From Smart Phones to Smart Students:

Learning versus Distraction with Smartphones in the Classroom

Zhe Deng, Aaron Cheng, Pedro Ferreira* and Paul Pavlov

*** Corresponding Author: pedrof@cmu.edu, www.pedro-ferreira.me**

Abstract

We collaborate with a vocational school in China to examine the effect of using smartphones in the classroom on the academic performance of students. We randomly allocate students taking lectures in Chinese verbal into four experimental conditions: (i) smartphones banned; (ii) smartphones allowed and used at will by students; (iii) smartphones allowed, used at will by students, and teachers asked students to use the devices to assist instruction; (iv) smartphones banned, and teachers asked students to use a paper-based aid to assist instruction. We measure the performance gain of students by the change in the scores they obtained in identical tests taken at the beginning and the end of the lectures. We find that allowing students to use smartphones during the lecture at will reduced performance gain compared to when smartphones were banned. However, allowing smartphones into the classroom and having teachers ask students to use them to assist instruction increased the performance gain. The performance gain of students using the paper-based aid was similar to that of the students who could not use smartphones. To unravel the underlying mechanisms that drive these effects, we use video feeds collected during our experimental lectures, allowing us to code the time students spent learning and distracted, with and without their smartphone. We show that the increase in performance gain when smartphones are used to assist instruction comes from students spending a larger percentage of the time learning during the lecture using the device and from the fact that the positive marginal effect associated with smartphone learning outweighs the negative marginal effect associated with smartphone distraction. Our findings contribute to the literature on technology-assisted learning and offer practical and policy implications that teachers and schools can follow to cautiously allow smartphones in the classroom to improve student success.

Keywords: smartphone policy, learning, distraction, academic performance, randomized controlled trial

1. Introduction

The penetration of smartphones increased worldwide during the last decade. In the U.S., and according to Statista, smartphone ownership increased from 33% in 2011 to 82% in 2022. (Ghose, 2017) discusses how this stimulated significant research on the role of smartphones in business settings. Yet, much less is known about how allowing these devices into classrooms may affect the performance of students. Furthermore, the research on the impact of ICTs on productivity yields mixed results (Acemoglu et al., 2014, Lee et al. 2018), perhaps because most papers study very aggregate economic levels (i.e., countries, industries, firms). Fewer studies examine the impact of ICTs on the performance of workers, and even fewer studies explore the relationship between ICTs and attention (Alavi and Leidner, 2001).

Extant research on the impact of ICTs on the academic performance of students also reports mixed effects, perhaps because many studies focus on the effect of investments in ICTs, subsidies to ICTs, and availability of ICTs in schools, instead of looking directly at how students use ICTs for learning purposes. Thus, the existing literature provides little help to anticipate how smartphones in classrooms may affect student performance. Analyzing the effect of smartphones in the academic environment should thus focus on student level studies that provide causal measures of how smartphones affect attention, i.e., how they induce learning and distraction, and how the tradeoff between these behaviors affects performance.

We contribute to close this research gap by running 2 Randomized Controlled Trials (RCTs) to test the effect of smartphones in the classroom on the performance of students. We collect individual-level data on how students use smartphones during lectures to identify why the observed effects arises. We ask three research questions? 1) *How does allowing smartphones in the classroom affect the time students spend learning versus distracted?* 2) *How does smartphone learning and distraction in the classroom affect the performance of students?* 3) *How do demographic factors moderate the observed effects?*

During our field experiments, students in a vocational school in China were randomly assigned to several experimental conditions during 11 lectures of Chinese verbal, namely: i) smartphones banned from the classroom; ii) smartphones allowed into the classroom and used at will by students; iii) smartphones allowed into the classroom, used at will by students with teachers asking them to use the devices to assist instruction; iv) smartphones banned from the classroom with teachers asking students to use a paper-based aid to assist instruction. Assisted instruction in condition iii) was implemented by having the teachers asking students to use a smartphone dictionary app (that we developed for this experiment) during the lecture to check the pronunciation, definition, and etymology of the words taught. In condition iv) assisted instruction was implemented by giving students a paper-based dictionary, with the same content as the smartphone dictionary app, and having teachers ask students to use it for the same purposes as in condition iii). All students took identical tests at the beginning and the end of the lecture, allowing us to measure performance gain as the difference between pre-lecture and post-lecture test scores.

The analysis of the experimental data obtained yields notable findings. First, allowing smartphones into the classroom without teachers asking students to use the devices to aid instruction decreased the performance gain, on average, by 27% of a standard deviation compared to banning smartphones from the classroom. Second, allowing smartphones into the classroom with teachers asking students to use the device to aid instruction increased the performance gain, on average, by 26% of a standard deviation. These findings immediately suggest that simply banning smartphones from classrooms may not be the best policy because the positive learning effects that may arise may outweigh the negative ones. In fact, we show that, at the average, a 1 percentage point (p.p.) increase in the percentage of learning time allocated to the smartphone (as opposed to learning time allocated to other activities) increases the performance of students by 2.5 points, whereas a 1 p.p. increase in the percentage of distraction time allocated to the smartphone (as opposed to distraction time allocated to other activities) reduces the performance of students by 0.13 points. We also find that students use the smartphones for learning purposes more when the teachers ask them to do so compared to when they do not. This increase in smartphone usage for learning purposes, albeit small in absolute terms for many students, combined with the productivity of smartphone learning, outweighs the negative impact of smartphone distraction.

Measuring the time allocated by students to learning and distraction, on and off the smartphone, was accomplished by teams of video coders who replayed the video feeds from all lectures in our experiments and annotated when students were learning (e.g., using the dictionary app) and distracted (e.g., dozing off, looking astray, taking to each other, playing games or on social media on the smartphone). In our second experiment, we also recorded the teachers during the lectures. Analyzing their facial emotions and facial posture shows that the same teacher taught similar lessons across all our experimental conditions, avoiding confounders such as teacher-side learning. In addition, condition iv) in our second experiment allows us to separate the effect of the content – the additional content provided to the students included in the dictionary app and in the paper-based dictionary – from the effect of the channel – smartphone vs. paper. Our results show that the performance gain with the paper-based dictionary is similar to that observed under the condition where smartphones were banned from the classroom, i.e., the paper-based dictionary did not increase the performance gain of students as the smartphone did.

Finally, our paper reports a rich set of heterogeneous effects from moderators such as the familiarity of the students with the content of the lectures, the familiarity of the students with ICT, their gender, whether they were born in a rural or urban area, and how well they perform in school in general terms.

Our work embodies several contributions to the literature. First, we are among the first to run RCTs in a real-life classroom setting to explore the effect of smartphones on the performance of students, which avoids the limitations of using observational studies, and allows us to claim causal effects. Second, we decompose the overall effect of smartphones on student performance into the positive effect associated with assisted learning and the negative effect associated with smartphone distraction. Third, we actively measure the time students spend learning and distracted, on and off the smartphones, during the lectures, in a novel way that avoids the problems in the prior literature with self-reported measures. Finally, our study offers empirical evidence that teachers and administrators can use to develop guidelines and laws to manage the introduction of smartphones into the classroom.

2. Research Background and Literature Review

2.1 Impact of ICTs on student performance

Information and Communication Technologies (ITCs) are now pervasive in schools. The UNESCO Institute for Statistics reports that in 2020 more than 87% of secondary schools worldwide had access to computers for pedagogical purposes (UNESCO, 2023). The U.S. Institute of Education Sciences reports that 98% of all schools in the U.S. have computers, with 45% of them indicating owning one computer per student. However, the prior literature shows mixed evidence of the impact of ICTs on educational outcomes. (Leuven et al., 2007) found that offering subsidies for ICTs in the Netherlands to schools with disadvantaged pupils hurt learning outcomes. (Fried 2008, Carter et al., 2017) also found that the introduction of computers in schools can hurt learning outcomes in the U.S. (Vigdor et al., 2014) and (Belo et al., 2014) found that broadband availability and usage in middle schools reduced student grades, in the U.S. and Portugal, respectively.

Several studies found no effect of ICTs on student performance. (Angrist and Lavy, 2002) found that increasing computer usage in primary school in Israel did not change test scores. (Goolsbee and Guryan, 2006) found that using technology in public schools in California did not change the performance of students. (Barrera-Osorio and Linden, 2009) showed that computer-assisted programs for language learning in Colombia had no effect on student learning. (Faber et al., 2015) found no effect of improved broadband speed on educational outcomes across students in primary and secondary schools in England.

Other studies found benefits from using ICTs for educational purposes. For example, (Barrow et al. 2009) showed that computer-aided instruction was more effective than traditional learning methods in several urban U.S. school districts. (Kumar and Mehra ,2018) and (Muralidharan et al., 2019) found positive effects of computer-based personalized learning during after-school programs for students in middle schools in India. (Machin et al., 2007) found that subsidies for ICTs given in several school districts in England increased the performance of students in English and Science (though not in Mathematics), and (Fairlie and London, 2012) found that investment in ICTs in a large community college in California also increased the performance of students. The studies described above focus on different settings and treatments, and thus it is hard to compare them. In any case, the variance in the results reported in the literature highlights the complexity associated with appropriately using ICTs to improve education.

2.2 Smartphone pervasiveness and bans in schools

Smartphone and computer ownership exhibit very different statistics. According to Statista, in the U.S. and by the end of 2022, 34% of adults 30 years old or more owned a desktop, 35% owned a laptop, and 48% owned a tablet, but 88% owned a smartphone. The gap to smartphone ownership is also considerable for individuals aged between 18 and 29, for whom these statistics were 28%, 37%, 36% and 87%, respectively. According to Statista, in the U.S. and in 2021, 31% of 8-year-old kids had a smartphone, and this statistic increased to 93% for 18-year-old kids.

These statistics combined with the mixed evidence on the impact of ICTs on the performance of students led to fierce discussions about whether schools should allow students to take smartphones into classrooms (Allen, 2017). Parents and teachers often oppose allowing smartphones in classrooms because of the potential adverse effects that may arise, such as cyber-bullying (Osborne, 2012), distraction (Kuznekoff et al., 2015, Beland and Murphy, 2016), access to unreliable information (Wineburg et al., 2016), cheating (Haller, 2017), and physical or mental distress (Ward et al., 2017, Resnick, 2019). On the other hand, students usually favor allowing smartphones into classrooms. They highlight that these devices can be effectively used during lectures to assist learning. The increasing number of educational apps for mobile phones that support learning and lecture management, by facilitating real-time interaction and collaboration among students and teachers, fuel these claims (Taylor and Francis, 2015). In addition, allowing students to use their smartphones to assist learning during lectures may relieve schools from expensive investments in ICTs, which are increasingly required today to improve student experience.

However, the concerns of teachers and parents have led several policymakers to ban smartphones from classrooms and sometimes from schools altogether. For instance, France passed a bill requiring all students aged 3-15 to leave their smartphones at home or keep them off if taken to school (Law No. 2018-698). Likewise, China published a national notice guiding K-12 students to limit access to mobile phones at school (JJTH No. 3-2021). Similar policies and measures have also been proposed and enacted in other countries, such as the U.K., India, and the U.S. According to the National Center for Education Statistics, 2020, 76% of schools in the U.S have already banned cellphones. According to (UNESCO, 2023), by the end of 2022, 13% of countries have introduced laws to ban mobile phones from schools and 14% of countries have policies and guidelines to do so. This split (27% Vs. 73%) calls for a better understanding of whether smartphones can be productively used to assist with instruction.

2.3 Smartphones and the performance of students

People required to multi-task take longer to finish tasks compared to the time it would take them to finish them in tandem. According to (Rubinstein et al., 2001) the slowdown arises due to the loss of time with switching and gets worse as tasks become more complex. (Foerde et al., 2006) show that people who multi-task can learn factual information but then have a harder time applyiing it to new situations. (Chen and Yan, 2016, Uncapher and Wagner, 2018, Wilmer et al., 2017) discuss how smartphones may also hinder the reward process, which is a key function for learning. Importantly, multi-tasking requires attentional capacity, reducing cognitive control (Chinchanachokchai et al., 2015, van dur Schuur et al., 2015), which may lead to distraction and errors (Courage et. al, 2015).

Negative effects of laptop multi-tasking on student performance have already been reported in prior literature (Hembrooke and Gay, 2003, Fried, 2008, Kraushaar and Novak, 2010, Wood et al., 2012, Zhang et al., 2015). Just like laptops, smartphones also incentivize multi-tasking, which increases the cognitive load on students (Mayer and Moreno, 2003) and hinders the cognitive processes required to learn (Judd, 2014, Lee et al., 2015). (Rosen et al., 2013) show that U.S. students who used Facebook and texted while studying registered lower GPA. (Junco, and Cotten, 2012) found that using Facebook and texting while doing schoolwork decreased the performance of U.S. college students. (Wood et al., 2012) found that students who used Facebook while attending lectures had lower scores compared to students who did not.

(Ramjan et. al, 2021) analyzed 27 studies in multiples countries all surveying nursing students. The authors found that 20 studies reported a negative effect of increased smartphone usage on learning in both the classroom setting and during clinical practicum. The nursing students reported that their smartphone use distracted them from engaging in the appropriate learning processes. In another example, (Amez and Baert, 2020) reviewed 23 studies and found that 18 of them found a negative association between smartphones and academic success. The other 5 studies found no statistically significant associations. However, only 4 of these studies tracked smartphone usage actively and only 6 used data on grades from teachers or the administration. All the other studies relied on usage data and grades from surveys where students self-reported these measures. However, this leads to inaccuracy and bias due to the students' tendency to under-report less-desirable experiences (Krumpal, 2013).

For this reason, (Kates et. al, 2018) offers a meta-analysis of 39 studies selected because they include some construct of mobile phone usage, such as number of text messages or calls sent/received or even of measurements of the time spent using the devices, and some objective measure of academic performance, such as test scores or GPA. In total, these studies covered about 150 thousand students across more than 6 countries and still found adverse effects of mobile phone usage in schools on student performance.

All the studies analyzed in the reviews described above are correlational and thus do not offer empirical evidence of causal effects. An empirical approach to improve upon simple correlational studies is to use longitudinal data. (Bjerre-Nielsen et. al, 2020) ran such a study monitoring 470 students at the Technical University of Denmark over 2 years using a mobile app to unobtrusively log their activity. Their results show a negative association between in-class smartphone usage and grades, even after controlling for several observed student characteristics. Namely, a 1 standard deviation increase in classroom smartphone usage decreased GPA by 18% of a standard deviation. The authors also show that the magnitude of this effect decreased to 6.1% with fixed effects, hinting at the fact that cross-sectional studies may have overestimated the negative effects of smartphones. Note, however, that this study did not separate the classroom usage of smartphones into learning and distraction.

(Amez et. al, 2023) collected longitudinal data on students' smartphone use and educational performance of students attending 11 different programs at Ghent University and the University of Antwerp, over 3 consecutive years. They surveyed students and merged their data with exam scores. Using a random effects model, they found that a 1 standard deviation increase in smartphone usage led to a reduction of 0.349 points (out of 20) in exam scores and to a decrease of 2.6 p.p. in the fraction of exams passed. Finally, (McDonald, 2013) measured the effects of in-class texting policies, namely mild texting policy (do not use so you can respect others), strict cell phone policy (lose points in final grade if caught texting), and no presented policy regarding texting. The author found a negative relationship between texting and final grades in all conditions even after controlling for student GPA and attendance.

On the other hand, smartphones may also improve the performance of students. In particular, mobility allows them to use the same information-intensive services (e.g., e-learning websites) that laptops offer but everywhere and anytime (Lepp et al., 2014). In short, smartphones provide access to information for study purposes at a rate that textbooks do not (Zhang et al., 2014). Smartphones can also benefit students by supporting individualized learning with predefined teaching materials, such as in museums and laboratories (e.g., Hall and Bannon, 2006). Social networking and messaging apps may also help to quickly share relevant study information among peers and students and teachers potentially leading to more efficient study and fruitful collaboration (Chen and Ji, 2015, Lepp et al., 2015).

The trained attention hypothesis argues that one can train attention like one does with other functions of the brain (Sohlberg and Mateer, 1987). Frequent multi-tasking can help in this regard because it promotes mental flexibility (Courage et al., 2015) and fileting of irrelevant information (Alzahabi and Becker, 2013, Ophir et al., 2009). As such, smartphones may also improve the performance of students. However, there is only scant empirical evidence to support these claims. Two large scale studies stand out. (Lin et al., 2021) surveyed more than 10 thousand college university students in China and found that using mobile learning and news apps increased academic performance, whereas playing video games and engaging in social media decreased academic performance. (Rabiu et. al, 2016) ran a survey on more than 6 thousand secondary school students in Nigeria and found a positive association between mobile phone usage and scores in both the mathematic achievement test and the English language achievement test.

3. Empirical Context and Setup

We conducted 2 RCTs in a school in China to study how allowing smartphones in the classroom affect the performance of students. This vocational school did not implement any mobile-related regulations prior to our experiments. The sub-sections below describe the design of our experiments.

3.1 First Experiment

At the time of our first experiment, there were 482 students in the class of 2019, all in their second year of studies. They were between 14 and 23 years old and enrolled in 8 different majors. They took core classes in a large classroom holding up to 125 students and elective courses in other small classrooms. The experiment was conducted in 11/2018 during 3 lectures in Chinese verbal.

Figure 1 depicts our experimental protocol. We randomly selected 125 students from the class of 2019 for each of 3 experimental conditions. In condition B, smartphones were Banned from the classroom. In the condition Ta, students were Allowed to use smartphones during the lecture as they wished. In condition Ti, students were allowed to use smartphones during the lecture as they wished, and the teacher asked them to use the devices to aid instruction. During this experiment the same teacher provided 3 identical lectures to all 375 students. In all such lectures, the teacher explained the pronunciation, definition, and etymology of several words and their logical functions in sentences from pre-identified pieces of literature, including folktale, drama, and poetry, in both traditional and modern Chinese. A traditional Chinese prose from the usual curriculum of this course, with which students were unlikely to be familiar, was used for these lectures. The teacher was asked to provide the same lecture at the same pace to the 3 groups of students in our experiment. The lectures took place on 3 consecutive weekdays, all at the same time of the day and always in the same classroom.

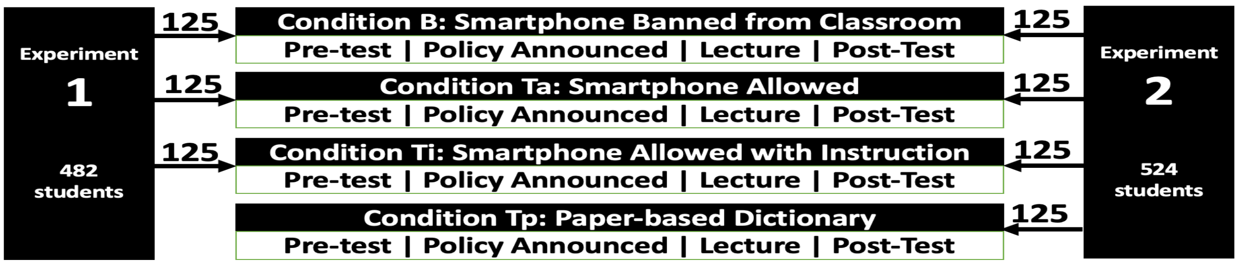


Figure 1. Experimental design 125 students were randomly selected into several conditions. All lectures started and ended with the same knowledge test to assess performance gain per condition.

We could not have all 3 lectures delivered simultaneously since there was only 1 teacher available for our experiment. However, teacher characteristics, such as style and experience, are likely to affect the performance of students. Therefore, conditions B, Ta and Ti were implemented in this order to minimize potential interference. Suppose students in different conditions chat about the lecture. Still, students in condition Ta are unlikely to anticipate that they could use smartphones during the lecture, given that the students who took the lecture before were not allowed to do so. Similarly, students in condition Ti could potentially anticipate that they could use smartphones during this lecture but were unlikely to anticipate being provided specific instructions for how to use the smartphone for learning purposes.

All 3 lectures were 90 minutes long and started with a pre-test (5 min), during which smartphones were turned off, to assess the student's knowledge of the words that would be taught. These tests included a series of multiple-choice questions similar to other quizzes taken in Chinese verbal, and the test scores were used to compute final grades. Next, the teacher announced the policy regarding smartphone use during the lecture. In condition B, students were requested to place their devices in a wall hang-up organizer at the back of the room, effectively preventing them from using the smartphone during the lecture. In condition Ta, students were told that they could use the smartphone during the lecture at will. In condition Ti, students were told that they could use the smartphone during the lecture at will and access a web-based dictionary app. The teacher demonstrated how to do so after the pre-test.

The figure in Appendix C shows a snapshot of the screen of this app. Students could use their smartphone at any time during the lecture to scan a QR code in the corner of their desk to access the app and search for words to learn about their pronunciation, meaning, and etymology. After the announcement of the smartphone policy, the teacher delivered the lecture. The only difference between condition Ta and Ti was that in the latter the teacher asked students to use the dictionary app every time a new word was introduced. At the end of the lecture, all students in all conditions took the same test as when the lecture started (although they did not know a-priori that there would be a test, much less one similar to the pre-test). We designed this experiment to ensure that all students across all conditions had little knowledge about the new words taught (as we will see from the significantly poor scores in the pre-tests). Therefore, our exercise may be considered as a representative case of providing students with new knowledge.

Finally, the large classroom where the experimental lectures took place was equipped with a video anti-cheating system, with 20 cameras mounted on the ceiling, which are part of a school-wide security system with an additional 36 cameras across the campus. This system is used regularly during examinations to ensure academic integrity. Students and parents consent to it. Students know about the cameras in the classroom where our experiment took place and that these cameras are routinely used. However, they were unaware that they were used during our experiment until the delayed debriefing that we provided them at the end of the experiment (detail about this procedure are provided in Appendix D). Several video coders looked at the video feeds provided by this unique setting to measure student attention by identifying the time students spent using their smartphones and not, learning and distracted.

3.2 Second Experiment

We ran a second experiment in November 2022 during an additional eight lectures in Chinese verbal. At that time there were 524 students in the class of 2024, all in their second year of studies. They were also between 14 and 23 years old and enrolled in 8 majors. The design of this experiment is similar to that of our first experiment, as shown in Figure 1. We randomly selected 125 students from the class of 2024 for each of 4 experimental conditions. The first 3 conditions mimic the 3 conditions in our first experiment. In the fourth condition (Tp), smartphones were banned from the classroom and the teacher asked them to use a paper-based dictionary for learning purposes.

