

Brandt, Alicia; Oskorouchi, Hamid R.; Sousa-Poza, Alfonso

Article — Published Version

The effect of SMS nudges on higher education performance

Empirical Economics

Provided in Cooperation with:

Springer Nature

Suggested Citation: Brandt, Alicia; Oskorouchi, Hamid R.; Sousa-Poza, Alfonso (2023) : The effect of SMS nudges on higher education performance, Empirical Economics, ISSN 1435-8921, Springer, Berlin, Heidelberg, Vol. 66, Iss. 5, pp. 2311-2334, <https://doi.org/10.1007/s00181-023-02516-5>

This Version is available at:

<https://hdl.handle.net/10419/306402>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>



The effect of SMS nudges on higher education performance

Alicia Brandt¹ · Hamid R. Oskorouchi^{1,2}  · Alfonso Sousa-Poza^{1,2}

Received: 22 February 2023 / Accepted: 9 October 2023 / Published online: 23 November 2023
© The Author(s) 2023

Abstract

Driven by higher education's challenges in maintaining student motivation and achievement during the recent pandemic-induced shift to online learning, we investigate the effectiveness of text messages as a nudging tool to increase academic performance. To do so, we use a nonplacebo randomized controlled trial in which the treatment group directly receives SMS texts that review lecture content and give deadline reminders, while the control group only has access to the same information on the course page. Our findings suggest that the reception of motivating SMS messages per se, rather than the content review, has a positive effect on examination outcomes.

Keywords Nudging · Text messaging · Teaching

JEL Classification I21 · I23

1 Introduction

Disruption from the recent COVID-19 pandemic, whose lockdown measures affected 220 million students worldwide (UNESCO 2021, Marinoni et al. 2020), rapidly transformed on-campus teaching into “emergency online education,” thereby creating challenges for students and lecturers alike (Marinoni et al. 2020). In particular, the temporal and spatial distances inherent in asynchronous lectures made it difficult to maintain student learning engagement across the entire semester (Roblyer et al. 2007), leading universities to promote course digitalization, and teaching and learning methods that go beyond the lecture format (Bhagat and Kim 2020). This necessity resulted in a growing demand from educational institutions for different tools and digital devices to facilitate online teaching (UNESCO 2021) of students forced to learn off campus.

✉ Hamid R. Oskorouchi
hamid.oskorouchi@uni-hohenheim.de

¹ Institute for Health Care & Public Management, University of Hohenheim, Stuttgart, Germany

² IZA, Bonn, Germany

One means of addressing this need is to support lectures with short message service (SMS) texts, an inexpensive and easily implemented technology that can assist both content learning and student motivation. Given students' generally constant access to mobile devices (Ziden et al. 2017), text messaging can convey course-related information like relevant reading material while encouraging consistent review of lecture subject matter (Breunig et al. 2021; Ziden et al. 2017). In doing so, it promotes time- and location-independent engagement with learning content (Gasaymeh and Aldalalah 2013). In higher education especially, such messages can even influence behavior and serve as a motivator by potentially nudging students in their educational decisions (Escueta et al. 2020). For example, timely notifications can increase student attention to specific coursework (Castleman and Meyer 2020) among students struggling to allot sufficient class preparation time to numerous different courses (Breunig et al. 2021).

In recent years, a handful of researchers have evaluated the impact of text messages on student academic performance using randomized controlled trials (RCT) that employ either motivational, course-related, or performance-oriented intervention content. The findings, however, are mixed: Whereas a few experiments indicate SMS-induced performance gains (e.g., Castleman and Meyer 2020) stemming from either a motivational effect (Gasaymeh and Aldalalah 2013) or improved study behavior (O'Connell and Lang 2018), several large-scale studies find no evidence for any effect on academic attainment (e.g., Breunig et al. 2021; Dobronyi et al. 2019; Gill et al. 2016). The researchers variously explain these null effects by low-touch and infrequent interactions (Breunig et al. 2021; Oreopoulos et al. 2020) or the ease of ignoring messages (Lavecchia et al. 2016). Whatever the attribution, this lack of consensus on text messaging's actual impact underscores the importance of further research (Oreopoulos et al. 2020), a need as yet met by only a handful of experiments. Our work is thus a useful expansion of this limited research on the impact of course-related SMS messages.

To test SMS message effectiveness for improving student academic performance, we assess whether and to what extent a simple SMS nudge influences final grades and whether content plays a role in this dynamic. We do so via an RCT in which 346 master's course participants receive regular study reminders and information on lecture content in messages formulated solely to regulate study behavior without changing the students' state of information or incentives. Rather, this nudge is designed to influence student learning engagement so it translates into improved performance, a treatment effect we measure by final examination and online test scores. We also explore the underlying mechanisms and potential transmission channels for any observed increase in performance by administering an endline survey of such subjective measures as time management and motivation.

Not only does this study provide additional evidence on the impact and decisiveness of course-related SMS messages in effectively nudging higher education students toward improved performance, it is one of only three we know of that use RCTs to assess these nudges' effects on class outcomes in a university setting. The first, a large-scale evaluative experiment at a Canadian university (Oreopoulos and Petronijevic 2018), uses SMS communications (either information about campus offers or messages of encouragement) to compare the influence of in-person coaching with personalized mobile coaching. Whereas in-person coaching appears to positively impact

GPA and grades, SMS coaching seems less promising. The authors therefore conclude (as do Dobronyi et al. 2019) that motivational SMS messages are not an effective tool to nudge students toward better academic achievement. These findings are confirmed in a study closely resembling our own (Breunig et al. 2021), which reports that simple weekly reminders on recommended readings for upcoming lectures do not enhance learning engagement of undergraduate students. Our study contrasts with this earlier work, however, in two important ways: First, our focus is on postgraduate students, a demographic that could arguably be more responsive to text messages than undergraduates; and second, our intervention took place during the COVID-19 pandemic, when lack of face-to-face student-faculty interaction raised the importance of the messaging medium itself. At the same time, it also differs methodologically from a handful of other pandemic era studies using RCTs to gauge the efficacy of various interventions (e.g., Hardt et al. 2020; Carlana and La Ferrara 2021), none of which employed a simple, cost-effective SMS nudge within a postgraduate setting.

2 Experimental design

Experimental participants. The participants for this experimental study were students in an English-language graduate course at the University of Hohenheim, Stuttgart (Germany), which covers the basic theory and application of various statistical methods for multivariate data analysis (MVDA) (see Tables 1 and 2 for a topical overview and program-dependent course characteristics, respectively). Although this course is open to students from different programs and faculties, most are pursuing a Master of Science in either Management or Education for Business and Economics. Participant recruitment occurred during the first week of the 2020/2021 winter semester, when lecturers announced the option to receive course information and useful reminders via SMS. Volunteers registered online by providing consent and a mobile phone number, as well as optionally filling out a short survey on basic demographic and academic characteristics (see consent form in B1). We then used a reproducible computerized procedure to randomize them into a treatment and nonplacebo control group.

