ANNENBERG
BROWN UNIVERSITY

EdWorkingPaper No. 21-466

Can nudging mentors weaken student support? Experimental evidence from a virtual communication intervention

Stacey L. Brockman
Mathematica Policy Research, Inc.

This paper presents results from a randomized trial of a nudge intervention designed to encourage and enhance virtual student support. During the 2019-20 school year, randomly selected mentors in a school-based mentoring program received monthly reminders with tips for communicating with youth via text, email, and phone. Unexpectedly, the results showed that although the informational reminders did not impact the frequency of mentors' outreach, they reduced the rate at which students reached out and responded to their mentors. Moreover, and possibly as a consequence, mentors who received the intervention felt less connected to students and less satisfied with their mentoring relationships, and treated students gained less from the mentoring program as a whole in terms of their personal and attitudinal growth. This study's findings add an important nuance to the evidence on how behavioral interventions in educational contexts operate. Although past studies find that reminder nudges can support individuals' engagement in discrete tasks, this evidence suggests that prescribing relational practices may be less effective. Thus, mentor supports must be carefully designed in order to yield the intended benefits for students.

VERSION: September 2021

Suggested citation: Brockman, Stacey L.. (2021). Can nudging mentors weaken student support? Experimental evidence from a virtual communication intervention. (EdWorkingPaper: 21-466). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/7bg3-dk82

Can nudging mentors weaken student support? Experimental evidence from a virtual communication intervention

Stacey L. Brockman

Mathematica Policy Research, Inc.

ORCID iD: 0000-0002-3695-4602

I would like to thank Brian Jacob, Robin Jacob, Lindsay Page, Awilda Rodriguez, and Matthew Ronfeldt for their valuable input on earlier drafts of this manuscript. I am also grateful for feedback from participants of the 2020 Association for Educational Finance and Policy (AEFP) annual conference and the Causal Inference in Education Research Seminar (CIERS) at the University of Michigan. Madeline Trumbauer provided excellent research assistance.

I received support from the Institute of Education Sciences, U.S. Department of Education (PR/Award # R305B1170015) to complete this work as part of a predoctoral training fellowship. Additional support was provided by the Rackham Graduate School at the University of Michigan. Correspondence concerning this article should be addressed to Stacey L. Brockman, Mathematica Policy Research, Inc., 220 East Huron Street, Suite 300, Ann Arbor, MI 48104-1912. Contact: slbrock@umich.edu.

Abstract

This paper presents results from a randomized trial of a nudge intervention designed to encourage and enhance virtual student support. During the 2019-20 school year, randomly selected mentors in a school-based mentoring program received monthly reminders with tips for communicating with youth via text, email, and phone. Unexpectedly, the results showed that although the informational reminders did not impact the frequency of mentors' outreach, they reduced the rate at which students reached out and responded to their mentors. Moreover, and possibly as a consequence, mentors who received the intervention felt less connected to students and less satisfied with their mentoring relationships, and treated students gained less from the mentoring program as a whole in terms of their personal and attitudinal growth. This study's findings add an important nuance to the evidence on how behavioral interventions in educational contexts operate. Although past studies find that reminder nudges can support individuals' engagement in discrete tasks, this evidence suggests that prescribing relational practices may be less effective. Thus, mentor supports must be carefully designed in order to yield the intended benefits for students.

Keywords: mentoring, virtual student support, behavioral intervention, randomized experiment

As many students continue to attend school remotely due to the COVID-19 pandemic, mentoring programs have a critical role to play in mitigating students' struggles. Typically, teachers, school counselors, nurses, social workers, and other school personnel play a critical role in identifying students in need of social-emotional supports. In fact, the vast majority of students who ever obtain mental health services do so through their school (Rones & Hoagwood, 2000). Social distancing requirements have further isolated youth from face-to-face connections with coaches, extended family, and afterschool staff. Trying fill these gaps in students' social support networks, mentoring programs and schools alike have launched initiatives to connect with students virtually (CASEL, 2020; Simonton, 2020). Volunteer mentors, perhaps more readily than overwhelmed teachers and school counselors, are well-positioned to offer students individualized attention and connection. Yet as mentoring organizations pivot to provide online mentoring, they require guidance about the best ways to prepare and support the adults responsible for connecting virtually with youth.

This paper presents results from a randomized trial of a mentoring intervention that aimed to enhance virtual communication. Although designed with a youth mentoring context in mind, the lessons learned may apply anywhere student support is offered virtually.

The Winning Futures career mentoring program supports youth in Metro Detroit high schools by pairing young people with adult mentors from the local business community. Through its career exploration curriculum and small group mentoring, the program aims to equip youth with the skills and attitudes necessary for success beyond high school. Although primarily an in-person program, Winning Futures leaders encouraged mentors to connect with students by email, text message, or phone at least monthly. Evidence from prior years indicated, however, that many pairs did not communicate much, or at all, outside of in-person meetings. To address this, during the 2019-20 school year, a randomly selected group of mentors received monthly messages reminding

them to communicate virtually with students. The messages contained discussion-starters that mentors could use to initiate conversations. Three reminders were sent approximately one month apart during Spring 2020. An end-of-program survey assessed the interventions' impacts.

Unexpectedly, the results showed that although the informational reminders did not make mentors any more (or less) communicative, they reduced the rate at which students responded to mentors' communication and also reduced the likelihood that students reached out to mentors on their own. Moreover, and possibly as a consequence, mentors who received the intervention gave worse ratings of their relationships with students, and students whose mentors received the intervention gained less from the mentoring program as a whole in terms of attitudinal growth.

Exploratory analyses investigated how reminding mentors to reach out could lead to worse ratings of mentoring relationships and student attitudes. The results revealed that students of treated mentors sent fewer text messages – students' main method of virtual outreach – and that they were less-responsive to their mentors' texts and calls. Meanwhile, there was suggestive evidence that the reminders may have shifted the mode of communication mentors used to reach out (towards emails, and away from text messages and phone calls) and may have updated mentors' understandings about the ideal frequency of virtual outreach (to about once per week). The findings underscore the importance of carefully attending to the content of reminder nudge messages.

Taken together, the results suggest that light-touch informational reminders can affect virtual communication, but that the direction of the effects is not unambiguously positive. This study's findings add an important nuance to the evidence on how behavioral interventions in educational contexts operate. Although past studies find that reminder nudges can support individuals' engagement in discrete tasks, prescribing relational practices may be less effective.

Prior Literature

Youth mentoring programs pair young people with caring adults who act as advocates, guides, and role models. In recent years, mentoring has become a common approach to supporting youth, with significant financial resources dedicated to programs and an estimated 2.5 million adults volunteering as mentors each year (Raposa, Dietz, & Rhodes, 2017). Experimental and quasi-experimental studies find that mentored youth perform better academically, receive fewer disciplinary referrals in school, self-report more positive perceptions of their abilities and college prospects (Herrera, Grossman, Kauh, & McMaken, 2011; Maynard, Kjellstrand, & Thompson, 2014; Woods & Preciado, 2016). Yet the magnitude of the measured effects of mentoring is often quite modest. Several meta-analyses report average effects of only about one fifth of a standard deviation on outcomes such as: emotional well-being, high-risk behavior, social competence, academic improvement, and employment success (DuBois, Holloway, Valentine, & Cooper, 2002; DuBois, Portillo, Rhodes, Silverthorn, & Valentine, 2011; Raposa et al., 2019). Effect sizes of around 0.2 standard deviations are typically considered small to medium for educational interventions (Cohen, 1988; Kraft, 2019). Thus, although mentoring can be an effective intervention, many programs, at least as currently implemented, appear to have room to improve. This raises the question, what can be done to strengthen mentoring programs to ensure they more often realize their promise?

Improving mentoring programs to enhance youth support

Prior research shows that the quality of mentoring relationships may be one, if not the, key mechanism through which mentoring programs impact youth (DuBois et al., 2002; Nakkula & Harris, 2014; Rhodes, Spencer, Keller, Liang, & Noam, 2006). For example, Bayer, Grossman, and DuBois (2015) conducted a reanalysis of data from the well-known experimental evaluation of the *Big Brothers Big Sisters* (BBBS) school-based mentoring program (Herrera et al., 2007). The authors examined whether mentees' feelings of closeness were a causal mechanism through which mentoring programs, like BBBS, impacted academic growth and found evidence to suggest that this

is the case. In fact, the authors found that the program's effects on every measure of academic improvement were mediated by students' sense of closeness to their mentor. In other words, the authors concluded that, unless mentoring pairs formed relationships that the youth would rate as "close," participation in the BBBS program had "no discernable effect" on students' end-of-year academic outcomes (Bayer et al., 2015, p. 425).

Given the importance of establishing strong mentoring bonds, studies have explored the factors associated with youths' and mentors' perceptions of higher quality mentoring relationships (Deutsch & Spencer, 2009; Parra, DuBois, Neville, Pugh-Lilly, & Povinelli, 2002); consistent, meaningful communication emerges as a key predictor. Findings from an exploration of the widely used *Check & Connect* mentoring program in 27 schools and across five states by L. Kern, Harrison, Custer, and Mehta (2019) are illustrative. The authors used surveys of mentored high schoolers and their mentors to gauged participants' perceptions of and preferences related to their mentoring relationships. Of all the factors considered, the topics discussed during mentoring sessions were the only one to significantly predict relationship quality. Specifically, discussions about school and future plans were perceived as helpful and predicted stronger ratings of the relationship among both students and mentors. Additionally, L. Kern et al. (2019) found that nearly one in five mentees and one in three mentors expressed a desire for more frequent meetings (more often than once a week). Together, their findings underscore the importance of fostering both meaningful and consistent exchanges. The authors, thus, concluded that programs leaders might consider approaches to increasing mentors' availability in order to increase their programs' overall effects.

Although increasing the frequency of mentor-mentee contact may be beneficial, many mentoring programs already face time and resource constraints. This may be especially true for school-based programs operating within the timeframe of a typical school day. As a result, finding time for additional in-person meetings may not be practical or even feasible. Moreover, the COVID-

19 pandemic has prevented many youth mentoring providers from interacting with youth face-to-face. Given these constraints, virtual communication may offer a promising alternative.

To summarize, prior research suggests that mentoring programs can bolster their effects by strengthening mentoring relationships and encouraging consistent communication. In light of the restrictions on in-person meetings currently faced by many programs, virtual methods of communication offer a promising avenue through which to pursue mentoring program improvements. The current study evaluates the effects of an intervention designed to strengthen mentor-mentee outreach via text messaging, email, and phone. In addition to contributing evidence on the value of and best practices related to virtual communication, the findings also extend prior research on the use of behavioral interventions, or "nudges," in educational contexts.

Behavioral interventions in education

Behavioral interventions are an increasingly common approach to supporting educational outcomes. Often using low-cost, light-touch strategies, these interventions nudge students and the adults that support them to take small steps that can lead to big differences in educational outcomes (Oreopoulos, 2020). Nudges tend to be most effective for individuals and in situations most constrained by the behavioral barriers that the interventions target (Damgaard & Nielsen, 2018). For example, several studies tested the effects of nudging graduating high schoolers and early college students to complete the requirements to enter, stay enrolled, and succeed in college (Bird, Castleman, Goodman, & Lamberton, 2017; Castleman & Page, 2015, 2016, 2017; Kizilcec, Schneider, Cohen, & McFarland, 2014). Results indicated that reminders are effective, especially when they concern discrete tasks (e.g., completing financial aid applications or contributing to an online forum).

Studies reveal particularly positive impacts of sending informational reminders to the adults responsible for supporting students' learning (e.g., parents and teachers). Individuals' attention is

limited, which means people are prone to forgetting to make key decisions and taking key actions. Parents and teachers are no different. Reminder nudges work by refocusing attention on the problem or task. They also have informational value, by reminding individuals of already-known information or providing easy access to new information. Informational reminders can be delivered as emails, text messages, letters, or by other modes of communication and typically include succinct messages with information relevant to students' educational situations or progress.

Studies of reminder nudges consistently find positive effects on parental engagement with subsequent positive impacts on students' school performance and skills (Damgaard & Nielsen, 2018; Oreopoulos, 2020). Reminders sent to parents have been shown to boost students' grades, courses passed, and school attendance (Bergman, 2015; Bergman & Chan, 2019; Robinson, Lee, Dearing, & Rogers, 2018). Bergman (2016), for example, showed that simply informing parents about how to use an online portal to track their student's performance improved students' grades. Rogers and Feller (2018) found that letters reminding parents about the importance of regular school attendance reduced both overall and chronic absenteeism.

Taken all together, existing evidence on the effects of behavioral interventions in educational contexts shows that even fairly light-touch nudges can sometimes yield at least modest gains. In particular, research on informational reminders paints a positive picture of their potential to both alter adults' behavior through increasing their engagement, and to improve student outcomes. Less clear, however, is whether these positive results will apply to relational practices and contexts, such as mentor-student communication and virtual student support. Discrete tasks and relational practices may differ in key ways that make nudges either more or less effective. Virtual communication may need to occur organically in order for it to feel authentic to students. On the other hand, reminder nudges may prove to be even more important in the context of virtual mentoring if the reminders help to sustain mentors' attention and engagement in the mentoring

relationship between in-person meetings. Either way, the extent to which existing evidence on reminding individuals to complete discrete tasks may, or may not, apply to relational practices remains an open question.

Current Study

Research context: Winning Futures school-based career mentoring program

This study took place in partnership with the Winning Futures (WF) career mentoring program which served students at seven Metro Detroit high schools during the 2019-20 school year. In WF, volunteer adult mentors from the local professional community guide students through a structured curriculum. Throughout the study, participants attended weekly in-person sessions in which they fostered relationships through discussions of students' educational and career goals.