During this experiment a new (male) teacher provided 4 identical lectures in Chinese verbal to all 524 students and, another new (female) teacher, provided another 4 identical lectures in other topics of Chinese verbal. The structure of all these lectures was similar to the structure of the lectures offered during our first experiment. As before, teachers were asked to provide the same lecture at the same pace to the 4 groups of students assigned to each of them in our experiment. The 4 lectures provided by the first teacher took place on 4 consecutive weekdays, all at the same time of the day and always within the same classroom. The second teacher followed the same protocol in the following week. Our experimental conditions B, Ta and Ti were implemented in the same order as in our first experiment. Condition Tp was added on the fourth day to mitigate concerns of interference. As such, even if students talked about using smartphones in the classroom, the students in the Tp condition were unlikely to anticipate that they would be asked to use a paper-based dictionary to aid instruction.

The figure in Appendix C shows a picture of the paper-based dictionary given to students in condition Tp as part of the announcement regarding smartphone policy. The handouts mimic the presentation of words offered to students through the smartphone app. Furthermore, the content printed in the handouts was the same as that uploaded to the dictionary app. Therefore, we can compare the performance gain of students between conditions Ti and Tp to isolate the effect of channel (smartphone vs. paper), which we are interested in this paper, from the effect of content. Finally, this experiment took place in the same large classroom where our first experiment also happened (thus exactly the same physical setting).

4. Data and Empirical strategy

4.1 Data Collected

We used 3 anonymized datasets in our paper: i) data provided by the school on student demographics; ii) scores in the pre-tests and post-tests; iii) video feeds of the lectures. Table 1 describes the key variables. We take the difference between the pre-test and post-test scores for each student to compute the performance gain during each lecture. Note again that the pre-test and post-test are the same and thus their scores are comparable. A group of experienced teachers in Chinese verbal designed this test with the following guidelines: i) the questions should inquire only about what was taught during the lecture; ii) the questions should be multiple-choice (to facilitate objective grading); (iii) a 100-point scale should be used, and the distribution of scores should approximate a normal distribution (that is, the test should not be exaggeratedly easy nor exaggerated hard). The teacher offering the lecture graded all tests from all students attending that lecture anonymously, thus ensuring grading consistency and anonymity.

Covariate	**Definition**
Experimental conditions	Dummary variables
Banned (B)	Smartphone banned
Allowed (Ta)	Smartphone used at wil
Instruction (Ti)	Smartphone used at will and for instruction
Paper (Tp)	Paper-based dictionary used for instruction
Performance outcomes	Test scores in 0-100 point scale
Pre-test score	Score in the test at the beginning of lecture
Post-test score	Score in the test at the end of lecture
Performance gain	Post-test score minus pre-test score

Covariate	**Definition**
Endogenous variables	Measured in seconds
Learning smartphone (LS)	Time learning using the smartphone
Distraction smartphone (DS)	Time distracted using the smartphone
Learning Other (LO)	Time learning not using the smartphone
Distraction other (DO)	Time distracted not using the smartphone
Note: control variables included in our work: gender, age, ethnics, local graduate, place born middle school graduate,schooling system, major	

Table 1. Key variables used in our empirical analyses.

We use video feeds from the lectures to compute the time that each student spent during the lecture on i) Learning using the Smartphone (LS); ii) Learning without using the smartphone (LO); Distraction using the Smartphone (DS); Distraction without using the smartphone (DO). Students in the videos were matched to those in our experiments using the seat number reported at the top of their answer key (reporting this in tests was a common practice at the school, and students were used to it).

Getting estimates for LS, LO, DS, DO from the video feeds was accomplished by local teams of video coders including faculty and staff at the school. We used 20 and 30 coders for the videos in the first and second experiments, respectively. Coders were trained to play the video clips, label events as LS, LO, DS, DO for each student, and estimate their duration by taking note of the corresponding start and end timestamps. The cameras in the classroom provided high-definition video streams allowing coders to check the screen of the students' smartphones and thus evaluate whether they were learning (e.g., using the dictionary app or other educational apps/websites) or distracted (e.g., playing games, on social media).

We assembled 180 and 480 30-minute video clips during our first and second experiments, respectively. Coders were split into groups of 3 or 4 and each coder in a group coded independently the videos randomly given to her. Agreement across coders for the same video was very high (the average inter-coder reliability statistics (Cohen's k) were 0.96 ($p < 0.001$) and 0.92 (p-value<0.01) for the first and second experiments, respectively), which provides evidence of robustness and consistency in how the behavior of students during our experiments was coded.

In addition, we had a camera pointing at the teacher in the lectures of our second experiment. The video feeds from this camera were sampled every 5 seconds to get a frame with the teacher's face. From her face, we got estimates for 7 facial emotions using software provided by the Open Computer Vision library (OpenCV) available at http://opencv.org. Roughly 10% of the frames in each lecture failed to report an emotion because the teacher may have, for example, moved around and been off camera. These data were complemented with data on the teacher's facial posture measured using state of the art computer vision algorithms, namely the HSEmotion package (High-Speed face Emotion recognition) from Andrey Savhenko: https://github.com/HSE-asavchenko/hsemotion, the GazeTracking package from Antoine Lame: https://github.com/antoinelame/GazeTracking, and https://developers.google.com/mediapipe. Table 2 describes the covariates used in our analyses.

Covariate
Facial emotions [0,1]
Angry
Disgust
Fear
Sad
Happy
Surprised
Neutral

Covariate	Definition
Facial Posture (coordinates and mm)	
Head pose x,y	Displacement of teachers' head
Head posEye gaze x,y	Point where teacher is looking at
Distance between lips	Proxy for whether teacher is speaking

Table 2. Covariates obtained from analyzing the teachers' video feeds.

3.2 Empirical Strategy

We followed an empirical strategy with 3 steps to analyze our data. First, we look at the effect of smartphone policy on post-test scores and student performance gain using

$$Y_j = \beta_0 + \beta_1 Ta_j + \beta_2 Ti_j + X'_j\beta_3 + \varepsilon_j \qquad \text{(Eq. 1)}$$

where the unit of analysis is a student, X_j is a vector of student characteristics, Ta_j and Ti_j are dummy variables indicating whether the student is in condition Ta or Ti (both 0 if the student is assigned to condition B, which is the baseline condition in all our regressions). Y_j represents the student's post-test test score or performance gain and ε_j is the idiosyncratic error term. This specification allows us to measure how the dependent variable changes from condition B to conditions Ta and Ti in causal terms, given that students were randomly allocated to experimental conditions. We also used this setup with the pre-test score as our dependent variable to show balance across conditions. We do not cluster the standard errors, given our randomized setup and all the student level controls included in our regressions. We report statistical significance using Bonferroni correct for multiple testing.

Second, we examine the effect of smartphone policy on how students use the device during the lectures using regressions similar to Eq. 1, but where Y_j represents the student's percentage of learning and distraction time allocated to the smartphone, %LS=LS/(LS+LO) and %DS=DS/(DS+DO), respectively. Also, in this case, this specification allows us to measure the causal effect of smartphone policy on these percentages relative to condition B. Third, we study at how %LS and %DS affect the performance gain of students using

$$\%DS_j = \alpha_0^{DS} + \alpha_1^{DS} Ta_j + \alpha_2^{DS} Ti_j + X'_j\alpha_3^{DS} + \epsilon_j^{DS} \qquad \text{(Eq. 2)}$$

$$\%LS_j = \alpha_0^{LS} + \alpha_1^{LS} Ta_j + \alpha_2^{LS} Ti_j + X'_j\alpha_3^{LS} + \epsilon_j^{LS} \qquad \text{Eq. 3)}$$

$$Y_j = \alpha_0 + \alpha_1 \widehat{\%DS_j} + \alpha_2 \widehat{\%LS_j} + X'_j\alpha_3 + \epsilon_j \qquad \text{(Eq. 4)}$$

The first two regressions above are the first stages for our IV analysis and represent the two regressions in the previous step of our empirical strategy. The third regression above is our second stage, allowing us to measure the Local Average Treatment Effect (LATE) of %LS and %DS on student performance gain (Angrist and Imbens 1995). Using LATE is customary when analyzing the effects of endogenous mechanisms using RCTs. This empirical strategy allows us to discuss why our results arise. For example, if the effects of Ta and Ti on %LS and %DS are relatively small in magnitude then students were unlikely motivated to use smartphones during the lectures to begin with. In particular, if %LS does not increase from Ta to Ti then teachers were unable to induce students to use the smartphone to aid instruction. If, however, the effects of %LS and %DS are relatively large in magnitude then students were likely motivated to use smartphones during the lectures. In this case, the effects obtained in the second stage identify how productively students use the devices for learning and distraction. The potential tradeoff between these two effects determines what happens to performance gain. Finally, if %LS and %DS do not change from B to Ta or Ti then our first stage regressions do not work as intended and we cannot draw conclusions about how %LS and %DS affect performance gain.

5. Analysis of the First Experiment

5.1 Summary Statistics

Table 3 provides descriptive statistics for the key data collected during our first experiment. The average pre-test and post-test scores across all conditions were 29 and 81 points, which indicate a significant performance gain. Students spent 88% of lecture time learning (4090/60=68 min) across all three conditions, most of it not on the smartphone. In the Ti condition, about 50% of the students used the smartphone for learning of which about 10% spent more than 16.6 doing so. However, and also in this condition, about 80% of the students spent time distracted on the smartphone. On average, 57% of their distraction time was spent on the smartphone. Students spent a similar amount of time learning across all our conditions. This finding is in line with the idea that students devoted as much attention to learning as they possibly could during the lectures, and, at least in our case, whether smartphones were banned or allowed into the classroom did not change this statistic. This may be due to their limited attention span preventing them from allocating more (maybe 100%) of the lecture time to learning (Kahneman 1973, pp. 10, Lachman et al. 1979, pp 186).

Students spent a similar amount of time on the smartphone (9 min) in conditions Ta and Ti. Therefore, we find that students also devoted a limited amount of time to the smartphones. Attention is a scarce resource allocated to different behaviors and devices depending on their "style" or "rhetoric" (e.g., direct manipulation, multimedia presentation, Lanham 2006, pp. xi). The nature of the human-machine interaction also seems to regulate the allocation of attention by humans to machines (Piccoli et al. 2001) rather than the nature of the content itself (learning vs. distraction in our case). In our setting, substituting no smartphone time for smartphone time is determined by the students' use of smartphones for learning plus distraction and is not affected by how they are prompted to do so by the teacher.

	B (N=125)	Ta (N=119)	Ti (N=123)	TRUE (N=367)
Pre-test score				
Mean (SD)	29.5 (22.0)	27.3 (21.5)	29.8 (21.6)	28.9 (21.7)
Median [Min, Max]	28.0 [0, 98.0]	21.0 [0, 94.0]	27.0 [0, 88.0]	24.0 [0, 98.0]
Post-test score				
Mean (SD)	82.1 (15.0)	73.6 (24.1)	86.8 (10.5)	80.9 (18.2)
Median [Min, Max]	87.0 [18.0, 100]	84.0 [4.00, 100]	87.0 [65.0, 100]	86.0 [4.00, 100]
Performance gain				
Mean (SD)	52.6 (21.6)	46.4 (23.8)	57.0 (16.8)	52.1 (21.3)
Median [Min, Max]	56.0 [-13.0, 91.0]	49.0 [-25.0, 88.0]	60.0 [12.0, 92.0]	55.0 [-25.0, 92.0]
Smartphone learning				
Mean (SD)	0 (0)	35.8 (117)	218 (332)	84.6 (224)
Median [Min, Max]	0 [0, 0]	0 [0, 520]	0 [0, 1450]	0 [0, 1450]
Smartphone distraction				
Mean (SD)	5.84 (30.4)	504 (628)	333 (494)	277 (502)
Median [Min, Max]	0 [0, 256]	293 [0, 3330]	125 [0, 3330]	44.0 [0, 3330]
Other learning				
Mean (SD)	4130 (583)	3990 (719)	3880 (750)	4000 (693)
Median [Min, Max]	4330 [2190, 4800]	4160 [1470, 4800]	4040 [847, 4800]	4160 [847, 4800]
Other distraction				
Mean (SD)	660 (581)	274 (402)	371 (584)	438 (554)
Median [Min, Max]	464 [0, 2610]	117 [0, 2180]	121 [0, 3860]	206 [0, 3860]
Total learning				
Mean (SD)	4130 (583)	4020 (691)	4100 (759)	4090 (680)
Median [Min, Max]	4330 [2190, 4800]	4190 [1470, 4800]	4400 [847, 4800]	4280 [847, 4800]
Total distraction				
Mean (SD)	665 (583)	779 (691)	704 (759)	715 (680)
Median [Min, Max]	470 [0, 2610]	608 [0, 3330]	404 [0, 3950]	521 [0, 3950]
Total smartphone				
Mean (SD)	5.84 (30.4)	540 (658)	551 (586)	362 (566)
Median [Min, Max]	0 [0, 256]	293 [0, 3330]	405 [0, 3330]	65.0 [0, 3330]
Total non-smartphone				
Mean (SD)	4790 (30.4)	4260 (658)	4250 (586)	4440 (566)
Median [Min, Max]	4800 [4540, 4800]	4510 [1470, 4800]	4400 [1470, 4800]	4740 [1470, 4800]

Table 3: Descriptive statistics for the key covariates used throughout our analyses.

5.2 Effects of Smartphone Policy

Table 4 shows our results. We observe that the pre-test scores are similar across all our experimental conditions providing evidence of balance. On average, allowing smartphones into the classroom without the teacher asking students to use them to aid instruction (Ta) reduces performance gain by 16.4% (-5.696/34.671) (27% of a 1 standard deviation), compared to when smartphones are not allowed into the classroom. On average, allowing smartphones into the classroom with teachers asking students to use them to aid instruction (Ti) increases performance gain by 16.0% (5.532/34.671) (26% of 1 standard deviation) compared to when smartphones were not allowed into the classroom.

	Dependent variable:					
	Pre-Treatment		Post-Treatment		Performance Gain	
	(1)	(2)	(3)	(4)	(5)	(6)
Ta	-2.210	-1.820	-8.465***	-7.516***	-6.255*	-5.696*
	(2.781)	(2.451)	(2.228)	(1.898)	(2.677)	(2.476)
Ti	0.276	-1.333	4.644	4.198*	4.368	5.532*
	(2.757)	(2.425)	(2.209)	(1.878)	(2.655)	(2.450)
Constant	29.496***	46.967	82.112***	81.638***	52.616***	34.671
	(1.942)	(24.698)	(1.556)	(19.126)	(1.869)	(24.946)
Controls	No	Yes	No	Yes	No	Yes
Ti-Ta	2.487	0.487	13.109***	11.715***	10.622***	11.228***
Observations	367	367	367	367	367	367
Adjusted R2	-0.003	0.240	0.083	0.351	0.036	0.196

Note: Statistical significance with Bonferroni correction (* p<0.1; ** p<0.05; *** p<0.01)

Table 4: Effect of smartphone policy on test scores and performance gain.

Table 5 shows the first stage results. Students did not use smartphones for learning purposes too much when the devices are allowed into the classroom and teachers did not ask them to use the devices to aid instruction. This changed when teachers did so. The percentage of learning time allocated to the smartphone increased by 3.8 p.p. from Ta to Ti. Students used smartphones for distraction significantly across both the Ta and Ti conditions. The percentage of distraction time allocated to the smartphone is 64% and 47% in the Ta and Ti conditions, respectively, and these statistics are statistically different.

	Dependent variable:			
	First-Stage %LS	First-Stage %DS	First-Stage %LS	First-Stage %DS
	(1)	(2)	(3)	(4)
Ta	0.011	0.591***	0.009	0.589***
	(0.006)	(0.040)	(0.006)	(0.039)
Ti	0.048***	0.489***	0.048***	0.486***
	(0.006)	(0.040)	(0.006)	(0.039)
Constant	-0.015	0.064	-0.035	-0.131
	(0.016)	(0.103)	(0.064)	(0.417)
Controls	No	No	Yes	Yes
Ti-Ta	0.037***	-0.103**	0.038***	-0.103**
Observations	359	359	359	359
Adjusted R2	0.161	0.414	0.170	0.447
Note:	Statistical significance with Bonferroni correction (* p<0.1; ** p<0.05; *** p<0.01)			

Table 5: Effect of smartphone policy on the % of learning and distraction time on the smartphone.

Table 6 shows the results obtained from our second stage regressions. First, our instruments are not weak (Stock and Yogo, 2005) as shown by the Cragg-Donald statistic. Second, the percentage of learning time allocated to the smartphone (%LS) increases performance gain. At the average of %LS, a 1 p.p. increase in LS increased performance gain by 2.5 points. Conversely, the percentage of distraction time allocated to the smartphone (%DS) decreases performance gain. At the average of %DS, a 1 p.p. increase in DS reduced performance gain by 0.13 points. Therefore, the percentage of learning time allocated to the smartphone had a significantly larger impact on performance gain than associated to distraction.

	Dependent variable:	
	Performance Gain	
	(1)	(2)
% Learning Smartphone (%LS)	231.444***	252.439***
	(71.062)	(68.660)
% Distraction Smartphone (%DS)	-13.620**	-12.892**
	(5.665)	(5.475)
% Learning (%L)	-3.555	-10.570
	(8.855)	(9.006)
Constant	55.699***	50.910
	(7.924)	(31.293)
Controls	No	Yes
Cragg-Donald Statistic	25.53>7.03	25.93>7.03
Observations	359	359
Adjusted R2	-0.194	-0.123
Note:	Statistical significance: * p<0.1; ** p<0.05; *** p<0.01	

Table 6: Effect of percentage of learning and distraction times using the smartphone on performance gain.

As a result, allowing smartphones into the classroom and having teachers ask students to use them for instruction can increase performance gain because even a small usage of the smartphone for learning purposes can overcome (2.5/0.13=20x) more usage for distraction purposes. We find that how students perform is related to the nature of learning and distraction, i.e., how much of these activities are performed using the smartphone versus not, or, in other words, how smartphone-intensive these activities are. The percentage of time learning during the lectures does not predict performance gain in our setting because it does not change significantly across our conditions, as shown by the summary statistics.

6. Analysis of the Second Experiment

6.1 Summary Statistics

	B (N=250)	Ta (N=238)	Ti (N=246)	Tp (N=244)	TRUE (N=978)
Pre-test score					
Mean (SD)	26.8 (21.7)	24.9 (21.6)	27.9 (21.9)	31.0 (23.1)	27.6 (22.1)
Median [Min, Max]	24.9 [0, 91.3]	16.6 [0, 91.3]	24.9 [0, 83.0]	24.9 [0, 91.3]	24.9 [0, 91.3]
Post-test score					
Mean (SD)	76.2 (15.5)	69.2 (24.4)	84.9 (11.0)	83.2 (15.6)	78.5 (18.3)
Median [Min, Max]	83.0 [8.30, 91.3]	78.9 [0, 99.6]	83.0 [58.1, 99.6]	91.3 [16.6, 99.6]	83.0 [0, 99.6]
Performance gain					
Mean (SD)	49.4 (21.7)	44.3 (24.5)	57.1 (16.9)	52.3 (21.8)	50.8 (21.8)
Median [Min, Max]	49.8 [-16.6, 91.3]	49.8 [-24.9, 91.3]	58.1 [16.6, 99.6]	58.1 [0, 99.6]	49.8 [-24.9, 99.6]
Smartphone learning					
Mean (SD)	0 (0)	40.7 (117)	220 (331)	0 (0)	65.3 (198)
Median [Min, Max]	0 [0, 0]	5.00 [0, 527]	6.00 [0, 1460]	0 [0, 0]	0 [0, 1460]
Smartphone distraction					
Mean (SD)	0 (0)	564 (626)	363 (493)	0 (0)	228 (463)
Median [Min, Max]	0 [0, 0]	355 [41.0, 3390]	152 [18.0, 3370]	0 [0, 0]	0 [0, 3390]
Other learning					
Mean (SD)	4230 (575)	3960 (714)	3870 (748)	4340 (517)	4100 (672)
Median [Min, Max]	4430 [2260, 4800]	4130 [1410, 4760]	4030 [842, 4770]	4560 [2530, 4800]	4290 [842, 4800]
Other distraction					
Mean (SD)	574 (575)	237 (387)	347 (579)	426 (517)	398 (535)
Median [Min, Max]	374 [0, 2540]	55.0 [0, 2140]	91.5 [0, 3830]	202 [0, 2260]	164 [0, 3830]
Total learning					
Mean (SD)	4230 (575)	4000 (685)	4090 (754)	4370 (517)	4170 (654)
Median [Min, Max]	4430 [2260, 4800]	4160 [1410, 4760]	4390 [844, 4780]	4600 [2540, 4800]	4390 [844, 4800]
Total distraction					
Mean (SD)	574 (575)	800 (685)	710 (754)	426 (517)	626 (654)
Median [Min, Max]	374 [0, 2540]	638 [41.0, 3390]	415 [22.0, 3960]	202 [0, 2260]	407 [0, 3960]
Total smartphone					
Mean (SD)	0 (0)	604 (656)	583 (585)	0 (0)	294 (528)
Median [Min, Max]	0 [0, 0]	360 [45.0, 3390]	435 [20.0, 3370]	0 [0, 0]	0 [0, 3390]
Total non-smartphone					
Mean (SD)	4800 (0)	4200 (656)	4220 (585)	4800 (0)	4510 (528)
Median [Min, Max]	4800 [4800, 4800]	4440 [1410, 4760]	4370 [1430, 4780]	4800 [4800, 4800]	4800 [1410, 4800]

Table 7: Descriptive statistics for key covariates used in our analyses.

Table 7 provides descriptive statistics for the key variables used in our second experiment. The average pre-test and post-test score across all conditions were 25 and 83 points, and the average post-treatment test score was 83 points, indicating again a significant performance gain. Students spent again 88% of the lecture time learning (4170/60=70 minutes) across all conditions, most of the time not on the smartphone. In the Ti condition, about 45% of the students used the smartphone to learn, of which about 10% spent more than 13.8 minutes learning on the smartphone. However, and also in this condition, all the students spent time distracted on their smartphones. On average, interestingly, 61% of their distraction time was on the smartphone. Students spent a similar amount of time learning across all conditions. The time spent by students on the smartphone (10 min) is also similar in conditions Ta and Ti.

6.2 Effect of Smartphone Policy

Table 8 shows our main results. On average allowing smartphones into the classroom without the teacher asking students to use them to aid instruction (Ta) reduced performance gain by 12.9% (-4.952/38.463, or 21% of a standard deviation) compared to when smartphones were not allowed into the classroom. On average, allowing smartphones into the classroom with teachers asking students to use them to aid instruction (Ti) increased performance gain by 23.2% (8.914/38.463, or 41% of 1 standard deviation) compared to when smartphones were not allowed into the classroom.

