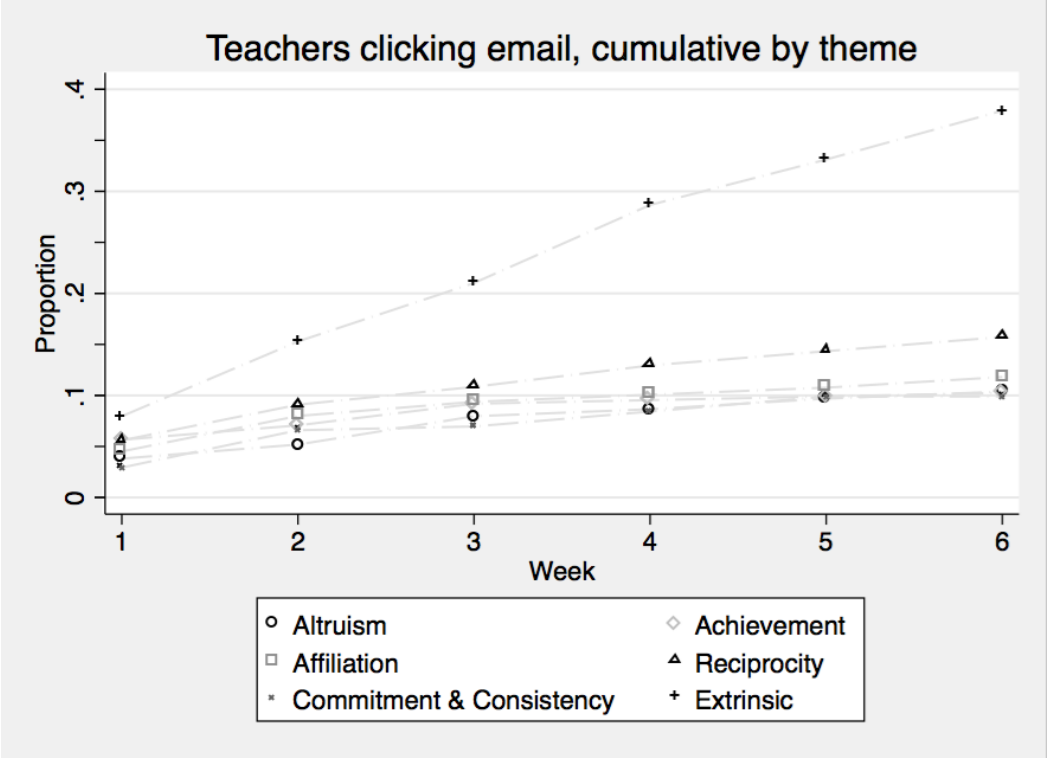


Figure 2c.