In addition to in-person class sessions, WF leaders encouraged mentors to communicate with their mentees via text message, email, or phone at least once per month. Feedback from prior cohorts suggested, however, that virtual communication was an area for organizational growth. Although students said they appreciated their mentors' messages, mentors reported low rates of response from students. Thus, WF leaders identified virtual communication as an area for program improvement during the 2019-20 school year.

Bolstering virtual outreach: Mentor nudge intervention

In an effort to strengthen virtual communication, a random subset of mentors were selected to receive additional encouragement from WF to reach out to mentees. Approximately monthly emails contained specific prompts intended to support productive mentoring exchanges. These informational reminders encouraged mentors to reach out by reminding them of WF's expectations for regular between-session contact and by suggesting ideas for how mentors might initiate conversations with mentees. Figure 1 displays an example of the mentor reminder messages.

Current study contributions

This paper fills several gaps in our collective knowledge about improving mentoring programs, virtual student support, and behavioral interventions as a means to achieve those two ends. This study experimentally tested the effects of a low-cost, light-touch mentoring intervention on virtual communication and its subsequent impacts on mentoring relationship quality and student attitudinal growth. These findings may be more valuable now than ever before, as students face additional challenges, yet their access to school-based supports are restricted. This study also fills a gap in our understanding of how to improve mentoring programs more broadly. Although prior research supports a link between mentor-mentee communication and relationship quality, I am not aware of any other studies that directly examine how the frequency of exchanges relate to participants' sense of connection. Finally, this is also the first study, to my knowledge, to apply a behavioral intervention in the context of a mentoring program. While past studies establish that informational reminders can support engagement in concrete educational tasks, less is known about how such interventions effect relational practices and the provision of student support.

Research questions

To fill the aforementioned gaps, this study presents the results of a randomized control trial (RCT) of an intervention designed to bolster virtual communication. The experiment evaluated the effects of mentor nudges (i.e., informational reminders) on several outcomes, including: mentor-student communication frequency, student responsiveness, mentoring relationship quality, and the student attitudes targeted by the mentoring program. The following questions guided the analysis:

1. Prior to receiving any informational reminders, was more frequent virtual outreach associated with differences in mentoring relationship quality or students' attitudes?
2. Does reminding mentors to reach out by providing structured communication prompts (i.e., informational reminders) increase the frequency of mentor-student communication?
3. Do informational reminders increase students' responsiveness to mentor communication?

4. Do informational reminders support stronger mentoring relationships?
5. Do students whose mentors received informational reminders benefit more in terms of attitudinal improvements (specifically, self-perceived self-efficacy, growth mindset, goal orientation, perseverance, and adult support)?

Methods

Randomization

Each WF adult volunteer mentored about three students, and thus the study used a clustered RCT design with students nested within mentors. Random assignment was conducted at the mentor-level, with half of all mentors randomly assigned to receive monthly informational reminders to reach out to students (treatment) and the other half assigned to "business as usual" (comparison). Stratification by school ensured roughly even numbers in the treatment and comparison groups at each site. To ensure balance across the groups in terms of pre-treatment student and mentor characteristics, I used rerandomization to assign mentors to the treatment group (Morgan & Rubin, 2012).

Participants

The experiment originally included all mentors and students participating in WF during the 2019-20 schoolyear across the seven schools where WF operated mentoring programs (N=494 students assigned to N=143 mentors). Due to low response rates on the end-of-program survey, which was administered online just after the shift to online mentoring due to the COVID-19 pandemic, the final analytic sample included only a subset of the original participants (N=192 students; N=102 treatment and N=90 comparison; and N=114 mentors; N=59 treatment and N=55 comparison). The abrupt shift to online mentoring disconnected many students from the WF program as a whole, and as a result, the overall survey response rate for students was 39 percent. Encouragingly, mentors' response rate was much higher (80 percent), and the differences in

response rates across the treatment and comparison groups were very small (3.3 percentage points for students, and 0.02 percentage points for mentors). This is reassuring because it means that whether or not students and mentors responded to the post-survey did not appear to be linked to their treatment group status. Furthermore, the rates of overall and differential non-response for mentors were far below the boundary established by the What Works Clearinghouse for interpretation under both cautious and optimistic assumptions, whereas the rates for students lay just beyond (What Works Clearinghouse, 2020) Thus, while students' low responses mean that conclusions from this study may not completely represent the original experimental sample, there was no reason to suspect that survey non-response weakened the internal validity of the findings, and it is still possible to draw causal claims about the impact of informational reminders.

Measures

To evaluate the effects of mentor reminders on virtual communication, I focused on two main outcomes: frequency of weekly virtual outreach and student responsiveness. Students' and mentors' reports for both outcomes were collected on a post-survey. For the purposes of multiple hypothesis testing considerations, described below, I treated the analyses of these two outcomes as confirmatory. I also explored effects on several additional outcomes, including: students' and mentors' ratings of their relationships, and the student attitudes targeted by the WF program (e.g., self-efficacy, goal orientation, and perseverance). For the purposes of multiple hypothesis testing considerations, I considered these analyses exploratory.

All of the measures I analyzed were self-reported by students and mentors on surveys. WF leaders administered a pre-program survey at the start of the school year (Sep 2019), a baseline survey at the beginning of the experiment (Jan 2020), and a post-survey administered at the end (Apr 2020). I mean-imputed all missing baseline measures. The section below describes the outcome

measures, and Appendix Table 1 displays sample items, information about internal consistency, and the dates of data collection.

Outreach frequency. Baseline and post-surveys asked students and mentors how often, during a typical week, they engaged in the following ways: received a text message, received an email, received a phone call, sent a text message, sent an email, gave a phone call. All frequency items used the same scale: 0 times, 1 time, 2 times, 3-5 times, and more than 5 times. It is worth noting that because the outreach frequency items only asked about communication during a typical week, it is possible that the measures might obscure intervention impacts on less-frequent communication (e.g., mentor or student outreach that occurred once or twice a month).

Mentors reported their outreach to and from each of their mentees separately. For each individual, I summed the ratings of outreach frequency across each virtual format to create scales that ranged from zero interactions in a given week up to 15 possible interactions (responses of "3-5 times" were coded as 4 and "5 or more interactions" were coded as 5). Before conducting analyses, I standardized all outcomes to the comparison group mean and standard deviation such that the estimated treatment effects represent effect sizes.

Student responsiveness. Students and mentors each also rated students' responsiveness to mentors' communication. Using the response categories "Never, Rarely, Sometimes, Always" students reported how often they respond when their mentor reached out via text, email, and phone call (separate items, averaged across communication methods to create a single composite variable). Mentors responded to parallel items about each mentee's responsiveness.

Mentoring relationship quality. Relationship quality in youth mentoring contexts refers to participants' sense of connection and shared goals. Prior research differentiates between two dimensions of quality: relational and instrumental (Nakkula & Harris, 2014). Relational quality refers to how the mentor and mentee feel about each other and their relationship. Subconstructs that

compose relational quality typically include relational satisfaction, feelings of closeness, relational compatibility, and relational competence. Instrumental quality, on the other hand, measures the extent to which the mentoring relationship is oriented towards the youth's growth. Although relational quality has received more attention in prior research, both are related to positive youth outcomes (Anderson, Christenson, Sinclair, & Lehr, 2004; Herrera, Sipe, & McClanahan, 2000).

Mentoring relationship quality (MRQ) was measured from both students' and mentors' perspectives using parallel surveys developed to assess MRQ from each participant's perspective. Mentors items came from the *Match Characteristics Questionnaire* (Harris & Nakkula, 2018a); student-rated items were adapted from the *Youth Mentoring Survey* (Harris & Nakkula, 2018b). These measures were originally developed using diverse samples of U.S. youth of all grade-levels and their mentors who participated in large mentoring organizations, and published results indicate that each scale had good internal consistency (Karcher, Nakkula, & Harris, 2005).

Students rated their sense of relational quality, or the degree to which they felt happy, close, and satisfied with their mentoring relationship. Relational quality was assessed by 14 items (baseline α=0.92; post-survey α=0.93) such as, "My mentor and I are close (very good friends)" and "My mentor makes me feel special." They also rated the instrumental quality of their relationships. Eight items assessed instrumental quality (baseline α=0.88; post-survey α=0.92) such as, "I talk with my mentor when I have problems or things that worry me" and "My mentor helps me get in less trouble (make better decisions, behave better, etc.)."

Mentor survey items measured two subconstructs: closeness and satisfaction. Four items (baseline α=0.85; post-survey α=0.82) assessed mentors' sense of closeness to their mentees, for example, "I feel like my mentee and I are good friends (buddies)" and "My mentee shows me how much he/she cares about me (says things, smiles, etc.)." Five items (baseline α=0.88; post-survey α=0.92) assessed mentors' satisfaction or sense of fulfillment in the relationship by how frequently

items such as these were true, "My mentee is willing to learn from me" and "I feel like I am making a difference in my mentee's life."

Student attitudes. The WF mentoring program as a whole aimed to improve students' attitudes towards and readiness for post-secondary education and the workforce. To assess WF's progress towards this goal, the post-surveys measured the following student attitudes: self-efficacy, growth mindset, goal orientation, perseverance, and adult support. The surveys used published, validated measures of each construct, and analyses indicated that each scale achieved good internal consistency (see Appendix Table 1).

Intervention timing and adaptations due to the COVID-19 pandemic

The intervention began in January 2020, about three months after the start of the WF program, thus giving students and mentors time to build initial connections. Appendix Figure 1 displays a timeline of the intervention and data collection. Three email messages were sent between January 2020 and before the closure of school buildings in March 2020 due to the COVID-19 pandemic. Two additional messages were sent following school closures. Although the effect of the final two reminders were not captured in the main analyses, they were included in robustness checks. To summarize: the main findings estimate the effect of three nudges sent prior to the pandemic using data collected shortly after the suspension of in-person mentoring. Additional surveys administered at the end of the program provide a check on the robustness of the main findings.

In addition to surveys, WF leaders helped to recruit and facilitate focus groups with students and mentors after the conclusion of the program in May 2020. Although these data were collected as part of a separate evaluation of the WF program as a whole, several questions asked specifically about interacting via virtual communication.

Empirical strategy

Descriptive analyses. Before estimating the interventions' impacts, I first investigated the baseline correlates of virtual outreach frequency. These analyses described the baseline levels of communication between students and mentors (in the absence of any intervention) and assessed the extent to which greater communication was related to the perceived benefits of mentoring. To do so, I estimated a series of ordinary least squares (OLS) regressions predicting the frequency of mentor and student weekly outreach as the outcome. The predictor variables included all of the baseline covariates available at the start of the experiment (i.e., mentor and student sociodemographic characteristics, student attitudes, and mentoring relationship quality). I estimated the correlation between outreach frequency and each baseline measure individually and all together.

Intervention impacts. I computed the main effects of the intervention by comparing outcomes for students whose mentors received reminders with those of students whose mentors did not. I estimated the effects using OLS regressions with robust standard errors clustered within mentor groups. I specified intent-to-treat (ITT) models of the following general form:

$$Y_{ijc} = \beta_0 + \beta_1 Z_j + X_{ijc} + \delta_c + \varepsilon_{ijc}$$

where for student i paired with mentor j in classroom c, Z_j indicates whether the student's mentor was assigned to receive informational reminders (Z_j=1) or not (Z_j=0); X_{ijs} is a vector of mentor and student covariates including an indicator for whether any baseline covariates were mean-imputed; δ_c is a classroom-fixed effect; and ε_{ijc} is an idiosyncratic error term. β_1 estimates the average treatment effect of mentor reminders. Alternative specifications included a random intercept, u_j, to account for clustering of students within mentors instead of clustered standard errors. For ease of interpretation, I standardized all outcome measures to the comparison group mean and standard deviation such that the estimated treatment effects represent effect sizes. Standardized outcomes included the following: average weekly frequency of mentor and student outreach by text, email, and phone combined (student-reported measure of mentor outreach, and

mentor-reported measure of student outreach); student responsiveness to mentors' outreach (both mentor-rated and student self-reported); student attitudes (self-efficacy, growth mindset, goal orientation, perseverance, and adult support); and mentoring relationship quality (student-perceived relational and instrumental quality, and mentor-perceived closeness and satisfaction).

Exploratory analyses. Finally, I conducted exploratory analyses to identify possible causal mechanisms through which mentor reminders impacted outreach frequency. I investigated the impact of the intervention on each mode of communication separately (text, email, and phone). I also explored whether the intervention's effects operated along the extensive versus intensive margin, that is, any communication versus none, or frequent versus infrequent communication. To do so, I recoded the outreach frequency and responsiveness scales as two binary variables. The first indicated no outreach or response versus any (0=no outreach/never respond; 1=any outreach/respond more often than "never"). The second variable indicated more frequent outreach and consistent responses versus less (0=fewer than two weekly touchpoints/responding less than "always"; 1=two or more weekly touchpoints/"always" responding to mentors' messages).

Randomization inference. As described above, I used rerandomization to select the treatment and comparison groups (Morgan & Rubin, 2012). While useful for ensuring balance on covariates, rerandomization changes the distribution of the test statistic such that traditional methods of inference must be adjusted or else they will result in overly conservative *p*-values (Athey & Imbens, 2017). Specifically, I used the user-written "ritest" STATA package for randomization inference to re-estimate the *p*-values from each of the analyses described above (Heß, 2017).

Multiple hypothesis testing. When the effects of an intervention on multiple outcomes are analyzed, it is important to account for the fact that the probability of detecting at least one statistically significant impact merely by chance increases with the number of outcomes considered. Although several possible approaches have been proposed (Kirk, 1994; Schochet, 2008), there is

currently little consensus on the most appropriate method for adjusting statistical tests to account for multiple outcomes. Therefore, following guidelines for quantitative methods used in social experiments, I did not explicitly adjust *p*-values to correct for multiple comparisons (Shadish, Cook, & Campbell, 2002). Instead, I carefully limited the number of outcomes in the analysis, and grouped research questions into confirmatory and exploratory categories (Schochet, 2008).