	Dependent variable:					
	Pre-Treatment		Post-Treatment		Performance Gain	
	(1)	(2)	(3)	(4)	(5)	(6)
Ta	-1.960	-1.249	-7.037***	-5.841***	-5.077**	-4.592**
	(1.965)	(1.729)	(1.607)	(1.379)	(1.922)	(1.799)
Ti	1.044	-0.933	8.696***	7.981***	7.652***	8.914***
	(1.949)	(1.682)	(1.594)	(1.341)	(1.906)	(1.750)
Constant	26.826***	26.378	76.227***	64.841***	49.402***	38.463*
	(1.373)	(16.554)	(1.122)	(13.202)	(1.342)	(17.227)
Controls	No	Yes	No	Yes	No	Yes
Ti-Ta	3.004	0.317	15.733***	13.822***	12.729***	13.506***
Observations	734	734	734	734	734	734
Adjusted R2	0.001	0.271	0.113	0.385	0.054	0.219
Note:	Statistical significance with Bonferroni correction (* p<0.1; ** p<0.05; *** p<0.01)					

Table 8: Effect of smartphone policy on test scores and performance gain.

Table 9 shows the results obtained from our first stage regressions. Students did not use smartphones for learning purposes too much when the devices were allowed into the classroom and teachers did not ask students to use them to aid instruction. This changed when teachers did so. The percentage of learning time allocated to the smartphone increased by 3.9 p.p. from Ta to Ti. Students use smartphones for distraction across both the Ta and Ti conditions. The percentage of distraction time allocated to the smartphone is 71% and 51% in the Ta and Ti conditions, respectively, and they are statistically different.

	Dependent variable:			
	First-Stage %LS (1)	First-Stage %DS (2)	First-Stage %LS (3)	First-Stage %DS (4)
Ta	0.012**	0.730***	0.014***	0.731***
	(0.005)	(0.026)	(0.005)	(0.026)
Ti	0.052***	0.619***	0.053***	0.611***
	(0.004)	(0.025)	(0.004)	(0.025)
Constant	-0.019	-0.268***	0.061	0.039
	(0.012)	(0.066)	(0.044)	(0.252)
Controls	No	No	Yes	Yes
Ti-Ta	0.04***	-0.112***	0.039***	-0.12***
Observations	712	712	712	712
Adjusted R2	0.173	0.569	0.218	0.593
Note:	Statistical significance with Bonferroni correction (* p<0.1; ** p<0.05; *** p<0.01)			

Table 9: Effect of smartphone on percentage of learning and distraction time on the smartphone

Table 10 shows the results obtained from our second stage regressions. Our instruments are not weak (Stock and Yogo, 2002) as shown by the Cragg-Donald statistic. The percentage of learning time allocated to the smartphone (%LS) increases performance gain. At the average of %LS, a 1 p.p. increase in LS increases the performance gain by 3.2 points. The percentage of distraction time allocated to the smartphone (%DS) decreases performance gain. At the average of %DS, a 1 p.p. increase in DS reduces performance gain by 0.11 points. As in our first experiment, allowing smartphones into the classroom and having teachers ask students to use them for instruction can increase performance gain because even a small usage of the device for learning can overcome (3.2/0.11=29x) more usage for distraction, even when the total learning and distraction times remain unchanged. Again, the percentage of time learning during the lectures does not predict performance gain because it did not change significantly across our conditions, as shown by the summary statistics. In sum, we observe similar results in the first and second experiments, which provides evidence of empirical replication, increasing our confidence in our results.

	Dependent variable:	
	Performance Gain	
	(1)	(2)
% Learning Smartphone (%LS)	292.369***	315.304***
	(51.722)	(52.061)
% Distraction Smartphone (%DS)	-10.472***	-11.252***
	(3.631)	(3.734)
% Learning (%L)	-5.067	-10.009
	(6.754)	(6.910)
Constant	52.782***	25.912
	(5.906)	(23.592)
Controls	No	Yes
Cragg-Donald Statistic	54.71>7.03	54.24>7.03
Observations	712	712
Adjusted R2	-0.335	-0.281
Note:	Statistical significance: * p<0.1; ** p<0.05; *** p<0.01	

Table 10: Effect of percentage of time learning and distraction on smartphone on performance gain.

6.3 Additional Results

6.3.1 The Effect of Channel Vs. Content

The second experiment included a fourth condition in which smartphones were not allowed into the classroom and teachers asked students to use a paper-based dictionary to aid instruction. This allows us to disentangle the effect of channel – smartphone vs. paper -- from the effect of the content given to students to aid instruction. The handouts given to students under condition Tp included the same content as that uploaded to the smartphone app used in condition Ti. Furthermore, in condition Tp, teachers were required to use the same prompts to ask students to use the paper-based dictionary as they did in condition Ti to ask them to use the smartphones to aid instruction.

The results obtained are shown in table 11. We find that student performance gain is similar in conditions B (smartphones are banned from the classroom) and Tp, and, therefore, the increase in student performance gain previously observed under condition Ti erodes with the paper-based dictionary. Note that the coefficients of Tp in the regressions of performance gain are slightly positive, and thus perhaps the paper-based dictionary was not completely useless, but highly statistically insignificant (these coefficients are roughly 3 times lower than those observed for Ti). This finding clarifies that the positive effects observed in condition Ti are associated to providing the content to students on the smartphone.

	Dependent variable:					
	Pre-Treatment		Post-Treatment		Performance Gain	
	(1)	(2)	(3)	(4)	(5)	(6)
Ta	-1.960	-1.707	-7.037***	-6.492***	-5.077*	-4.785*
	(1.999)	(1.899)	(1.562)	(1.431)	(1.934)	(1.875)
Ti	1.044	-0.236	8.696***	8.273***	7.652***	8.509***
	(1.982)	(1.859)	(1.549)	(1.401)	(1.918)	(1.835)
Tp	4.129	3.465	7.011***	6.219***	2.882	2.753
	(1.986)	(1.912)	(1.552)	(1.441)	(1.922)	(1.888)
Constant	26.826***	37.872*	76.227***	66.539***	49.402***	28.667
	(1.396)	(15.241)	(1.091)	(11.482)	(1.351)	(15.046)
Controls	No	Yes	No	Yes	No	Yes
Tp-Ti	3.004	0.317	15.733***	13.822***	12.729***	13.506***
Observations	978	978	978	978	978	978
Adjusted R2	0.007	0.141	0.112	0.286	0.041	0.137
Note:	Statistical significance with Bonferroni correction (* p<0.1; ** p<0.05; *** p<0.01)					

Table 11: Effect of smartphone policy on test scores and performance gain.

Table 12 shows that the results above arise because students did not use the paper-based dictionary for too long when asked by teachers to do so, as opposed to what happens with the smartphone. Likely, the paper-based dictionary is not attractive enough and teachers had a hard time to get students to use it.

	Dependent variable:			
	%LT	%DT	%LT	%DT
	(1)	(2)	(3)	(4)
Ta	0.012***	0.728***	0.013***	0.728***
	(0.004)	(0.022)	(0.004)	(0.022)
Ti	0.052***	0.618***	0.052***	0.612***
	(0.004)	(0.022)	(0.004)	(0.022)
Tp	0.006	-0.010	0.004	-0.016
	(0.004)	(0.022)	(0.004)	(0.023)
Constant	-0.015	-0.223***	0.036	-0.040
	(0.009)	(0.053)	(0.032)	(0.184)
Controls	No	No	Yes	Yes
Tp-Ti	-0.046***	-0.628***	-0.049***	-0.628***
Observations	951	951	951	951
Adjusted R2	0.187	0.665	0.214	0.678
Note:	Statistical significance with Bonferroni correction (* p<0.1; ** p<0.05; *** p<0.01)			

Table 12: Effect of smartphone policy on percentage of learning and distraction time on smartphone.

We note that the regressions above include all 4 conditions (compared to the condition when smartphones are banned from the classroom). Therefore, the time spent on the paper-based dictionary in condition Tp is used in lieu of the time spent on the smartphone under Ti. We call this new covariate "time on the tool", the tool being the smartphone under Ti and the paper-based dictionary under Tp. We assume that the time students spent distracted with the paper-based dictionary was zero. This seems a reasonable assumption given that it does not offer the opportunities for distraction that smartphones do.

6.3.2 Teacher Learning Across Experimental Conditions

Condition Ti in our experiments took place after condition Ta, and the teachers used in our experiments could have behaved differently in Ta and Ti, potentially due to learning occurring during Ta. The potential positive effects of such teacher-side learning could confound our results. We asked teachers to teach exactly the same lecture at the same pace to all groups of students in our experiment, but this does not guarantee that they were able to do so. We address this concern by comparing the facial emotions and facial posture of the two teachers used in our second experiment while they were delivering the lectures. The boxplots in Appendix B show that all these covariates are statistically similar on average across the corresponding lectures. Yet, two different teachers could register similar average facial emotions or facial posture and still deliver very different lectures. For example, one teacher could be very happy at the beginning of the lecture and not so happy towards the end, whereas the other teacher could feel very happy towards the end of the lecture but less so at the beginning.

We address this concern by comparing lecture segments. We choose two moments in time during the lecture, at random, and compare the average of the facial emotions and of the facial posture covariates for each teacher between conditions Ta and Ti only between the two moments. We obtained a p-value for each covariate indicating whether it is different, on average, between Ta and Ti during the lecture segment considered. We then bootstrapped this exercise and repeat the analysis 10000 times. We collect 10000 p-values for each facial emotion and facial posture covariate. We then check how many times this p-value is above 0.05. A high percentage of p-values above 0.05 indicates that the covariate is similar for this teacher across lecture segments.

Table 13 shows the results obtained for both teachers used in our second experiment. All percentages are high indicating that these teachers taught similar lectures under condition Ta and condition Ti. Therefore, both teachers taught similar lessons in these two experimental conditions, easing our concerns that there might have been teacher learning helping them deliver a better lecture in condition Ti, an effect that would then carry over to our results.

Covariate	teacher	angry	disgust	fear	happy	sad	surprise	neutral	head pose x	head pose y	dist bet lips	eye gaze x	eye gaze y
Percentage of p-values>0.05	1	99%	100%	99%	87%	98%	61%	86%	100%	100%	96%	100%	99%
	2	100%	100%	100%	84%	97%	60%	87%	100%	100%	98%	100%	98%

Table 13: Percentage of p-values>0.05 when comparing a teacher across conditions Ta and Ti.

7. Analyses of Heterogeneous Effects

Table 14 shows the moderators analyzed. Application Exam Score (AES) is the score that students obtained in the exam they took to apply for schools. This exam covers a significant breath of topics, including math, Chinese verbal and writing, English, and science. We use AES to proxy how good students are in general, which is a strategy commonly used in the education literature (e.g., Belo, Ferreira, and Telang, 2013). Born in rural indicates students that were born in a Rural Hukou. Hukou is a system of household registration used in mainland China. The classification of Hukous into rural and urban is explain at https://www.gov.cn/guoqing/2005-09/13/content_5043917.htm. IT Major indicates students enrolled in a major IT, which include computerized accounting, e-commerce, computer animation and computer graphic design, accounting and information systems, accounting and graphic design, digital design, and computer science. Table 14 shows that these moderators are highly independent of each other and thus pick up different sub-populations of students for each analysis performed below.

Table 14: Correlation table for moderators used for heterogeneous effects.

We start by showing table 15 below that replicates the results for all students across both experiments. This table shows: 1) how conditions Ta and Ti affect performance gain, in rows 1-3, columns 1 and 2 for the first and second experiments, respectively; 2) how %LS and %DS affect performance gain, in rows 4 and 5, columns 1 and 2 for the first and second experiments, respectively; 3) how conditions Ta and Ti affect the percentage of smartphone time allocated to learning and distraction , in columns 3 and 4 for the first experiment, and in columns 5 and 6 for the second experiment. The color code is dark and light shades for statistically significant negative and positive effects, respectively (effects that are statistically not significant have no shade). This color code allows for immediately noting the replication of results from experiment 1 to 2. All full regression results are included in Appendix A.

Population	All											
Experiment	1		2		1		1		2		2	
Dependent Variable	Performance Gain				%LS		%DS		%LS		%DS	
Covariate	(1)		(2)		(3)		(4)		(5)		(6)	
Ta	-5.696	*	-4.592	**	0.009		0.589	***	0.014	***	0.731	***
Ti	5.532	*	8.914	***	0.048	***	0.486	***	0.053	***	0.611	***
Ti-Ta	11.228	***	13.506	***	0.038	***	-0.103	**	0.039	***	-0.120	***
%LS	252.439	***	315.304	***								
%DS	-12.892	**	-11.252	***								

Table 15: Summary of all results obtained during our first experiment.

7.1 Heterogeneous effects on the familiarity with the content lecture

The questions in the pre-test and post-test are the same. Therefore, the score in the pre-test measures the students' prior knowledge of the topic that will be taught during the lecture, and thus captures how familiar students already are with it. Table 16 shows the results obtained for students less and more familiar with the content of the lecture, respectively (i.e., students in the 1st and 4th quartiles of the distribution of pre-test scores (<16 and >34 points)). It is immediate from this table that the results for the students less familiar with the content of the lecture align with the results for all students in both experiments. Also, we can observe that both types of students reallocate the percentage of their learning and distraction time when the smartphone is allowed into the classroom similarly.

Population	Less Familiary with Lecture Content											
Experiment	1		2		1		1		2		2	
Dependent Variable	Performance Gain				%LS		%DS		%LS		%DS	
Covariate	(1)		(2)		(3)		(4)		(5)		(6)	
Ta	-15.725	***	-13.933	***	0.005		0.594	***	0.017		0.747	***
Ti	8.836	**	10.677	***	0.051	***	0.477	***	0.052	***	0.591	***
Ti-Ta	24.561	***	26.144	***	0.046	***	-0.157		0.035	***	-0.156	***
%LS	343.148	***	505.451	***								
%DS	-29.948	***	-22.706	***								
Population	More Familiary with Lecture Content											
Experiment	1		2		1		1		2		2	
Dependent Variable	Performance Gain				%LS		%DS		%LS		%DS	
Covariate	(1)		(2)		(3)		(4)		(5)		(6)	
Ta	-0.297		-2.858		0.011		0.635	***	0.018	*	0.770	***
Ti	6.532	*	8.767	***	0.044	***	0.479	***	0.052	***	0.610	***
Ti-Ta	9.501	**	11.625	***	0.032	**	-0.117		0.034	***	-0.160	***
%LS	194.206	*	257.147	***								
%DS	-9.308		-10.163									

Table 16: Results for students less and more familiar with the content of the lecture.

We also note the following: 1) in both experiments, the performance gain of the students who are less familiar with the content of the lecture increases in Ti and reduces in Ta and, and the performance gain of the students more familiar with the content of the lecture increases in Ti and does not change in Ta; the percentage of learning time allocated to the smartphone increases and percentage of distraction time allocated to the smartphone decreases, from Ta to Ti, for both types of students (although the latter is not statistically significant in the first experiment but similar in magnitude to what happens in the second experiment, where we have more observations and thus more statistical power); in both experiments, the performance gain of both types of students increases with the percentage of learning time allocated to the smartphone in both experiments, but the percentage of distraction time allocated to the smartphone reduces the performance gain of the students less familiar with the content of the lecture and does not change the performance of the students more familiar with the lecture content.

Allowing smartphones into the classroom may put the students that are less familiar with the content of the lecture at a disadvantage, unless teachers ask students to use the device to aid instruction, in which case the increase in performance gain in Ti is similar in magnitude for both types of students. This result does not arise because students fail to use smartphones Ti for learning purposes. Instead, the students who are less familiar with the content of the lecture are at a disadvantage because smartphone distraction hurts their performance gain more compared to the students that are more familiar with the lecture content.

7.2 Heterogeneous effects on prior student performance

According to (Cohen and Levinthal, 1990), absorptive capacity is highly correlated to knowledge accumulated. If using the smartphone to aid instruction increases performance gain, then this effect may be more pronounced for better students. To test this hypothesis, we collected data on the AES for the students in our first experiment and split them into "under-performing" and "high-performing" according to whether they belong in the 1^{st} or 4^{th} quartiles of the distribution of AES. Table 17 shows results for each group of students, separately. Both types of students reallocate the percentage of their learning and distraction time when the smartphone was allowed into the classroom similarly. However, these groups of students register different performance gains in Ti, namely: 1) the students with low AES increase performance gain in Ti, while the students with high AES do not, and the performance gain of both types of students does not change in Ta; 2) both types of students increase the percentage of learning time on the smartphone and reduce the percentage of distraction time on the smartphone from Ta to Ti; 3) the percentage of learning time on the smartphone increases the performance gain of the students with low AES but not that of those with high AES, and the percentage of distraction time on the smartphone does not change the performance gain of both types of students.

Population	Low AES Score					
Experiment	1	2	1	1	2	2
Dependent Variable	Performance Gain		%LS	%DS	%LS	%DS
Covariate	(1)	(2)	(3)	(4)	(5)	(6)
Ta	-0.696		0.005	0.689 ***		
Ti	15.62 ***		0.037 ***	0.546 ***		
Ti-Ta	16.316 ***		0.029 ***	-0.143 *		
%LS	503.271 ***					
%DS	-1.513					
Population	**High AES Score**					
Experiment	1	2	1	1	2	2
Dependent Variable	Performance Gain		%LS	%DS	%LS	%DS
Covariate	(1)	(2)	(3)	(4)	(5)	(6)
Ta	-4.696		0.030 **	0.701 ***		
Ti	-0.94		0.062 ***	0.428 ***		
Ti-Ta	3.756		0.032 **	-0.273 ***		
%LS	65.824					
%DS	-9.15					

Table 17: Results for students with low and high Application Exam Score.

Hence, we find that allowing smartphones into the classroom may help “under-performing” students when the teacher asked them to use the device to aid instruction, which may help close the performance gap. This result does not arise because students fail to use smartphones to learn under condition Ti. Instead, this result arises because the percentage of learning time allocated to the smartphone is significantly more productive for the students with low AES (the average post-quiz score for the students with high AES is 78.81 points and, thus, there is still plenty of room for improvement).

7.3 Heterogeneous effects on IT vs. non-IT majors (familiarity with the technology)

The ability of students to take advantage of ICTs may also introduce heterogeneity in our setting. Table 18 shows our results for students enrolled in IT and non-IT majors, used here to proxy familiarity with the smartphone. In the first experiment, we observe that: 1) the performance gain increases in Ti and reduces in Ta for students in IT majors and it does not change in both Ti and Ta for students in non-IT majors; 2) both types of students increase and decrease the percentage of learning and distraction time allocated to the smartphone from Ta to Ti, respectively (although the change in DS is relatively small and thus not statistically significant for IT majors); 3) the percentage of learning and distraction time allocated to the smartphone increases and decreases the performance gain of the students in IT majors, respectively, and do not change the performance gain of students in non-IT majors.

Population	IT Major											
Experiment	1		2		1		1		2		2	
Dependent Variable	Performance Gain				%LS		%DS		%LS		%DS	
Covariate	(1)		(2)		(3)		(4)		(5)		(6)	
Ta	-11.251	*	-12.111	***	-0.004		0.451	***	0.006		0.626	***
Ti	15.528	***	17.195	***	0.033	***	0.415	***	0.037	***	0.535	***
Ti-Ta	26.779	***	29.306	***	0.037	***	-0.036		0.031	***	-0.090	
%LS	696.455	***	813.614	***								
%DS	-23.728	*	-23.813	***								
Population	non-IT Major											
Experiment	1		2		1		1		2		2	
Dependent Variable	Performance Gain				%LS		%DS		%LS		%DS	
Covariate	(1)		(2)		(3)		(4)		(5)		(6)	
Ta	-2.582		-1.17		0.015		0.653	***	0.016	**	0.772	***
Ti	0.447		4.614	*	0.056	***	0.520	***	0.059	***	0.651	***
Ti-Ta	3.029		5.784	**	0.041	***	-0.132	**	0.044	***	-0.121	***
%LS	62.975		119.164	*								
%DS	-7.379		-4.07									

Table 18: Results for students enrolled in IT and non-IT majors.

Therefore, students in IT-majors are at an advantage when teachers ask students to use smartphones to aid instruction. Conversely, these students are at a disadvantage when smartphones are allowed into the classroom, but teachers do not ask students to use the device to aid instruction. These results do not arise because students fail to use smartphones to learn in Ti. These results arise because students in IT-majors use smartphones more productively. These students are likely more familiar with how to use IT, and smartphones in particular. This may allow them to use these devices more productively for both learning and distraction. Conversely, the students in non-IT majors may have a harder time using these devices for both learning and distraction, resulting in changes in their performance gain in Ta and Ti.

In the second experiment, we observe that: 1) the performance gain increases in Ti for both types of students, it reduces in Ta for students in IT majors, and it does not change for students in non-IT majors in either condition; 2) both types of students increase and decrease the percentage of learning and distraction time allocated to the smartphone from Ta to Ti, respectively (although the change in DS is still not statistically significant for students in IT majors although larger than in the first experiment); 3) the percentage of learning time allocated to the smartphone increases the performance gain of both types of students; the percentage of distraction time allocated to the smartphone decreases the performance gain of students in IT majors and does not change the performance gain of students in non-IT majors.

Thus, note that the performance gain of students in non-IT majors increases from 0.48 points to 4.614 points in condition Ti from the first to the second experiment. Yet, it is still the case that the performance gain of students in IT majors increases much more (3.7x =17.20/4.61). Again, this result does not arise because students fail to use smartphones to learn in Ti. This result arises because the students in non-IT majors doubled their productivity of smartphone learning time from the first to the second experiment.

Our partner school went to remote online education between our first and second experiments due to COVID-19. During COVID-19, all students were required to use IT, in particular laptops, tablets and smartphones, to participate in lectures. This may have propelled students in non-IT majors to become more familiar with how to use IT for learning purposes productively, which would explain why these students started benefiting from using smartphones to aid instruction in 2022. This is a hypothesis that we propose for future research, as we do not have the appropriate data to test it in our context. Also, note that remote online education hit all students in our experiments at the same time, thus it would be hard to identify a causal effect in this setting. In any case, the literature shows that COVID-19 led many students to use e-learning platforms (Wang et al., 2020), which increased their digital acumen (Zhu and Liu, 2020, Al Mazrooei et al., 2022. Still, we reiterate that there is still a significant gap in 2022 in the increase of the performance gain (in condition Ti) between students in IT and non-IT majors.

7.4 Heterogeneous effects on rurality

The research on digital divide shows that students from rural and urban areas have different levels of access to IT and benefit differently from it (Forman et al., 2005, Lythreatis et. al, 2022). Therefore, rurality may also play a role in our findings. To test this hypothesis, we split students according to whether they were born in a rural or urban area. Table 19 shows that in the first experiment, we observe that: 1) the performance gain of the students born in rural areas decreases in Ta and does not change in Ti, and the performance gain of the students born in urban areas does not change in Ta and increases in Ti; 2) both types of students increase and decrease the percentage of their learning and distraction time allocated to the smartphone (although DS is not statistically significant in the first experiment but similar in magnitude to what happens in the second experiment, where we have more observations and thus more statistical power); 3) the percentage of learning and distraction time allocated to the smartphone increases and does not change the performance gain of the students born in urban areas, respectively, and the percentage of learning and distraction time allocated to the smartphone does not change and decreases the performance gain of the students born in rural areas, respectively.