Table 1 MVDA topics

Work package (WP)	Topic	Time for preparation	SMS treatment
WP1	Introduction	1 week	No
WP2	Examining your data	2 weeks	Yes
WP3	Multiple regression	2 weeks	Yes
WP4	Logistic regression	1 week	Yes
WP5	Factor analysis	3 weeks	Yes
WP6	ANOVA	1 week	Yes
WP7	Machine learning	1 week	Yes

Author compilation based on the 2020–2021 MVDA syllabus

Table 2 Interprogram differences in the final examination

Study program	Module	Credits	Examination duration	Maximum achievable points		Number of topics covered
				Exam	Online tests	
Management	Compulsory	6 (180 h)	90 min	90	9	7
International Business and Economics	Elective					
Bioeconomy, Agribusiness						
Education for Business and Education	Compulsory	4 (120 h)	60 min	60	6	5

Author compilation based on the 2020–2021 MVDA syllabus

Experimental setting. As at many other institutions, the COVID pandemic forced the University of Hohenheim to switch to a fully online teaching modality for the entire winter semester 2020/2021. Under this modality, the course comprised 15 asynchronous lectures grouped into seven work packages, which culminated in a live digital wrap-up session in which students could ask the lecturer questions and receive clarifications. In addition to posting videos of the recorded lectures and related materials on a weekly basis, the university's ILIAS e-learning platform hosted an online forum for lecturer-student communication, as well as online tests for each different work package. These tests consisted of 12–20 multiple choice questions that resembled those on the final examination, for which it also provided simulations.

Although both the online tests and the final examination required 50% correct answers for a pass, the written examination determined the bulk of the final course grade, with successful completion of the online tests offering an extra 10% points (meaning a possible 110% maximum points, but only 100% of the points needed for the highest grade). This chance for extra credit was intended to encourage students to remain up-to-date throughout the semester while also compensating for the potentially negative effects of pandemic teaching modalities or the missed opportunity for additional encouragement in the control group.

We graph the dynamics between the 2019 winter semester MVDA course and the SMS intervention in Fig. 1. On the right-hand side, we outline the course's primary teaching modalities which, together with course design, supplementary materials, and lecturer experience and enthusiasm, inherently influence student motivation and time management. Importantly, all the above-mentioned factors ultimately affect the level of knowledge acquired during the course and thus academic performance. Because a large share of performance variation, however, is also attributable to individual demographics, registration for the experiment included voluntary collection of relevant student characteristics to check for pretreatment balance. In addition, as the flowchart

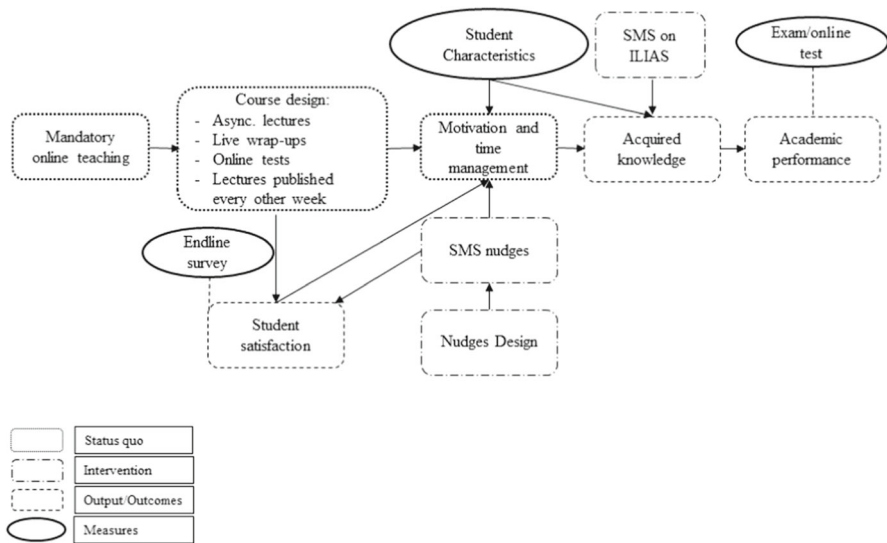


Fig. 1 SMS Intervention Flowchart

illustrates, we administered an endline survey of student satisfaction with the course modalities to assess the post-treatment balance.

Experimental procedure. Based on the documented efficiency of SMS nudge interventions in other contexts (Goh et al. 2012), we designed our SMS messages to boost student motivation and course involvement through direct engagement via an informal medium (the smartphone) outside of scheduled online lectures and forums. We thus coordinated the message content and topics with weekly course lectures, practical sessions, and online tests.¹ To ensure experimental fairness, these nudges provided no additional information (i.e., the SMS contained only the same information as on the lecture slides) and their content was uploaded weekly onto the ILIAS e-learning platform, thereby treating the control group with online exposure only. Because all students, regardless of treatment assignment, had to log into the e-learning platform frequently to (re)watch the weekly posted lectures, slides, online tests, and live wrap-up sessions, the control group had easy access to the SMS content.

Although this content could admittedly have been shared across various student social media groups, its disclosure was necessary to reduce any discomfort among control group members (i.e., reverse engagement bias) and to allow disentanglement of any motivational treatment effect from improved academic outcomes due simply to greater information availability (c.f. Heckman et al. 2000). We thereby not only

¹ Table C1 presents the full set of SMS; here, we report some randomly chosen examples: SMS 2: "Examining your data: Do not simply delete outliers! It could improve your multivariate analysis, but at risk of limiting generalizability. You can use e.g., Mahalanobis D2 to identify the outliers."; SMS 8: "In the next two weeks we recommend you work on "Regression Analysis". Prepare yourself for the next online test in which up to 2 extra points can be collected. The test will be online from 11.12. until 19.12., 23:59."; SMS 14: "To identify multicollinearity, you can use the tolerance measure (rather than a simple correlation matrix). A high tolerance value indicates a small degree of multicollinearity.".

isolated the SMS messages' motivational effect but mitigated any potential reverse engagement bias among the control group observations. As an additional guard against biases, we also kept the SMS content hidden from the lecturers who designed the final examination and graded the online tests.

In evaluating the research outcomes, we apply the minimum detectable effect (MDE) size reported in our original study protocol submitted to the Ethics Committee of the University of Hohenheim in May 2020, derived from an a priori power analysis using the standard deviation for a sample of 270 MVDA examination grades from the previous winter semester (17.3 points). In this latter, by setting a power ($1 - \beta$) of 0.80 and a statistical significance threshold (α) of 0.05, we obtained an MDE of 4 points (or 0.23 SD units).

To minimize the treatment effect standard error, we randomized the days and times that these participants received single text messages based on conventional study hours (Monday to Saturday, 8AM-7PM) (Ziden et al. 2017). We had previously determined the optimal treatment frequency by means of a pilot study implemented in the summer semester 2020,² which identified three texts per week (i.e., 34 SMS messages over the semester) as appropriate for increasing engagement while avoiding information fatigue. Likewise, to increase message effectiveness, we paid close attention to content framing and wording (Ajzenman and López Bóo 2019), avoiding repetitive sentences, using sufficient linguistic variety, and selecting terms that communicate a sense of urgency or reward (e.g., “last chance” or “successfully”), especially in deadline reminders.