Results

Baseline equivalence and descriptive statistics

I checked to ensure that the treatment and comparison groups were balanced across the following covariates: student and mentor gender, student and mentor race, students' free/reduced price lunch status, students' parental education level, mentors' years of mentoring experience, students' attitudes at the start of the mentoring program, and students' and mentors' baseline reports of communication and relational quality. Table 1 presents these descriptive statistics for the entire analytic sample and assesses the baseline equivalence of the treatment and comparison groups. Most baseline characteristics of students and mentors assigned to the treatment and comparison groups were similar and only a four were outside the range that the What Works Clearinghouse considers acceptable with statistical adjustment ($0.05 \leq$ Effect size ≤ 0.25). Specifically, treated mentors were less-likely to be Black; treated mentors and students reported lower levels of mentoring relationship quality; and students of treated mentors self-reported having lower levels of a growth mindset.

With a comprehensive set of baseline covariates, testing for balance on each increases the probability of Type I error. I therefore assessed the joint significance of all baseline imbalances as well. Based on the results (F-statistic = 1.204, $p = 0.273$), I failed to reject the null hypothesis of baseline equivalence. Although there were scattered, modest imbalances on individual covariates, this is not unexpected, and all analyses controlled for all baseline measures displayed in Table 1.

Virtual communication at baseline and correlates of communication frequency

I first examined whether more frequent virtual communication at baseline was associated with any differences in mentoring relationship quality ratings or measures of students' attitudes. These analyses had two objectives. First, to describe the extent to which mentors and students were already using virtual methods of communication, even in the absence of informational reminders. Second, these analyses confirmed that more frequent communication was, in fact, correlated with stronger perceptions of mentoring relationships.

Figure 2 displays the frequencies with which students and mentors reported sending and receiving text messages, emails, and phone calls during a typical week. As these figures show, baseline levels of virtual communication were fairly low. For example, 65 percent of mentors reported receiving zero text messages from their mentees during a typical week, and 95 percent and 97 percent reported receiving zero emails and phone calls, respectively. Mentor outreach as reported by students was somewhat higher: only 42 percent of students reported never receiving text messages from their mentors during a typical week, and 33 percent of students reported that their mentors sent two or more texts. Among both students and mentors, text messaging was the most common form of communication, and very few respondents reported ever making or receiving phone calls (94 percent of mentors and 93 percent of students self-reported never placing any calls).

As Figure 2 shows, students and mentors both self-reported reaching out more often than their counterparts reported receiving communication. For instance, 59 percent of students reported never sending texts compared to 65 percent of mentors who reported never receiving texts. Mentor outreach followed a similar trend: only 31 percent of mentors reported never texting, whereas 42 percent of students reported never receiving texts. Given the possibility that social desirability bias influenced how individuals' self-reported communication frequency, the main impact analyses concentrated on outreach received (i.e., students' reports of messages received from mentors and vice versa).

Table 2 presents the results of descriptive analyses of the factors correlated with more frequent virtual outreach prior to the intervention. Higher quality mentoring relationships were associated with more frequent virtual communication. For example, for each one point increase in students' ratings of their relationship with mentors (scale ranged from 1-4), they also reported receiving an average of one additional touchpoint every other week (including texts, emails, or phone calls). I observed a similar trend for mentors. For each one point increase in mentors' ratings of relationship quality (scale ranged from 1-6), mentors reported receiving about one additional check-in from students every three to four weeks.

Other factors were less predictive of communication frequency. On the whole, students' baseline attitudinal measures were not associated with either higher or lower levels of outreach. Most of sociodemographic background characteristics were unrelated to outreach with the exceptions of students' free and reduced price lunch (FRPL) status (students' receipt of FRPL was associated with higher levels of mentor outreach) and mentoring experience (students paired with more experienced mentors reported receiving slightly less communication).

Taken all together, these results show that even though the WF organization asked mentors to reach out beyond in-person meetings at least monthly, many pairs never used virtual communication to connect prior to the intervention. Importantly, these findings also show that pairs with closer relationships tended to communicate virtually more often, suggesting that enhancing virtual outreach may be a promising approach to strengthening relationships and increasing the benefits of mentoring that youth receive. With that in mind, I turn next to the results of the intervention intended to bolster virtual communication through informational mentor reminders.

Intervention impacts on outreach frequency and student responsiveness

Table 3 presents the impact of mentor reminders on mentor outreach, student outreach, and student responsiveness in Panel A. Impact estimates are reported in effect sizes. I found that the

intervention had a null effect on mentor outreach: students whose mentors received the reminders did not report receiving any more or less frequent communication (*ES* = 0.05, SE = 0.15, p = 0.741). Student outreach, however, was negatively affected. Treated mentors' reports of virtual communication from students were just over one quarter of a standard deviation lower, a small to moderate effect (*ES* = -0.27, SE = 0.12, p = 0.05). In practical terms, whereas comparison group mentors reported receiving about four or five touchpoints per month (an average of one per week), treated mentors reported that their students only reached out about three times per month. This result aligns with the impact of the intervention on student responsiveness, which was also negative. On a 4-point scale ranging from 1=Never to 4=Always, students' rated the frequency of their responses about one third of a standard deviation, about half a scaled-point (0.6-point), lower if their mentors received reminders (*ES* = -0.33, SE = 0.13, p = 0.035). Interestingly, treated mentors' perceptions of students responsiveness were unaffected (*ES* = 0.10, SE = 0.14, p = 0.499). To summarize, the main results indicated that although mentors' outreach to and perceptions of students were unchanged, reminding mentors negatively affected students' outreach and students' responsiveness to mentors' virtual communication.

Panel B of Table 3 presents the impact of mentor reminders on students' and mentors' perceptions of their mentoring relationships. I found that receiving informational reminders had a moderate, negative effect on mentors' perceptions of their relationships with students. Treated mentors' ratings of relational closeness were about one quarter of a standard deviation, or about 0.25-points lower, compared to an average score of 4.5 on a 1-6 scale among mentors in the comparison group (*ES* = -0.26, SE = 0.11, p = 0.029). Mentors' perceptions of their relational satisfaction were also worse, about one third of a standard deviation, or about 0.4-points, lower compared to the comparison group average score of 4.4 (*ES* = -0.33, SE = 0.10, p = 0.006). I also found suggestive evidence that the reminders negatively affected students' ratings of both the

relational and instrumental quality of their relationships, although it was not possible to rule out a null effect (relational quality: *ES* = -0.12 , SE = 0.12, p = 0.25; instrumental quality: *ES* = -0.14, SE = 0.12, p = 0.224).

One possible interpretation of these findings is that the messages heightened mentors' awareness of the level of communication (or lack thereof) between themselves and students, leading to lower levels of perceived closeness and satisfaction. As the next set of results reveals, however, this hypothesis is unlikely to account for the full effects of the intervention as other student outcomes were also negative affected.

Intervention impacts on student attitudes

In addition to the negative effects on students' outreach frequency and responsiveness, I also found that students of treated mentors gained less from the WF mentoring program as a whole in terms of their attitudinal growth. Table 4 shows the effects of the intervention on students' self-reported attitudes: self-efficacy, growth mindset, goal orientation, perseverance, and adult support. Although there were no effects on growth mindset and adult support, I found that students rated their self-efficacy, goal orientation, and perseverance about one fifth to one quarter of a standard deviation lower if their mentors received reminders (self-efficacy: *ES* = -0.28, SE = 0.12, p = 0.017; goal orientation: *ES* = -0.2, SE = 0.11, p = 0.43; perseverance: *ES* = -0.21, SE = 0.12, p = 0.036). In slightly more practical terms, the impact on self-efficacy, for example, indicated that the intervention reduced students' self-ratings by about 0.2-points, compared to a comparison group average of 4.3 on a 1-5 scale. Although a fairly small effect, the impacts on student attitudes were not negligible, especially given the low intensity of the intervention (a mere three reminder emails sent to mentors).

Exploratory analyses: Possible causal mechanisms

In an attempt to shed light on how reminding mentors to reach out could lead to worse ratings of mentoring relationships and student attitudes, I conducted several exploratory analyses to

identify possible causal mechanisms. Since I hypothesized that changes in virtual communication (the subject of the mentor reminders) likely explained any subsequent results, these analyses aimed to describe in more detail exactly how pairs' virtual communication was impacted.

To begin, I examined treatment effects on virtual outreach and student responsiveness via each mode of communication separately (text, email, and phone calls). Table 5 displays the effect of the nudges on mentor- and student-initiated outreach via text, email, and phone (Panel A), as well as the effects on students' and mentors' perceptions of the rate at which students responded to each mode of communication (Panel B). Figure 3 depicts several selected findings graphically.

Additionally, I explored whether the treatment effects operated on the intensive margin (any vs. no communication) or extensive margin (frequent vs. less-frequent communication). The goal of this approach was to determine whether the overall null effects on mentors' outreach and negative impacts on students' outreach and responses might be driven by changes in either the number of pairs who ever communicated versus never communicated, or who communicated frequently (e.g., two or more touchpoints per week, or reported that students' "always" responded) versus less-frequently. Table 6 presents the results of linear probability models estimating the treatment effects along these two margins.

Looking first at Table 5, the results shed some light on how the nudges impacted virtual outreach. First, the negative impacts on students' outreach overall were due to their decreased use of text messaging. Compared to students in the comparison group, from whom mentors reported receiving an average of one text per week, treated mentors reported receiving about 0.5 fewer texts per week (*ES*=-0.34, SE=0.18, *p*=0.1). Figure 3 shows the number of text messages mentors reported receiving in an average week at baseline and at the time of the post-survey, illustrating the lower level of outreach among treated students.

In addition to sending fewer texts, treated students perceived themselves as less-responsive to their mentors' text messages. Comparison group students self-rated their responses to mentors' text messages with an average score of 3.2 (between 3=Sometimes and 4=Always), whereas treated students' ratings were an average of 0.4-points lower, for an average score of about 2.8 (between 2=Rarely and 3=Sometimes). Taken together, the results related to students' outreach and response suggest that something about the reminders (which mentors received via email) led students to respond less-often to their mentors' text messages and to initiate communication via text less-often.

I also found suggestive evidence that the intervention may have enhanced pairs' communication via email. Although the exploratory results were consistent with the main finding of null effects on mentors' outreach, there was a positive, though non-significant, impact on mentors' email frequency. Students in the comparison group reported receiving an average of 0.6 emails per week, or about 2-3 per month (assuming 4.5 weeks per month). Treated students, on the other hand, reported about 0.4 more emails per week, or an average of 4-5 emails per month. The bottom panel of Figure 3 depicts these differences across treatment statuses. A possible explanation of this finding is that as mentors received the reminders, which were sent via email, the messages prompted them to immediately reach out to students (thus increasing their outreach via email). Although it was not possible to rule out a null effect on mentors' email frequency, the possibility that the nudges enhanced pairs' use of email was also supported by a significant, and positive, impact on mentors' perceptions of students' responses to emails (ES = 0.33, SE = 0.13, p=0.005). This positive effect was especially notable because treated mentors perceived their students as less-responsive to phone calls (null effects on texts), and treated students also saw themselves as less-responsive to text message and calls. It is worth noting that although the effect on student responsiveness to emails was positive, the magnitude of the impact was still fairly small. Compared to non-nudged mentors, whose average rating of students' responses was 1.2 (somewhere between 1=Never and 2=Rarely),

treated mentors gave their students' an average rating of 1.4 (only slightly closer to "Rarely" responding to emails).

To summarize, students whose mentors received reminders were less likely to reach out via text message, which was the primary mode of communication used by students. In addition, students of nudged mentors perceived themselves as less-responsive to mentors' communication, with significant effects on responses to texts and phone calls. Meanwhile, there was suggestive evidence that nudged mentors may have sent more emails (although the impacts were small and non-significant), and there was evidence of a small, but significant, improvement in students' responses to mentors' emails (although again, the practical significance of the effects was small). If virtual communication generally, or perhaps emails in particular, between nudged mentors and students was less effective, or if students felt their communication with mentors was less personal or meaningful, this could explain why students reported being less-responsive and less-likely to initiate communication themselves.

The estimated treatment effects along the intensive versus extensive margin, displayed in Table 6, provide some additional insight into how the reminders impacted communication. Encouragingly, the results do not indicate that the intervention eliminated all virtual outreach. Comparing the impacts on receiving any vs. no communication with the effect on less-frequent vs. more-frequent communication, the changes were larger and more often significant along the latter margin. This suggests that the reminders did not cut off all outreach, but that they may have reduced the urgency with which both parties, students in particular, responded. In other words, treated students did not stop reaching out altogether, but merely reached out and responded less often.

Robustness checks

The results presented above were robust to several checks. First, given that survey non-response due to the COVID-19 pandemic altered the sample of students with outcomes data

available for analysis, I explored whether any baseline covariates predicted responding to the post-survey (results available upon request). I found that higher levels of parental education was the only significant predictor of an increased likelihood of responding. In addition, I examined how the baseline imbalances present in the analytic sample compared to those in the original experimental sample and the sample of students with complete mentors survey responses. Appendix Tables 2 and 3 show how the treatment and comparison groups differed in these alternate samples. Although the differences across treatment statuses were smaller within the original sample, the pattern of differences matches the analytic sample. In both cases, the treatment group reported somewhat lower levels of MRQ and communication frequency at baseline. The fact that these imbalances were present to a degree at baseline makes it less-likely that imbalances in the analytic sample were caused by differential survey non-response. All together, these checks suggest that changes in the sample due to survey non-response were unlikely to explain the estimated treatment effects.