Therefore, in our first experiment, allowing smartphones into the classroom benefited the students born in urban areas at the expense of students born in rural areas both when teachers ask and do not ask students to use the smartphone to aid instruction. As in our previous examples, this result does not arise because students fail to use smartphones to learn in Ti. Instead, this result arises because students born in urban areas use smartphones for learning more productively (and less so for distraction).

Population	Not Born in Rural Area					
Experiment	1	2	1	1	2	2
Dependent Variabler	Performance Gain		%LS	%DS	%LS	%DS
Covariate	(1)	(2)	(3)	(4)	(5)	(6)
Ta	-0.643	-3.183	0.012	0.506 ***	0.018 **	0.775 ***
Ti	13.192 ***	8.329 **	0.048 ***	0.383 ***	0.044 ***	0.605 ***
Ti-Ta	13.835 ***	11.512 ***	0.036 ***	-0.123	0.026 ***	-0.170 ***
%LS	435.104 ***	375.634 **				
%DS	-18.97	-11.322				
Population	**Born in Rural Area**					
Experiment	1	2	1	1	2	2
Dependent Variable	Performance Gain		%LS	%DS	%LS	%DS
Covariate	(1)	(2)	(3)	(4)	(5)	(6)
Ta	-6.968 *	-5.755 **	0.008	0.641 ***	0.011	0.710 ***
Ti	0.279	8.786 ***	0.048 ***	0.553 ***	0.058 ***	0.616 ***
Ti-Ta	7.247 *	14.541 ***	0.040 ***	-0.088	0.047 ***	-0.094 **
%LS	142.576	295.183 ***				
%DS	-12.294 *	-11.533 **				

Table 19: Results for students born in rural and urban areas.

In the second experiment, we observe that: 1) the performance gain of both types of students increases similarly in Ti, and the performance gain reduces and does not change in Ta for the students born in rural and urban areas, respectively; 2) both types of students increase and decrease the percentage of learning and distraction time allocated to the smartphone from Ta to Ti, respectively; 3) the percentage of learning time allocated to the smartphone increases the performance gain of both types of students; the percentage of distraction time allocated to the smartphone decreases the performance gain of students born in rural areas and does not change the performance gain of the students born in urban areas.

Combing all results, we find that allowing smartphones into the classroom may put the students born in rural areas at a disadvantage, unless teachers ask students to use the smartphones to aid instruction, in which case both types of students increase their performance gain similarly. As before, this result does not arise because students fail to use smartphones to learn in Ti. Instead, this result arises because the students born in urban areas increased their productivity associated to using the smartphone for learning purposes from the first to the second experiment. Our results for the heterogeneity in terms or rurality are similar to the ones observed before for students in IT vs. non-IT majors but significantly stronger in magnitude, which leads us to raise a hypothesis for future research similar to the one stated above. In this case, we hypothesize that COVID-19 may have forced people in more rural areas to become better acquainted with IT to communicate and combat isolation and, in what matters to us, to engage in learning activities.
Heterogeneous effects on gender: Research also shows that females and males use IT differently and, thus, gender may also moderate our results (Venkatesh et al., 2000, Venkatesh et al., 2003, Acilar and Sæbø, 2023).

Table 20 presents the results obtained with respect to this potential source of heterogeneity. We can immediately see that the results for males are aligned with the results for our general population of students, but the results for females do not. Namely: 1) in both experiments, the performance gain of male students increases in Ti and reduces in Ta, whereas the performance gain of female students does not change in either condition; 2) in both experiments, both male and female students increase and decrease the percentage of their learning and distraction time allocated to the smartphone (although the latter is not statistically significant in the first experiment but similar in magnitude to what happens in the second experiment, where we have more students and thus more statistical power); in both experiments, the percentage of learning and distraction time allocated to the smartphone increases and reduces the performance gain of male students and it does not change the performance gain of female students.

Population	Male					
Experiment	1	2	1	1	2	2
Dependent Variabler	Performance Gain		%LS	%DS	%LS	%DS
Covariate	(1)	(2)	(3)	(4)	(5)	(6)
Ta	-14.071 ***	-11.766 ***	0.007	0.569 ***	0.015 *	0.715 ***
Ti	14.214 ***	17.572 ***	0.044 ***	0.449 ***	0.048 ***	0.575 ***
Ti-Ta	282.85 ***	29.337 ***	0.037 ***	-0.119	0.033 ***	-0.141 ***
%LS	632.347 ***	766.188 ***				
%DS	-33.287 ***	-31.051 ***				
Population	**Female**					
Experiment	1	2	1	1	2	2
Dependent Variable	Performance Gain		%LS	%DS	%LS	%DS
Covariate	(1)	(2)	(3)	(4)	(5)	(6)
Ta	0.375	1.696	0.011	0.603 ***	0.013	0.704 ***
Ti	0.196	3.821	0.050 ***	0.509 ***	0.056 ***	0.633 ***
Ti-Ta	-0.179	2.125	0.039 ***	-0.095	0.043 ***	-0.108 ***
%LS	22.411	82.272				
%DS	-0.496	-10.431				

Table 20: Results for female and male students.

In sum, we find that allowing smartphones into the classroom may put female students at a disadvantage when teachers ask students to use the smartphones to aid instruction. Conversely, allowing smartphones into the classroom puts male students at a disadvantage when teachers do not ask the students to use the smartphones to aid instruction. However, in this case, the advantage of female students is only that their performance gain does not reduce. As in previous examples, these results do not arise because students fail to use smartphones to learn in Ti. These results arise because male students use smartphones more productively for both learning and distraction.

Furthermore, these results remain unchanged after the COVID-19. We must add that, like all other results that we report in this paper, our findings arise in the specific context that we study. We partner with a vocational school in China and all our experimental lectures covered content in Chinese verbal. In any case, we believe that these results on the heterogeneous effects of gender call for additional research to find ways to extend the benefits of allowing smartphones into classrooms to female students.

8. Discussion and Conclusions

Smartphones are now pervasive. Students are likely to carry one with them at all times and thus take them into the classroom. However, both researchers and practitioners (students, parents, teachers, school managers, and policymakers) are doubtful about the effect of smartphones on educational outcomes. This doubt results in a significant lack of consensus about how to manage smartphones in educational settings.

We ran 2 RCT where smartphones were banned from the classroom in our baseline condition. In the second condition, smartphones were allowed into the classroom and used at will by students. In the third condition, they were allowed into the classroom, used at will by students, and the teacher asked them to use a smartphone to aid instruction. In fourth condition of our second experiment, smartphones were banned from the classroom and teachers asked students to use a paper-based aid was used. We find that smartphones in the classroom change the behavior of students and, consequently, their performance. On average, banning smartphones from the classroom increases their performance compared to allowing them to be used at the students' will. However, the students' performance increases even further if teachers ask students to use the devices to aid instruction. Using the paper-based aid did not change the performance of students. In our case, performance was measured by the difference in the scores obtained in a pre-test and in a post-test taken by students at the beginning and end of lectures in Chinese verbal. Assisted instruction was achieved by having the teacher prompt students to use a smartphone dictionary app (developed for this experiment) in one condition and a paper-based dictionary in another condition during the lecture to check the pronunciation, definition, and etymology of several Chinese words.

The prior research in educational settings has been unable to track organic distraction and learning at the student level productively. Instead, the video feeds of the lectures in our experiments allowed us to measure the time students spent learning and distracted, on and off the smartphone. Using these measurements, we find that the time that students spent distracted and learning was roughly the same across all experimental conditions. Hence, in our setting, the total amount of time students spent distracted vs. learning during lectures does not help predict performance. Our work shows that whether distraction Vs. learning take place with or without the smartphone is what matters for performance. When the teacher asked students to use the smartphone during the lecture to aid instruction, the percentage of time that students spent learning on the smartphone, increased significantly, raising performance. The percentage of the time that students spent distracted with the smartphone also increased but the productivity of smartphone learning outweighs the negative effect of smartphone distraction.

Notably, we find multiple heterogeneous effects. Namely, allowing smartphones into the classroom may: 1) put the students that are less familiar with the content of the lecture at a disadvantage because their smartphone distraction hurts their performance gain more, unless teachers ask students to use the devices to aid instruction; 2) benefit more “under-performing” students when the teacher asks them to use the device to aid instruction, because smartphone learning is much more productive for them, which helps close the performance gap across students; 3) put students in IT-majors at an advantage when teachers ask students to use the smartphones to aid instruction because these students use smartphones more productively for learning purposes, despite the slight catch up by students in non-IT majors after COVID-19 times; 4) have benefited the students not born in rural areas and not those born in rural areas before COVID-19 times because the former students use the smartphone for learning purposes more productively; however, the latter students caught up after COVID-19 when their productivity from using smartphones for learning purposes increased; 5) put female students at a disadvantage when teachers ask the students to use the smartphones to aid instruction because male students use the smartphone more productively for both learning purposes; furthermore, these results remain unchanged after COVID-19.

Our work offers insights for policy making. All our results support allowing smartphones into classrooms as long as: 1) teachers ask students to use the device for learning purposes, even if they spend more time distracted than learning with the device; 2) school principals add programs to help students in non-IT majors and females to take advantage of the smartphone for learning purposes. Hence, our results may encourage stakeholders in education to focus on developing smartphone apps with teachers to improve technology-assisted learning by making sure that the teachers can appropriately manage the learning versus distraction tradeoff introduced by those apps in the classroom. Our paper shows, by example, that one may allow technology, in our case smartphones, into the classroom and, at the same time, increase students' performance. Also, the richness of the heterogeneous effects that we find may support the idea of letting different schools enact different guidelines according to the mix of students that they serve.

Finally, our work suffers from some limitations. First, whether smartphones can help instruction depends on the subject studied and context. Our results are for Chinese verbal and for a particular dictionary app. Results may change based on the subject and the specific smartphone app used. Also, all lectures in our experiments were 90 minutes long, so we cannot evaluate how the results change for shorter (or longer) lectures. However, contextualization and generalization are not mutually exclusive (Cheng et al. 2016). For example, our study features an app that we developed with minimal effort for this experiment. Therefore, it is likely that the effect of smartphone-assisted learning can only become stronger with better apps professionally developed by the education industry. Also, our work is about the effect of smartphones, but other devices, such as tablets, are also increasingly used by youngsters. Thus, learning about their effect on educational outcomes is also essential. While it is hard to generalize across devices, we still believe that our findings may apply without much change to the case of tablets because, in general, tablets and smartphones allow for similar applications and thus provide similar opportunities for both learning and distraction.

Also, we are unable to offer evidence on the long-term effects of smartphones in classrooms on students' performance. Measuring long-term effects would be challenging in our setting. Interference across experimental conditions could easily develop over time, and we cannot track smartphone usage outside the classroom. In any case, it would be interesting to know whether using smartphones in classrooms exhibits any cumulative learning effects over time and whether such effects could carry over to adjacent learning settings (e.g., at home). We encourage future research to develop the appropriate experimental designs and measurement tools to validate our findings more broadly and, hopefully, offer long-term findings on how to enhance student learning using technology.

References

Acemoglu, Daron, David Autor, David Dorn, Gordon H. Hanson, and Brendan Price. 2014. "Return of the Solow Paradox? IT, Productivity, and Employment in US Manufacturing." American Economic Review, 104 (5): 394-99.

Acilar, Ali & Sæbø, Øystein. (2021). Towards Understanding the Gender Digital Divide: A Systematic Literature Review. Global Knowledge Memory and Communication.

Al Mazrooei, A.K., Hatem Almaki, S., Gunda, M. 2022. A systematic review of K–12 education responses to emergency remote teaching during the COVID-19 pandemic. Int Rev Educ 68, 811–841.

Alavi M, Leidner DE (2001) Research Commentary: Technology-Mediated Learning—A Call for Greater Depth and Breadth of Research. *Information Systems Research* 12(1):1–10.

Allen KA (2017) Mobile phones in the classroom: A helpful or harmful hindrance? Psychlopaedia.

Hasler M (2020) Dynamic Attention Behavior Under Return Predictability. *Management Science* 66(7):2906–2928.

Alzahabi, R., & Becker, M. W. (2013). The association between media multitasking, task-switching, and dual-task performance. *Journal of Experimental Psychology: Human Perception and Performance, 39*(5), 1485-1495.

Amez, S., & Baert, S. (2020). Smartphone use and academic performance : a literature review. INTERNATIONAL JOURNAL OF EDUCATIONAL RESEARCH, 103.

Amez, S., Vujić, S., De Marez, L., & Baert, S. (2023). Smartphone use and academic performance: First evidence from longitudinal data. New Media & Society, 25(3), 584–608.

Angrist J, Lavy V (2002) New Evidence on Classroom Computers and Pupil Learning. *The Economic Journal* 112(482):735–765.

Angrist JD, Imbens GW (1995) Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity. *Journal of the American Statistical Association* 90(430):431–442.

Barrera-Osorio F, Linden LL (2009) The use and misuse of computers in education: evidence from a randomized experiment in Colombia. World Bank Policy Research Working Paper 4836.

Barrow L, Markman L, Rouse CE (2009) Technology's Edge: The Educational Benefits of Computer-Aided Instruction. *American Economic Journal: Economic Policy* 1(1):52–74.

Beland LP, Murphy R (2016) Communication: Technology, distraction & student performance. *Labour Economics* 41:61–76.

Belo R, Ferreira P, Telang R (2014) Broadband in School: Impact on Student Performance. *Management Science* 60(2):265–282.

Bjerre-Nielsen, A., Andersen, A., Minor, K., & Lassen, D. D. (2020). The Negative Effect of Smartphone Use on Academic Performance May Be Overestimated: Evidence From a 2-Year Panel Study. Psychological Science, 31(11), 1351–1362.

Carter SP, Greenberg K, Walker MS (2017) The impact of computer usage on academic performance: Evidence from a randomized trial at the United States Military Academy. *Economics of Education Review* 56:118–132.

Chen, Q., & Yan, Z. (2016). Does multitasking with mobile phones affect learning? A review. *Computers in Human Behavior, 54,* 34–42.

Chen, R. S., & Ji, C. H. (2015). Investigating the relationship between thinking style and personal electronic device use and its implications for academic performance. Computers in Human Behavior, 52, 177-183.

Chinchanachokchai, S., Duff, B. R. L., & Sar, S. (2015). The effect of multitasking on time perception, enjoyment, and ad evaluation. *Computers in Human Behavior*, *45*, 185-191.

Cohen WM, Levinthal DA (1990) Absorptive Capacity: A New Perspective on Learning and Innovation. *Administrative Science Quarterly* 35(1):128.

Courage, Mary & Bakhtiar, Aishah & Fitzpatrick, Cheryll & Kenny, Sophie & Brandeau, Katie. (2015). Growing up multitasking: The costs and benefits for cognitive development. Developmental Review. 35.

Duchowski AT (2017) *Eye Tracking Methodology* (Springer International Publishing, Cham).

Faber B, Sanchis-Guarner R, Weinhardt F (2015) *ICT and Education: Evidence from Student Home Addresses*

Fairlie RW, London RA (2012) The effects of home computers on educational outcomes: Evidence from a field experiment with community college students. *The Economic Journal* 122(561):727–753.

Foerde K, Knowlton BJ, Poldrack RA. Modulation of competing memory systems by distraction. Proc Natl Acad Sci U S A. 2006 Aug 1;103(31):11778-83.

Forman, C., Goldfarb, A., & Greenstein, S. (2005). How did location affect adoption of the commercial Internet? Global village vs. urban leadership. Journal of Urban Economics, 58(3), 389-420.

Fried CB (2008) In-class laptop use and its effects on student learning. *Computers & Education* 50(3):906–914.

Ghose A (2017) *Tap: Unlocking the Mobile Economy* (MIT Press).

Goolsbee A, Guryan J (2006) The Impact of Internet Subsidies in Public Schools. *Review of Economics and Statistics* 88(2):336–347.

Hall, T. and Bannon, L. (2006), Designing ubiquitous computing to enhance children's learning in museums. Journal of Computer Assisted Learning, 22: 231-243.

Haller S (2017) Study: 1 in 3 students use cellphones to cheat on exams.

Judd, T. (2014). Making sense of multitasking: The role of Facebook. Computers & Education, 70, 194-202.

Junco, R., & Cotten, S. R. (2012). No A 4 U: The relationship between multitasking and academic performance. *Computers & Education, 59*(2), 505–514.

Kates, Aaron & Wu, Huang & Coryn, Chris. (2018). The effects of mobile phone use on academic performance: A meta-analysis. Computers & Education. 127.

Kraushaar, J.M. & Novak, David. (2010). Examining the Effects of Student Multitasking with Laptops during the Lecture. Journal of Information Systems Education. 21. 241-251.

Krumpal, I. (2013). Determinants of social desirability bias in sensitive surveys: A literature review. *Quality & Quantity: International Journal of Methodology, 47*(4), 2025–2047.

Kumar A, Mehra A (2018) *Remedying Education with Personalized Homework: Evidence from a Randomized Field Experiment in India* (Social Science Research Network, Rochester, NY).

Kuznekoff JH, Munz S, Titsworth S (2015) Mobile Phones in the Classroom: Examining the Effects of Texting, Twitter, and Message Content on Student Learning. *Communication Education* 64(3):344–365.

Lachman R, Lachman JL, Butterfield EC (2015) *Cognitive Psychology and Information Processing: An Introduction*

Lee G, Shao BBM, Vinze A (2018) The Role of ICT as a Double-Edged Sword in Fostering Societal Transformations. *Journal of the Association for Information Systems* 19(3):209–246.

Lee, Jennifer & Lipsmeyer, Lin & Robertson, Tip. (2012). The impact of media multitasking on learning. Learning. Media and Technology. 94-104.

Lepp, A., Barkley, J. E., & Karpinski, A. C. (2014). The relationship between cell phone use, academic performance, anxiety, and satisfaction with life in college students. *Computers in Human Behavior, 31,* 343–350.

Lepp, A., Barkley, J. E., & Karpinski, A. C. (2015). The relationship between cell phone use and academic performance in a sample of US college students. Sage Open, 5(1).

Leuven E, Lindahl M, Oosterbeek H, Webbink D (2007) The Effect of Extra Funding for Disadvantaged Pupils on Achievement. *Review of Economics and Statistics* 89(4):721–736.

Lin, Y., Liu, Y., Fan, W., Tuunainen, V. K., & Deng, S. (2021). Revisiting the relationship between smartphone use and academic performance: A large-scale study. Computers in Human Behavior, 122.

Lythreatis, S., Singh, S. K., & El-Kassar, A-N. (2022). The digital divide: A review and future research agenda. *Technological Forecasting and Social Change*, *175*.

S, McNally S, Silva O (2007) New Technology in Schools: Is There a Payoff?. *The Economic Journal* 117(522):1145–1167.

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist, 38*(1), 43–52.

McDonald, S. (2013). The effects and predictor value of in-class texting behavior on final course grades. College Student Journal, 47(1), 34-40.

Muralidharan K, Singh A, Ganimian AJ (2019) Disrupting Education? Experimental Evidence on Technology-Aided Instruction in India. *American Economic Review* 109(4):1426–1460.

Ophir E, Nass C, Wagner AD. Cognitive control in media multitaskers. Proc Natl Acad Sci U S A. 2009 Sep 15;106(37):15583-7.

Osborne C (2012) Cyberbullying increases in line with mobile phone usage? (infographic). *ZDNet*.

Piccoli G, Ahmad R, Ives B (2001) Web-based virtual learning environments: A research framework and a preliminary assessment of effectiveness in basic IT skills training. *MIS Q.* 25(4):401–426.

Piccoli, G., Ahmad, R., & Ives, B. (2001). Web-Based Virtual Learning Environments: A Research Framework and a Preliminary Assessment of Effectiveness in Basic IT Skills Training. MIS Quarterly, 25(4), 401–426.

Potter RE, Balthazard P (2004) The Role of Individual Memory and Attention Processes during Electronic Brainstorming. *MIS Quarterly* 28(4):621–643.

Rabiu, H., Muhammed, A. I., Umaru, Y., & Ahmed, H. T. (2016). Impact Of Mobile Phone Usage On Academic Performance Among Secondary School Students In Taraba State, Nigeria. European Scientific Journal, 12(1), 466.

Ramjan LM, Salamonson Y, Batt S, Kong A, McGrath B, Richards G, Roach D, Wall P, Crawford R. The negative impact of smartphone usage on nursing students: An integrative literature review. Nurse Educ Today. 2021

Resnick B (2019) Have smartphones really destroyed a generation? We don't know. *Vox*.

Rosen, L. D., Mark Carrier, L., & Cheever, N. A. (2013). Facebook and texting made me do it: Media-induced task-switching while studying. Computers in Human Behavior, 29(3), 948–958.

Rubinstein JS, Meyer DE, Evans JE. Executive control of cognitive processes in task switching. J Exp Psychol Hum Percept Perform. 2001 Aug;27(4):763-97.

Sohlberg, M. M., & Mateer, C. A. (1987). Effectiveness of an attention-training program. *Journal of Clinical and Experimental Neuropsychology, 9*(2), 117–130.

Stock J, Yogo M. Testing for Weak Instruments in Linear IV Regression. In: Andrews DWK Identification and Inference for Econometric Models. New York: Cambridge University Press ; 2005. pp. 80-108.

Taylor & Francis (2015) How do mobile devices in the classroom impact student learning?

Taylor K (2015) Ban on Cellphones in New York City Schools to Be Lifted. The New York Times (January 6)

Uncapher, M. R., & Wagner, A. D. (2018). Minds and brains of media multitaskers: Current findings and future directions. *PNAS, 115*(40), 9889–9896.

UNESCO (2023). Global Education Monitoring Report: Technology in education: A tool on whose terms? Paris

Van Der Schuur, W. A., Baumgartner, S. E., Sumter, S. R., & Valkenburg, P. M. (2015). The consequences of media multitasking for youth: A review. Computers in Human Behavior, 53, 204-215.

Venkatesh, V., & Morris, M. G. (2000). Why don't men ever stop to ask for directions? Gender, social influence, and their role in technology acceptance and usage behavior. MIS Quarterly, 24(1), 115-139.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. MIS Quarterly, 27(3), 425-478.