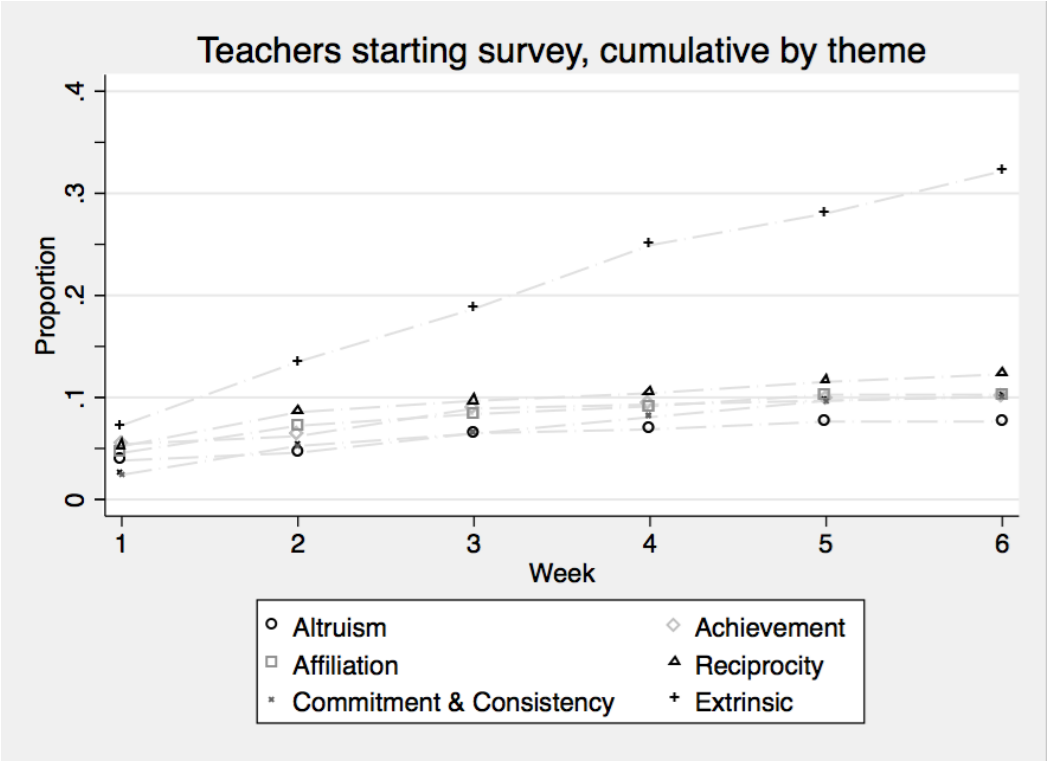
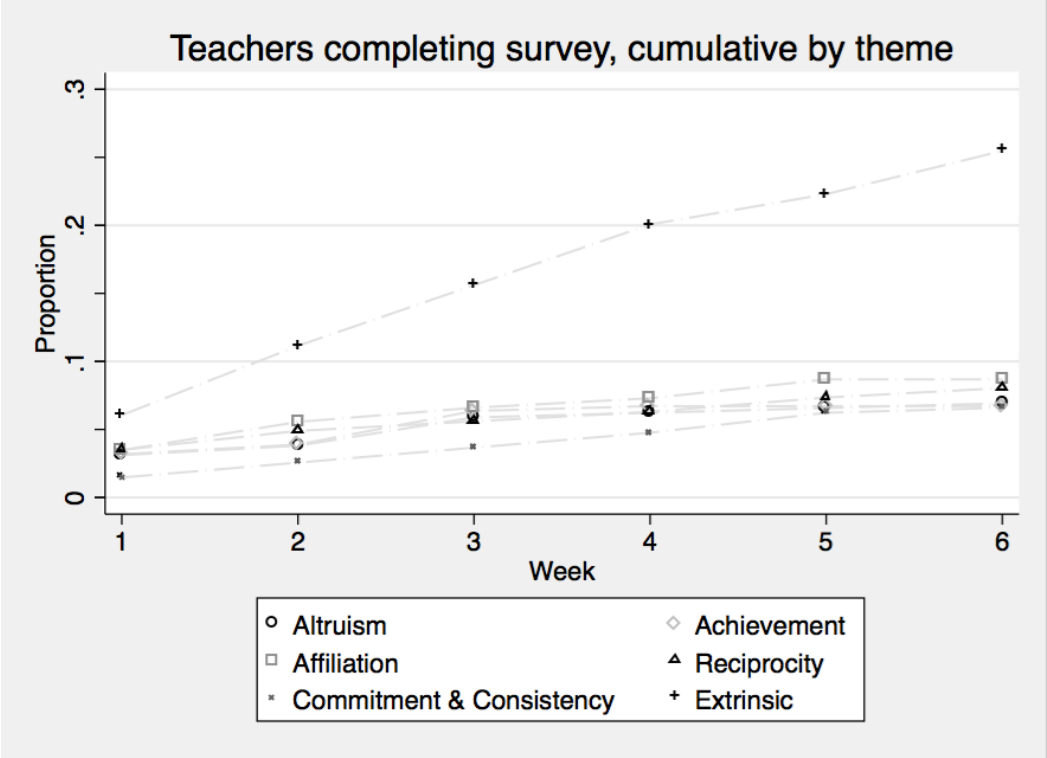


Figure 2d.



Main Tables

Table 1, Panel A. Week 1 by Teacher, Proportion Opening Email and Clicking, Starting and Completing Survey

Theme	Open (%)	Click (%)		Start (%)		Complete (%)	N	
Achievement	61.97	22.54		19.72		12.68	71	
Affiliation	50.85	22.03		20.34		16.95	59	
Reciprocity	60.76	20.25		17.72		12.66	79	
Commitment and Consistency	58.00	16.00		12.00		8.00	50	
Extrinsic Reward	67.80	42.37	***	35.59	**	32.20	59	**
Altruism[1]	56.94	15.28		13.89		12.50	72	
Total	59.49	22.82		19.74		15.64	390	

[1] Reference Category

Denominator calculated as the total number of teachers sent each theme in week 1.