Next, to account for the clustered nature of the data (students nested within mentor groups), I estimated models that included a random-effect for mentors rather than using OLS with clustered robust standard errors. The results were both statistically and qualitatively very similar to those presented in Tables 3 and 4. Appendix Tables 4 and 5 display results of this alternate specification.

In the exploratory analyses, I used linear probability models to investigate the effects along the margins of any outreach/response (vs. no outreach/response) and in terms of frequent outreach/always responding (vs. less consistent outreach/responses). Given the binary nature of these outcomes, logistic regression was possibly more appropriate approach to estimating the treatment effects. Appendix Table 6 displays the results using logistic regression, which were qualitatively similar to the main findings reported in Table 5.

Finally, I estimated the effect of the mentor reminders on communication frequency and mentoring relationship quality using data from two additional student surveys. The first was a mid-

point survey administered in late February and early March 2020, just before schools closed due to COVID-19. WF representatives sent two mentor reminders prior to administering the mid-point survey, which was collected during an in-person WF session at which mentors were not present. Secondly, students responded to an online survey following the conclusion of the WF program (May 2020). This survey captured the effect of all five mentor reminders. Before analyzing the effects using data from these two additional surveys, I first examined whether baseline covariates were balanced across the treatment and comparison groups for the samples of students who responded to each survey and I concluded that these samples were balanced (results available upon request).

Appendix Table 7 presents the estimated treatment effects at the time of the mid-point survey and final survey. Although there are some differences in statistical significance, likely due to changes in the underlying sample sizes, the results are qualitatively similar to the main findings. These similar results lend further support to the findings of the main analyses.

Limitations

While this study contributed several key findings about supporting virtual communication via mentor reminders, several limitations warrant consideration. First, all outcome measures relied on students' and mentors' self-reported perceptions of their communication, relationships, and attitudes, which could be biased by individuals' tendencies to report socially desirable responses or by their ability to accurately recall communication during a typical week. In addition, because the outreach frequency items asked about communication during a typical week, it is possible that the measures obscured effects on communication that occurred less than weekly. For example, if a mentor pair sent only one or two text messages per month, it is unclear how they would respond to items asking about outreach during a typical week.

Another limitation was the high rate of missing outcome data for students due to the challenge of engaging study participants during the COVID-19 pandemic. Although the internal

validity of the findings was not affected by differential response rates across treatment statuses, low overall response rates could impact the study's external generalizability. The reduced sample size may also have limited the precision with which true intervention effects could be detected. Finally, this study took place within a particular mentoring program. Thus, whether or not the findings would hold in other contexts is unknown.

Perhaps the most important limitation is that the study was not designed to explore specific mechanisms that could explain the impact of the intervention. Future research, thus, is needed to better understand how informational reminders impact mentors. In particular, future studies could test whether altering the content or format of mentor reminders leads to different results.

Finally, this study revealed that students' use of virtual communication remained quite low, regardless of whether their mentors were reminded to reach out. Future research might explore whether directing nudges at students, rather than mentors, holds more promise. Other approaches to bolstering student-mentor communication also warrant further exploration.

Discussion

To summarize the main findings: reminding mentors to reach out to students did not increase (or decrease) mentors' use of virtual communication overall, and may have even slightly increased their outreach via email. The intervention did, however, reduce the rate at which students reached out, specifically via text message. Students whose mentors received reminders also perceived themselves as less responsive to mentors' outreach. Taken together, this suggests that something about the informational reminders reduced students' engagement in virtual communication.

The results also showed that reminding mentors weakened students' and mentors' perceptions of the quality of their relationships. For students, the effects were small and statistically non-significant. For mentors, however, the negative effects were moderately sized and significant. Given that changes in virtual communication, the subject of the reminders, were the most likely

mechanism through which the intervention impacted any subsequent outcomes, the reduction in students' outreach may have been to blame. The fact that students reached out less, combined with how the reminders may have heightened mentors' attention to the communication they were – or were not – receiving, perhaps discouraged mentors and led them to feel that their relationships with students were less effective, close, and satisfying. A similar process could have occurred for students: less-frequent communication, perhaps as a result of less-satisfying virtual interactions with mentors, may have weakened students' views of their mentoring relationships as warm and productive.

Although mentors' feelings of discouragement could explain the intervention's negative effects on relationship quality, the fact that students benefited less from the mentoring program as a whole suggests that the reminders likely impacted more than just mentors' perceptions. Some aspect of the intervention also led students to rate their self-efficacy, goal-oriented motivation, and perseverance worse than their peers whose mentors did not receive the reminders.

One possible explanation is that receiving external reminders and using prescribed discussion prompts undermined the authenticity of mentors' communication. This intervention assumed that simply because there was a positive correlation between communication frequency and relational quality, manipulating the level of frequency would produce a corresponding change in relationships. But this might not be the case, as evidence from other contexts makes clear. For example, Loewenstein, Krishnamurti, Kopsic, and McDonald (2015) set out to test the causality of another, and widely observed, positive correlation: the link between sexual frequency and happiness. Although considerable research has observed that sexual frequency is a strong positive predictor of self-reported happiness, prior studies could not determine whether increased frequency leads, causally, to an increase in well-being. To test this assumption, the authors designed an experiment in which they recruited married couples who were willing to change their patterns of sexual behavior, and randomly assigned half to receive instructions to double their sexual intercourse frequency. They

found that the treatment successfully induced behavioral changes, but that, unexpectedly, increased sexual frequency did not increase happiness, and in fact, the reverse was true. Treated couples reported lower levels of mood (the daily measure of happiness), enjoyment, and desire for sex. The authors explored various mechanisms, and found that the treatment itself (i.e., directive instructions) undermined couples' intrinsic motivation to have sex. Ultimately, Loewenstein and colleagues (2015) concluded that their experimental design obscured their ability to test the causal relationship between sexual frequency and happiness. They did, however, determine that prescribing relational engagement can backfire as an approach to increasing couples' connection or happiness.

Connecting the findings from Loewenstein et al. (2015) to the present study, there are some similarities, albeit in pursuit of different behavioral changes. Individuals in both situations received directive instructions to engage in a relational practice (in this case, reaching out by text, email, or phone calls). Although the frequency with which pairs engaged in these relational practices was correlated with stronger ratings of their relationship quality, both experiments revealed that simply directing pairs to increase their frequency did not improve the relationships, and in fact, the reverse occurred. Rather than fostering connection, it appeared that receiving directive instructions may have actually undermined the quality or authenticity of pairs' engagement.

This finding represents an important nuance in our understanding of how behavioral interventions operate. Although prior evidence shows that reminder nudges can support individuals' completion of discrete tasks (e.g., checking their child's grades online, or completing financial aid paperwork; Bergman & Chan, 2019; Castleman & Page, 2016), prescribing relational practices may be less effective. The current study contributes evidence that relationships may be best developed organically and that directive instructions can make relational connections feel less-authentic.

Although the current study cannot uncover exactly what happened in this setting, a possible scenario is that the discussion-starters recommended within the reminders made mentors' messages

feel overly-structured or less-genuine to students. If so, this could account for the decreases in students' responses and outreach. Less-effective communication, and subsequent feelings of disconnection, could also explain the negative impacts of the reminders on relationship quality. Lower quality communication and weaker relationships might also lead students to gain less from the mentoring program as a whole.

Students' and mentors' comments during focus groups support the hypothesis that strong relationships may be a necessary precursor to consistent virtual communication, rather than the other way around. Asked about students' overall low response rates to mentors' outreach, one student said: "[Students] probably don't respond right away because they don't have, [or] they might not get a bond with their mentor…[Students respond] if they connect on a personal level or have fun together." Reflecting on the challenges of remote mentoring during the pandemic, one mentor said, "If students don't have to participate in school, so to speak, they're certainly not going to be 'all hands on deck' when it comes to communicating with someone they didn't get to have a great relationship with all the way through, like they were supposed to."

More evidence is needed to fully understand the unexpected finding that encouraging communication actually diminished mentor-student connections, with subsequent negative impacts on mentoring relationships and student attitudes. This study does, however, make several conclusions clear. First is that virtual communication and relationship quality are correlated, although the possible causality and direction of the association remains uncertain. Second, it is clear that light-touch informational reminders can change virtual communication patterns, but that they must be carefully designed in order to yield the intended benefits for students. Lastly, unlike nudges targeting discrete tasks, reminders may be less effective at encouraging relational practices. As this study shows, educational leaders must tread carefully as they take steps to encourage virtual

outreach. Particularly now, as remote learning separates many students from supportive school personnel, it is vital to learn and thoughtfully apply best practices for supporting students virtually.

Conclusion

While many are looking forward to resuming in-person mentoring when it is safe to do so, some aspects of virtual student support seem likely to remain. For adults whose roles involve student support, multiple points of contact and flexibility are likely a good thing, and to the extent that technology can help with that, virtual outreach should continue. This is why, perhaps now more than ever before, it is important to identify best practices for supporting students using virtual communication. This study begins to address this gap by assessing the effectiveness of an intervention designed to enhance mentors-students virtual outreach. Not only does this study add to our collective understanding about remote student support, it also extends existing research on behavioral interventions by showing that the kind of informational reminders proven effective for discrete tasks may operate somewhat differently when targeting relational practices. Importantly, this study shows that although a light-touch intervention, such as reminder emails, can have small effects on virtual communication, the results are not unambiguously positive. This study provides suggestive evidence that reminding mentors and prescribing discussion prompts may change the quality of their virtual outreach in ways that actually make students less responsive to their support. Further research is needed to explore how mentor reminders may change either the content or the quality of their outreach and whether, perhaps, virtual communication interventions would be more effective if directed at students.

References

Anderson, A. R., Christenson, S. L., Sinclair, M. F., & Lehr, C. A. (2004). Check & Connect: The importance of relationships for promoting engagement with school. *Journal of School Psychology, 42*(2), 95-113.

Athey, S., & Imbens, G. W. (2017). Chapter 3 - The Econometrics of Randomized Experimentsa. In A. V. Banerjee & E. Duflo (Eds.), *Handbook of Economic Field Experiments* (Vol. 1, pp. 73-140): North-Holland.

Bayer, A., Grossman, J. B., & DuBois, D. L. (2015). Using volunteer mentors to improve the academic outcomes of underserved students: The role of relationships. *Journal of Community Psychology, 43*(4), 408-429. doi:10.1002/jcop.21693

Bergman, P. (2015). Parent-Child information frictions and human capital investment: Evidence from a field experiment. *CESifo Working Paper Series* (No. 5391). Retrieved from https://ssrn.com/abstract=2622034

Bergman, P. (2016). Technology adoption in education: Usage, spillovers and student achievement. *CESifo Working Paper Series*(No. 6101). Retrieved from https://ssrn.com/abstract=2866866

Bergman, P., & Chan, E. W. (2019). Leveraging Parents through Low-Cost Technology: The Impact of High-Frequency Information on Student Achievement. *Journal of Human Resources*(56), 125-158.

Bird, K., Castleman, B. L., Goodman, J. S., & Lamberton, C. (2017). Nudging at a national scale: Experimental evidence from a FAFSA completion campaign. *EdPolicyWorks Working Paper Series.* (No. 54. March 2017). Retrieved from https://curry.virginia.edu/sites/default/files/files/EdPolicyWorks_files/55_Nudging_at_a_National_Scale.pdf

CASEL. (2020). *An initial guide to leveraging the power of social emotional learning as you prepare to reopen and renew your school community*. Retrieved from https://casel.org/wp-content/uploads/2020/05/CASEL_Leveraging-SEL-as-You-Prepare-to-Reopen-and-Renew.pdf

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

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. (2017). Parental Influences on Postsecondary Decision Making: Evidence From a Text Messaging Experiment. *Educational Evaluation and Policy Analysis, 39*(2), 361-377.

Chen, G., Gully, S. M., & Eden, D. (2001). Validation of a new general self-efficacy scale. *Organizational Research Methods, 4*(1), 62-83. doi:10.1177/109442810141004

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, N.J.: Lawrence Erlbaum.

Damgaard, M. T., & Nielsen, H. S. (2018). Nudging in education. *Economics of Education Review, 64*, 313-342.

De Castella, K., & Byrne, D. (2015). My intelligence may be more malleable than yours: the revised implicit theories of intelligence (self-theory) scale is a better predictor of achievement, motivation, and student disengagement. *European Journal of Psychology of Education, 30*(3), 245-267. doi:10.1007/s10212-015-0244-y

Deutsch, N. L., & Spencer, R. (2009). Capturing the magic: Assessing the quality of youth mentoring relationships. *New Directions for Youth Development, 2009*(121), 47-70. doi:10.1002/yd.296

DuBois, D. L., Holloway, B. E., Valentine, J. C., & Cooper, H. (2002). Effectiveness of mentoring programs for youth: A meta-analytic review. *American Journal of Community Psychology, 30*(2), 157-197. doi:10.1023/a:1014628810714

DuBois, D. L., Portillo, N., Rhodes, J. E., Silverthorn, N., & Valentine, J. C. (2011). How effective are mentoring programs for youth? A systematic assessment of the evidence. *Psychological Science in the Public Interest, 12*(2), 57-91. doi:10.1177/1529100611414806

Harris, J. T., & Nakkula, M. J. (2018a). *Match Characteristics Questionnaire (MCQ). Unpublished measure.* Fairfax, VA: Applied Research Consulting.

Harris, J. T., & Nakkula, M. J. (2018b). *Youth Mentoring Survey (YMS). Unpublished measure.* Fairfax, VA: Applied Research Consulting.

Herrera, C., Grossman, J. B., Kauh, T. J., Feldman, A. F., McMaken, J., & Jucovy, L. Z. (2007). *Making a Difference in Schools: The Big Brothers Big Sisters School-Based Mentoring Impact Study.* Philadelphia, PA: Public/Private Ventures Report.

Herrera, C., Grossman, J. B., Kauh, T. J., & McMaken, J. (2011). Mentoring in schools: An impact study of Big Brothers Big Sisters school-based mentoring. *Child Development, 82*(1), 346-361.