Treatment evaluation. We evaluate the intervention treatment effect based on points earned on the final examination (with and without inclusion of the 10% extra credit for online tests). Because the first examination session occurred immediately after the lecture phase ends, whereas the second took place 3 months later at the start of the next semester, the resulting dataset allows comparative analyses between the short- and longer-term effects (see Figure B2 and Table C1 for the project timeline and SMS content, respectively).

3 Data and identification strategy

3.1 Data

Electronic survey data. Beginning in the second week of the winter semester 2020/2021, (November 2020), we collected baseline and follow-up information via ILIAS online survey from the 393³ students enrolled in the MVDA graduate class,

² The treatment was implemented between April 27 and July 24, 2020, in a German-language University of Hohenheim master's course on the economics of agriculture and food. Because of a smaller course enrolment, only 40 students registered for the project and were randomly assigned into treatment and control groups.

³ This number refers to students enrolled in the online course page; the actual number that sat the end-of-term examination was 362.

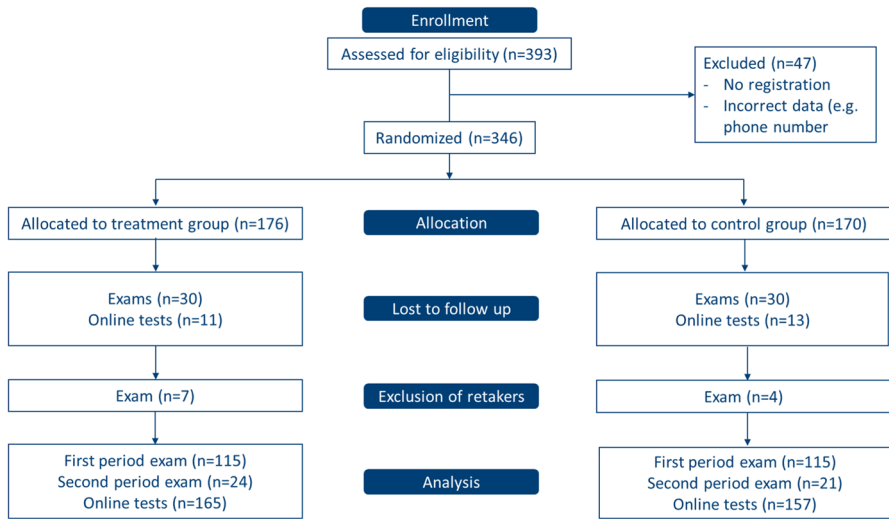


Fig. 2 Flow diagram of sampling process

346⁴ of whom signed up for the SMS message experiment and provided such demographic data as gender, age, and program of study. Randomization of this sample (see Fig. 2) yielded treatment and control groups of 176 and 170 participants, respectively.

In the last weeks of semester before the final examination, we encouraged students to complete a second short online survey on their perceptions of lecture modalities; more specifically, on their experience with the course environment and their own study behavior. Those who did numbered 152 observations (75 and 77 from treatment and control groups, respectively), a significantly smaller sample than the original. Nonetheless, given the strict anonymity of the official student evaluations, the second survey served as a post-treatment balance test. It also collected information on outcomes for which we expected no relation with the motivational SMS nudges, including dummy and Likert scale measures of satisfaction with asynchronous lectures, experience of technical issues, and willingness for teaching to be online in a post-pandemic scenario. Rather than any direct questions about SMS reception, the survey included an open-ended item inviting general comments on the lecture modalities. The only respondent to suggest improved SMS delivery was a control group member who had received no messages.

Administrative data. As previously outlined, we proxy student performance by final examination and online test grades obtained from the official university examination registry. Because students from different masters' courses must fulfill slightly different requirements to pass the final, we rescaled our two main outcome variables (final examination and online test points) on a 0–100 range for comparability. Additionally, we created two longitudinal datasets on the grades of single examination and online test questions, respectively. Depending on the level of difficulty, the questions were

⁴ The original sample numbered 350 students, but we excluded four observations for reporting nonexistent telephone numbers.

worth between 1 and 4 points, however, for comparability we rescaled them to a semi-continuous measure (from 0 to 1). However, in all regression analyses on examination and online test points, we present treatment effects in SD units. Lastly, we mark each question as either treated or not, based on whether the associated SMS message contains useful information for answering it correctly.

Descriptive statistics. For the final sample of 275 participants whose results are analytically relevant, the average age is 24, approximately two thirds are female, and most (83%) are first semester enrollees in the master's in management program (see Table 3, Panel A). As column 5 shows, the baseline characteristics of the treatment and control groups are balanced (at $\alpha = 0.05$) both individually and jointly (F test p value = 0.465). Of all those enrolled in the course, 43% took part in the endline survey (Panel B). Finally, in these longitudinal data, the average (rescaled) score per question in the examination and online tests was 0.79 and 0.67 points, respectively, with 16% and 14% of the questions classified as treated (Panels C and D).

3.2 Identification strategy

In the non-placebo randomized control trial used to estimate the motivational SMS message impact on student academic outcomes, although the randomly selected treatment group cannot contain never-takers,⁵ the treatment coefficient must be interpreted as intention to treat (ITT) effects. In fact, some students might have decided to non-randomly opt out of treatment by simply ignoring the SMS nudges. Likewise, the overall sample cannot contain always-takers because no student in the control group ever receives a message. In fact, because take-up of interventions like ours cannot be forced on the students, an ITT effect is the effect of interest. Nonetheless, as long as both groups' semester-long access to SMS content on the ILIAS platform ensured their equal satisfaction with course modalities (i.e., the assumption of no reverse engagement bias holds), the treatment impacts are likely to isolate a purely motivational effect, as subsequently tested using post-treatment measures of course satisfaction.

For the remainder of this discussion, we formalize the different estimation strategies based on the type of data analyzed (i.e., whether cross-sectional or longitudinal). We begin with Eq. 1, which formalizes the OLS estimation strategy for the set of regressions that use cross-sectional data to measure the motivational message effect on the final examination including (10% extra) online test points. Here, Y_i denotes either the examination points, T_i^{group} is a dichotomous indicator taking value one if an observation belongs to the treatment group, and zero otherwise, X_i is a vector of demographic and academic control variables at the individual level, and ε_i is an idiosyncratic error term.

When the dependent variable is the probability of passing the final examination (with or without 10% extra points), we employ the same specification as in Eq. 1 but estimate it using logistic regression. To maximize statistical power, we again pool the question-level data, which considerably increases the sample size and enables robustness testing of our estimates using individual-level cross-sectional data. Statistical

⁵ No student requested to be excluded from the treatment after having enrolled into the experiment.