Table 1, Panel B. By Teacher, Proportion Opening Email and Clicking, Starting and Completing Survey

Theme	Open (%)	N		Click (%)	N		Start (%)	N		Complete (%)	N
Achievement	49.12	283		10.25	283		10.08	258		6.71	283
Affiliation	48.96	288		11.81	288		10.27	263		8.68	288
Reciprocity	53.85	286		15.73	286	+	12.27	269	+	8.04	286
Commitment and Consistency	49.45	273		9.89	273		10.08	248		6.59	273
Extrinsic Reward	66.24	314	***	37.90	314	***	32.18	289	***	25.48	314
Altruism[1]	48.10	289		10.38	289		7.63	262		6.92	289
Any theme	87.44	390		61.79	390		57.44	390		47.44	390

[1] Reference Category

Denominator calculated as the number of teachers sent each theme.

Table 1, Panel C. By Theme, Proportion Opening Email and Clicking, Starting and Completing Survey

Theme	Open (%)	N		Click (%)	N		Start (%)	N		Complete (%)	N
Achievement	36.91	550		5.27	550		5.24	496		3.45	550
Affiliation	36.83	562		6.41	562		5.26	513		4.45	562
Reciprocity	41.08	555	+	8.83	555	*	6.40	516	+	4.14	555
Commitment and Consistency	37.94	535		5.61	535		5.20	481		3.36	535
Extrinsic Reward	50.78	577	***	22.01	577	***	17.78	523	***	13.86	577
Altruism[1]	35.56	568		5.46	568		3.88	515		3.52	568
Total	39.92	3347		9.02	3347		7.36	3044		5.53	3347

[1] Reference Category

Denominator calculated as the total emails sent by theme.

$^{+}$ $p \leq .10$, * $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$

For open survey, click survey, and complete survey, N represents the number of teachers that were sent the motivational theme and had not yet completed the survey.

For start survey, N represents the number of teachers that were sent the motivational theme and had not yet already started the survey.

Table 2. Percent of teachers who started and completed the survey, by theme[1]

	Start		Complete	
Theme	%	N	%	N
Achievement	11.61	26	10.27	19
Affiliation	12.05	27	13.51	25
Reciprocity	14.73	33	12.43	23
Commitment and Consistency	11.16	25	9.73	18
Extrinsic Reward	41.52	93	43.24	80
Altruism[1]	8.93	20	10.81	20
All starters (or completers)	100.00	224	100.00	185
As a proportion of all teachers	57.44	390	47.44	390

[1] Of teachers who started (or completed) the survey.

Table 3, Panel A. Longitudinal Logistic Regression Model Predicting Teacher Response, 1st week only[1]

Theme	Open Email	Click Survey	Start Survey	Complete Survey
Achievement	1.29	1.88	1.59	1.08
Affiliation	0.43	1.98	1.51	1.34
Reciprocity	0.93	1.96	1.45	1.12
Commitment and Consistency	0.97	1.10	0.87	0.62
Extrinsic Reward	2.21	5.49 ***	3.17 **	3.20 **
Treatment	0.73	0.63	0.61 +	0.59 +
2nd email (Tuesday)	0.70 +	0.55 *	0.60 *	0.62 +
Constant	3.62	0.02 **	0.05 **	0.08 *
N	741	741	730	741

+ p≤.10, * p≤.05, ** p≤.01, *** p≤.001

[1] Model reports odds ratio. Teacher controls include subject, if the teacher taught English to speakers of other languages, an indicator for special education teacher, and years of experience. Sample size denotes the number of emails sent. 390 teachers are included in the models.