Herrera, C., Sipe, C. L., & McClanahan, W. S. (2000). *Mentoring School-Age Children: Relationship Development in Community-Based and School-Based Programs.* Philadelphia, PA: Public/Private Ventures.

Heß, S. (2017). Randomization inference with Stata: A guide and software. *Stata Journal, 17*(3), 630-651. Retrieved from http://www.stata-journal.com/article.html?article=st0489

Karcher, M. J., Nakkula, M. J., & Harris, J. (2005). Developmental Mentoring Match Characteristics: Correspondence between Mentors' and Mentees' Assessments of Relationship Quality. *Journal of Primary Prevention, 26*(2), 93-110. doi:10.1007/s10935-005-1847-x

Kern, L., Harrison, J. R., Custer, B. E., & Mehta, P. D. (2019). Factors That Enhance the Quality of Relationships Between Mentors and Mentees During Check & Connect. *Behavioral Disorders, 44*(3), 148-161. doi:10.1177/0198742918779791

Kern, M. L., Benson, L., Steinberg, E. A., & Steinberg, L. (2016). The EPOCH measure of adolescent well-being. *Psychological assessment, 28*(5), 586-597. doi:10.1037/pas0000201

Kirk, R. E. (1994). *Experimental Design: Procedures for the Behavioral Sciences*. Belmont, CA: Brooks/Cole Pub. Co.

Kizilcec, R. F., Schneider, E., Cohen, G. L., & McFarland, D. A. (2014). Encouraging forum participation in online courses with collectivist, individualist and neutral motivational framings. *eLearning Papers, 37*, 13-22.

Kraft, M. A. (2019). *Interpreting Effect Sizes of Education Interventions. (EdWorkingPaper: 19-10)*. Retrieved from Annenberg Institute at Brown University: http://www.edworkingpapers.com/ai19-10

Lippman, L. H., Anderson Moore, K., Guzman, L., Ryberg, R., McIntosh, H., Ramos, M. F., . . . Kuhfeld, M. (2014). *Flourishing Children*: Springer.

Loewenstein, G., Krishnamurti, T., Kopsic, J., & McDonald, D. (2015). Does Increased Sexual Frequency Enhance Happiness? *Journal of Economic Behavior & Organization, 116*, 206-218.

Maynard, B. R., Kjellstrand, E. K., & Thompson, A. M. (2014). Effects of Check and Connect on Attendance, Behavior, and Academics: A Randomized Effectiveness Trial. *Research on Social Work Practice, 24*(3), 296-309. doi:10.1177/1049731513497804

Morgan, K. L., & Rubin, D. B. (2012). Rerandomization to improve covariate balance in experiments. *Annals of Statistics, 40*(2), 1263-1282. doi:10.1214/12-AOS1008

Nakkula, M. J., & Harris, J. T. (2014). Chapter 4: Assessing Mentoring Relationships. In D. L. DuBois & M. J. Karcher (Eds.), *Handbook of Youth Mentoring* (Second ed., pp. 45-62). Thousand Oaks, CA: SAGE.

Oreopoulos, P. (2020). *Promises and limitations of nudging in education. (EdWorkingPaper: 20-296).* Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/5kqf-7j06

Parra, G. R., DuBois, D. L., Neville, H. A., Pugh-Lilly, A. O., & Povinelli, N. (2002). Mentoring relationships for youth: Investigation of a process-oriented model. *Journal of Community Psychology, 30*(4), 367-388. doi:10.1002/jcop.10016

Raposa, E. B., Dietz, N., & Rhodes, J. E. (2017). Trends in Volunteer Mentoring in the United States: Analysis of a Decade of Census Survey Data. *American Journal of Community Psychology, 59*(1-2), 3-14. doi:10.1002/ajcp.12117

Raposa, E. B., Rhodes, J., Stams, G. J. J. M., Card, N., Burton, S., Schwartz, S., . . . Hussain, S. (2019). The effects of youth mentoring programs: A meta-analysis of outcome studies. *Journal of Youth and Adolescence, 48*(3), 423-443. doi:10.1007/s10964-019-00982-8

Rhodes, J. E., Spencer, R., Keller, T. E., Liang, B., & Noam, G. (2006). A model for the influence of mentoring relationships on youth development. *Journal of Community Psychology, 34*(6), 691-707. doi:10.1002/jcop.20124

Robinson, C. D., Lee, M. G., Dearing, E., & Rogers, T. (2018). Reducing student absenteeism in the early grades by targeting parental beliefs. *American Educational Research Journal, 55*(6), 1163-1192.

Rogers, T., & Feller, A. (2018). Reducing student absences at scale by targeting parents' misbeliefs. *Nature Human Behaviour, 2*(5), 335-342. doi:10.1038/s41562-018-0328-1

Rones, M., & Hoagwood, K. (2000). School-Based Mental Health Services: A Research Review. *Clinical Child and Family Psychology Review, 3*(4), 223-241. doi:10.1023/A:1026425104386

Schochet, P. Z. (2008). *Technical methods report: Guidelines for multiple testing in impact evaluations (NCEE 2008-4018)*. Washington, D.C.: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin Company.

Simonton, S. (2020). MENTOR Portal Lets Youth Organizations Go Virtual, Shift Mentoring Online [Press release]. Retrieved from https://www.evidencebasedmentoring.org/mentor-portal-lets-youth-organizations-go-virtual-shift-mentoring-online/

WestEd. (2008). California Healthy Kids Survey (CHKS). Retrieved from https://www.wested.org/project/california-healthy-kids-survey-chks/

What Works Clearinghouse. (2020). *What Works Clearinghouse Standards Handbook, Version 4.1*. U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, What Works Clearinghouse. Washington, DC.

Woods, C. S., & Preciado, M. (2016). Student–Mentor relationships and students' college attitudes. *Journal of Education for Students Placed at Risk (JESPAR), 21*(2), 90-103.

Figure 1. Mentor Reminder #1, sent: Monday, Feb 3, 2020

Dear Mentor,

Have you been in touch with your mentees outside of class recently? With the new semester just getting started, your support can give your mentees confidence to put their best effort forward. Take a moment today to reach out by email, text, or phone, and remember to follow-up to your mentees' responses!

Ask your mentee:

- Light, easy questions to respond to:
  - What's new since we last saw each other?
  - Happy Monday! What are you looking forward to this week?
  - How is the start to your week going?
- Questions that show you care:
  - Are you looking forward to the new semester?
  - What are you most proud of about last semester?
  - Is there anything you want to improve at this semester? How can I support you?

Thanks for being a mentor! Your support means so much to your mentees!

Table 1. Summary statistics and treatment-comparison group balance on student and mentor baseline characteristics

Variable	Min	Max	Overall mean/prop.	SD	Treatment group mean/prop.	Comparison group mean/prop.	Difference	Effect size of diff.	*p*-value of diff.
Student characteristics									
Female	0	1	0.51	0.50	0.48	0.54	-0.06	-0.11	0.454
Black	0	1	0.59	0.49	0.60	0.58	0.02	0.05	0.747
White	0	1	0.10	0.30	0.13	0.07	0.06	0.20	0.187
Other race	0	1	0.31	0.46	0.27	0.35	-0.08	-0.18	0.226
Receives FRPL	0	1	0.81	0.39	0.78	0.84	-0.06	-0.15	0.310
Mother has a BA degree	0	1	0.43	0.50	0.45	0.39	0.06	0.12	0.435
Father has a BA degree	0	1	0.30	0.46	0.29	0.32	-0.03	-0.07	0.677
Either parent has a BA degree	0	1	0.48	0.46	0.49	0.47	0.02	0.04	0.783
Mentor characteristics									
Female	0	1	0.51	0.50	0.48	0.53	-0.05	-0.11	0.467
Black	0	1	0.40	0.49	0.32	0.48	-0.15*	-0.31	0.029
White	0	1	0.46	0.50	0.50	0.41	0.09	0.18	0.219
Other race	0	1	0.15	0.35	0.18	0.11	0.07	0.18	0.202
Experience mentoring	0	17	2.46	3.12	2.73	2.15	0.58	0.19	0.199
Self-reported student attitudes									
Self-efficacy	1	5	3.90	0.69	3.89	3.91	-0.02	-0.03	0.835
Growth mindset	1	4	2.05	0.74	1.95	2.17	-0.22*	-0.30	0.040
Goal orientation	1	5	3.90	0.70	3.86	3.95	-0.09	-0.13	0.388
Perseverance	1	5	3.78	0.83	3.76	3.80	-0.05	-0.06	0.704
Adult Support	1	4	3.42	0.72	3.38	3.47	-0.09	-0.12	0.401
Mentoring relationship quality									
Student: relational quality	1	4	2.96	0.68	2.90	3.04	-0.14	-0.20	0.185
Student: instrumental quality	1	4	2.74	0.74	2.64	2.84	-0.19~	-0.26	0.080
Mentor: closeness	1	6	4.47	0.79	4.37	4.60	-0.24	-0.30	0.139
Mentor: satisfaction	1	6	4.09	1.01	3.99	4.22	-0.23	-0.23	0.267
Virtual communication frequency									
Mentor outreach frequency: student-reported†	0	12	1.81	2.21	1.71	1.93	-0.21	-0.10	0.519
Student outreach frequency: mentor-reported†	0	4	0.61	1.00	0.53	0.74	0	-0.21	0.294
Student responsiveness: student-reported	1	4	2.08	1.04	2.02	2.15	-0.13	-0.13	0.400
Student responsiveness: mentor-reported	1	4	1.56	0.64	1.51	1.64	0	-0.21	0.293
Student sample size			192		102	90	*F*-test for joint significance		1.204
Mentor sample size			114		59	55	*p*-value of F-test		0.273

Notes: Table displays summary statistics (min, max, mean/proportion, and standard deviation) of student and mentor characteristics measured at baseline for the sample of students and mentors included in main analytic sample. Results of analyses to assess baseline balance across treatment statuses also displayed. Balance assessed in terms of each covariate individually, as well as all covariates jointly. †Outreach frequency measures sum the number of contacts via text, email, and phone received in an average week.

Figure 2. Frequency of mentor outreach and student outreach (by text, email, phone) at baseline. Respondents reported how often in an average week they recalled sending/receiving a point of contact via each method.