Vigdor JL, Ladd HF, Martinez E (2014) Scaling the Digital Divide: Home Computer Technology and Student Achievement. *Econ Inq* 52(3):1103–1119.
Wallsten S (2015) What Are We Not Doing When We Are Online? *Economic Analysis of the Digital Economy*.
Wang, C., Pan, R., Wan, X., Tan, Y., Xu, L., McIntyre, R. S., Choo, F. N., Tran, B., Ho, R., Sharma, V. K., & Ho, C. 2020. A longitudinal study on the mental health of general population during the COVID-19 epidemic in China. Brain, behavior, and immunity, 87, 40–48.
Ward AF, Duke K, Gneezy A, Bos MW (2017) Brain Drain: The Mere Presence of One's Own Smartphone Reduces Available Cognitive Capacity. *Journal of the Association for Consumer Research* 2(2):140–154.
Wilmer HH, Sherman LE, Chein JM. Smartphones and Cognition: A Review of Research Exploring the Links between Mobile Technology Habits and Cognitive Functioning. Front Psychol. 2017
Wineburg S, McGrew S, Breakstone J, Ortega T (2016) Evaluating information: The cornerstone of civic online reasoning. *Stanford Digital Repository. Retrieved January* 8:2018.
Wood, E., Zivcakova, L., Gentile, P., Archer, K., De Pasquale, D., & Nosko, A. (2012). Examining the impact of off-task multi-tasking with technology on real-time classroom learning. *Computers & Education, 58*(1), 365–374.
Zhang MW, Ho CS, Ho RC. Methodology of development and students' perceptions of a psychiatry educational smartphone application. Technol Health Care. 2014;22(6):847-55.
Zhang, T., Sun, X., Chai, Y., & Aghajan, H. (2015). A look at task-switching and multi-tasking behaviors: From the perspective of the computer usage among a large number of people. Computers in human behavior, 49, 237-244.
Zhu, X., & Liu, J. 2020. Education in and After Covid-19: Immediate Responses and Long-Term Visions. Postdigital Science and Education, 2(3), 695–699

Appendices

Appendix A: Full Regression Results

Heterogeneous effects on the familiarity with the content of the lecture

1st experiment: students **more** familiar with lecture content

	Dependent variable:					
	Pre-Treatment	Pre-Treatment	Post-Treatment	Post-Treatment	Performance Gain	Performance Gain
	OLS	OLS	OLS	OLS	felm	OLS
	(1)	(2)	(3)	(4)	(5)	(6)
Ta	2.168	2.915	0.596	-0.054	-1.571	-2.970
	(3.708)	(3.664)	(1.989)	(1.928)	(3.652)	(3.877)
Ti	-0.717	-2.179	5.504***	4.352**	6.221*	6.532*
	(3.423)	(3.409)	(1.836)	(1.794)	(3.372)	(3.608)
Constant	54.317***	108.229***	88.585***	108.987***	34.268***	0.757
	(2.476)	(36.803)	(1.328)	(19.369)	(2.439)	(38.947)
Controls	No	Yes	No	Yes	No	Yes
Ti-Ta	-2.885	-5.095	4.907**	4.407*	7.792	9.501**
Observations	119	119	119	119	119	119
Adjusted R2	-0.012	0.169	0.069	0.263	0.030	0.080
Note:	Statistical significance with Bonferroni correction (* p<0.1; ** p<0.05; *** p<0.01)					

1st experiment: students **less** familiar with lecture content

	Dependent variable:					
	Pre-Treatment	Pre-Treatment	Post-Treatment	Post-Treatment	Performance Gain	Performance Gain
	OLS	OLS	OLS	OLS	felm	OLS
	(1)	(2)	(3)	(4)	(5)	(6)
Ta	0.757	0.194	-15.417***	-15.531***	-16.174***	-15.725***
	(1.049)	(1.053)	(4.043)	(3.587)	(4.071)	(3.650)
Ti	-0.146	-0.537	3.769	8.299**	3.914	8.836**
	(1.066)	(1.143)	(4.110)	(3.893)	(4.139)	(3.962)
Constant	7.243***	-12.795	75.622***	89.636**	68.378***	102.431***
	(0.773)	(10.578)	(2.980)	(36.019)	(3.001)	(36.656)
Controls	No	Yes	No	Yes	No	Yes
Ti-Ta	-0.902	-0.731	19.186***	23.83***	20.088	24.561***
Observations	122	122	122	122	122	122
Adjusted R2	-0.009	0.072	0.170	0.404	0.182	0.400
Note:	Statistical significance with Bonferroni correction (* p<0.1; ** p<0.05; *** p<0.01)					

1st experiment: effect of smartphone learning and distraction on performance gain

	Dependent variable:	
	Performance Gain (1)	Performance Gain (2)
% Learning Smartphone (%LS)	210.988*	194.206*
	(113.905)	(118.005)
% Distraction Smartphone (%DS)	-8.147	-9.308
	(9.172)	(8.277)
% Learning (%L)	10.942	-5.575
	(13.863)	(14.921)
% Learning Smartphone (%LS) X Less Familiar	132.160	194.013
	(160.109)	(179.330)
% Distraction Smartphone (%DS) X Less Familiar	-21.801*	-21.554*
	(12.680)	(12.191)
% Learning (%L) X Less Familiar	-4.055	5.668
	(18.148)	(18.382)
Less Familiar	38.132**	30.162*
	(15.867)	(16.872)
Constant	24.747**	47.970
	(12.439)	(39.744)
Controls	No	Yes
% Learning Smartphone for Less Familiar	343.148***	388.219***
% Distraction Smartphone for Less Familiar	-29.948***	-30.862***
Observations	237	237
Adjusted R2	0.110	0.102
Note:	Statistical significance: * p<0.1; ** p<0.05; *** p<0.01	

1st experiment: effect of smartphone policy on smartphone learning and distraction times

	Dependent variable:			
	First-Stage %LS (1)	First-Stage %DS (2)	First-Stage %LS (3)	First-Stage %DS (4)
Ta	0.018	0.612***	0.011	0.635***
	(0.011)	(0.074)	(0.012)	(0.075)
Ti	0.051***	0.516***	0.044***	0.479***
	(0.011)	(0.069)	(0.011)	(0.070)
Ta Less Familiar	-0.011	0.001	-0.007	-0.041
	(0.016)	(0.102)	(0.016)	(0.105)
Ti Less Familiar	0.001	-0.058	0.007	-0.002
	(0.016)	(0.100)	(0.016)	(0.105)
Less Familiar	0.001	-0.003	-0.008	-0.013
	(0.011)	(0.072)	(0.012)	(0.080)
Constant	-0.022	-0.019	-0.065	-0.453
	(0.020)	(0.130)	(0.086)	(0.553)
Controls	No	No	Yes	Yes
Ti-Ta	0.032**	-0.096	0.032**	-0.157
Ta for Less Familiar	0.007	0.613***	0.005	0.594***
Ti for Less Familiar	0.051***	0.458***	0.051***	0.477***
Ti-Ta for Less Familiar	0.044***	-0.156*	0.046***	-0.117
Observations	237	237	237	237
Adjusted R2	0.156	0.400	0.186	0.421
Note:	Statistical significance with Bonferroni correction (* p<0.1; ** p<0.05; *** p<0.01)			

2nd experiment: students **more** familiar with lecture content

	Dependent variable:					
	Pre-Treatment	Post-Treatment	Performance Gain	Pre-Treatment	Post-Treatment	Performance Gain
	(1)	(2)	(3)	(4)	(5)	(6)
Ta	3.935	4.316	1.231	1.459	-2.704	-2.858
	(2.474)	(2.389)	(1.572)	(1.450)	(2.455)	(2.501)
Ti	1.284	-0.113	9.721***	8.654***	8.438***	8.767***
	(2.256)	(2.164)	(1.433)	(1.314)	(2.238)	(2.266)
Constant	48.516***	65.038**	82.230***	84.841***	33.713***	19.803
	(1.615)	(22.736)	(1.026)	(13.803)	(1.602)	(23.805)
Controls	No	Yes	No	Yes	No	Yes
Ti-Ta	-2.651	-4.43	8.49***	7.195***	11.141***	11.625***
Observations	271	271	271	271	271	271
Adjusted R2	0.002	0.178	0.160	0.368	0.078	0.154
Note:	Statistical significance with Bonferroni correction (* p<0.1; ** p<0.05; *** p<0.01)					

2nd experiment: students **less** familiar with lecture content

	Dependent variable:					
	Pre-Treatment	Post-Treatment	Performance Gain	Pre-Treatment	Post-Treatment	Performance Gain
	(1)	(2)	(3)	(4)	(5)	(6)
Ta	0.547	0.559	-13.933***	-14.908***	-14.480***	-15.467***
	(0.680)	(0.708)	(3.007)	(2.655)	(3.049)	(2.734)
Ti	-0.169	-0.105	7.094*	10.572***	7.263*	10.677***
	(0.692)	(0.736)	(3.061)	(2.761)	(3.104)	(2.843)
Constant	4.210***	-3.041	70.009***	93.178***	65.799***	96.219***
	(0.501)	(6.656)	(2.216)	(24.959)	(2.247)	(25.694)
Controls	No	Yes	No	Yes	No	Yes
Ti-Ta	-0.717	-0.663	21.027***	25.48***	21.744***	26.144***
Observations	227	227	227	227	227	227
Adjusted R2	-0.003	0.060	0.186	0.452	0.192	0.440
Note:	Statistical significance with Bonferroni correction (* p<0.1; ** p<0.05; *** p<0.01)					

2nd experiment: effect of smartphone learning and distraction on performance

	Dependent variable:	
	Performance Gain	Performance Gain
	(1)	(2)
% Learning Smartphone (%LS)	270.616***	257.147***
	(87.827)	(92.816)
% Distraction Smartphone (%DS)	-8.713	-10.163
	(6.555)	(6.378)
% Learning (%L)	7.751	-6.801
	(12.084)	(12.869)
% Learning Smartphone (%LS) Under Performing	234.836*	302.752**
	(129.790)	(148.271)
% Distraction Smartphone (%DS) Under Performing	-13.993	-17.909**
	(8.885)	(8.694)
% Learning (%L) X Under Performing	3.178	17.730
	(16.062)	(16.849)
Under Performing	29.179**	19.563
	(13.775)	(14.852)
Constant	25.701**	31.484
	(10.808)	(31.299)
Controls	No	Yes
% Learning Smartphone Under Performing	505.451***	559.899***
% Distraction Smartphone Under Performing	-22.706***	-28.071***
Observations	479	479
Adjusted R2	-0.138	-0.135
Note:	Statistical significance: * p<0.1; ** p<0.05; *** p<0.01	

2nd experiment: effect of smartphone policy on smartphone learning and distraction times

	Dependent variable:			
	First-Stage %LS	First-Stage %DS	First-Stage %LS	First-Stage %DS
	(1)	(2)	(3)	(4)
Ta	0.021**	0.760***	0.018*	0.770***
	(0.008)	(0.044)	(0.008)	(0.045)
Ti	0.057***	0.646***	0.052***	0.610***
	(0.007)	(0.040)	(0.007)	(0.041)
Ta Less Familiar	-0.012	-0.020	-0.001	-0.023
	(0.012)	(0.065)	(0.012)	(0.066)
Ti Less Familiar	-0.013	-0.087	0.0003	-0.019
	(0.011)	(0.063)	(0.011)	(0.066)
Less Familiar	0.002	0.027	-0.011	-0.003
	(0.008)	(0.047)	(0.009)	(0.051)
Constant	-0.029	-0.405***	0.045	-0.177
	(0.015)	(0.083)	(0.057)	(0.329)
Controls	No	No	Yes	Yes
Ti-Ta	0.036***	-0.114**	0.034***	-0.16***
Ta for Less Familiar	0.009	0.74***	0.017	0.747***
Ti for Less Familiar	0.044***	0.559***	0.052***	0.591***
Ti-Ta for Less Familiar	0.035***	-0.181***	0.035***	-0.156***
Observations	479	479	479	479
Adjusted R2	0.167	0.567	0.239	0.580
Note:	Statistical significance with Bonferroni correction (* p<0.1; ** p<0.05; *** p<0.01)			

Heterogeneous effects on prior performance (AES)

1st experiment: **"under-performing"** students

	Dependent variable:					
	Pre-Treatment	Pre-Treatment	Post-Treatment	Post-Treatment	Performance Gain	Performance Gain
	OLS	OLS	OLS	OLS	felm	OLS
	(1)	(2)	(3)	(4)	(5)	(6)
Ta	-2.537	-3.423	-6.050	-4.120	-3.513	-0.696
	(4.039)	(3.912)	(5.212)	(4.872)	(4.935)	(5.094)
Ti	-3.886	-4.138	9.386	11.481*	13.273**	15.620***
	(3.991)	(3.797)	(5.150)	(4.729)	(4.876)	(4.944)
Constant	21.000***	47.067	71.318***	64.021	50.318***	16.954
	(3.259)	(38.504)	(4.205)	(47.948)	(3.981)	(50.131)
Controls	No	Yes	No	Yes	No	Yes
Ti-Ta	-1.35	-0.715	15.436***	15.601***	16.786	16.316***
Observations	107	107	107	107	107	107
Adjusted R2	-0.010	0.165	0.095	0.303	0.135	0.187
Note:	Statistical significance with Bonferroni correction (* p<0.1; ** p<0.05; *** p<0.01)					

1st experiment: **"high-performing"** students

	Dependent variable:					
	Pre-Treatment	Pre-Treatment	Post-Treatment	Post-Treatment	Performance Gain	Performance Gain
	OLS	OLS	OLS	OLS	felm	OLS
	(1)	(2)	(3)	(4)	(5)	(6)
Ta	5.295	5.122	-1.652	0.426	-6.947	-4.696
	(5.608)	(5.660)	(2.844)	(2.541)	(5.155)	(5.052)
Ti	10.734	4.205	5.123	3.265	-5.610	-0.940
	(5.887)	(5.547)	(2.985)	(2.491)	(5.412)	(4.951)
Constant	38.159***	36.480	87.591***	12.400	49.432***	-24.080
	(3.671)	(62.577)	(1.862)	(28.098)	(3.375)	(55.855)
Controls	No	Yes	No	Yes	No	Yes
Ti-Ta	5.438	-0.916	6.775	2.839	1.337	3.756
Observations	105	105	105	105	105	105
Adjusted R2	0.013	0.193	0.027	0.376	0.001	0.230
Note:	Statistical significance with Bonferroni correction (* p<0.1; ** p<0.05; *** p<0.01)					

1st experiment: effect of smartphone learning and distraction on performance

	Dependent variable:	
	Performance Gain (1)	Performance Gain (2)
% Learning Smartphone (%LS)	-23.093	65.824
	(145.675)	(120.799)
% Distraction Smartphone (%DS)	-9.446	-9.150
	(11.101)	(10.455)
% Learning (%L)	-16.136	-20.391
	(18.683)	(17.800)
% Learning Smartphone (%LS) X Under Performing	503.871**	437.447**
	(204.820)	(189.425)
% Distraction Smartphone (%DS) X Under Performing	0.523	7.637
	(15.105)	(13.967)
% Learning (%L) X Under Performing	9.627	3.735
	(24.125)	(23.484)
Under Performing	-7.275	-5.343
	(21.347)	(20.729)
Constant	63.199***	39.629
	(16.668)	(42.287)
Controls	No	Yes
% Learning Smartphone for Under Performing	480.777***	503.271***
% Distraction Smartphone for Under Performing	-8.923	-1.513
Observations	210	210
Adjusted R2	-0.212	-0.078
Note:	Statistical significance: * p<0.1; ** p<0.05; *** p<0.01	

1st experiment: effect of smartphone policy on distraction and learning times

	Dependent variable:			
	First-Stage %LS (1)	First-Stage %DS (2)	First-Stage %LS (3)	First-Stage %DS (4)
Ta	0.030**	0.688***	0.030**	0.701***
	(0.011)	(0.071)	(0.011)	(0.073)
Ti	0.056***	0.429***	0.062***	0.428***
	(0.011)	(0.075)	(0.012)	(0.075)
Ta Under Performing	-0.026	-0.047	-0.025	-0.013
	(0.016)	(0.108)	(0.017)	(0.109)
Ti Under Performing	-0.019	0.090	-0.028	0.118
	(0.017)	(0.110)	(0.017)	(0.109)
Under Performing	0.0002	0.004	0.001	-0.048
	(0.012)	(0.081)	(0.014)	(0.088)
Constant	-0.011	-0.122	0.035	-0.189
	(0.021)	(0.137)	(0.083)	(0.538)
Controls	No	No	Yes	Yes
Ti-Ta	0.026	-0.259***	0.032**	-0.273***
Ta for Under Performing	0.004	0.641***	0.005	0.689***
Ti for Under Performing	0.037***	0.519***	0.037***	0.546***
Ti-Ta for Under Performing	0.033***	-0.122	0.029**	-0.143*
Observations	210	210	210	210
Adjusted R2	0.140	0.439	0.173	0.480
Note:	Statistical significance with Bonferroni correction (* p<0.1; ** p<0.05; *** p<0.01)			

Heterogeneity on IT Vs. non-IT majors (proxy for familiarity with smartphones)

1st experiment: students in non-IT majors

	Dependent variable:					
	Pre-Treatment	Pre-Treatment	Post-Treatment	Post-Treatment	Performance Gain	Performance Gain
	OLS	OLS	OLS	OLS	felm	OLS
	(1)	(2)	(3)	(4)	(5)	(6)
Ta	1.088	-0.440	-1.444	-3.022	-2.532	-2.582
	(3.204)	(3.071)	(2.348)	(2.026)	(2.889)	(2.953)
Ti	3.798	3.152	4.047	3.599	0.249	0.447
	(3.215)	(3.065)	(2.356)	(2.022)	(2.899)	(2.948)
Constant	24.553***	32.440	82.200***	57.568**	57.647***	25.128
	(2.217)	(33.120)	(1.625)	(21.846)	(1.999)	(31.852)
Controls	No	Yes	No	Yes	No	Yes
Ti-Ta	2.71	3.592	5.49*	6.621***	2.781	3.029
Observations	240	240	240	240	240	240
Adjusted R2	-0.002	0.117	0.015	0.297	-0.004	-0.005
Note:	Statistical significance with Bonferroni correction (* p<0.1; ** p<0.05; *** p<0.01)					

1st experiment: students in IT majors

	Dependent variable:					
	Pre-Treatment	Pre-Treatment	Post-Treatment	Post-Treatment	Performance Gain	Performance Gain
	OLS	OLS	OLS	OLS	felm	OLS
	(1)	(2)	(3)	(4)	(5)	(6)
Ta	-9.585	-6.587	-21.803***	-17.837***	-12.218**	-11.251*
	(5.091)	(4.392)	(4.256)	(4.297)	(4.650)	(4.806)
Ti	-7.848	-11.271**	5.684	4.257	13.532**	15.528***
	(4.953)	(3.860)	(4.140)	(3.776)	(4.524)	(4.224)
Constant	40.000***	82.216*	81.925***	90.126**	41.925***	7.911
	(3.622)	(36.763)	(3.028)	(35.968)	(3.308)	(40.232)
Controls	No	Yes	No	Yes	No	Yes
Ti-Ta	1.738	-4.684	27.487***	22.095***	25.749	26.779***
Observations	127	127	127	127	127	127
Adjusted R2	0.016	0.437	0.270	0.428	0.197	0.341
Note:	Statistical significance with Bonferroni correction (* p<0.1; ** p<0.05; *** p<0.01)					

1st experiment: effect of smartphone learning and distraction on student performance

	Dependent variable:	
	Performance Gain (1)	Performance Gain (2)
% Learning Smartphone (%LS)	698.423***	696.455***
	(149.222)	(150.054)
% Distraction Smartphone (%DS)	-27.215**	-23.728*
	(12.344)	(12.633)
% Learning (%L)	-15.484	-12.618
	(15.393)	(16.022)
% Learning Smartphone (%LS) X non-IT Major	-642.586***	-633.480***
	(171.348)	(174.670)
% Distraction Smartphone (%DS) X non-IT Major	21.943	16.350
	(14.082)	(14.582)
% Learning (%L) X non-IT Major	6.956	1.555
	(19.486)	(20.429)
non-IT Major	9.913	16.824
	(17.472)	(20.697)
Constant	55.133***	7.866
	(13.727)	(36.449)
Controls	No	Yes
% Learning Smartphone for non-IT Major	55.837	62.975
% Distraction Smartphone for non-IT Major	-5.273	-7.379
Observations	359	359
Adjusted R2	-0.327	-0.306
Note:	Statistical significance: * p<0.1; ** p<0.05; *** p<0.01	

1st experiment: effect of smartphone policy on smartphone learning and distraction

	Dependent variable:			
	First-Stage %LS (1)	First-Stage %DS (2)	First-Stage %LS (3)	First-Stage %DS (4)
Ta	0.002	0.456***	-0.004	0.451***
	(0.010)	(0.068)	(0.011)	(0.072)
Ti	0.035***	0.414***	0.033***	0.415***
	(0.010)	(0.065)	(0.010)	(0.066)
Ta non-IT Major	0.014	0.202**	0.019	0.202*
	(0.013)	(0.083)	(0.014)	(0.088)
Ti non-IT Major	0.020	0.121	0.022	0.105
	(0.012)	(0.081)	(0.013)	(0.082)
non-IT Major	-0.0004	0.004	-0.010	-0.196
	(0.009)	(0.058)	(0.016)	(0.103)
Constant	-0.011	0.102	-0.034	-0.065
	(0.016)	(0.107)	(0.064)	(0.412)
Controls	No	No	Yes	Yes
Ti-Ta	0.034***	-0.042	0.037***	-0.036
Ta for non-IT Major	0.015	0.658***	0.015	0.653***
Ti for non-IT Major	0.056***	0.535***	0.056***	0.52***
Ti-Ta for non-IT Major	0.041***	-0.123**	0.04***	-0.132**
Observations	359	359	359	359
Adjusted R2	0.171	0.436	0.173	0.452
Note:	Statistical significance with Bonferroni correction (* p<0.1; ** p<0.05; *** p<0.01)			

2nd experiment: students in **non-IT majors**

	Dependent variable:					
	Pre-Treatment (1)	Post-Treatment (2)	Performance Gain (3)	Pre-Treatment (4)	Post-Treatment (5)	Performance Gain (6)
Ta	1.230	-0.785	0.075	-1.954	-1.155	-1.170
	(2.289)	(2.130)	(1.703)	(1.452)	(2.084)	(2.102)
Ti	4.279	2.623	8.288***	7.237***	4.008	4.614*
	(2.297)	(2.133)	(1.709)	(1.454)	(2.091)	(2.105)
Constant	21.914***	30.023	76.168***	63.944***	54.254***	33.920
	(1.598)	(21.221)	(1.189)	(14.467)	(1.455)	(20.943)
Controls	No	Yes	No	Yes	No	Yes
Ti-Ta	3.049	3.407	8.212***	9.191***	5.163**	5.784**
Observations	474	474	474	474	474	474
Adjusted R2	0.003	0.168	0.057	0.339	0.010	0.029
Note:	Statistical significance with Bonferroni correction (* p<0.1; ** p<0.05; *** p<0.01)					