Table 3, Panel B. Longitudinal Logistic Regression Model Predicting Teacher Response[1]

Theme	Open Email	Click Survey	Start Survey	Complete Survey
Achievement	1.13	0.98	1.54	1.00
Affiliation	1.10	1.25	1.52	1.32
Reciprocity	1.45 *	1.81 *	1.93 +	1.21
Commitment and Consistency	1.18	1.10	1.72	1.05
Extrinsic Reward	2.86 ***	6.19 ***	7.95 ***	5.11 ***
Treatment	0.85	0.65 *	0.56 *	0.52 **
Week	1.18	0.80	1.10	0.85
Quadratic Week	0.96 +	1.02	0.99	1.01
2nd email (Tuesday)	0.96	0.89	0.91	0.63 **
Constant	0.77	0.06 ***	0.01 ***	0.04 ***
ICC	0.59	0.25	0.44	0.20
N	3,347	3,347	3,044	3,347

+ p≤.10, * p≤.05, ** p≤.01, *** p≤.001

[1] Model reports odds ratio. Teacher controls include subject, if the teacher taught English to speakers of other languages, an indicator for special education teacher, and years of experience. Sample size denotes the number of emails sent. 390 teachers are included in the models.

Table 4. Longitudinal Logistic Regression Model Predicting Teacher Response, conditional on not responding during first week[1]

	Open Email		Click Survey		Start Survey		Complete Survey	
Theme	M1	M2	M1	M2	M1	M2	M1	M2
Achievement	0.84	0.80	0.62	0.55	1.22	1.04	0.89	0.68
Affiliation	1.11	1.03	1.01	0.87	1.42	1.17	1.35	1.09
Reciprocity	1.45 +	1.51 +	1.85 +	1.73 +	2.07 +	1.83	1.26	1.03
Commitment and Consistency	1.21	1.19	1.05	0.95	1.88	1.66	1.23	0.98
Extrinsic Reward	3.33 ***	3.52 ***	6.66 ***	6.43 ***	9.18 ***	9.79 ***	6.00 ***	6.25 ***
Lag	--		--		--		--	
Achievement	--	0.75	--	0.65	--	0.60	--	0.44 *
Affiliation	--	0.72	--	0.49 *	--	0.45 *	--	0.39 *
Reciprocity	--	1.13	--	0.79	--	0.66	--	0.41 *
Commitment and Consistency	--	0.95	--	0.74	--	0.79	--	0.63
Extrinsic Reward	--	1.19	--	0.88	--	1.08	--	1.01
Treatment	0.95	0.94	0.68 +	0.66 *	0.69	0.64	0.53 *	0.44 *
Week	1.88 *	1.91 *	0.84	0.90	1.35	1.81	1.47	1.96
Quadratic Week	0.91 **	0.91 **	1.01	1.00	0.96	0.94	0.94	0.92
2nd email (Tuesday)	1.05	1.05	1.07	1.08	0.95	0.89	0.61 *	0.66 +
Constant	0.24 +	0.26	0.04 ***	0.05 **	0.01 **	0.01 *	0.01 **	0.01 *
ICC	0.60	0.60	0.24	0.24	0.32	0.44	0.17	0.36
N	2,606	2,606	2,606	2,606	2,314 [2]	2,314 [2]	2,606	2,606

+ p≤.10, * p≤.05, ** p≤.01, *** p≤.001

[1] Model reports odds ratio. Teacher controls include subject, if the teacher taught English to speakers of other languages, an indicator for special education teacher, and years of experience. Sample size denotes the number of emails sent. 329 teachers are included in the models.

[2] N represents the number of emails sent to teachers that had not yet already started the survey. 313 teachers are included in the analysis.

Appendix Table 1. Longitudinal Logistic Regression Model Predicting Teacher Response with Reciprocity as the Reference Category, conditional on not responding during first week[1]

Theme	Open Email	Click Survey	Start Survey	Complete Survey
Altruism	0.66 +	0.58 +	0.55	0.97
Achievement	0.53 **	0.32 **	0.57	0.66
Affiliation	0.68 +	0.50 *	0.64	1.06
Commitment and Consistency	0.79	0.55 +	0.91	0.95
Extrinsic Reward	2.34 ***	3.73 ***	5.36 ***	6.05 ***
Lag	--	--	--	--
Altruism	0.88	1.26	1.52	2.43 *
Achievement	0.66 *	0.82	0.91	1.08
Affiliation	0.63 *	0.61	0.68	0.95
Commitment and Consistency	0.84	0.93	1.20	1.53
Extrinsic Reward	1.05	1.11	1.63	2.46 +
Treatment	0.94	0.66 *	0.64	0.44 *
Week	1.91 *	0.90	1.81	1.96
Quadratic Week	0.91 **	1.00	0.94	0.92
2nd email (Tuesday)	1.05	1.08	1.03	0.66 +
Constant	0.44	0.07 **	0.01 *	0.00 **
ICC	0.60	0.24	0.44	0.36
N	2,606	2,606	2,314 [2]	2,606