Table 3 Descriptive statistics: full sample

Panel A					
Baseline survey and outcomes (first examination session)					
	(1) Mean	(2) SD	(3) Min	(4) Max	(5) T-C Difference (<i>p</i> value)
Examination points	66.511	15.595	6.667	97.778	–
Passing the examination (with extra points)	0.891	0.312	0	1	–
Extra points	9.263	1.411	1.111	10	–
Treatment	0.505	0.501	0	1	–
Female (ref. male)	0.676	0.469	0	1	0.912
Age (years)	24.051	1.830	20	32	0.617
Management (ref. any other)	0.825	0.380	0	1	0.505
Business Education (ref. any other)	0.153	0.360	0	1	0.745
Semester	1.596	1.101	1	7	0.083
Compulsory module (ref. eligible)	0.989	0.104	0	1	0.306
6 ECTS (ref. 4 ECTS)	0.815	0.389	0	1	0.621
First sitting (ref. second/third sitting)	0.945	0.228	0	1	0.396
F test <i>p</i> -value					0.465

Table 3 (continued)

Panel A					
Baseline survey and outcomes (first examination session)					
	(1)	(2)	(3)	(4)	(5)
	Mean	SD	Min	Max	T-C Difference (<i>p</i> value)
Observations	275				
Panel B					
Follow-up survey (full sample)					
	(1)	(2)	(3)	(4)	(5)
	Mean	SD	Min	Max	T-C Difference (<i>p</i> -value)
Survey participation (<i>N</i> = 430)	0.43	0.50	0	1	0.746
Satisfaction with online teaching	0.83	0.37	0	1	0.104
Technical issues	0.52	0.50	0	1	0.472
Continuation of online teaching	0.88	0.32	0	1	0.694
Observations	187				
Panel C					
Examination question data					
	Mean	SD	Min	Max	T-C Difference (<i>p</i> value)
Points per question	0.789	0.401	0	1	–

Table 3 (continued)

Panel C					
Examination question data					
	Mean	SD	Min	Max	T-C Difference (<i>p</i> value)
Treated question	0.157	0.363	0	1	–
Observations	10,580				
Panel D					
Online test data					
	Mean	SD	Min	Max	T-C Difference (<i>p</i> value)
Points per question	0.669	0.423	0	1	–
Treated question	0.138	0.345	0	1	–
Observations	23,598				

Column 5 reports the *p* value of the pretreatment differences between treatment and control groups when the treatment dummy variable is regressed on all pretreatment characteristics. **p* < 0.10, ***p* < 0.05, ****p* < 0.01

inference for estimates from all models account for heteroskedasticity by correcting the standard errors with a sandwich estimator or correcting for individual-level clustering when using pooled cross-sectional data.

$$Y_i = \alpha_0 + \alpha_1 T_i^{\text{group}} + \sum_{j=1}^8 \sum_{i=1}^N \alpha_j X_i + \varepsilon_i \quad (1)$$

When we employ longitudinal data at question level, we estimate a two-way fixed effect OLS regression in which the treatment assignment is at the question level. Because our experiment is not designed as a cluster RCT, controlling for question fixed effects is essential to avoid any bias in the treatment estimates related to question characteristics. In fact, as earlier noted, we intentionally keep instructors blind to our question design, which itself maximizes treatment effectiveness by targeting relevant subject topics. This estimation strategy, formally outlined in Eq. 2, corresponds to a two-group generalized difference-in-differences (DID) approach in which the coefficient of $T_{i,q}^{\text{question}}$ can be interpreted as an ITT. The standard errors' estimated accounts for heteroskedasticity by clustering the standard errors at the individual level.

$$Y_{i,q} = \alpha_0 + \alpha_1 T_{i,q}^{\text{question}} + \varphi_i + \zeta_q + \varepsilon_{i,q} \quad (2)$$

The strategies outlined above also enable robustness testing of our RCT design as follows: First, if the treatment is properly randomized, the inclusion of control variables should not significantly change the point estimates, a constancy we can test by running bivariate regressions on Eqs. 1 and 2. Second, according to the a priori statistical power analysis, our cross-sectional data can correctly detect a statistically significant effect at the 5% significance level 80% of the time when the MDE is greater than 0.23 SD units (4 points). Conversely, the longitudinal samples of examination and online test questions themselves have an even greater ability to correctly detect a true statistically significant ITT, with a sample size of 11,362 examination questions-students yielding an MDE of 0.018 points (0.045 SD units; $1 - \beta = 0.80$; $\alpha = 0.05$; $\text{SD} = 0.40$). However, it should be noted that the latter MDE calculation was performed ex-post because we did not have access to prior information on these longitudinal samples.

3.3 Potential bias

Although the treatment and control groups are statistically balanced on demographic and academic characteristics (see Table 3), which implies that the conditional independence assumption holds, biased treatment effect estimates could result from several mechanisms, whose threats to causality, limitations, and solutions we discuss below:

Selective attrition. Because the German higher education system allows students to sit the final examination⁶ either a few days after the last lecture or 2–3 months later, students might self-select into a specific session. In the case of MVDA, most students choose the first sitting available, with those selecting the later session showing

⁶ In the German system, students cannot take the same examination more than twice per semester.

lower average academic performance, even after we exclude students who previously failed. Obviously, if such self-selection is based on background characteristics and the pretreatment balance does not hold, then the first session sample estimates cannot be considered unbiased. However, according to the descriptive statistics and difference in means for the first session subsample (Table C2) and the estimated treatment effect on the probability of sitting in the first session examination (Table C3), such is not the case. It should also be noted that 60 of the students enrolled in the experiment did not sit the examination in either session, but these observations are both balanced in their pretreatment characteristics and equally distributed between treatment and control groups.

Reverse engagement bias. If exclusion from receiving the SMS nudges has any effect on the control group's academic performance (reverse engagement bias), these observations would not represent a valid counterfactual (because possibly negatively affected by the intervention). In fact, were this the case, the estimated ITT would not resemble the effect of the preprotocol plan. However, determining the bias direction for this scenario is far from simple given that exclusion could have either decreased their motivation (overstated ITT) or pushed them to study harder to compensate for the lack of treatment (understated ITT). Our experimental design thus tries to eliminate both possibilities by providing the SMS content to the whole sample via the e-learning platform. The appropriateness of this strategy is supported by the fact that in the (anonymous) lecturer evaluations, only one control group student reported not receiving any SMS messages.

To rule out the possibility of reverse engagement bias, we conduct a series of analyses that alternatively regress three different post-treatment measures of course satisfaction and a dummy for endline survey participation on the treatment group dummy variable and the full set of control variables. As Table 4 shows, the treatment variable coefficient never differs statistically from zero, confirming the improbability of any reverse engagement bias. Moreover, a supplemental test on whether the control observations' awareness of their status makes them work harder to compensate the lack of treatment exposure—offers no evidence of compensation for not receiving the SMS directly. For this test, because the online forum and live wrap-ups provided the sole opportunity to ask lecturers for clarification, we assess the treatment effect on the probability of posting in the forum at least once (see Table C4).