(we acknowledge the lack of power in row 2, column 5, although it the direction of the effect is clear, this effect increased from 0.25 and 0.48 in the first experiment to 4.01 and 4.61 in the second experiment, and, furthermore, it is statistically significant in column 6 with controls)

2nd experiment: students in **IT majors**

	Dependent variable:					
	Pre-Treatment (1)	Post-Treatment (2)	Performance Gain (3)	Pre-Treatment (4)	Post-Treatment (5)	Performance Gain (6)
Ta	-8.053*	-4.056	-20.569***	-16.168***	-12.516***	-12.111***
	(3.511)	(2.967)	(3.020)	(2.910)	(3.319)	(3.343)
Ti	-5.518	-8.843***	9.366***	8.351***	14.884***	17.195***
	(3.412)	(2.633)	(2.935)	(2.582)	(3.226)	(2.967)
Constant	36.192***	44.083	76.341***	53.774*	40.149***	9.691
	(2.453)	(25.415)	(2.110)	(24.922)	(2.319)	(28.635)
Controls	No	Yes	No	Yes	No	Yes
Ti-Ta	2.535	-4.787	29.935***	24.519***	27.4***	29.306***
Observations	260	260	260	260	260	260
Adjusted R2	0.014	0.432	0.284	0.464	0.210	0.354
Note:	Statistical significance with Bonferroni correction (* p<0.1; ** p<0.05; *** p<0.01)					

2^{nd} experiment: effect of smartphone learning and distraction on student performance

	Dependent variable:	
	Performance Gain	Performance Gain
	(1)	(2)
% Learning Smartphone (%LS)	766.596***	813.614***
	(109.474)	(121.690)
% Distraction Smartphone (%DS)	-21.864***	-23.813***
	(6.839)	(7.797)
% Learning (%L)	-13.459	0.469
	(11.241)	(12.325)
% Learning Smartphone (%LS) non-IT Major	-654.003***	-694.449***
	(124.570)	(136.256)
% Distraction Smartphone (%DS) non-IT Major	18.796**	19.743**
	(8.162)	(9.107)
% Learning (%L) X non-IT Major	4.765	-13.095
	(14.383)	(15.506)
non-IT Major	10.855	27.771*
	(12.406)	(15.779)
Constant	50.163***	5.066
	(9.444)	(27.386)
Controls	No	Yes
% Learning Smartphone non-IT Major	112.592.	119.164.
% Distraction Smartphone non-IT Major	-3.068	-4.07
Observations	712	712
Adjusted R2	-0.417	-0.470
Note:	Statistical significance: * p<0.1; ** p<0.05; *** p<0.01	

2^{nd} experiment: effect of smartphone policy on smartphone learning and distraction

	Dependent variable:			
	First-Stage %LS	First-Stage %DS	First-Stage %LS	First-Stage %DS
	(1)	(2)	(3)	(4)
Ta	0.003	0.622***	0.006	0.626***
	(0.008)	(0.043)	(0.008)	(0.045)
Ti	0.036***	0.533***	0.037***	0.535***
	(0.007)	(0.042)	(0.007)	(0.042)
Ta non-IT Major	0.014	0.164***	0.010	0.147**
	(0.009)	(0.053)	(0.010)	(0.055)
Ti non-IT Major	0.025**	0.135**	0.022*	0.116*
	(0.009)	(0.052)	(0.009)	(0.052)
non-IT Major	-0.001	-0.011	-0.020	0.011
	(0.007)	(0.038)	(0.013)	(0.072)
Constant	-0.014	-0.228***	0.069	0.085
	(0.012)	(0.068)	(0.045)	(0.252)
Controls	No	No	Yes	Yes
Ti-Ta	0.033***	-0.088*	0.031***	-0.09
Ta for non-IT Major	0.017***	0.785***	0.016**	0.772***
Ti for non-IT Major	0.061***	0.668***	0.059***	0.651***
Ti-Ta for non-IT Major	0.044***	-0.117***	0.044***	-0.121***
Observations	712	712	712	712
Adjusted R2	0.190	0.584	0.197	0.592
Note:	Statistical significance with Bonferroni correction (* p<0.1; ** p<0.05; *** p<0.01)			

Heterogenous Effects on Rurality

1st experiment: students born in **urban areas**

	Dependent variable:					
	Pre-Treatment	Pre-Treatment	Post-Treatment	Post-Treatment	Performance Gain	Performance Gain
	OLS	OLS	OLS	OLS	felm	OLS
	(1)	(2)	(3)	(4)	(5)	(6)
Ta	-1.868	-0.692	-5.262	-1.336	-3.394	-0.643
	(5.137)	(4.338)	(3.631)	(3.257)	(4.624)	(4.122)
Ti	-9.729	-9.779*	3.208	3.413	12.938**	13.192***
	(4.931)	(4.209)	(3.486)	(3.160)	(4.439)	(3.999)
Constant	35.917***	14.246	83.750***	203.608***	47.833***	189.362***
	(3.487)	(56.555)	(2.465)	(42.464)	(3.139)	(53.731)
Controls	No	Yes	No	Yes	No	Yes
Ti-Ta	-7.861	-9.087	8.471*	4.749	16.332	13.835***
Observations	137	137	137	137	137	137
Adjusted R2	0.017	0.380	0.025	0.307	0.084	0.357
Note:	Statistical significance with Bonferroni correction (* p<0.1; ** p<0.05; *** p<0.01)					

1st experiment: students born in **rural areas**

	Dependent variable:					
	Pre-Treatment	Pre-Treatment	Post-Treatment	Post-Treatment	Performance Gain	Performance Gain
	OLS	OLS	OLS	OLS	felm	OLS
	(1)	(2)	(3)	(4)	(5)	(6)
Ta	-1.763	-3.440	-9.988***	-10.408***	-8.226**	-6.968*
	(3.132)	(2.930)	(2.809)	(2.373)	(3.243)	(3.110)
Ti	6.573	3.523	5.536	3.801	-1.037	0.279
	(3.163)	(2.969)	(2.836)	(2.404)	(3.275)	(3.152)
Constant	25.494***	59.017*	81.091***	61.540**	55.597***	2.523
	(2.222)	(26.986)	(1.992)	(21.849)	(2.301)	(28.644)
Controls	No	Yes	No	Yes	No	Yes
Ti-Ta	8.336**	6.962*	15.524***	14.209***	7.188	7.247*
Observations	230	230	230	230	230	230
Adjusted R2	0.024	0.190	0.113	0.399	0.024	0.148
Note:	Statistical significance with Bonferroni correction (* p<0.1; ** p<0.05; *** p<0.01)					

1st experiment: effect of smartphone learning and distraction on student performance

	Dependent variable:	
	Performance Gain (1)	Performance Gain (2)
% Learning Smartphone (%LS)	127.414	142.576
	(96.583)	(98.264)
% Distraction Smartphone (%DS)	-12.929*	-12.294*
	(7.026)	(6.790)
% Learning (%L)	-1.883	-11.052
	(12.201)	(12.541)
% Learning Smartphone (%LS) X Not Born in Rural	297.906*	292.528*
	(160.816)	(166.494)
% Distraction Smartphone (%DS) X Not Born in Rural	-6.111	-6.675
	(14.225)	(13.825)
% Learning (%L) X Not Born in Rural	-4.175	0.133
	(19.759)	(20.037)
Not Born in Rural	-4.115	-6.398
	(17.558)	(17.719)
Constant	57.248***	46.853
	(10.991)	(36.641)
Controls	No	Yes
% Learning Smartphone for Not Born in Rural	425.32***	435.104***
% Distraction Smartphone for Not Born in Rural	-19.041	-18.97
Observations	359	359
Adjusted R2	-0.405	-0.347
Note:	Statistical significance: * p<0.1; ** p<0.05; *** p<0.01	

1st experiment: effect of smartphone policy on smartphone learning and distraction

	Dependent variable:			
	First-Stage %LS (1)	First-Stage %DS (2)	First-Stage %LS (3)	First-Stage %DS (4)
Ta	0.009	0.642***	0.008	0.641***
	(0.008)	(0.049)	(0.008)	(0.050)
Ti	0.048***	0.567***	0.048***	0.553***
	(0.008)	(0.050)	(0.008)	(0.050)
Ta Not Born in Rural	0.006	-0.144	0.004	-0.135
	(0.013)	(0.081)	(0.013)	(0.082)
Ti Not Born in Rural	-0.001	-0.198**	-0.0003	-0.170
	(0.012)	(0.079)	(0.013)	(0.081)
Not Born in Rural	0.00004	-0.005	0.002	-0.017
	(0.009)	(0.056)	(0.009)	(0.057)
Constant	-0.015	0.054	-0.037	-0.081
	(0.016)	(0.103)	(0.065)	(0.416)
Controls	No	No	Yes	Yes
Ti-Ta	0.04***	-0.076	0.04***	-0.088
Ta for Not Born in Rural	0.014	0.498***	0.012	0.506***
Ti for Not Born in Rural	0.048***	0.369***	0.048***	0.383***
Ti-Ta for Not Born in Rural	0.033***	-0.129	0.036***	-0.123
Observations	359	359	359	359
Adjusted R2	0.154	0.440	0.166	0.451
Note:	Statistical significance with Bonferroni correction (* p<0.1; ** p<0.05; *** p<0.01)			

2nd experiment: results for students born in **rural areas**

	Dependent variable:					
	Pre-Treatment	Post-Treatment	Performance Gain	Pre-Treatment	Post-Treatment	Performance Gain
	(1)	(2)	(3)	(4)	(5)	(6)
Ta	-2.408	-0.058	-4.902	-3.241	-2.494	-3.183
	(3.382)	(3.085)	(2.449)	(2.149)	(3.061)	(2.965)
Ti	-1.107	-2.606	5.718*	5.723**	6.824*	8.329**
	(3.419)	(3.031)	(2.475)	(2.111)	(3.094)	(2.913)
Constant	29.603***	26.733	77.836***	55.368**	48.232***	28.636
	(2.417)	(31.150)	(1.750)	(21.700)	(2.188)	(29.944)
Controls	No	Yes	No	Yes	No	Yes
Ti-Ta	1.301	-2.548	10.619***	8.964***	9.318***	11.512***
Observations	274	274	274	274	274	274
Adjusted R2	-0.005	0.271	0.058	0.368	0.028	0.205

Note: Statistical significance with Bonferroni correction (* p<0.1; ** p<0.05; *** p<0.01)

2nd experiment: results for students born in **urban areas**

	Dependent variable:					
	Pre-Treatment	Post-Treatment	Performance Gain	Pre-Treatment	Post-Treatment	Performance Gain
	(1)	(2)	(3)	(4)	(5)	(6)
Ta	-1.919	-1.409	-8.577***	-7.164***	-6.657**	-5.755**
	(2.402)	(2.096)	(2.100)	(1.788)	(2.470)	(2.305)
Ti	2.244	0.422	10.391***	9.208***	8.147***	8.786***
	(2.353)	(2.040)	(2.057)	(1.740)	(2.420)	(2.243)
Constant	25.263***	25.676	75.322***	71.271***	50.059***	45.595
	(1.653)	(19.622)	(1.445)	(16.741)	(1.700)	(21.576)
Controls	No	Yes	No	Yes	No	Yes
Ti-Ta	4.163	1.831	18.968***	16.372***	14.804***	14.541***
Observations	460	460	460	460	460	460
Adjusted R2	0.002	0.284	0.147	0.417	0.068	0.236

Note: Statistical significance with Bonferroni correction (* p<0.1; ** p<0.05; *** p<0.01)

2^{nd} **experiment:** effect of smartphone learning and distraction on student performance

	Dependent variable:	
	Performance Gain (1)	Performance Gain (2)
% Learning Smartphone (%LS)	127.414	142.576
	(96.583)	(98.264)
% Distraction Smartphone (%DS)	-12.929*	-12.294*
	(7.026)	(6.790)
% Learning (%L)	-1.883	-11.052
	(12.201)	(12.541)
% Learning Smartphone (%LS) X Not Born in Rural	297.906*	292.528*
	(160.816)	(166.494)
% Distraction Smartphone (%DS) X Not Born in Rural	-6.111	-6.675
	(14.225)	(13.825)
% Learning (%L) X Not Born in Rural	-4.175	0.133
	(19.759)	(20.037)
Not Born in Rural	-4.115	-6.398
	(17.558)	(17.719)
Constant	57.248***	46.853
	(10.991)	(36.641)
Controls	No	Yes
% Learning Smartphone for Not Born in Rural	425.32***	435.104***
% Distraction Smartphone for Not Born in Rural	-19.041	-18.97
Observations	359	359
Adjusted R2	-0.405	-0.347
Note:		Statistical significance: * p<0.1; ** p<0.05; *** p<0.01

2^{nd} **experiment:** effect of smartphone policy on smartphone learning and distraction

	Dependent variable:			
	First-Stage %LS (1)	First-Stage %DS (2)	First-Stage %LS (3)	First-Stage %DS (4)
Ta	0.010	0.709***	0.011	0.710***
	(0.006)	(0.032)	(0.006)	(0.032)
Ti	0.057***	0.624***	0.058***	0.616***
	(0.006)	(0.032)	(0.006)	(0.032)
Ta Not Born in Rural	0.005	0.053	0.007	0.065
	(0.009)	(0.053)	(0.009)	(0.052)
Ti Not Born in Rural	-0.014	-0.014	-0.014	-0.011
	(0.009)	(0.053)	(0.009)	(0.052)
Not Born in Rural	0.001	0.010	0.001	0.014
	(0.007)	(0.038)	(0.007)	(0.038)
Constant	-0.017	-0.275***	0.068	0.123
	(0.012)	(0.069)	(0.045)	(0.256)
Controls	No	No	Yes	Yes
Ti-Ta	0.047***	-0.085**	0.047***	-0.094**
Ta for Not Born in Rural	0.015	0.762***	0.018**	0.775***
Ti for Not Born in Rural	0.043***	0.61***	0.044***	0.605***
Ti-Ta for Not Born in Rural	0.028***	-0.152***	0.026***	-0.17***
Observations	712	712	712	712
Adjusted R2	0.175	0.569	0.220	0.588
Note:	Statistical significance with Bonferroni correction (* p<0.1; ** p<0.05; *** p<0.01)			

Heterogeneous Effects on Gender

1st experiment: **male** students

	Dependent variable:					
	Pre-Treatment	Pre-Treatment	Post-Treatment	Post-Treatment	Performance Gain	Performance Gain
	OLS	OLS	OLS	OLS	felm	OLS
	(1)	(2)	(3)	(4)	(5)	(6)
Ta	-1.843	-5.004	-17.882***	-19.075***	-16.039***	-14.071***
	(4.157)	(3.588)	(3.895)	(3.755)	(4.030)	(3.752)
Ti	-7.896	-9.049**	5.129	5.165	13.025***	14.214***
	(4.357)	(3.719)	(4.082)	(3.892)	(4.223)	(3.889)
Constant	25.918***	34.781	75.939***	66.655	50.020***	31.874
	(2.997)	(35.716)	(2.808)	(37.376)	(2.905)	(37.342)
Controls	No	Yes	No	Yes	No	Yes
Ti-Ta	-6.053	-4.045	23.012***	24.24***	29.064	28.285***
Observations	146	146	146	146	146	146
Adjusted R2	0.011	0.334	0.197	0.325	0.248	0.411
Note:	Statistical significance with Bonferroni correction (* p<0.1; ** p<0.05; *** p<0.01)					

1st experiment: **female** students

	Dependent variable:					
	Pre-Treatment	Pre-Treatment	Post-Treatment	Post-Treatment	Performance Gain	Performance Gain
	OLS	OLS	OLS	OLS	felm	OLS
	(1)	(2)	(3)	(4)	(5)	(6)
Ta	-1.939	-0.226	0.075	0.150	2.014	0.375
	(3.552)	(3.399)	(1.851)	(1.785)	(3.267)	(3.187)
Ti	4.514	3.499	3.832*	3.696*	-0.682	0.196
	(3.392)	(3.223)	(1.768)	(1.692)	(3.120)	(3.022)
Constant	31.803***	79.381*	86.092***	122.806***	54.289***	43.425
	(2.422)	(35.788)	(1.262)	(18.791)	(2.227)	(33.556)
Controls	No	Yes	No	Yes	No	Yes
Ti-Ta	6.453	3.725	3.757	3.546	-2.695	-0.179
Observations	221	221	221	221	221	221
Adjusted R2	0.007	0.153	0.018	0.149	-0.006	0.108
Note:	Statistical significance with Bonferroni correction (* p<0.1; ** p<0.05; *** p<0.01)					

1st experiment: effect of smartphone learning and distraction on student performance

	Dependent variable:	
	Performance Gain	Performance Gain
	(1)	(2)
% Learning Smartphone (%LS)	-51.105	22.411
	(97.550)	(93.072)
% Distraction Smartphone (%DS)	4.409	-0.496
	(7.863)	(7.642)
% Learning (%L)	0.330	-7.945
	(14.294)	(14.046)
% Learning Smartphone (%LS) X Male	700.284***	609.937***
	(160.978)	(157.919)
% Distraction Smartphone (%DS) X Male	-42.672***	-32.791**
	(13.008)	(12.725)
% Learning (%L) X Male	-29.232	-15.460
	(20.264)	(20.036)
Male	20.913	10.752
	(18.118)	(17.961)
Constant	53.855***	46.568
	(13.009)	(34.855)
Controls	No	Yes
% Learning Smartphone for Male	649.179***	632.347***
% Distraction Smartphone for Male	-38.262***	-33.287***
Observations	359	359
Adjusted R2	-0.473	-0.338
Note:	Statistical significance: * p<0.1; ** p<0.05; *** p<0.01	

1st experiment: effect of smartphone policy on smartphone learning and distraction

	Dependent variable:			
	First-Stage %LS	First-Stage %DS	First-Stage %LS	First-Stage %DS
	(1)	(2)	(3)	(4)
Ta	0.013	0.630***	0.011	0.603***
	(0.008)	(0.052)	(0.008)	(0.052)
Ti	0.051***	0.517***	0.050***	0.509***
	(0.008)	(0.050)	(0.008)	(0.050)
Ta Male	-0.006	-0.093	-0.004	-0.035
	(0.012)	(0.081)	(0.013)	(0.082)
Ti Male	-0.007	-0.079	-0.006	-0.060
	(0.012)	(0.082)	(0.013)	(0.081)
Male	0.0004	-0.001	0.004	-0.016
	(0.009)	(0.057)	(0.009)	(0.058)
Constant	-0.012	0.101	-0.037	-0.085
	(0.017)	(0.109)	(0.065)	(0.421)
Controls	No	No	Yes	Yes
Ti-Ta	0.037***	-0.113*	0.039***	-0.095
Ta for Male	0.007	0.537***	0.007	0.569***
Ti for Male	0.044***	0.437***	0.044***	0.449***
Ti-Ta for Male	0.036***	-0.1	0.037***	-0.119
Observations	359	359	359	359
Adjusted R2	0.156	0.416	0.166	0.444
Note:	Statistical significance with Bonferroni correction (* p<0.1; ** p<0.05; *** p<0.01)			

2nd **experiment**: **male** students

	Dependent variable:					
	Pre-Treatment	Pre-Treatment	Post-Treatment	Post-Treatment	Performance Gain	Performance Gain
	OLS	OLS	OLS	OLS	felm	OLS
	(1)	(2)	(3)	(4)	(5)	(6)
Ta	-1.843	-5.004	-17.882***	-19.075***	-16.039***	-14.071***
	(4.157)	(3.588)	(3.895)	(3.755)	(4.030)	(3.752)
Ti	-7.896	-9.049**	5.129	5.165	13.025***	14.214***
	(4.357)	(3.719)	(4.082)	(3.892)	(4.223)	(3.889)
Constant	25.918***	34.781	75.939***	66.655	50.020***	31.874
	(2.997)	(35.716)	(2.808)	(37.376)	(2.905)	(37.342)
Controls	No	Yes	No	Yes	No	Yes
Ti-Ta	-6.053	-4.045	23.012***	24.24***	29.064	28.285***
Observations	146	146	146	146	146	146
Adjusted R2	0.011	0.334	0.197	0.325	0.248	0.411
Note:	Statistical significance with Bonferroni correction (* p<0.1; ** p<0.05; *** p<0.01)					

2nd **experiment**: **female** students

	Dependent variable:					
	Pre-Treatment	Post-Treatment	Performance Gain	Pre-Treatment	Post-Treatment	Performance Gain
	(1)	(2)	(3)	(4)	(5)	(6)
Ta	-1.438	-0.542	1.716	1.154	3.154	1.696
	(2.521)	(2.361)	(1.367)	(1.319)	(2.366)	(2.305)
Ti	5.251*	3.313	7.986***	7.134***	2.734	3.821
	(2.408)	(2.210)	(1.306)	(1.235)	(2.260)	(2.159)
Constant	29.105***	41.509	80.215***	88.694***	51.111***	47.184*
	(1.719)	(21.612)	(0.932)	(12.078)	(1.613)	(21.105)
Controls	No	Yes	No	Yes	No	Yes
Ti-Ta	6.689**	3.855	6.27***	5.98***	-0.419	2.125
Observations	442	442	442	442	442	442
Adjusted R2	0.014	0.198	0.082	0.207	0.0004	0.120
Note:	Statistical significance with Bonferroni correction (* p<0.1; ** p<0.05; *** p<0.01)					

2nd experiment: effect of smartphone learning and distraction on student performance

	Dependent variable:	
	Performance Gain	Performance Gain
	(1)	(2)
% Learning Smartphone (%LS)	11.057	82.727
	(66.697)	(65.462)
% Distraction Smartphone (%DS)	5.843	0.620
	(4.952)	(5.024)
% Learning (%L)	-5.179	-10.431
	(10.773)	(11.033)
% Learning Smartphone (%LS) Male	742.709***	683.460***
	(117.993)	(120.510)
% Distraction Smartphone (%DS) Male	-37.766***	-31.671***
	(7.936)	(8.181)
% Learning (%L) X Male	-13.607	-4.566
	(14.734)	(15.165)
Male	8.951	0.803
	(12.907)	(13.285)
Constant	54.022***	35.905
	(9.575)	(26.365)
Controls	No	Yes
% Learning Smartphone Male	753.766***	766.188***
% Distraction Smartphone Male	-31.923***	-31.051***
Observations	712	712
Adjusted R2	-0.532	-0.512
Note:	Statistical significance: * p<0.1; ** p<0.05; *** p<0.01	