+ $p \leq .10$, * $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$

[1] Model reports odds ratio. Teacher controls include subject, if the teacher taught English to speakers of other languages, an indicator for special education teacher, and years of experience. Sample size denotes the number of emails sent. 329 teachers are included in the models.

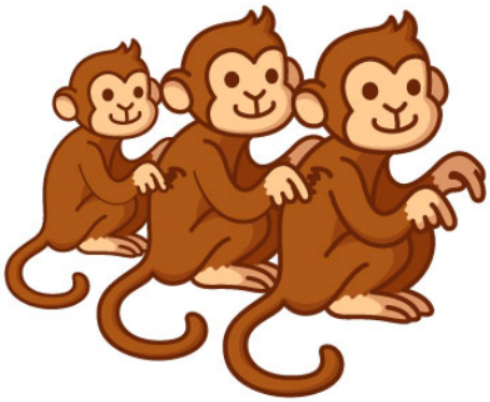
[2] N represents the number of emails sent to teachers that had not yet already started the survey. 313 teachers are included in the analysis.

Appendix Figures

A1. Reciprocity appeal for treatment teachers

From: OGAP <cpre@gse.upenn.edu>
Subject: Help us help you - complete your OGAP survey today!
Preheader:
Reply: cpre@gse.upenn.edu

CPRE | CONSORTIUM *for* POLICY RESEARCH *in* EDUCATION Graduate School of Education Penn GSE



You can help us help you improve your students' lives by completing your OGAP survey. Click the link below!

The Ongoing Assessment Project (OGAP) supports mathematics instruction in your school to help teachers meet the Pennsylvania Core Standards in mathematics.

Because we know your time is valuable, we will provide you with a $25 Amazon gift card for completing your survey!

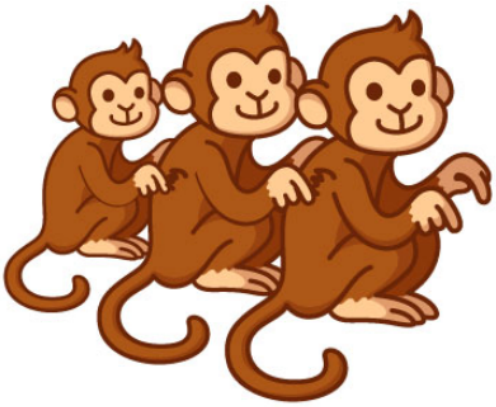
Please use your unique code Teacher ID to enter the survey!

CLICK HERE TO TAKE THE OGAP SURVEY

A2. Reciprocity appeal for control teachers

From: CPRE <cpre@gse.upenn.edu>
Subject: Help us help you - complete your NSF survey today!
Preheader:
Reply: cpre@gse.upenn.edu

CPRE | CONSORTIUM *for* POLICY RESEARCH *in* EDUCATION | Graduate School of Education Penn GSE



You can help us help you improve your students' lives by completing your NSF survey. Click the link below!

The National Science Foundation (NSF) supports mathematics instruction in the School District of Philadelphia to help teachers meet the Pennsylvania Core Standards in mathematics.

Because we know your time is valuable, we will provide you with a $25 Amazon gift card for completing your survey!

Please use your unique code Teacher ID to enter the survey!

CLICK HERE TO TAKE THE NSF SURVEY

A3. Extrinsic Rewards appeal for treatment teachers

From: OGAP <cpre@gse.upenn.edu>
Subject: Receive your $25 reward for completing your OGAP survey today!
Preheader:
Reply: cpre@gse.upenn.edu

CPRE | CONSORTIUM *for* POLICY RESEARCH *in* EDUCATION Graduate School of Education Penn GSE

Have summer reading plans? We'll help foot the bill. Receive a $25 amazon gift card for completing your OGAP survey today!