4 Results and discussion

Our econometric approach combines OLS and logit assessments of the treatment effect on academic outcomes and a series of generalized difference-in-differences (DID) models that measure the efficacy of matching SMS message texts to course content:

OLS regressions. To identify the causal effects of motivational SMS messages on the number of final examination points earned, we employ a series of OLS regressions that exploit the random experimental variation in the treatment variable. We report the outcomes first for the first session sample only (February 2021) and then for a pooled sample of the first and second sessions (July 2021) with retakers excluded (see

Table 4 Effect of SMS on the course satisfaction measures: OLS, full sample

	Dependent variable			
	Survey participation (1)	Satisfaction with online teaching (2)	Technical issues (3)	Continuation of online teaching (4)
Treatment	– 0.019 (0.06)	0.102 (0.06)	– 0.058 (0.08)	0.021 (0.05)
Female (ref. male)	0.120* (0.06)	– 0.056 (0.07)	– 0.042 (0.09)	– 0.049 (0.06)
Age	– 0.048*** (0.02)	– 0.013 (0.03)	0.005 (0.03)	– 0.004 (0.03)
Business Education	0.265 (0.19)	0.550** (0.24)	– 0.239*** (0.09)	– 0.096 (0.09)
Management	0.151 (0.19)	0.285 (0.29)	– 0.775*** (0.11)	0.136 (0.21)
Current semester	– 0.004 (0.03)	0.054* (0.03)	– 0.081** (0.04)	0.044* (0.02)
Compulsory module	– 0.201 (0.23)	– 0.145 (0.29)	0.088 (0.12)	– 0.283 (0.22)
90 min examination	0.076 (0.14)	0.293 (0.18)	0.331*** (0.10)	– 0.191 (0.20)
First attempt	0.105 (0.11)	0.084 (0.16)	0.283** (0.13)	– 0.101*** (0.03)
N	297	152	152	151
F <i>p</i> value	0.000	0.037	0.000	0.070
R-sqr	0.070	0.077	0.110	0.051

The results are based on data from the endline MVDA course satisfaction survey. Standard errors robust to heteroskedasticity are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Tables 5 and 6, respectively). The first column of each table lists the OLS estimates for a basic model with no control variables, after which we stepwise include demographic (column 2), academic (column 3), and examination-related (column 4) control variables.

Because the inclusion of additional controls is likely to affect the ITT effect size by a significant amount when the treatment (cor) relates with any observables and/or with its linked unobservables (i.e., in the presence of severe bias), this progressive inclusion tests the stability of our ITT estimates. At the same time, given the ability demonstrated in our a priori power analyses to correctly detect a statistically significant effect 80% of the time only when the true ITT is larger than 0.23 SD units (4 points),

Table 5 Effect of SMS on examination points: OLS, first session sample

	(1)	(2)	(3)	(4)
Treatment	0.203* (0.12)	0.213* (0.11)	0.255** (0.11)	0.255** (0.11)
Female (ref. male)		– 0.240** (0.12)	– 0.213* (0.12)	– 0.213* (0.12)
Age (years)		– 0.132*** (0.03)	– 0.107*** (0.03)	– 0.107*** (0.03)
Management (ref. any other)			1.138*** (0.42)	1.117*** (0.42)
Business Education (ref. any other)			0.776* (0.43)	0.777* (0.44)
Semester			– 0.038 (0.05)	– 0.039 (0.07)
Compulsory module (ref. eligible)			– 1.317*** (0.24)	– 1.297*** (0.29)
6 ECTS (ref. 4 ECTS)				0.027 (0.24)
First sitting (ref. second/third sitting)				– 0.029 (0.28)
N	230	230	230	230
F p-value	0.080	0.000	0.000	0.000
R-sqr	0.013	0.091	0.146	0.146

Dependent variable: examination points (in SD units) excluding the extra points granted during the semester. The results are obtained using the first session sample only. Standard errors robust to heteroskedasticity are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

the inclusion of a set of meaningful control variables improves the precision of the ITT point estimates, allowing us to draw more statistically valid conclusions.

According to the first session estimates (Table 5), receiving the motivational SMS message improves academic performance by between 0.20 and 0.25 SD units (3.38 and 4.34 points), depending on specification. Although all four ITT coefficients are statistically significant at the 0.05 level, we can confidently state that only when the ITT estimate exceeds the MDE of 0.23 SD units (4 points; columns 3 & 4) are the treatment effects positive and statistically different from zero. The effects for the full sample of experimental participants who sat the final examination, however, are significantly smaller (approximately half the MDE) and statistically nonsignificant (see Table 6), which we interpret to mean a nontrivial amount of treatment heterogeneity between the first and second session examinees. In fact, even if we use a dummy indicator to limit the potential bias from self-selection into the second session, the treatment effect is likely to vanish in the longer term (about 3 months).

Table 6 Effect of SMS on examination points: OLS, full sample

	(1)	(2)	(3)	(4)
Treatment	0.128 (0.11)	0.138 (0.11)	0.171 (0.11)	0.172 (0.11)
Female (ref. male)		– 0.235* (0.12)	– 0.231* (0.12)	– 0.229* (0.12)
Age (years)		– 0.121*** (0.03)	– 0.097*** (0.03)	– 0.097*** (0.03)
Management (ref. any other)			1.032** (0.42)	0.933** (0.42)
Business Education (ref. any other)			0.816* (0.42)	0.822* (0.43)
Semester			– 0.049 (0.05)	– 0.046 (0.07)
Compulsory module (ref. eligible)			– 1.251*** (0.30)	– 1.147*** (0.33)
6 ECTS (ref. 4 ECTS)				0.131 (0.25)
First session (ref. second/third session)				– 0.122 (0.23)
N	275	275	275	275
F p-value	0.261	0.000	0.000	0.000
R-sqr	0.005	0.065	0.094	0.096

Dependent variable: examination points (in SD units) excluding the extra points granted during the semester. The results are from the first and second session samples jointly. Standard errors robust to heteroskedasticity are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Given the treatment aim of boosting student motivation to keep up with lectures across the semester, this diminishment would assumedly apply most to students already planning to focus first on other subjects and sit the MVDA examination in the second session. Even if the much smaller second session sample of 45 observations precludes direct statistical testing of this hypothesis, the treatment appears to have no statistically significant effect on the probability of sitting the first versus the second examination session (see Table C3).

Logit regressions. We next ensure the quality of the above ITT estimates⁷ by using a logit maximum likelihood estimator to regress a dichotomous variable for passing the final examination on the treatment indicator. We again estimate this model for both the first session (Table 7) and the pooled cross-sectional (Table 8) samples while accounting or ignoring (top vs. bottom panels) the extra points from online tests (0–10).

⁷ The ITT interpretation of logit estimates is valid only under the assumption of treatment homogeneity. It should be noted that using OLS instead of logit produces coefficients almost identical to the logit marginal effects (at means) – estimates not reported and available upon request.