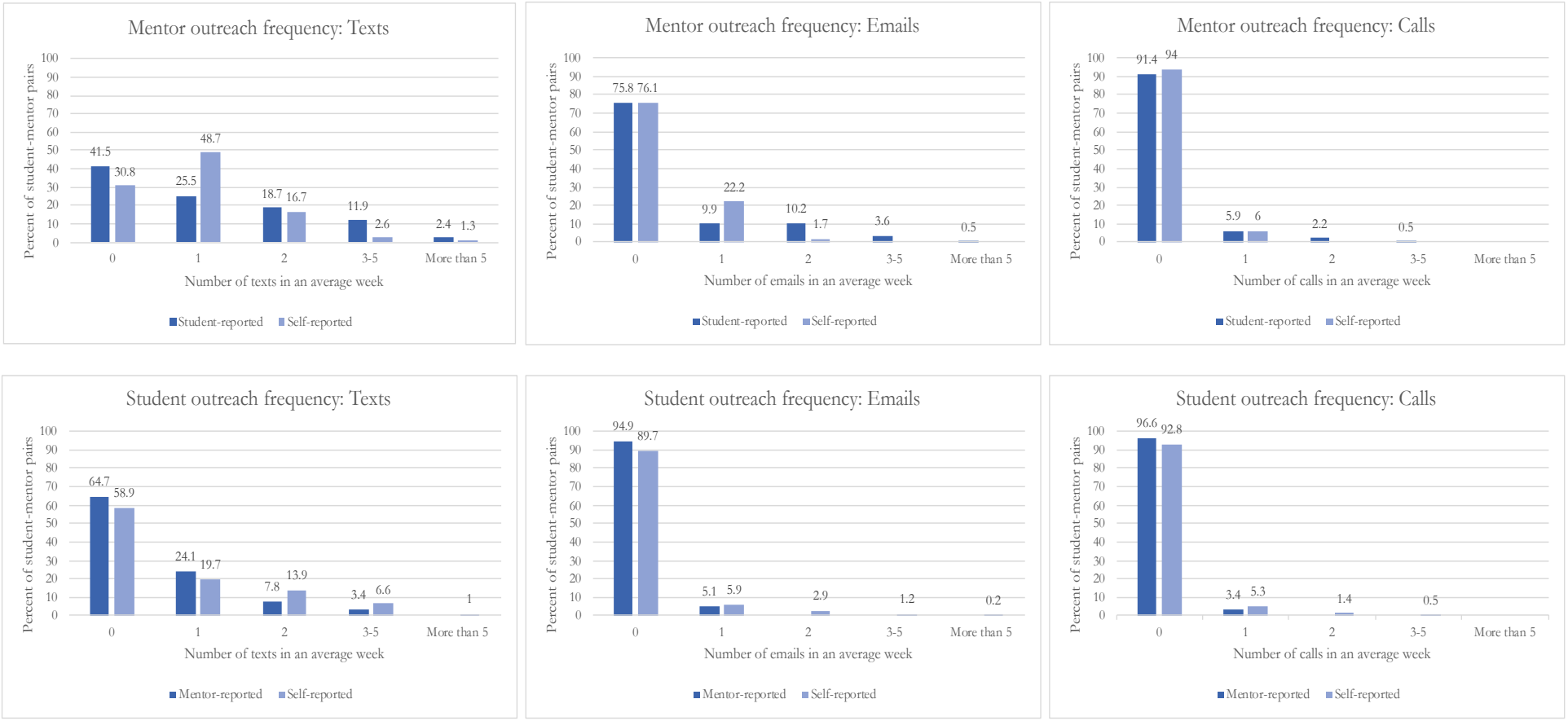


Table 2. Predictors of more frequent virtual communication at baseline

Dependent variable:	Mentor outreach frequency		Student outreach frequency	
	(1)	(2)	(3)	(4)
Mentor characteristics				
Female	0.23	-0.06	0.21	0.25
	(0.25)	(0.25)	(0.19)	(0.22)
Black	0.21	0.09	0.23	0.23
	(0.28)	(0.27)	(0.19)	(0.19)
Experience mentoring	-0.10*	-0.10*	0.06	0.02
	(0.04)	(0.04)	(0.06)	(0.06)
Avg. interest in WF program	-0.14	-0.45	-0.13	-0.40
	(0.40)	(0.41)	(0.23)	(0.24)
Student characteristics				
Receives FRPL	0.64*	0.56*	0.11	0.05
	(0.28)	(0.27)	(0.19)	(0.17)
Either parent has a BA degree	-0.02	-0.02	0.01	-0.05
	(0.24)	(0.24)	(0.14)	(0.15)
Student attitudes				
Self-efficacy	-0.13	-0.36~	-0.04	-0.07
	(0.15)	(0.21)	(0.06)	(0.05)
Growth mindset	-0.04	-0.08	0.05	0.03
	(0.13)	(0.14)	(0.06)	(0.05)
Goal orientation	0.22*	0.18	0.00	-0.00
	(0.11)	(0.14)	(0.07)	(0.08)
Perseverance	0.15	0.14	0.04	0.12
	(0.11)	(0.14)	(0.06)	(0.07)
Adult Support	0.07	0.01	-0.03	-0.07
	(0.13)	(0.12)	(0.07)	(0.07)
Mentoring relationship quality				
Student: Relational quality	0.56***	0.06	0.20**	0.16
	(0.09)	(0.23)	(0.06)	(0.13)
Student: Instrumental quality	0.65***	0.59*	0.17*	-0.05
	(0.12)	(0.27)	(0.07)	(0.15)
Mentor: Relational closeness	0.14	0.12	0.23**	-0.01
	(0.15)	(0.26)	(0.08)	(0.14)
Mentor: Relational satisfaction	0.13	-0.11	0.32***	0.31~
	(0.18)	(0.31)	(0.09)	(0.18)
Constant	Varies	3.27**	Varies	1.07~
		(1.25)		(0.56)
Classroom FE	x	x	x	x
SE clustered within mentors	x	x	x	x
Observations	Varies	415	Varies	234
R-squared	Varies	0.20	Varies	0.29

Notes: Table displays results of OLS regressions predicting mentor outreach frequency and student outreach frequency as a function of baseline covariates. Results in columns (1) and (3) show the estimated correlation between each predictor and the outcome by itself, and columns (2) and (4) shows correlations when all predictors are included in the model together. Outreach frequency measures summarize the number of texts, emails, and calls received during an average week. Analyses used non-imputed versions of outcomes variables in their original, non-standardized units. Robust standard errors, clustered within mentor groups are displayed in parentheses. *** p<0.001, ** p<0.01, * p<0.05, ~ p<0.1

Table 3. Effects of mentor reminders on outreach frequency, student responsiveness, and mentoring relationship quality

Panel A: Virtual outreach frequency and student responsiveness

	Student-reported		Mentor-reported	
Dependent variable:	Mentor outreach frequency	Student responsiveness	Student outreach frequency	Student responsiveness
Treatment effect	0.05	-0.33*	-0.27*	0.10
	(0.15)	(0.13)	(0.12)	(0.14)
Comparison group. mean	3.00	2.65	1.05	1.73
Comparison group SD	2.78	1.09	1.30	0.66
Observations	192	192	365	365
R-squared	0.30	0.37	0.30	0.35

Panel B: Mentoring relationship quality

	Student-reported		Mentor-reported	
Dependent variable:	Relational quality	Instrumental quality	Relational closeness	Relational satisfaction
Treatment effect	-0.12	-0.14	-0.26*	-0.33**
	(0.12)	(0.12)	(0.11)	(0.10)
Comparison group mean	3.06	2.96	4.46	4.41
Comparison group SD	0.74	0.80	0.83	1.15
Observations	192	192	365	365
R-squared	0.54	0.57	0.44	0.51

Notes: Table displays results of OLS regressions estimating treatment effects on virtual outreach frequency, student responsiveness to mentor outreach, and mentoring relationship quality. Outreach frequency measures summarize the number of texts, emails, and calls received during an average week. Student responsiveness gauges how students respond when their mentor reached out via text, email, and phone call (separate items, 1=Never to 4=Always; averaged across communication methods to create a single composite variable). Estimated impacts are reported in effect sizes: outcomes are standardized to the comparison group mean and SD. Robust SE clustered within mentor groups are displayed in parentheses. Comparison group mean and SD in original (unstandardized) units are displayed. Control variables included: classroom fixed-effects, mentor characteristics, student characteristics, baseline measures of student attitudes, mentoring relationship quality, virtual outreach, and student responsiveness. For observations where baseline values were missing, the overall item mean was imputed. *p*-values computed using randomization inference ("ritest" package in Stata). *** p<0.001, ** p<0.01, * p<0.05, ~ p<0.1

Table 4. Effect of mentor reminders on students' self-reported attitudes

Dependent variable:	Self-efficacy	Growth mindset	Goal orientation	Perseverance	Adult support
Treatment effect	-0.28*	-0.03	-0.2*	-0.21*	0.00
	(0.12)	(0.12)	(0.11)	(0.12)	(0.14)
Comparison group mean	4.27	2.17	4.07	3.97	3.49
Comparison group SD	0.60	0.83	0.68	0.78	0.62
Observations	192	192	192	192	192
R-squared	0.43	0.50	0.51	0.43	0.36

Notes: Table displays results of OLS regressions estimating treatment effects on students' self-reported attitudes. Estimated impacts are reported in effect sizes: outcomes are standardized to the comparison group mean and SD. Robust SE clustered within mentor groups are displayed in parentheses. Comparison group mean and SD in original (unstandardized) units are displayed. Control variables included: classroom fixed-effects, mentor characteristics, student characteristics, baseline measures of student attitudes, mentoring relationship quality, virtual outreach, and student responsiveness. For observations where baseline values were missing, the overall item mean was imputed. *p*-values computed using randomization inference ("ritest" package in Stata). *** p<0.001, ** p<0.01, * p<0.05, ~ p<0.1

Table 5. Effects of mentor reminders on outreach by and student responsiveness to texts, emails, and phone calls

Panel A: Outreach frequency

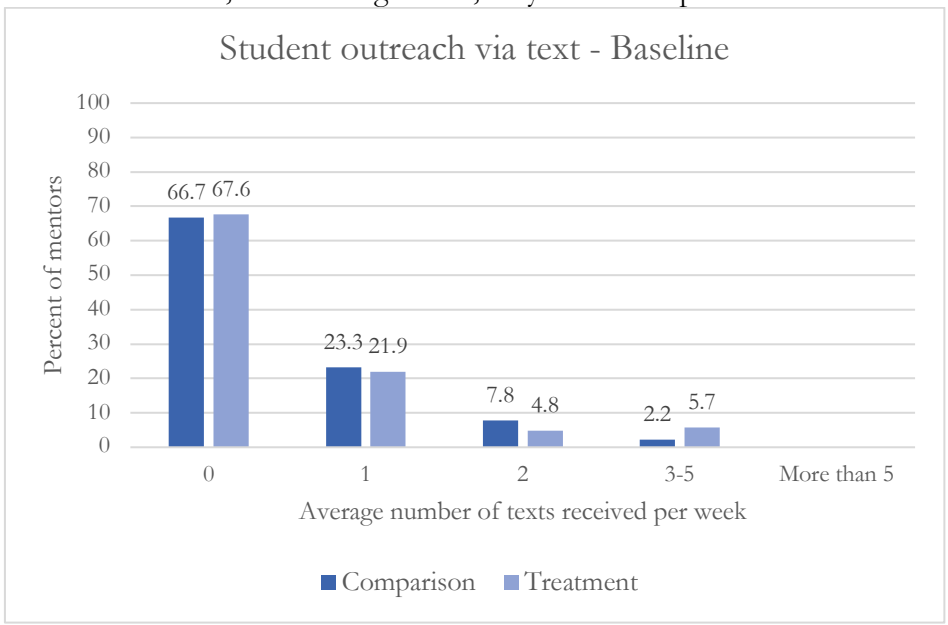
	Student-reported			Mentor-reported		
Dependent variable:	Mentor outreach frequency: Texts	Mentor outreach frequency: Emails	Mentor outreach frequency: Calls	Student outreach frequency: Texts	Student outreach frequency: Emails	Student outreach frequency: Calls
Treatment effect	0.102	0.349	-0.13	-0.34~	0.055	-0.04
	(0.24)	(0.23)	(0.15)	(0.18)	(0.05)	(0.04)
	[1.86, 1.60]	[0.61, 1.15]	[0.52, 1.00]	[0.93, 1.17]	[0.06, 0.29]	[0.05, 0.34]

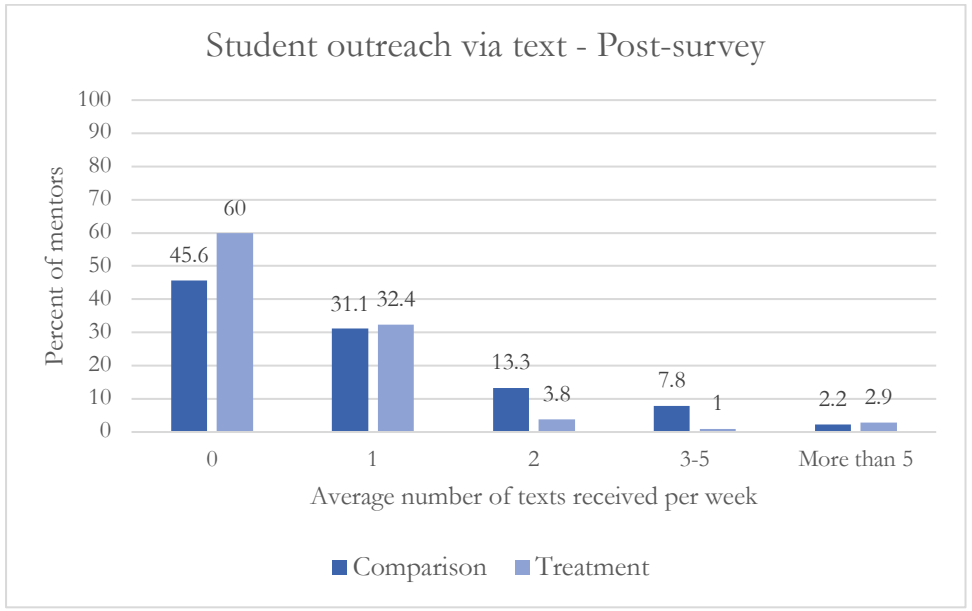
Panel B: Student responsiveness

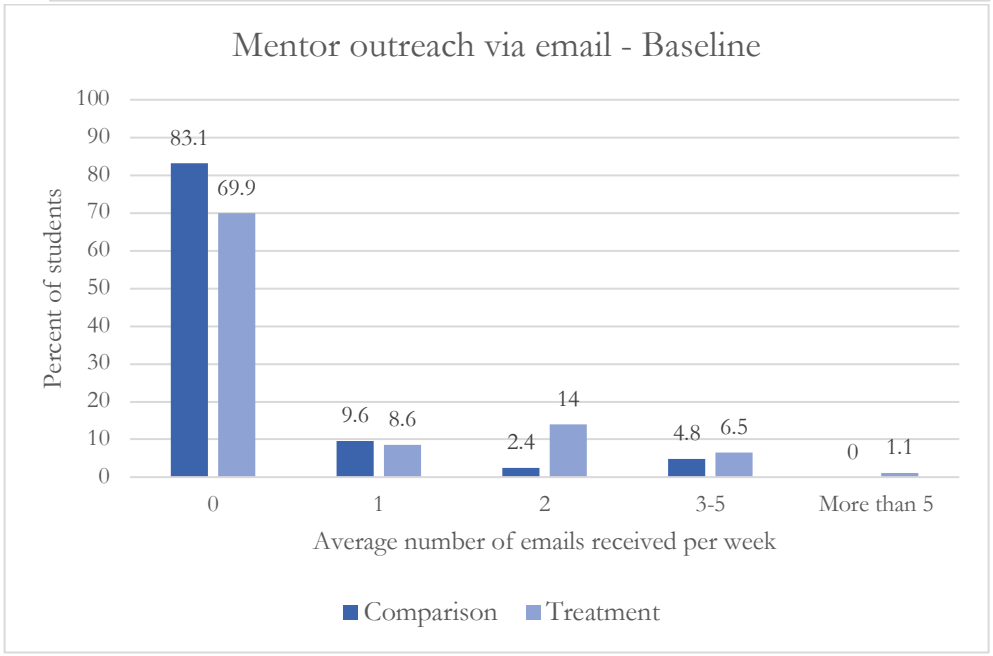
	Student-reported			Mentor-reported		
Dependent variable:	Student responsiveness to texts	Student responsiveness to emails	Student responsiveness to calls	Student responsiveness to texts	Student responsiveness to emails	Student responsiveness to calls
Treatment effect	-0.38*	-0.35	-0.43*	-0.02	0.329**	-0.31~
	(0.16)	(0.21)	(0.22)	(0.22)	(0.13)	(0.15)
	[3.24, 1.12]	[2.25, 1.31]	[2.44, 1.38]	[2.57, 1.12]	[1.21, 0.58]	[1.31, 0.77]

Notes: Table displays results of OLS regressions estimating treatment effects on outreach frequency (Panel A) and student responsiveness (Panel B) with results for text messaging, emails, and phone calls displayed separately. Each coefficient displayed is the estimated effect from a separate regression. Outreach variables measured the frequency of texts, emails, or calls (separate items) received during an average week. Student responsiveness gauges how often students responded when their mentor reached out via text, email, or phone call (separate items, original scale: 1=Never to 4=Always). Estimated impacts are reported in effect sizes: outcomes are standardized to the comparison group mean and SD. Robust SE clustered within mentor groups are displayed in parentheses. Comparison group means and SD in original units are displayed in square brackets. Control variables included: classroom fixed-effects, mentor characteristics, student characteristics, baseline measures of student attitudes, mentoring relationship quality, virtual outreach, and student responsiveness. For observations where baseline values were missing, the overall item mean was imputed. *p*-values computed using randomization inference ("ritest" package in Stata). *** p<0.001, ** p<0.01, * p<0.05, ~ p<0.1

Figure 3. Frequency of student outreach via text (top) and mentor outreach via email (bottom) at baseline and post-survey. Respondents reported how often, in an average week, they received a point of contact.