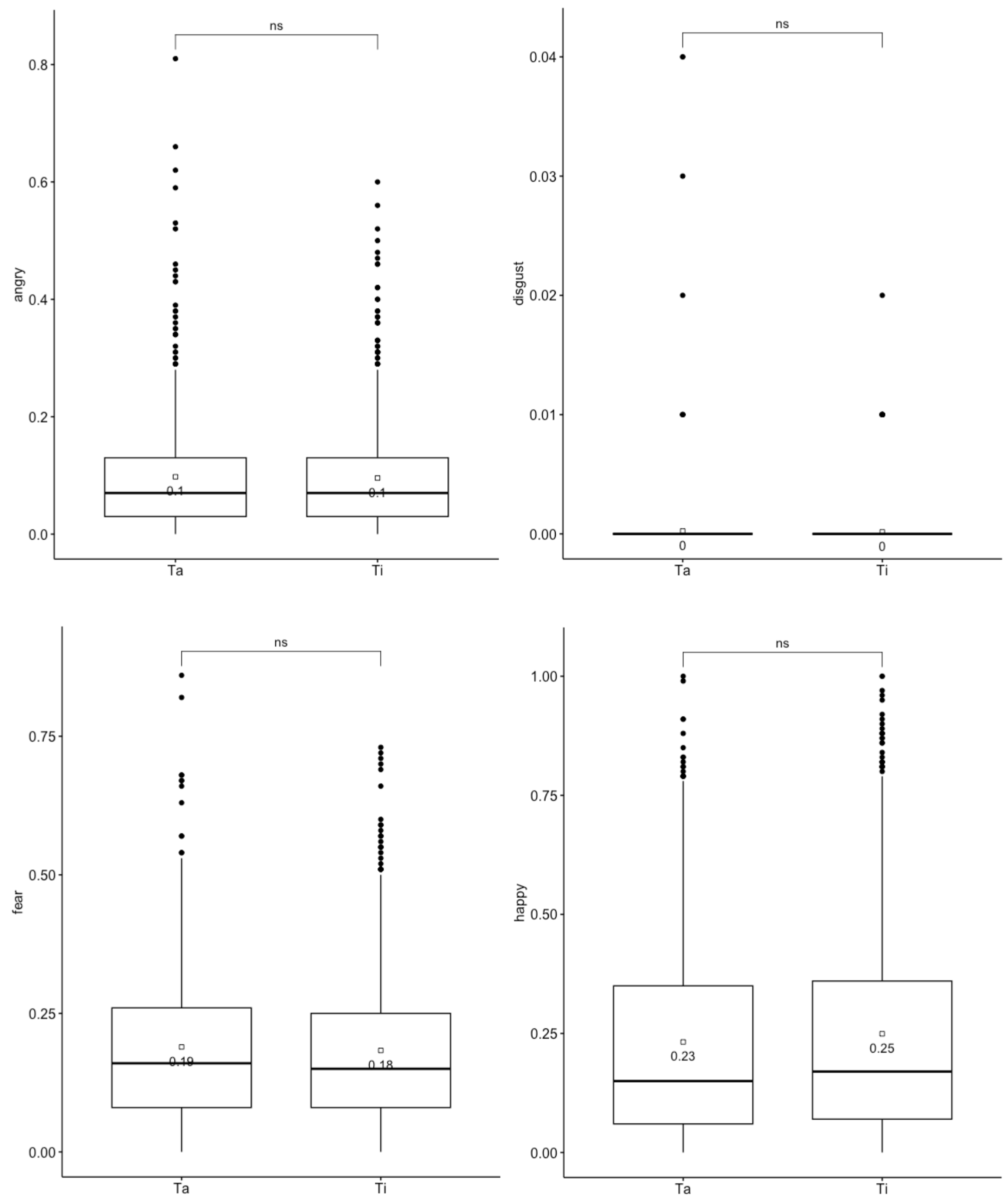
2nd experiment: effect of smartphone policy on smartphone learning and distraction

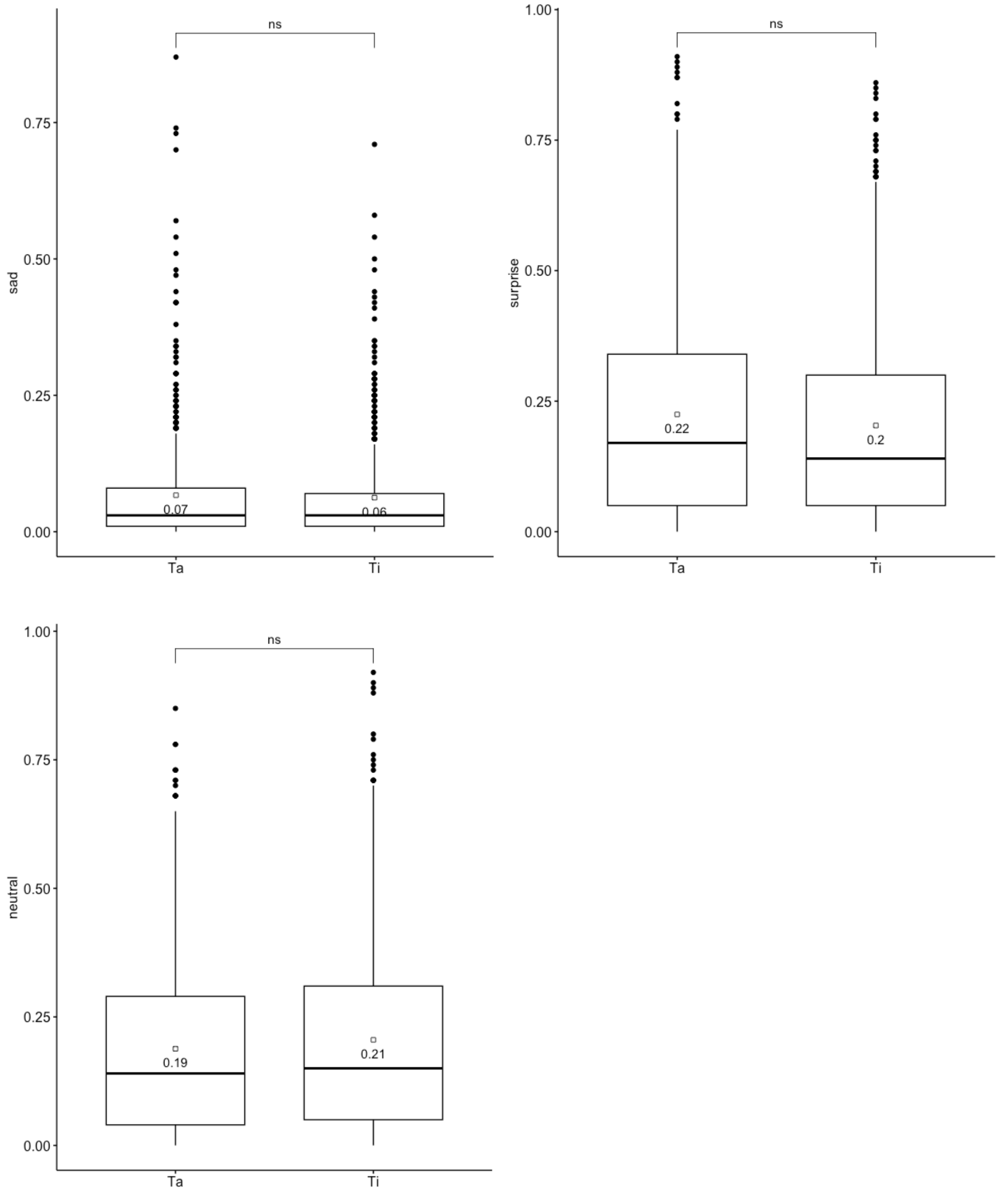
	Dependent variable:			
	First-Stage %LS	First-Stage %DS	First-Stage %LS	First-Stage %DS
	(1)	(2)	(3)	(4)
Ta	0.015**	0.770***	0.013	0.740***
	(0.006)	(0.034)	(0.006)	(0.034)
Ti	0.056***	0.647***	0.056***	0.633***
	(0.006)	(0.032)	(0.006)	(0.032)
Ta Male	-0.007	-0.092	0.002	-0.025
	(0.009)	(0.051)	(0.009)	(0.052)
Ti Male	-0.012	-0.081	-0.008	-0.058
	(0.009)	(0.052)	(0.009)	(0.052)
Male	0.0003	0.004	0.004	0.001
	(0.007)	(0.037)	(0.007)	(0.037)
Constant	-0.014	-0.237***	0.060	0.068
	(0.012)	(0.070)	(0.045)	(0.253)
Controls	No	No	Yes	Yes
Ti-Ta	0.041***	-0.122***	0.043***	-0.108***
Ta for Male	0.008	0.678***	0.015*	0.715***
Ti for Male	0.044***	0.566***	0.048***	0.575***
Ti-Ta for Male	0.035***	-0.112**	0.033***	-0.141***
Observations	712	712	712	712
Adjusted R2	0.174	0.573	0.217	0.592
Note:	Statistical significance with Bonferroni correction (* p<0.1; ** p<0.05; *** p<0.01)			

Appendix B: Teachers' Facial Emotions and Facial Posture Covariates

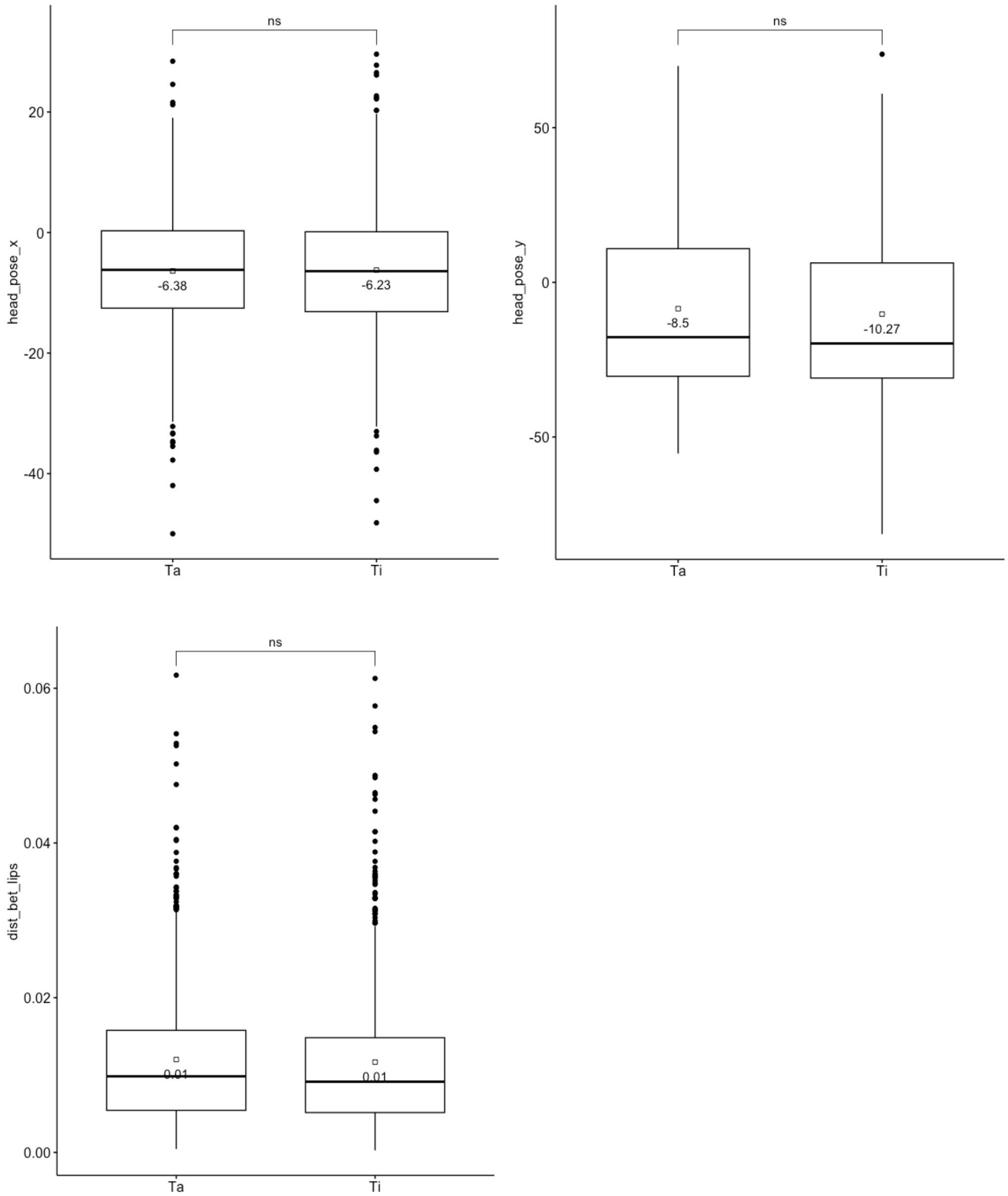
In these boxplots, "ns" stands for a difference that is not statistically significant (according to t-tests of means, stars would appear in lieu of "ns" if the differences were statistically significant).

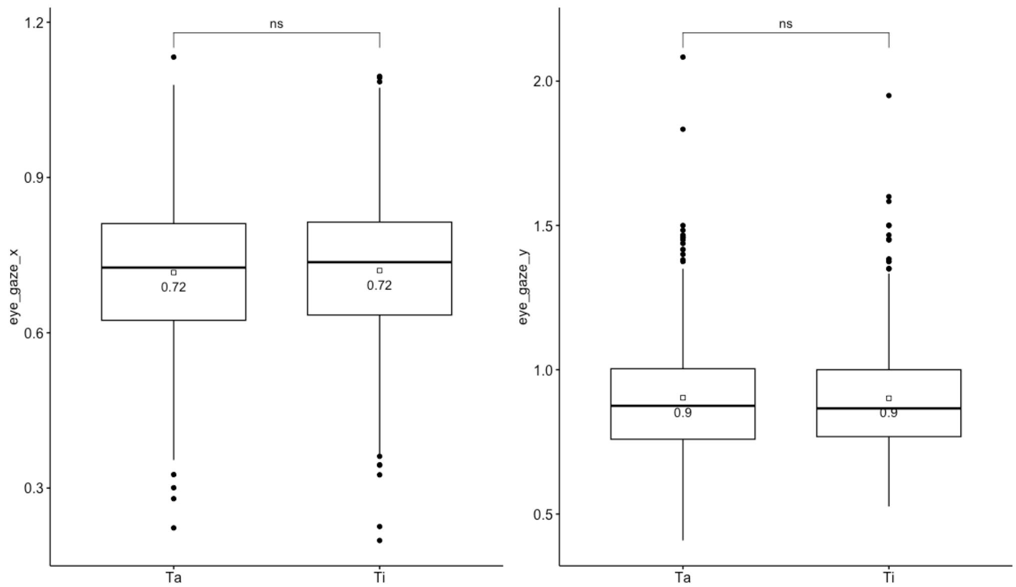
Facial emotions for the teacher used in the first week of our second experiment.



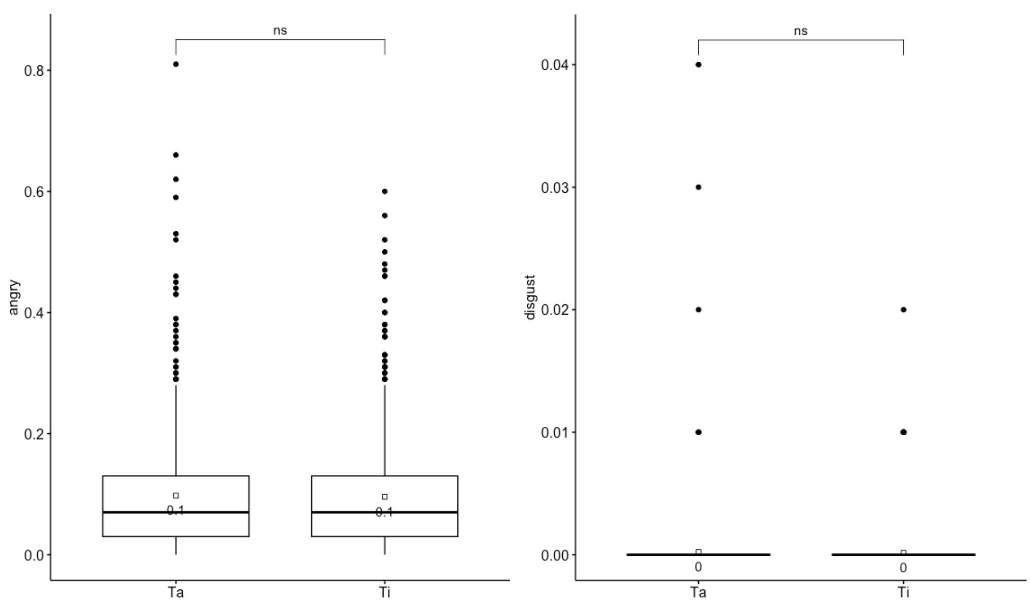

ns
sad
0.75
0.50
0.25
0.00
0.07
0.06
Ta
Ti
ns
surprise
1.00
0.75
0.50
0.25
0.00
0.22
0.2
Ta
Ti
ns
neutral
1.00
0.75
0.50
0.25
0.00
0.19
0.21
Ta
Ti

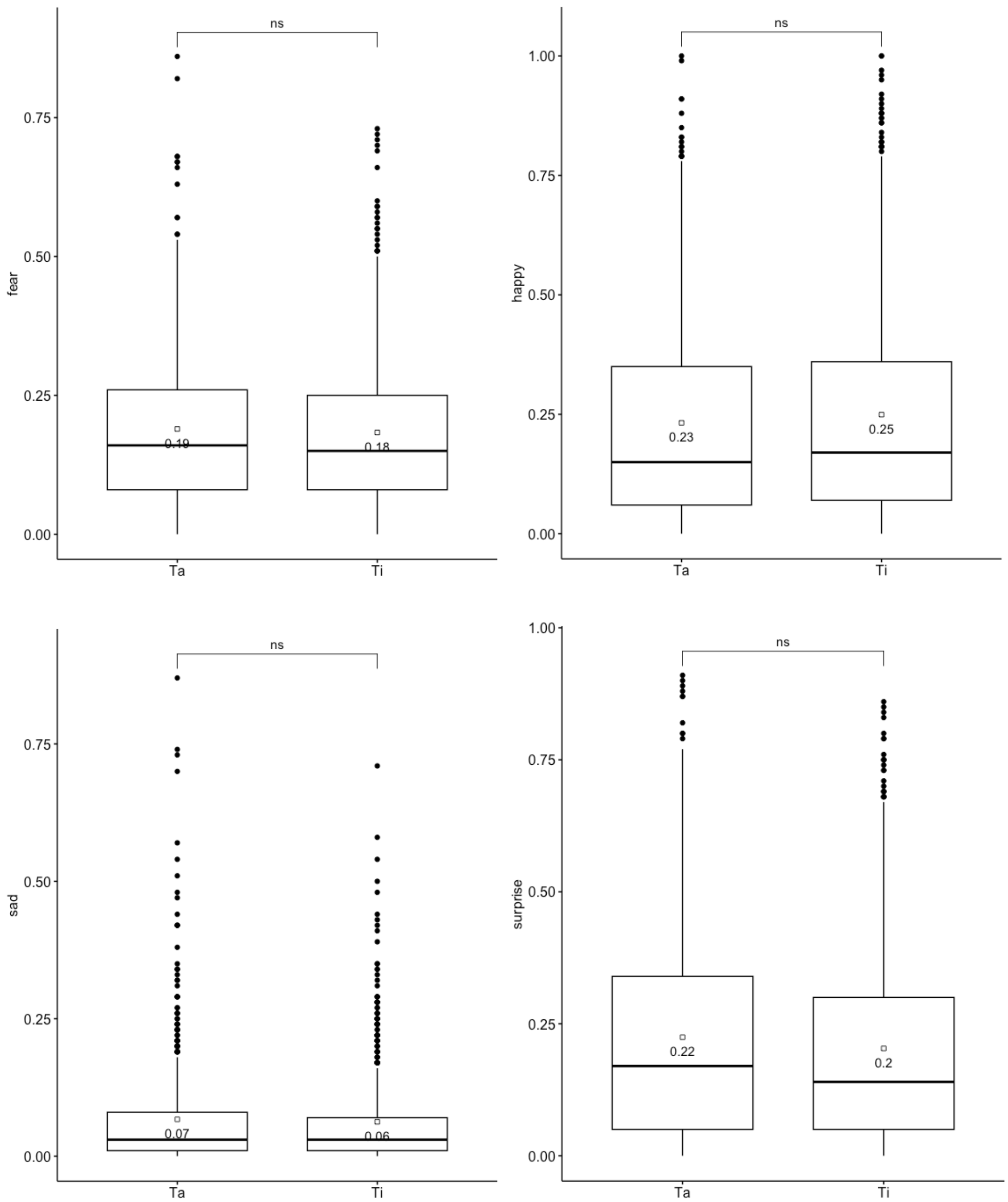
Facial posture covariates for the teacher used in the first week of our second experiment.



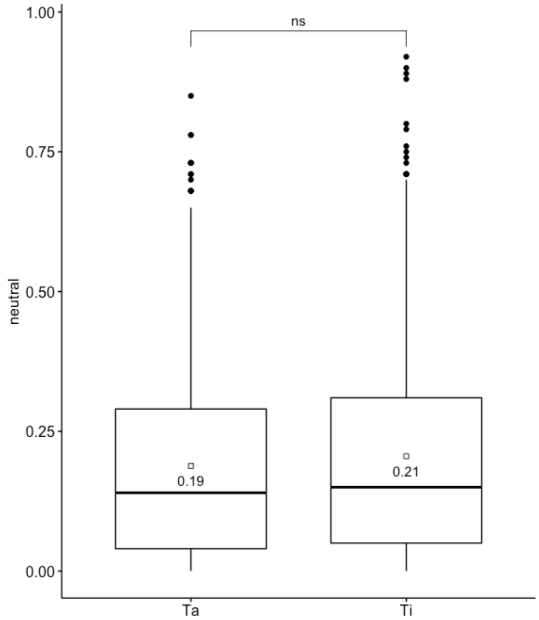


Facial emotions for the teacher used in the second week of our second experiment.