Table 7 Effect of SMS on passing the examination: logit, first session sample

	Log-odds of passing the examination (without extra points)			
	(1)	(2)	(3)	(4)
Treatment	0.941*	0.925*	1.121*	1.191**
	(0.55)	(0.56)	(0.60)	(0.60)
Demographic controls	No	Yes	Yes	Yes
Academic controls	No	No	Yes	Yes
Examination history controls	No	No	No	Yes
Extra Points	Yes	Yes	Yes	Yes
Marginal effects (at means)				
Treatment	0.060*	0.056*	0.063**	0.065**
	(0.03)	(0.03)	(0.03)	(0.03)
N	230	230	228	228
	Log-odds of passing the examination (with extra points)			
	(1)	(2)	(3)	(4)
Treatment	0.070*	0.070*	0.077**	0.078**
	(0.04)	(0.04)	(0.04)	(0.04)
Demographic controls	No	Yes	Yes	Yes
Academic controls	No	No	Yes	Yes
Examination history controls	No	No	No	Yes
Extra Points	No	No	No	No
Marginal effects (at means)				
Treatment	0.068*	0.067*	0.074**	0.076**
	(0.03)	(0.04)	(0.03)	(0.03)
N	230	230	228	228

Dependent variable: the log-odds of passing the examination with extra points included. Marginal effects (at means) represent percentage points changes in probability. The results, obtained using the first session sample only, are merely exploratory because the regressions were not declared in the study protocol. Standard errors robust to heteroskedasticity are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

As no analysis using graded online tests and related extra points as outcome variables was included in our initial experimental protocol, we view the associated results as exploratory.

According to these new estimations, the treatment has a statistically significant positive effect only on the first session sample. More specifically, the marginal effects at mean values reported in each panel suggest that being treated increases the probability of passing the examination by about 6.5 percentage points (pp) (Table 7, column 4) when extra points are included and 7.6 pp (Table 7, column 8) when not.

Table 8 Effect of SMS on passing the examination: logit, full sample

	Log-odds of passing the examination (without extra points)			
	(1)	(2)	(3)	(4)
Treatment	0.325 (0.39)	0.318 (0.40)	0.384 (0.41)	0.447 (0.40)
Demographic controls	No	Yes	Yes	Yes
Academic controls	No	No	Yes	Yes
Examination history controls	No	No	No	Yes
Extra Points	Yes	Yes	Yes	Yes
Marginal effects (at means)				
Treatment	0.031 (0.04)	0.030 (0.04)	0.036 (0.04)	0.041 (0.04)
N	275	275	272	272
	Log-odds of passing the exam (with extra points)			
	(5)	(6)	(7)	(8)
Treatment	0.039 (0.04)	0.041 (0.04)	0.042 (0.04)	0.045 (0.04)
Demographic controls	No	Yes	Yes	Yes
Academic controls	No	No	Yes	Yes
Examination history controls	No	No	No	Yes
Extra Points	No	No	No	No
Marginal effects (at means)				
Treatment	0.039 (0.04)	0.041 (0.04)	0.042 (0.04)	0.045 (0.04)
N	275	275	275	275

Dependent variable: the log-odds of passing the examination with extra points included. Marginal effects (at means) represent percentage points changes in probability. The results, obtained using the first and second session samples jointly, are merely exploratory because the regressions were not declared in the study protocol. Standard errors robust to heteroskedasticity reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Not only do these results demonstrate that the positive ITT effect is not solely dependent on the extra tests points, but the slightly larger treatment effect when they are omitted conforms to our expectation that the treatment, by stimulating more regular study, tends to benefit formal examinations more than open-book online tests. Conversely, we do not detect any positive treatment effect when using the pooled cross-sectional sample (see Table 8). We conclude that the SMS treatment has had an

economically meaningful effect on the probability of passing the examination for the first examination session students, but due to limited statistical power we are unable to distinguish if the positive ITT estimate for the full sample is statistically relevant.

Pooled OLS. To address any concerns regarding statistical power, especially when using the two-session sample, we replicate the analysis reported in Tables 5 and 6 employing data at the student-question level. In addition to significantly increasing the sample size (11,362 and 8,902 observations), this replication exploits the random variation in the treatment variable. In all models, the dependent variable takes a continuous value between zero and one, and all specifications control not only for demographic, academic, and examination characteristics, but also for question fixed effects.

The results of these pooled OLS regressions reveal that, on average, treatment with the motivational messages results in a 0.05 (full sample) or 0.07 (first session sample) SD units increase in points earned on each question (Table 9, columns 1 & 3), with both estimates highly statistically significant once corrected for heteroskedasticity. Although these results confirm some degree of effect heterogeneity between students in both sessions, they also underscore the likelihood that the statistical nonsignificance of the full sample cross-sectional results is due to insufficient statistical power to detect an effect smaller than the MDE.

Exploiting this same empirical approach, we also conduct an exploratory analysis into the ITT effect of SMS message reception on the online test points per question (in SD units) for course content that is more technical. That is, rather than differentiating

Table 9 Effect of SMS on examination questions: OLS

	Full sample		First session sample	
	(1)	(2)	(3)	(4)
Treatment	0.050*** (0.02)		0.068*** (0.02)	
SMS on question topic		− 0.048 (0.04)		− 0.047 (0.04)
Demographic controls	Yes	No	Yes	No
Academic controls	Yes	No	Yes	No
Individual FE	No	Yes	No	Yes
Question FE	Yes	Yes	Yes	Yes
N	11,362	11,362	8,902	8,902
F p value	0.000	0.000	0.000	0.000
R-sqr. (within)		0.192		0.184
R-sqr	0.186		0.176	

Dependent variable: points per examination question (SD units). The results are merely exploratory because the regressions were not declared in the study protocol. Standard errors robust to heteroskedasticity (columns 1 & 3), and clustered at the individual level (columns 2 & 4) are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

between examination sessions, we estimate the ITT for the full set of online tests (Table 9, column 1) versus the subset of tests related to OLS, logit, ANOVA, and factor analysis (column 3). As in the previous results for examination performance, the treatment increases the points earned by 0.03 and 0.02 SD units, but with only the first estimate being statistically significant. Compared to the effects on examination question points, these smaller ITTs are in line with our previously stated assumption that the treatment is only marginally effective in the context of an open-book test.

Difference in differences. Our final analysis exploits the longitudinal nature of the examination and online test question dataset by estimating a series of difference in differences (DID) that test the point gain when any of the SMS messages is closely related to a specific question. This treatment dummy variable varies at the question level for those in the treatment group but remains always equal to zero for those in the control group. In addition to the full set of controls, these models also include two sets of fixed effects that correct for individual- and question-level heterogeneity. Interestingly, when the message and examination question are on the same topic, we observe a null effect (Table 9, columns 3 & 4), but when considering online test question and SMS messages on the same topic, the students gain an average of 0.045 or 0.065 SD units (Table 10, columns 2 & 4, respectively). These results, together with those when the treatment is defined by group (e.g., Tables 5 and 6), confirm that the SMS messages have a beneficial effect on longer-term academic performance, but through motivation rather than the provision of additional information. We can draw this conclusion not only because the SMS message content matched that presented in

Table 10 Effect of SMS on online test questions: OLS

	All online tests		Online tests #2–5	
	(1)	(2)	(3)	(4)
Treatment	0.030*** (0.01)		0.021 (0.01)	
SMS on question topic		0.045* (0.03)		0.065* (0.03)
Demographic controls	Yes	No	Yes	No
Academic controls	Yes	No	Yes	No
Individual FE	No	Yes	No	Yes
Question FE	Yes	Yes	Yes	Yes
N	23,598	23,598	18,184	18,184
F p value	0.000	0.000	0.000	0.000
R-sqr. (within)		0.254		0.255
R-sqr	0.246		0.244	

Dependent variable: points per online test question (SD units). The results are merely exploratory because the regressions were not declared in the study protocol. Standard errors robust to heteroskedasticity (columns 1 & 3), and clustered at the individual level (columns 2 & 4) are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

lectures (and reported in the course slides), but our experimental design provided the same information to every student via the e-learning platform.