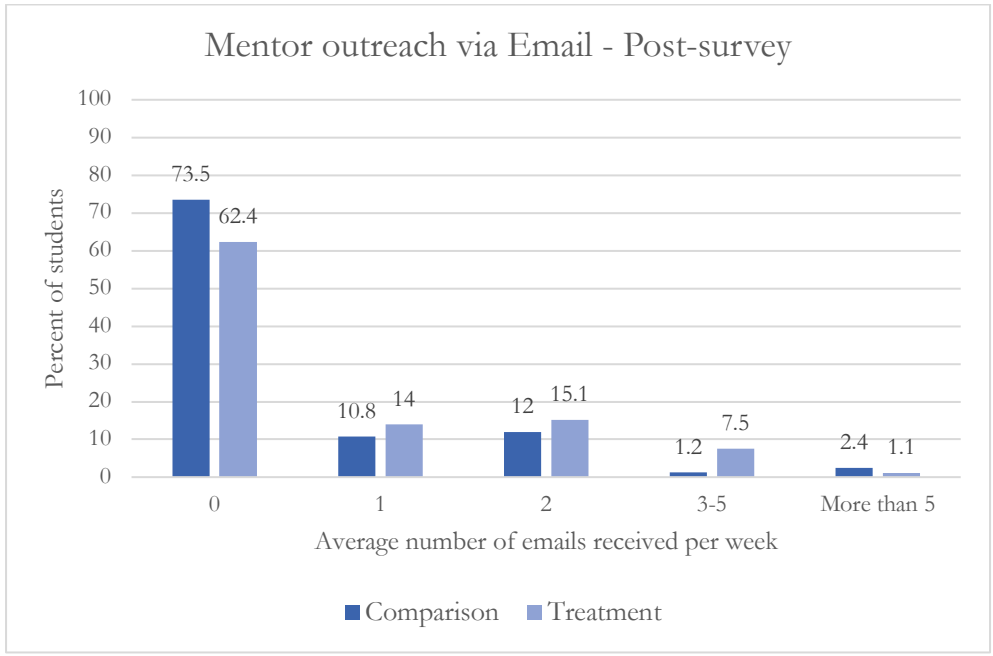


Table 6. Effects of mentor reminders on outreach frequency and student responsiveness along intensive and extensive margins

Dependent variable:	Student Reported		Mentor Reported	
	Mentor outreach frequency	Student responsiveness	Student outreach frequency	Student responsiveness
Outcome = Any outreach, ever responded (vs. No outreach, never responded)	-0.03 (0.07) [0.8, 0.40]	0.036 (0.05) [0.84, 0.36]	-0.11 (0.07) [0.57, 0.49]	0 (0.05) [0.77, 0.41]
Outcome = Received 2+ touchpoints, always responded (vs. Received <2 touchpoints, did not always respond)	0.071 (0.08) [0.61, 0.49]	-0.2** (0.06) [0.36, 0.48]	-0.12* (0.05) [0.25, 0.43]	0.024 (0.02) [0.03, 0.18]

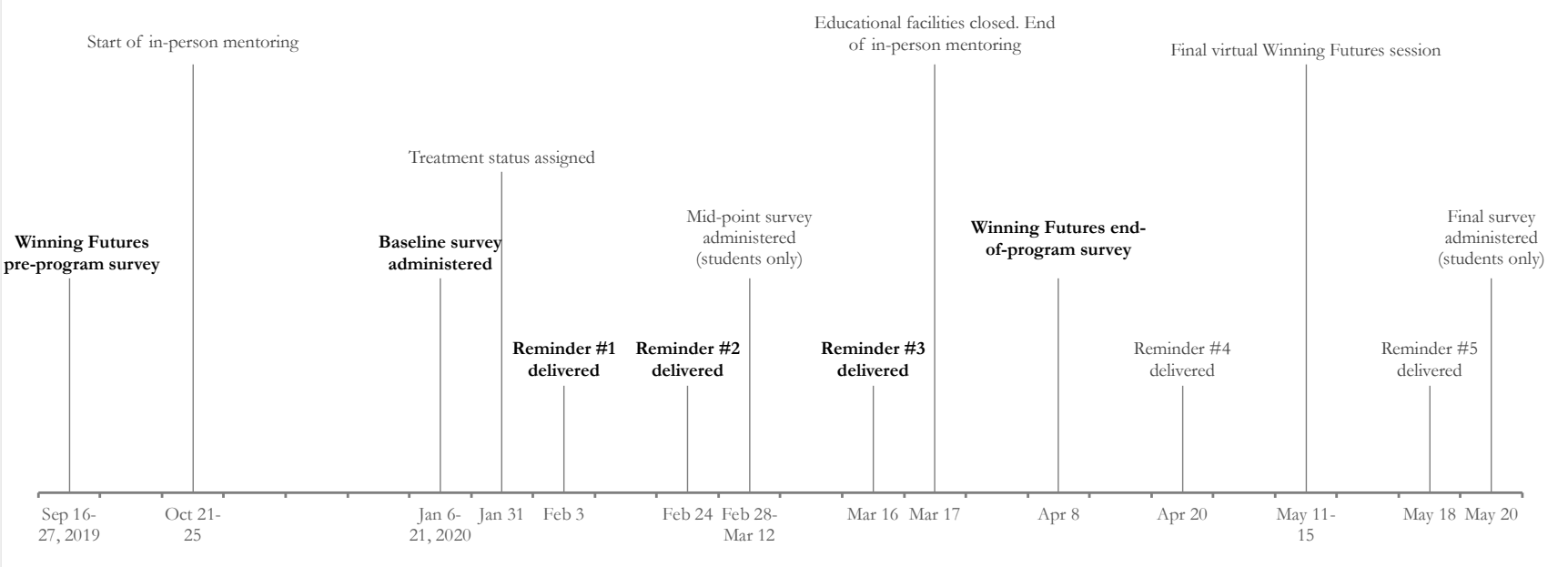
Notes: Table displays results of linear probability models using OLS regressions to estimate treatment effects on virtual outreach frequency and student responsiveness to mentor outreach. Outreach measures indicate the frequency of texts, emails, and calls received during an average week. Student responsiveness gauges how students respond when their mentor reached out via text, email, and phone call. Binary outcomes were estimated along two margins: Any outreach/response (=1) vs. none (=0) (top row), and Frequent/Consistent outreach and response (=1) vs. less-frequent/less-consistent (=0) (bottom row). Estimated impacts represent percentage point changes in the outcomes. Robust SE clustered within mentor groups are displayed in parentheses. Comparison group means and SD in original units are displayed in square brackets. Control variables included: classroom fixed-effects, mentor characteristics, student characteristics, baseline measures of student attitudes, mentoring relationship quality, virtual outreach, and student responsiveness. For observations where baseline values were missing, the overall item mean was imputed. *p*-values computed using randomization inference ("ritest" package in Stata).
*** p<0.001, ** p<0.01, * p<0.05, ~ p<0.1

Appendix Table 1. Measures of virtual communication, mentoring relationship quality, and student attitudes

Outcome, description	Measure, sample items, internal consistency (for validated survey measures), and date collected
Virtual communication	
Mentor/student outreach: Average number of times communicated via text, email, and phone	Items: "In an average week, how many times do you: Receive a text sent by your mentor (mentee)? Receive an email sent by your mentor (mentee)? Send a text to your mentor (mentee)? Send an email to your mentor (mentee)?" (Scale: 0, 1, 2, 3-5, More than 5) Collected: Mentors: Jan-20, Apr-20; Students: Jan-20, Feb-20, Apr-20, May-20
Student responsiveness: Frequency with which mentors/students perceive students' response to mentors' outreach	Items: Student version: "When your mentor sends a text/sends an email/gives a call, how often do you respond?" (4-point scale: Never … Always) Mentor version: "When you send a text/send an email/give a call, how often does your mentee respond?" (4-point scale: Never … Always) Collected: Mentors: Jan-20, Apr-20; Students: Jan-20, Feb-20, Apr-20, May-20
Mentoring relationship quality	
Relational quality (students' perceptions): Student feels happy, close, satisfied with relationship (relational); How much the student is open to support and perceives benefit from it (instrumental)	Sample items: Relational quality - "My mentor really cares about me … My mentor knows what is going on in my life." Instrumental quality - "I am doing better at school because of my mentor's help … My mentor listens to me better than other adults do." (Scale: 1=Not at all true … 4=Very true) Source: Youth Mentoring Survey (Harris & Nakkula, 2018b) Int. consistency: *Current study* – Relational quality (14 items): Baseline survey α=0.92; Post-survey α=0.93; Instrumental quality (8 items): Baseline survey α=0.88; Post-survey α=0.92. *Published* – In a sample of U.S. youth of all grade-levels, alpha for relational quality scale was α=0.84, alpha for instrumental quality scale was α=0.76. Collected: Jan-20, Feb-20, Apr-20, May-20
Relational quality (mentors' perceptions): Mentor feels close with student (closeness); Mentor's sense of fulfillment in the relationship (satisfaction)	Sample items: Closeness - "I feel like my mentee and I have a strong bond (are close or deeply connected) … I can trust what my mentee tells me." Satisfaction - "I feel like the mentoring relationship is getting stronger … I feel unsure that my mentee is getting enough out of our mentoring relationship." (Scale: 1=Never… 6=Often) Source: Match Characteristics Questionnaire (Harris & Nakkula, 2018a) Int. consistency: *Current study* – Relational closeness (4 items): Baseline survey α=0.85; Post-survey α=0.82; Relational satisfaction (5 items): Baseline survey α=0.88; Post-survey α=0.92. *Published* – Samples of mentors of youth across large mentoring organizations in the U.S., alpha for closeness scale was α=0.82, alpha for satisfaction scale was α=0.85. Collected: Jan-20, Apr-20

Outcome, description	Measure, sample items, internal consistency (for validated survey measures), and date collected
Student attitudes	
Self-efficacy: Youths' sense of confidence in and progress towards personal goals	Sample items: For each sentence, decide how much you disagree or agree with it: "I will be able to achieve most of the goals that I have set for myself," and "When facing difficult tasks, I am certain that I will accomplish them." (Scale: Strongly disagree, Disagree, Neutral, Agree, Strongly agree) Source: New General Self-Efficacy Scale (Chen, Gully, & Eden, 2001). Int. consistency: *Current study* – 7 items measured self-efficacy: Baseline survey α=0.91; Post-survey α=0.9 *Published* – In two samples of U.S. undergraduate students, alpha levels ranged from α = .86 to .90 Collected: Sep-19, Apr-20
Growth mindset: Extent to which youth see intelligence as malleable.	Sample items: Mark the choice that shows how much you agree with it: "You can learn new things, but you can't really change your basic intelligence," and "You have a certain amount of intelligence and you really can't do much to change it." (Scale: Strongly disagree, Disagree, Agree, Strongly agree) Source: Revised Implicit Theories of Intelligence (Self-Theory) Scale (De Castella & Byrne, 2015). Int. consistency: *Current study* – 3 items measured growth mindset: Baseline survey α=0.82; Post-survey α=0.88 *Published* – Sample of *N*=680 Australian young people (grades 10-12) alpha level for the Entity Beliefs Subscale was α=.87. Collected: Sep-19, Apr-20
Goal orientation: Frequency and approach to planning for goals.	Sample items: Indicate how much these statements describe you: "I have goals in my life," and "If I set goals, I take action to reach them." (Scale: Not at all like me … Exactly like me) Source: Flourishing Children Project, goal orientation subscale (Lippman et al., 2014). Int. consistency: *Current study* – 7 items measured goal orientation: Baseline survey α=0.86; Post-survey α=0.85 *Published* – Sample of *N*=680 Australian young people (grades 10-12) alpha level for the Entity Beliefs Subscale was α=.87. Diverse sample of U.S. youth, alpha for two goal orientation sub-scales was α=0.88. Collected: Sep-19, Apr-20
Perseverance: Youths' belief that they can complete school-related tasks.	Sample items: Indicate how much each statement describes you: "I finish whatever I begin," and "I keep at my school work until I am done with it." (Scale: Almost never, Sometimes, Often, Very often, Almost always) Source: EPOCH measure of adolescent well-being-perseverance scale (M. L. Kern, Benson, Steinberg, & Steinberg, 2016). Int. consistency: *Current study* – 4 items measured perseverance: Baseline survey α=0.81; Post-survey α=0.79 *Published* – Sample of *N*=680 Australian young people (grades 10-12) alpha level for the Entity Beliefs Subscale was α=.87. Across multiple samples of U.S. youth alpha levels ranged from α=0.72 to 0.85. Collected: Sep-19, Apr-20
Adult support Level of support from caring adults outside of school	Sample items: How true are these statements for an adult outside of your family or school: "There is an adult who always wants you to do your best," and "There is an adult who believes that you will be a success" (Scale: Not at all true … Very much true) Source: California Healthy Kids Survey: Resilience & Youth Development Module (WestEd, 2008) Int. consistency: *Current study* – 6 items measured adult support: Baseline survey α=0.89; Post-survey α=0.87 *Published* – No published studies of internal consistency available Collected: Sep-19, Apr-20