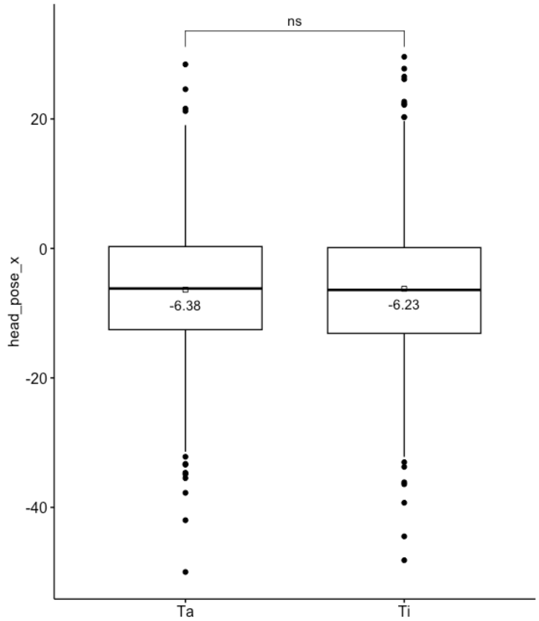


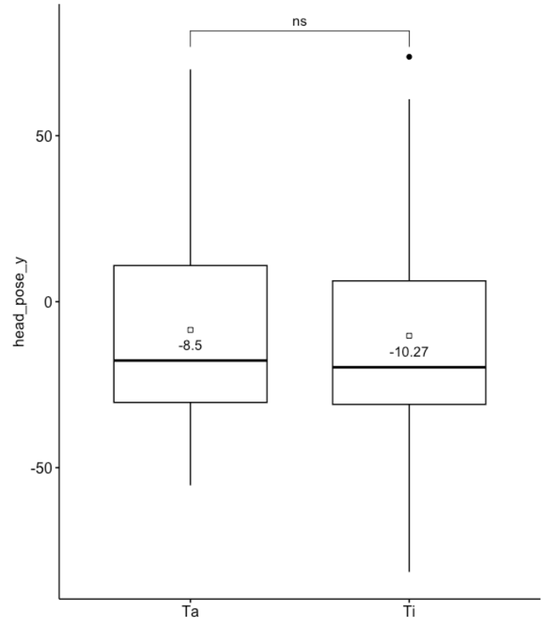


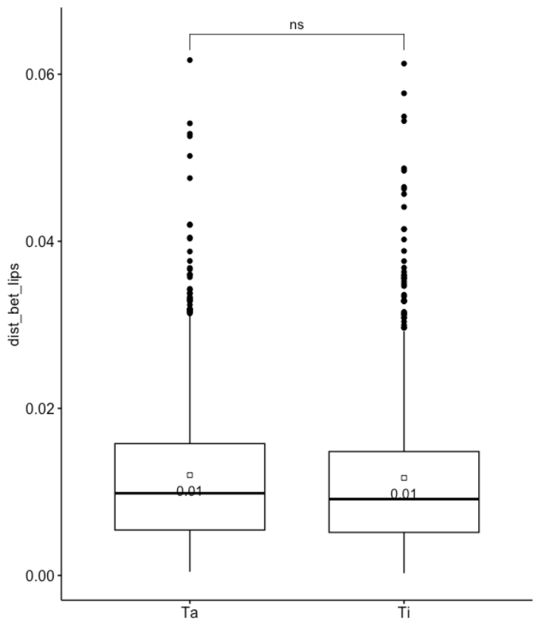
ns
fear
0.00
0.25
0.50
0.75
0.19
0.18
Ta
Ti
ns
happy
0.00
0.25
0.50
0.75
1.00
0.23
0.25
Ta
Ti
ns
sad
0.00
0.25
0.50
0.75
0.07
0.06
Ta
Ti
ns
surprise
0.00
0.25
0.50
0.75
1.00
0.22
0.2
Ta
Ti



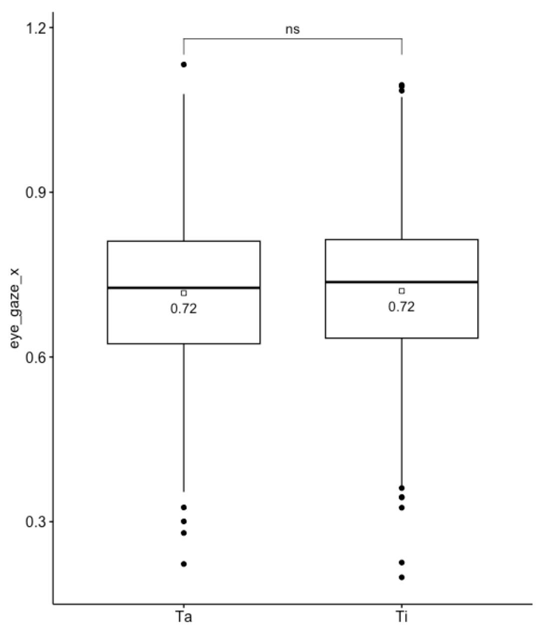
Facial posture covariates for the teacher used in the second week of our second experiment.



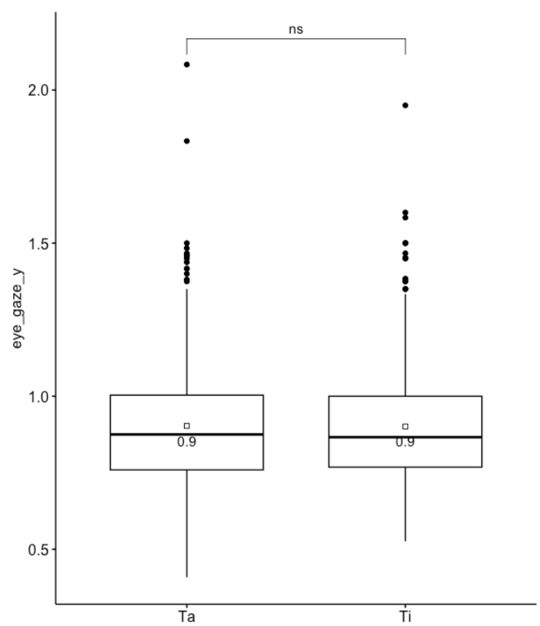




ns
dist_bet_lips
0.06
0.04
0.02
0.00
0.01
0.01
Ta
Ti



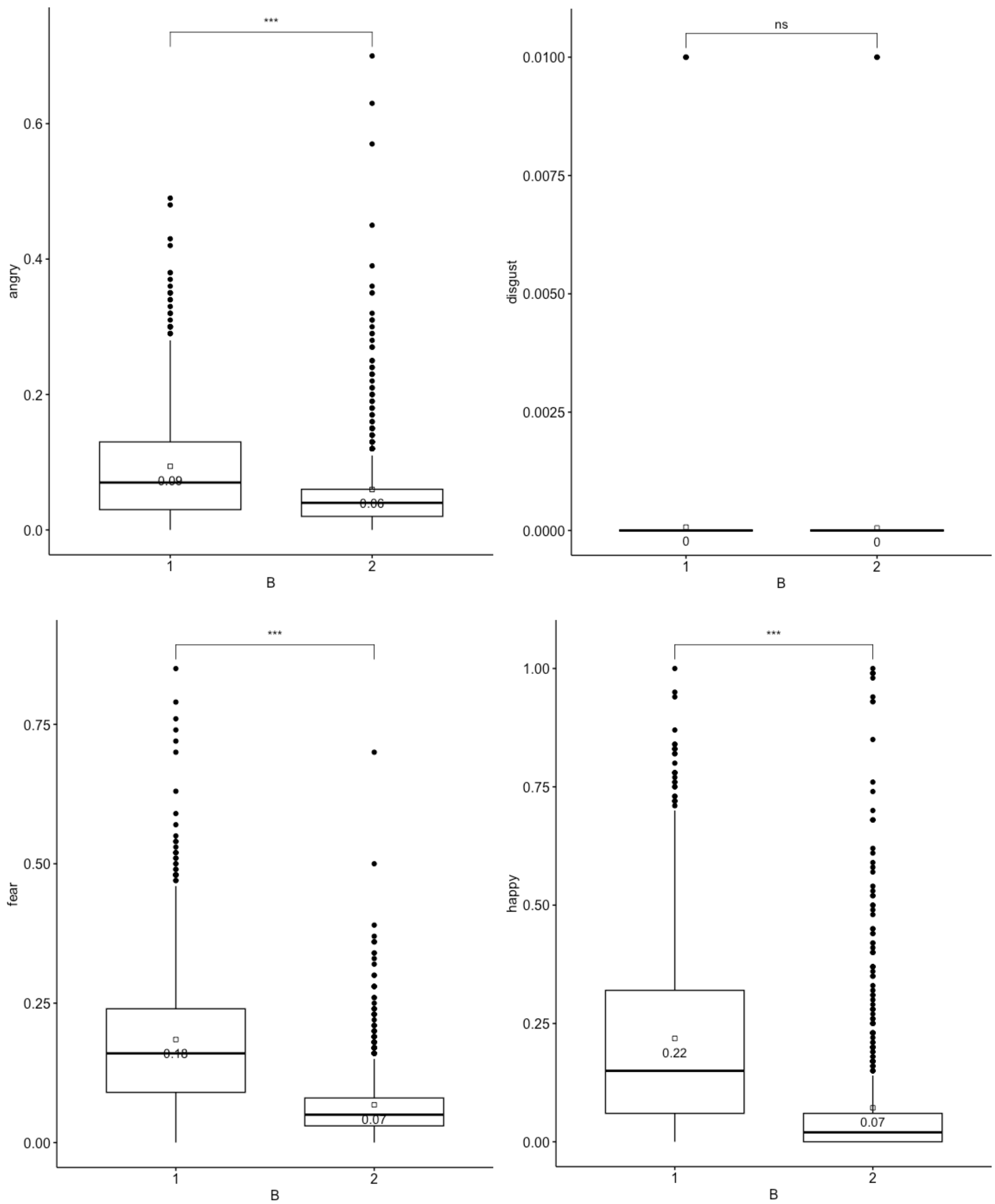
ns
eye_gaze_x
1.2
0.9
0.6
0.3
0.72
0.72
Ta
Ti

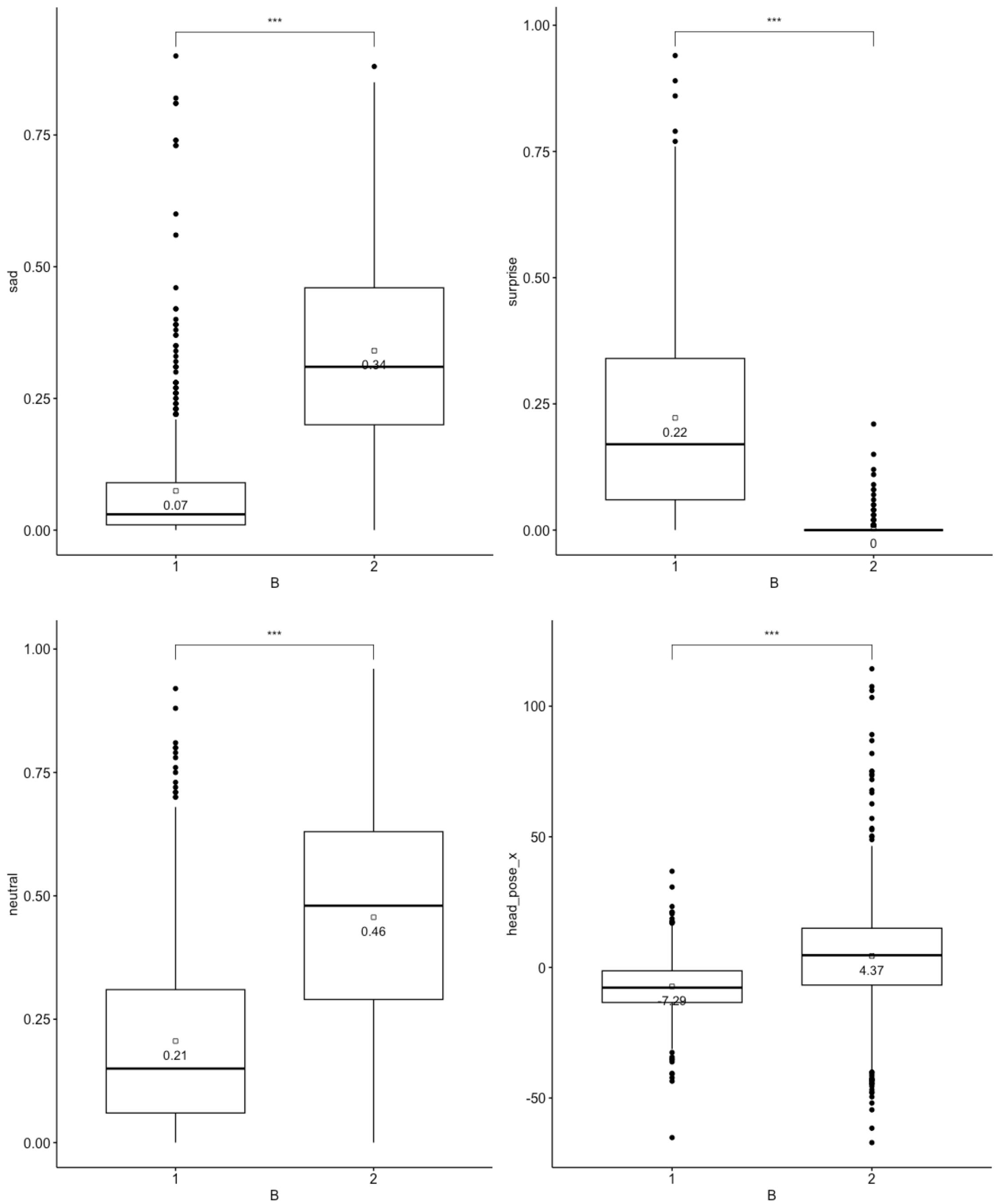


ns
eye_gaze_y
2.0
1.5
1.0
0.5
0.9
0.9
Ta
Ti

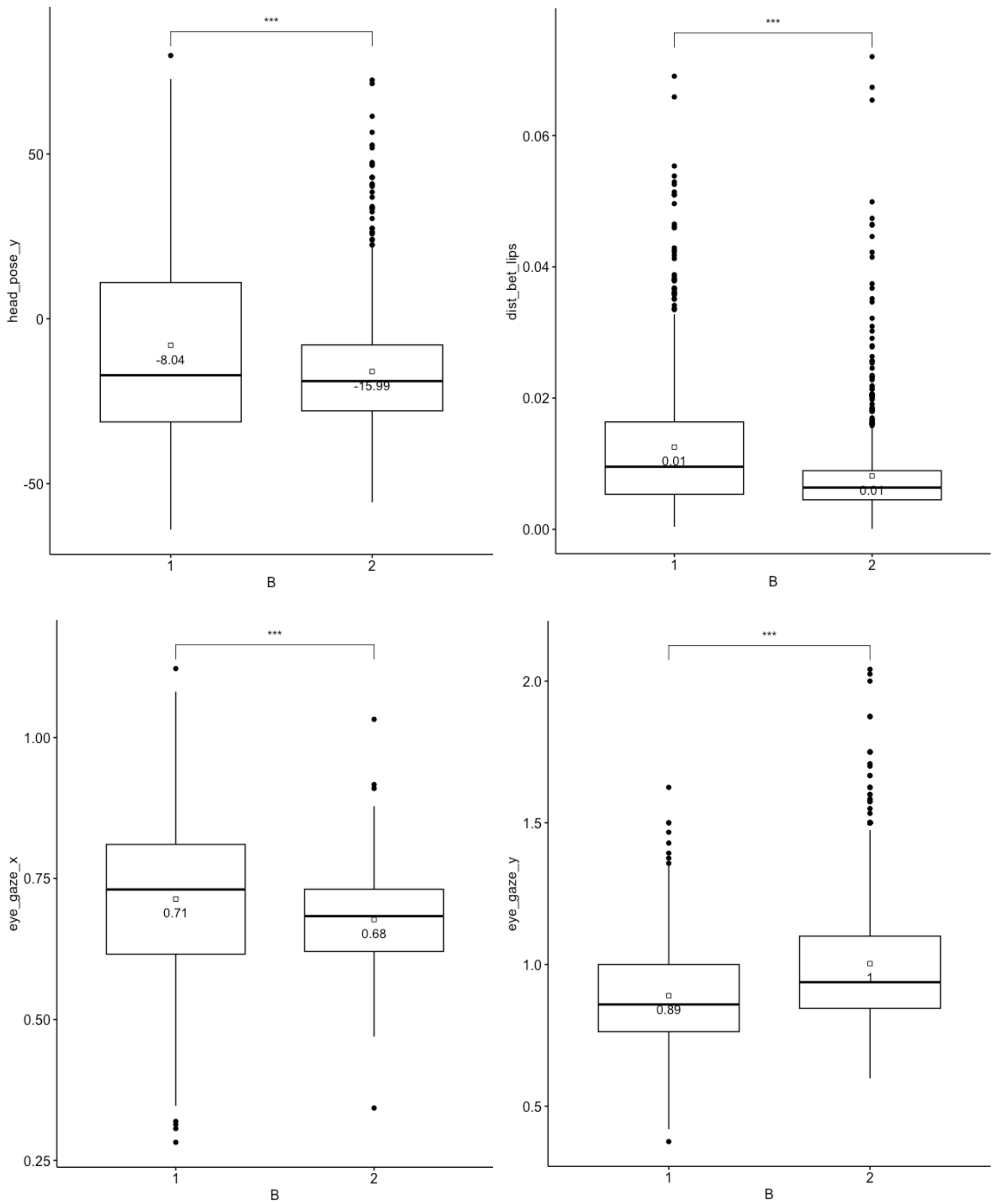
Comparison of Facial Emotions and Facial Posture Covariates in the Second Experiment

Comparing teacher 1 and teacher 2, in the horizontal axis:





***
sad
0.00
0.25
0.50
0.75
0.07
0.34
1
2
B
***
surprise
0.00
0.25
0.50
0.75
1.00
0.22
0
1
2
B
***
neutral
0.00
0.25
0.50
0.75
1.00
0.21
0.46
1
2
B
***
head_pose_x
-50
0
50
100
-7.29
4.37
1
2
B



Statistics when comparing 10000 synced lecture segments across the 2 teachers:

Covariate	angry	disgust	fear	happy	sad	surprise	neutral	head pose x	head pose y	dist bet lips	eye gaze x	eye gaze y
p-values>0.05	3%	100%	0%	0%	0%	0%	0%	1%	35%	3%	16%	22%

(disgust is the same across teachers because there is none presented in the lectures)

Appendix C. Snapshots of the educational aids offered as part of our experiments

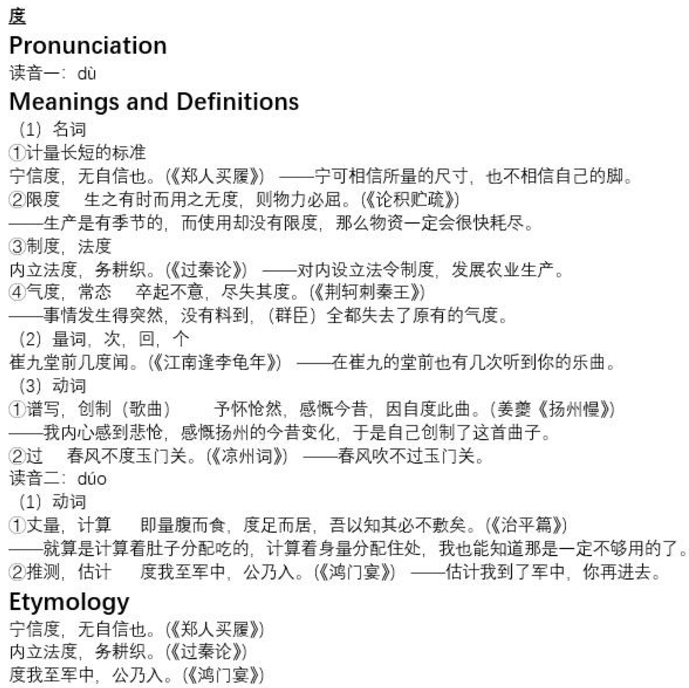


Figure 2. Screenshot of the dictionary app and a picture of the paper-based dictionary.

Appendix D. IRB Protocol with Delayed Debriefing Scripts

Our experiment was conducted after approval by the Institutional Review Board of a US-based University (de-identified here for review purposes) as presenting no more than minimal risk to participants because:

1) Students in the school that partnered with us knew of the video cameras in the classroom; these cameras were routinely used in this school for anti-cheating purposes;

2) After our experiment, students were told that their lectures were recorded and that the corresponding video feeds would be used to measure the time they spent distracted or learning, on and off the smartphone;

3) Students were also told that collecting these data would be done by faculty and staff at the school and that researchers would only get access to de-identified individual-level data. Furthermore, the video feeds would not be shared with researchers and would only be kept at the local school;

4) Students could opt out and withdraw their data from our research study. In this case, the observations of such a student would be deleted from our de-identified datasets. To date, no student or family required their data to be withdrawn;

5) Parents and guardians of students under 18 were also included in the communications referred to above. The results of our study were shared with school administrators, teachers, staff, students, and their families and were well received in all cases;

Below, we provide a translation of the script used for the delayed debriefing:
Thank you for participating in this lecture! We were trying to run an experiment to determine which policy can benefit you the most in terms of your learning performance. Your participation is greatly appreciated! You were subject to 1 of 3 smartphone use policies: banning smartphones from the classroom, allowing smartphones into the classroom, and allowing smartphones into the classroom and using them to assist with instruction. We needed to collect data on the time that each student allocated to mobile learning, mobile distraction, other mobile-unrelated learning, and other mobile-unrelated distraction. To achieve this, we captured your behavior using the anti-cheating video system in the large classroom when you took the Chinese verbal class.

Please note that we did not disclose the experiment setting when you took the lecture only to obtain the necessary unconditioned response. The scores you got in the tests you answered during this lecture will not count towards your final grade. All your data will be stored in the Teaching Department of the school. Only de-identified data will be used for further analysis. Please note that we are only interested in investigating the effect of the three policies above, and our goal is to identify the best policy for further classroom protocol. Therefore, we will not disclose the data to any of your teachers or advisors. Likewise, you will not be judged by any school faculty or staff for the behaviors recorded in the data. The school thanks you for your participation in helping us formulate a better classroom policy!

Now that you know the purpose of our study and are fully informed, you may decide that you do not want your data used in this research. If you would like your data removed from the study and permanently deleted, please contact me and leave me with your name and student ID. The data administrator in the Teaching Department will delete the corresponding data before any further analysis starts.

Please do not disclose any information about our research procedures and/or hypotheses to anyone who might participate in this study in the future, as this could affect the study results. If you have any questions or concerns regarding this study, its purpose or procedures, or if you have a research-related question, please feel free to contact [insert name(s), email address(es), and phone number(s)]. If you feel upset after completing the study or find that some questions or aspects of this study trigger distress, talking with a qualified clinician may help. If you feel you would like assistance, please contact our school clinic for psychological/mental health services at [email address(es) and phone number(s)].