The Ongoing Assessment Project (OGAP) supports mathematics instruction in your school to help teachers meet the Pennsylvania Core Standards in mathematics.

Click the link below to complete your survey and get your $25 gift card!

Please use your unique code Teacher ID to enter the survey!

CLICK HERE TO TAKE THE OGAP SURVEY

A4. Extrinsic Rewards appeal for control teachers

From: CPRE <cpre@gse.upenn.edu>
Subject: Receive your $25 reward for completing your NSF survey today!
Preheader:
Reply: cpre@gse.upenn.edu

CPRE | CONSORTIUM *for* POLICY RESEARCH *in* EDUCATION — Graduate School of Education Penn GSE

Have summer reading plans? We'll help foot the bill. Receive a $25 amazon gift card for completing your NSF survey today!

The National Science Foundation (NSF) supports mathematics instruction in the School District of Philadelphia to help teachers meet the Pennsylvania Core Standards in mathematics.

Click the link below to complete your survey and get your $25 gift card!

Please use your unique code Teacher ID to enter the survey!

CLICK HERE TO TAKE THE NSF SURVEY

A5. Achievement appeal for treatment teachers

From: OGAP <cpre@gse.upenn.edu>
Subject: Cross the finish line by completing your OGAP survey today!
Preheader:
Reply: cpre@gse.upenn.edu

CPRE | CONSORTIUM *for* POLICY RESEARCH *in* EDUCATION Graduate School of Education Penn GSE

Help reach the goal of improving education and complete your OGAP survey linked to below!

The Ongoing Assessment Project (OGAP) supports mathematics instruction in your school to help teachers meet the Pennsylvania Core Standards in mathematics.

Because we know your time is valuable, we will provide you with a $25 Amazon gift card for completing your survey!

Please use your unique code Teacher ID to enter the survey!

CLICK HERE TO TAKE THE OGAP SURVEY

A6. Affiliation appeal for control teachers

From: CPRE <cpre@gse.upenn.edu>
Subject: Join your peers and complete your NSF survey today!
Preheader:
Reply: cpre@gse.upenn.edu

CPRE | CONSORTIUM *for* POLICY RESEARCH *in* EDUCATION — Penn Graduate School of Education GSE

Last year a majority of your peers filled out their NSF survey. Join the trend and click the link below to complete yours today!

The National Science Foundation (NSF) supports mathematics instruction in the School District of Philadelphia to help teachers meet the Pennsylvania Core Standards in mathematics.

Because we know your time is valuable, we will provide you with a $25 Amazon gift card for completing your survey!

Please use your unique code Teacher ID to enter the survey!

CLICK HERE TO TAKE THE NSF SURVEY

A7. Commitment and Consistency appeal for treatment teachers

From: OGAP <cpre@gse.upenn.edu>
Subject: Demonstrate your commitment to education and complete your OGAP survey today!
Preheader:
Reply: cpre@gse.upenn.edu

CPRE | CONSORTIUM *for* POLICY RESEARCH *in* EDUCATION Graduate School of Education Penn GSE

Tend to the seed of a child's education and complete your OGAP survey by clicking the link below!

The Ongoing Assessment Project (OGAP) supports mathematics instruction in your school to help teachers meet the Pennsylvania Core Standards in mathematics.

Because we know your time is valuable, we will provide you with a $25 Amazon gift card for completing your survey!

Please use your unique code Teacher ID to enter the survey!

CLICK HERE TO TAKE THE OGAP SURVEY

A8. Altruism appeal for control teachers

From: CPRE <cpre@gse.upenn.edu>
Subject: Show your giving nature by completing your NSF survey today!
Preheader:
Reply: cpre@gse.upenn.edu

CPRE | CONSORTIUM *for* POLICY RESEARCH *in* EDUCATION — Penn Graduate School of Education GSE

Demonstrate your care for children by filling out your NSF survey linked to below!

The National Science Foundation (NSF) supports mathematics instruction in the School District of Philadelphia to help teachers meet the Pennsylvania Core Standards in mathematics.

Because we know your time is valuable, we will provide you with a $25 Amazon gift card for completing your survey!

Please use your unique code Teacher ID to enter the survey!

CLICK HERE TO TAKE THE NSF SURVEY