Appendix Figure 1. Study timeline: data collection and intervention implementation



Notes: Bolded items indicate interventions and measures assessed in main analyses

Appendix Table 2. Summary statistics and baseline balance assessment for students and mentors in the original experimental sample

Variable	Min	Max	Overall mean/prop.	SD	Treatment group mean/prop.	Comparison group mean/prop.	Difference	Effect size of diff.	*p*-value of diff.
Student characteristics									
Female	0	1	0.45	0.50	0.44	0.46	-0.02	-0.05	0.612
Black	0	1	0.62	0.49	0.65	0.59	0.06	0.13	0.161
White	0	1	0.11	0.31	0.10	0.12	-0.02	-0.06	0.490
Other race	0	1	0.28	0.45	0.25	0.30	-0.04	-0.10	0.294
Receives FRPL	0	1	0.80	0.39	0.78	0.82	-0.04	-0.10	0.265
Mother has a BA degree	0	1	0.39	0.49	0.42	0.35	0.07	0.14	0.162
Father has a BA degree	0	1	0.26	0.44	0.26	0.27	-0.02	-0.04	0.716
Either parent has a BA degree	0	1	0.44	0.45	0.45	0.43	0.02	0.05	0.604
Mentor characteristics									
Female	0	1	0.45	0.50	0.45	0.45	0.00	0.00	0.964
Black	0	1	0.38	0.49	0.37	0.39	-0.02	-0.03	0.725
White	0	1	0.48	0.50	0.48	0.48	0.01	0.02	0.843
Other race	0	1	0.12	0.33	0.13	0.12	0.01	0.02	0.810
Experience mentoring	0	17	2.32	2.78	2.49	2.14	0.34	0.12	0.173
Student attitudes									
Self-efficacy	1	5	3.86	0.70	3.85	3.87	-0.02	-0.03	0.765
Growth mindset	1	4	2.14	0.74	2.11	2.17	-0.07	-0.09	0.335
Goal orientation	1	5	3.82	0.76	3.82	3.82	-0.01	-0.01	0.936
Perseverance	1	5	3.67	0.84	3.70	3.65	0.06	0.07	0.476
Adult Support	1	4	3.43	0.71	3.42	3.43	-0.01	-0.01	0.890
Mentoring relationship quality									
Student: relational quality	1	4	2.86	0.68	2.82	2.91	-0.08	-0.12	0.205
Student: instrumental quality	1	4	2.66	0.72	2.62	2.71	-0.10	-0.14	0.165
Mentor: closeness	1	6	4.37	0.90	4.36	4.39	-0.03	-0.03	0.793
Mentor: satisfaction	1	6	3.97	1.07	3.93	4.00	-0.07	-0.07	0.615
Virtual communication frequency									
Mentor outreach frequency	0	12	1.81	2.14	1.66	1.97	-0.31	-0.15	0.139
Student outreach frequency	1	4	2.01	1.02	1.98	2.04	-0.06	-0.06	0.556
Student responsiveness: student-reported	0	4	0.62	0.98	0.56	0.68	0.00	-0.13	0.339
Student responsiveness: mentor-reported	1	4	1.52	0.63	1.52	1.52	0.00	-0.01	0.954
Student sample size			494		252	242	*F*-test for joint significance		1.205
Mentor sample size			143		74	69	*p*-value of F-test		0.281

Notes: Table displays summary statistics (min, max, mean/proportion, and standard deviation) of student and mentor characteristics measured at baseline for the sample of students and mentors included in original experimental sample. Results of analyses to assess baseline balance across treatment statuses also displayed. Balance assessed in terms of each covariate individually, as well as all covariates jointly.

Appendix Table 3. Summary statistics and baseline balance assessment for sample of students whose mentors responded to the post-survey

Variable	Min	Max	Overall mean/prop.	SD	Treatment group mean/prop.	Comparison group mean/prop.	Difference	Effect size of diff.	*p*-value of diff.
Student characteristics									
Female	0	1	0.45	0.50	0.42	0.47	-0.05	-0.09	0.374
Black	0	1	0.58	0.49	0.61	0.53	0.08	0.16	0.136
White	0	1	0.12	0.33	0.11	0.13	-0.02	-0.07	0.533
Other race	0	1	0.30	0.46	0.27	0.33	-0.06	-0.12	0.247
Receives FRPL	0	1	0.79	0.41	0.77	0.80	-0.02	-0.06	0.583
Mother has a BA degree	0	1	0.38	0.49	0.42	0.33	0.09	0.19	0.101
Father has a BA degree	0	1	0.26	0.44	0.26	0.27	-0.01	-0.03	0.841
Either parent has a BA degree	0	1	0.44	0.45	0.46	0.41	0.04	0.10	0.357
Mentor characteristics									
Female	0	1	0.44	0.50	0.42	0.47	-0.04	-0.09	0.416
Black	0	1	0.34	0.47	0.29	0.40	-0.11*	-0.23	0.031
White	0	1	0.51	0.50	0.54	0.48	0.06	0.13	0.219
Other race	0	1	0.13	0.34	0.15	0.10	0.04	0.13	0.202
Experience mentoring	0	17	2.45	2.94	2.69	2.18	0.51~	0.17	0.100
Student attitudes									
Self-efficacy	1	5	3.85	0.68	3.84	3.87	-0.02	-0.04	0.734
Growth mindset	1	4	2.15	0.72	2.14	2.16	-0.02	-0.02	0.828
Goal orientation	1	5	3.81	0.76	3.79	3.83	-0.04	-0.06	0.589
Perseverance	1	5	3.66	0.82	3.67	3.65	0.02	0.02	0.824
Adult Support	1	4	3.42	0.71	3.39	3.46	-0.07	-0.10	0.341
Mentoring relationship quality									
Student: relational quality	1	4	2.84	0.66	2.78	2.92	-0.14~	-0.21	0.053
Student: instrumental quality	1	4	2.64	0.71	2.58	2.71	-0.12	-0.17	0.118
Mentor: closeness	1	6	4.40	0.93	4.39	4.41	-0.02	-0.02	0.868
Mentor: satisfaction	1	6	4.01	1.12	3.95	4.08	-0.13	-0.12	0.398
Virtual communication frequency									
Mentor outreach frequency	0	12	1.80	2.14	1.60	2.03	-0.43~	-0.20	0.071
Student outreach frequency	0	4	0.60	1.01	0.57	0.63	0	-0.06	0.696
Student responsiveness: student-reported	1	4	2.04	1.04	1.97	2.13	-0.16	-0.16	0.160
Student responsiveness: mentor-reported	1	4	1.53	0.66	1.53	1.54	0	0.00	0.981
Student sample size			366		194	172	*F*-test for joint significance		1.084
Mentor sample size			120		64	56	*p*-value of F-test		0.386

Notes: Table displays summary statistics (min, max, mean/proportion, and standard deviation) of student and mentor characteristics measured at baseline for the sample of students whose mentors responded to the post-survey. Results of analyses to assess baseline balance across treatment statuses also displayed. Balance assessed in terms of each covariate individually, as well as all covariates jointly.

Appendix Table 4. Alternative model specification estimating main treatment effects on virtual communication and mentoring relationship quality

Panel A: Communication frequency and student responsiveness

	Student-reported		Mentor-reported	
Dependent variable:	Mentor outreach frequency	Student responsiveness	Student outreach frequency	Student responsiveness
Treatment effect	0.07	-0.33*	-0.29*	0.07
	(0.15)	(0.15)	(0.14)	(0.15)
Comparison unstd. mean	3.00	2.65	1.05	1.73
Comparison unstd.SD	2.78	1.09	1.30	0.66
Observations	192	192	365	365

Panel B: Mentoring relationship quality

	Student-reported		Mentor-reported	
Dependent variable:	Relational quality	Instrumental quality	Relational closeness	Relational satisfaction
Treatment effect	-0.13	-0.15	-0.26*	-0.35**
	(0.13)	(0.13)	(0.11)	(0.11)
Comparison unstd. mean	3.06	2.96	4.46	4.41
Comparison unstd. SD	0.74	0.80	0.83	1.15
Observations	192	192	365	365

Notes: Table displays results of regressions estimating treatment effects on virtual outreach frequency, student responsiveness to mentor outreach, and mentoring relationship quality. Mentor-random effects accounted for clustering of students within mentor-groups. Additional control variables included: classroom fixed-effects, mentor characteristics, student characteristics, baseline measures of student attitudes, mentoring relationship quality, virtual outreach, and student responsiveness. For observations where baseline values were missing, overall item mean imputed. Estimated impacts reported in effect sizes (outcomes standardized to the comparison group mean and SD). Comparison group mean and SD in original (unstandardized) units displayed.
*** p<0.001, ** p<0.01, * p<0.05, ~ p<0.1

Appendix Table 5. Alternative model specification estimating main treatment effects on student self-reported attitudes

Dependent variable:	Self-efficacy	Growth mindset	Goal orientation	Perseverance	Adult support
Treatment effect	-0.29*	-0.07	-0.21~	-0.22~	-0.03
	(0.13)	(0.15)	(0.12)	(0.12)	(0.15)
Comparison unstd. mean	4.27	2.17	4.07	3.97	3.49
Comparison unstd. SD	0.60	0.83	0.68	0.78	0.62
Observations	192	192	192	192	192

Notes: Table displays results of regressions estimating treatment effects on students' self-reported attitudes.. Mentor-random effects accounted for clustering of students within mentor-groups. Additional control variables included: classroom fixed-effects, mentor characteristics, student characteristics, baseline measures of student attitudes, mentoring relationship quality, virtual outreach, and student responsiveness. For observations where baseline values were missing, overall item mean imputed. Estimated impacts reported in effect sizes (outcomes standardized to the comparison group mean and SD). Comparison group mean and SD in original (unstandardized) units displayed. *** p<0.001, ** p<0.01, * p<0.05, ~ p<0.1

Appendix Table 6. Alternative models estimating treatment effects along intensive and extensive margins using logistic regression

Dependent variable:	Student Reported		Mentor Reported	
	Mentor outreach frequency	Student responsiveness	Student outreach frequency	Student responsiveness
Outcome = Any outreach, ever responded (vs. No outreach, never responded)	0.64 (0.48) [0.8, 0.40]	0.44 (0.57) [0.84, 0.36]	0.53 (0.18) [0.57, 0.49]	1.08 (0.40) [0.77, 0.41]
Outcome = Received 2+ touchpoints, always responded (vs. Received <2 touchpoints, did not always respond)	1.42 (0.68) [0.61, 0.49]	0.23** (0.12) [0.36, 0.48]	0.29* (0.14) [0.25, 0.43]	0.00 (0.00) [0.03, 0.18]

Notes: Table displays results of logistic regressions estimating treatment effects on virtual outreach frequency and student responsiveness to mentor outreach. Outreach measures indicate the frequency of texts, emails, and calls received during an average week. Student responsiveness gauges how students respond when their mentor reached out via text, email, and phone call. Binary outcomes were estimated along two margins: Any outreach/response (=1) vs. none (=0) (top row), and Frequent/Consistent outreach and response (=1) vs. less-frequent/less-consistent (=0) (bottom row). Estimated impacts reported as odds ratios. SE clustered within mentor groups are displayed in parentheses. Comparison group means and SD in original units are displayed in square brackets. Control variables included: classroom fixed-effects, mentor characteristics, student characteristics, baseline measures of student attitudes, mentoring relationship quality, virtual outreach, and student responsiveness. For observations where baseline values were missing, the overall item mean was imputed. *p*-values computed using randomization inference ("ritest" package in Stata). *** p<0.001, ** p<0.01, * p<0.05, ~ p<0.1

Appendix Table 7. Effects of mentor reminders as assessed at the experiment mid-point and conclusion

Panel A: Virtual communication frequency and student responsiveness

	Student-reported on mid-point survey		Student-reported on final survey	
Dependent variable:	Mentor outreach frequency	Student responsiveness	Mentor outreach frequency	Student responsiveness
Treatment effect	-0.12	-0.13	-0.04	0.00
	(0.11)	(0.10)	(0.25)	(0.21)
Comparison unstd. mean	2.64	2.20	3.84	2.73
Comparison unstd. SD	2.99	1.06	3.49	1.08
Observations	292	292	124	124
R-squared	0.39	0.44	0.39	0.29

Panel B: Mentoring relationship quality

	Student-reported on mid-point survey		Student-reported on final survey	
Dependent variable:	Relational quality	Instrumental quality	Relational quality	Instrumental quality
Treatment effect	-0.05	-0.14	-0.10	-0.01
	(0.09)	(0.08)	(0.18)	(0.16)
Comparison unstd. mean	3.01	2.93	3.22	3.13
Comparison unstd. SD	0.72	0.75	0.61	0.68
Observations	292	292	124	124
R-squared	0.63	0.60	0.44	0.50

Notes: Table displays results of OLS regressions estimating treatment effects on virtual outreach frequency, student responsiveness to mentor outreach, and mentoring relationship quality as assessed by student responses to the mid-point and final surveys. Robust SE clustered within mentor groups displayed in parentheses. Control variables included: classroom fixed-effects, mentor characteristics, student characteristics, baseline measures of student attitudes, mentoring relationship quality, virtual outreach, and student responsiveness. For observations where baseline values were missing, overall item mean imputed. Estimated impacts reported in effect sizes (outcomes standardized to the comparison group mean and SD). Comparison group mean and SD in original (unstandardized) units displayed. *p*-values computed using randomization inference ("ritest" package in Stata). *** p<0.001, ** p<0.01, * p<0.05, ~ p<0.1