Because the 4-point treatment effect of the motivational nudge corresponds to only a small behavioral change but at a marginal cost (Smith et al. 2018), we lastly examine the results from the perspective of cost-effectiveness. Given a total expenditure for the entire SMS project of approximately €328—which includes initial costs for the mobile phone, SMS dispatch, and personnel time designing, organizing, and implementing the intervention—the amount per student in a treatment group of 176 is €1.86, relatively inexpensive compared to other behavioral interventions in higher education. Statistically, based on a treatment effect of 0.25 *SD*, this sum is equal to about €7.44 per one *SD* increase in performance. Clearly, programs such as tutoring (Carlana and La Ferrara 2021), mentoring (Hardt, Nagler, and Rincke 2020), or SMS coaching (Oreopoulos et al. 2020) entail higher expenses due to staff training costs. Hence, given the comparable or somewhat more pronounced effect on academic performance compared to the effort in other studies, our intervention seemingly exhibits high cost-effectiveness, an assumption supported by Kraft's (2020) cost-effectiveness analysis in the context of educational interventions.

5 Conclusions

By using an RCT experimental framework to investigate the causal effect of motivational SMS nudges on student academic performance, we confirm that one-way communication between instructors and students via SMS outside scheduled lecture hours is an effective tool for improving the performance on both online tests and the final examination. This positive treatment effect is in line with that documented by earlier studies in different contexts (e.g., Angrist et al. 2022; So 2016). The experiment's nonplacebo design also provides evidence for the mechanism underlying the treatment effect; namely, greater student engagement and motivation, as opposed to the availability of additional information (as in Paloyo et al. 2016). We also admit, however, that based on the 30% and 40% reduction in estimated ITT among students that sat the examination around 3 months after the last course lecture, the treatment effect is likely to fade away with time.

We acknowledge several limitations in interpreting the study results that could potentially affect its broader applicability (external validity). Most notably, at least three different explanations are possible for our intervention yielding positive treatment outcomes not observable in other studies, one of them general; the others more contextually specific. First, and most obvious, the postgraduate participants in our study may have responded differently to the text messaging than their undergraduate counterparts in other studies. Second, the intervention occurred during the severest COVID-19 lockdowns when the mandate against face-to-face contact likely raised the importance that students attached to direct SMS communication from the course coordinator. Third, because student-professor interactions in German universities are more formal than those in many other countries (e.g., the US), the cultural context could have significantly affected the impact of messaging directly from a faculty member. Finally, it should also be noted that our study's external validity is negatively affected

by the fact that about 12% of the students enrolled in our course did not agree to participate in the experiment. Nevertheless, our research findings remain valuable to higher education lecturers and policy makers because of text messaging's potential as an extremely low-cost supplementary learning tool that is easily implemented and reaches all class members regardless of personal characteristics. Of particular note, our isolation of the texts' motivational effect by exclusion of any information not covered in the lecture implies the possible enhancement of message effectiveness by inclusion of extra information (e.g., background on the topic). Similarly, the relatively primitive nature of the simple SMS text may imply that substituting SMS with an over-the-top Internet service able to deliver audio, video, and other media images might significantly improve the messaging experience and thereby its effectiveness (So 2016). Even personalization of the simple text message might improve motivational effectiveness, although doing so would also be more costly.

Whereas all such conjectures offer valuable opportunities for future research, our own findings confirm that supplementing traditional lectures with even a seemingly primitive communicational medium built on simple, motivational, course-focused messages can efficiently and cost effectively improve student motivation and engagement. Not only is this potential especially important when online teaching is unavoidable, but it holds great pedagogical promise for on-campus teaching.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00181-023-02516-5>.

Acknowledgements We are extremely grateful to our colleagues from the Institute of Health Care and Public Management of the Hohenheim University for their precious support in implementing this experiment. Finally, we thank the editor and two anonymous reviewers for their useful comments and suggestions on an earlier version of this manuscript.

Funding Open Access funding enabled and organized by Projekt DEAL

Declarations

Conflict of interest None.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ajzenman N, Florencia LB (2019) Lessons from behavioral economics to improve treatment adherence in parenting programs: an application to SMS. IZA Discussion Paper No. 12808
- Angrist N, Bergman P, Matsheng M (2022) Experimental evidence on learning using low-tech when school is out. *Nat Hum Behav* 6(7):941–950. <https://doi.org/10.1038/s41562-022-01381-z>

- Bhagat S, Kim DJ (2020) Higher education amidst COVID-19: challenges and silver lining. *Inf Syst Manag* 37(4):366–371. <https://doi.org/10.1080/10580530.2020.1824040>
- Bhatti Y, Dahlgaard JO, Hansen JH, Hansen KM (2017) Moving the campaign from the front door to the front pocket: field experimental evidence on the effect of phrasing and timing of text messages on voter turnout. *J Election Public Opin Parties* 27(3):291–310. <https://doi.org/10.1080/17457289.2016.1270288>
- Breunig C, Klüser KJ, Yang Q (2021) Can students be encouraged to read? Experimental evidence from a large lecture. *Eur Polit Sci*. <https://doi.org/10.1057/s41304-021-00351-9>
- Carlana M, Eliana LF (2021) Apart but connected: online tutoring and student outcomes during the COVID-19 pandemic. CEPR Discussion Paper No. DP15761
- Castleman BL, Meyer KE (2020) Can text message nudges improve academic outcomes in college? Evidence from a West Virginia Initiative. *Rev High Educ* 43(4):1125–1165. <https://doi.org/10.1353/rhe.2020.0015>
- Chow CK, Redfern J, Hillis GS, Thakkar J, Santo K, Hackett ML, Jan S, Graves N, de Keizer L, Barry T, Bompont S, Stepien S, Whittaker R, Rodgers A, Thiagalingam A (2015) Effect of lifestyle-focused text messaging on risk factor modification in patients with coronary heart disease: a randomized clinical trial. *JAMA* 314(12):1255–1263. <https://doi.org/10.1001/jama.2015.10945>
- Dale A, Strauss A (2009) Don't forget to vote: text message reminders as a mobilization tool. *Am J Polit Sci* 53(4):787–804. <https://doi.org/10.1111/j.1540-5907.2009.00401.x>
- Damgaard MT, Nielsen HS (2018) Nudging in education. *Econ Educ Rev* 64:313–342. <https://doi.org/10.1016/j.econedurev.2018.03.008>
- Dobronyi CR, Oreopoulos P, Petronijevic U (2019) Goal setting, academic reminders, and college success: a large-scale field experiment. *J Res Educ Effect* 12(1):38–66. <https://doi.org/10.1080/19345747.2018.1517849>
- Downer SR, Meara JG, Da Costa AC (2005) Use of SMS text messaging to improve outpatient attendance. *Med J Aust* 183(7):366–368. <https://doi.org/10.5694/j.1326-5377.2005.tb07085.x>
- Escueta M, Nickow AJ, Oreopoulos P, Quan V (2020) Upgrading education with technology: insights from experimental research. *Journal of Economic Literature* 58(4):897–996. <https://doi.org/10.1257/jel.20191507>
- Hardt D, Markus N, Johannes R (2020) Can peer mentoring improve online teaching effectiveness? An RCT during the Covid-19 pandemic. CESifo Working Paper No. 8671
- Gasaymeh A-MM, Aldalalah OM (2013) The impact of using SMS as learning support tool on students' learning. *Int Edu Stud (IES)* 6(10):112–123. <https://doi.org/10.5539/ies.v6n10p112>
- Gill CJ, Ngoc BL, Halim N, Viet HN, Williams AL, van Tan N, McNabb M, Ngoc LTT, Falconer A, Ha HAP, Rohr J, Hoang H, Michiel J, Thanh TNT, Bird L, Vu HP, Yeshitla M, van Nhu H, Sabin L (2016) The mCME Project: a randomized controlled trial of an SMS-based continuing medical education intervention for improving medical knowledge among Vietnamese community based physicians' assistants. *PLoS ONE* 11(11):1–19. <https://doi.org/10.1371/journal.pone.0166293>
- Goh T-T, Seet B-C, Chen N-S (2012) The impact of persuasive SMS on students' self-regulated learning. *Br J Edu Technol* 43(4):624–640. <https://doi.org/10.1111/j.1467-8535.2011.01236.x>
- Gurol-Urganci I, de Jongh T, Vodopivec-Jamsek V, Atun R, Car J (2013) Mobile phone messaging reminders for attendance at healthcare appointments. *Cochrane Database Syst Rev*. <https://doi.org/10.1002/14651858.CD007458.pub3>
- Heckman J, Hohmann N, Smith J, Khoo M (2000) Substitution and dropout bias in social experiments: a study of an influential social experiment. *Q J Econ* 115(2):651–694. <https://doi.org/10.1162/003355300554764>
- Jones G, Edwards G, Reid A (2009) How can mobile SMS communication support and enhance a first year undergraduate learning environment? *ALT-J* 17(3):201–218. <https://doi.org/10.1080/09687760903247625>
- Kraft MA (2020) Interpreting effect sizes of education interventions. *Educ Res* 49(4):241–253. <https://doi.org/10.3102/0013189x20912798>
- Kuznekoff JH, Munz S, Titsworth S (2015) Mobile phones in the classroom: examining the effects of texting, twitter, and message content on student learning. *Commun Educ* 64(3):344–365. <https://doi.org/10.1080/03634523.2015.1038727>
- Lavecchia AM, Liu H, Oreopoulos P (2016) Behavioral economics of education: progress and possibilities. *Handbook of the Economics of Education*, Elsevier, Amsterdam

- Marinoni G, Van't Land H, Jensen T (2020) The impact of COVID-19 on global higher education. *Int Higher Edu* 102:7–9
- Montes JM, Medina E, Gomez-Beneyto M, Maurino J (2012) A short message service (SMS)-based strategy for enhancing adherence to antipsychotic medication in schizophrenia. *Psychiatry Res* 200:89–95. <https://doi.org/10.1016/j.psychres.2012.07.034>
- Nkhoma CA, Thomas S, Nkhoma MZ, Sriratanaviriyakul N, Truong TH, Vo HX (2018) Measuring the impact of out-of-class communication through instant messaging. *Edu Train* 60(4):318–334. <https://doi.org/10.1108/et-12-2017-0196>
- O'Connell SD, Lang G (2018) Can personalized nudges improve learning in hybrid classes? Experimental evidence from an introductory undergraduate course. *J Res Technol Educ* 50(2):105–119. <https://doi.org/10.1080/15391523.2017.1408438>
- Oreopoulos P, Petronijevic U (2018) Student coaching: How far can technology go? *J Human Resour* 53(2):299–329. <https://doi.org/10.3368/jhr.53.2.1216-8439R>
- Oreopoulos P, Petronijevic U, Logel C, Beattie G (2020) Improving non-academic student outcomes using online and text-message coaching. *J Econ Behav Organ* 171:342–360. <https://doi.org/10.1016/j.jebo.2020.01.009>
- Paloyo AR, Rogan S, Siminski P (2016) The effect of supplemental instruction on academic performance: an encouragement design experiment. *Econ Educ Rev* 55:57–69. <https://doi.org/10.1016/j.econedurev.2016.08.005>
- Rau P-L, Gao Q, Li-Mei Wu (2008) Using mobile communication technology in high school education: motivation, pressure, and learning performance. *Comput Educ* 50:1–22. <https://doi.org/10.1016/j.compedu.2006.03.008>
- Roblyer MD, Freeman J, Donaldson MB, Maddox M (2007) A comparison of outcomes of virtual school courses offered in synchronous and asynchronous formats. *Internet Higher Edu* 10(4):261–268. <https://doi.org/10.1016/j.iheduc.2007.08.003>
- Sanders M, Elspeth K, Raj C, Michael L, Elizabeth L, Xian-Zhi S (2019) Using text reminders to increase attendance and attainment: evidence from a field experiment. <https://doi.org/10.2139/ssrn.3349116>
- Smith BO, White DR, Kuzyk PC, Tierney JE (2018) Improved grade outcomes with an e-mailed “grade nudge.” *J Econ Educ* 49(1):1–7. <https://doi.org/10.1080/00220485.2017.1397570>
- So S (2016) Mobile instant messaging support for teaching and learning in higher education. *Internet Higher Edu* 31:32–42. <https://doi.org/10.1016/j.iheduc.2016.06.001>
- UNESCO (2021) COVID-19: reopening and reimagining universities, survey on higher education through the UNESCO National Commissions. UNESCO, Paris, France
- Wang K, Wang C, Xi L, Zhang Y, Ouyang Y, Lou H, Zhang W, Zhang L (2014) A randomized controlled trial to assess adherence to allergic rhinitis treatment following a daily short message service (SMS) via the mobile phone. *Int Arch Allergy Immunol* 163:51–58. <https://doi.org/10.1159/000356317>
- Ziden AA, Rosli M, Gunasegaran T, Azizan SN (2017) Perceptions and experience in mobile learning via SMS. *Int J Interact Mob Technol* 11(1):116

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.