To Honor or Dishonor Student Choices? The Impact of Self-Regulation on Instructional Methods and Learning Outcomes

Gillian Gold (gilliang@andrew.cmu.edu)
Michael W. Asher (masher@andrew.cmu.edu)
Paulo F. Carvalho (pcarvalh@cs.cmu.edu)
Human-Computer Interaction Institute, Carnegie Mellon University
5000 Forbes Avenue, Pittsburgh, PA 15213

Abstract

Self-regulated learning involves making decisions about how to study, but students often choose suboptimal strategies. Our experiment investigated how preferences for instructional methods—video, practice, or both combined—affect learning outcomes. We randomly assigned 130 participants to either receive their preferred method of instruction (honoring initial choice) or a different method (dishonoring initial choice). Contrary to previous research showing preferences for lectures, our participants initially selected practice-based approaches. However, when asked again after instruction, the majority of participants chose the combined approach. This shift in preferences suggests that students may overvalue comprehensive approaches, even when practice alone was equally effective and reduced instructional time by 66%. Whether preferences were honored or dishonored did not significantly affect performance or efficiency, thus control over instructional methods may be less important than the methods themselves. Based on our findings, future research should focus on guiding students to utilize practice to optimize learning efficiency.

Keywords: self-regulated learning; instructional methods; practice with feedback; learning efficiency

Introduction

Consider a scenario in which two students are preparing for the same exam. The first student is given the autonomy to choose how they prepare: they may choose to watch videos, engage in practice problems, or use a combination of both. The second student is instructed on how to prepare, being told exactly which method to follow. How will the degree of control they have over their learning process affect their choices and outcomes? Will the first student, with the freedom to choose, select a method that optimizes their learning? Conversely, does the second student's lack of choice undermine their ability to self-regulate, potentially leading to a less effective learning experience? These questions aim to uncover whether control in self-regulated decisions is an essential part of effective learning.

Prior Research

Self-regulation is a key component of effective learning, influencing how students choose to study, what instructional methods they rely on, and how they manage their time (Bjork et al., 2013). Students' decisions are often influenced by self-efficacy, or their belief in their own ability to succeed in specific tasks. Confident students tend to engage with more challenging material, set ambitious goals, and invest more effort, whereas students with low self-efficacy might avoid challenging tasks or stick to methods that do not require as much effort, even if they are less effective (Schunk, 1990). These choices can have significant implications for academic success (Zimmerman, 1990).

Self-regulated learning involves processes like planning, monitoring, and evaluating one's own learning activities. However, evidence shows that students often fail to make choices that align with the most effective learning strategies (Dunlosky & Rawson, 2012; Hartwig & Dunlosky, 2012).

One key area where self-regulation comes into play is in selecting between different instructional methods, such as the choice between active practice versus passive lecture. Testing (Roediger & Karpicke, 2006) and active learning (Freeman et al., 2014) have been shown to be effective additions to passive instruction, and most instructors rely on a combination of the two approaches (Capone, 2022). However, students often express a preference for lectures and avoid opportunities for active practice (Deslauriers et al., 2019; Dunlosky & Rawson, 2012).

Further, in a series of studies, Asher et al. (2025) found that participants who engaged in practice-based learning with feedback performed just as effectively and learned 15% more efficiently than those who also received the combination of lecture and practice. Nonetheless, when asked at the end of the study which form of instruction would be best to prepare for a test, participants preferred video, selecting either lecture or a combination of lecture and practice.

The lack of calibration to accurately judge the most effective instructional methods raises several important questions about the interplay of instructional methods and students' preferences.

First, it raises questions about the impact of dishonoring students' preferences on learning outcomes. A large body of research points to the importance of respecting students' autonomy (Ryan & Deci, 2000) and allowing them to make choices about their educational experiences (Patall et al., 2008). If an instructor ignores students' preferences for lecture and chooses to rely exclusively on active learning, the literature on practice testing might suggest this will lead to better learning outcomes. On the other hand, does the act of denying students' choices undermine their motivation and ability to self-regulate? This issue is not merely a theoretical concern; it also has practical implications for the allocation

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*

of instructional time and ultimately students' academic performance.

Second, it is important to examine how different forms of instruction might actually change a student's preferences. In prior research on instructional preferences, a common approach is to assess students' preferences for instructional methods after they have engaged in some instruction (Carvalho et al., 2017; Deslauriers et al., 2019). However, it is possible that these preferences might differ from what students would have chosen before instruction. A successful practice session might be able to convince a student that practice alone can be helpful. Or, when assigned a lecture, students might infer that this is the best way to learn. Thus, it is important to explore whether pre-instruction preferences differ from post-instruction preferences.

To test these questions, we conducted an experiment with an honor/dishonor paradigm, in which students were either assigned to their preferred study method (honored) or assigned to a method that differs from their preference (dishonored). This experimental approach allowed us to measure how honoring or dishonoring preferences affects learning outcomes and self-efficacy (Tullis et al., 2018). We also investigated preferences prior to any instruction, the effect of baseline self-efficacy on these choices, and how these preferences change following instruction.

If students' post-instruction choices reflect a general understanding of learning processes, we would predict that learners in our study will be more likely to initially select instructional methods incorporating a video, as suggested by Asher et al. (2025). However, if learners' post-instruction preferences are influenced by recent experiences, we might see them select methods incorporating more practice. Further, if control over instructional methods results in better learning outcomes, learners whose preferences are honored should show increased efficiency and higher test performance. Conversely, if granting control does not influence learning outcomes, then assigned instructional methods, rather than being honored or dishonored, will predict learning outcomes. Given that prior research suggests students are not well-calibrated when assessing the effectiveness of their own learning strategies (Deslauriers et al., 2019; Dunlosky & Rawson, 2012), we would expect that our learners choose the same method as they do prior to instruction.

By exploring the role of self-regulation, this study will provide valuable insights into how students make decisions about instructional strategies and the effectiveness of practice with feedback in different contexts.

Methods

Participants

A total of 146 participants were recruited through Prolific. The sample size was determined to detect effects $f > .3$ with 80% power, which we determined based on prior research with the same materials (Asher et al., 2025). Participants were randomly assigned to either the honor (N = 78) or dishonor (N = 68) regulation conditions. The experiment took about 30 minutes, and participants were paid $4.80.

We excluded any participant that did not finish the experiment (8 participants in each condition). The final sample consisted of 130 total participants, including both the honor (N = 70) and dishonor condition (N = 60).

Procedure

At the start of the study, all participants were told that they would receive instruction followed by a final test. Next, they endorsed three statements about their self-efficacy in statistics, arts, and science (e.g., "I'm good at statistics"), on a scale from 0 (strongly disagree) to 100 (strongly agree). All participants were then asked to select how they would like to prepare for a final test on statistics: (1) by watching a short video, (2) by watching a short video and completing practice activities with feedback, or (3) by completing practice activities with feedback.

Next, all participants proceeded to a lesson about measures of central tendency that used one of these three methods of instruction, and they were randomly assigned to either have their preference "honored" or "dishonored." Participants in the honor condition engaged in their selected instructional method (video, practice, or combined), while participants in the dishonor condition were randomly assigned a different method than their preference.

After the lesson, all participants completed a series of trivia questions. Participants proceeded to the testing portion of the study, which consisted of an assessment about statistics. After the test, the same three questions about self-efficacy were presented again. Finally, participants were reminded of the instruction they had just received and then asked which form of instruction would have prepared them best for the test (video, practice, or a combination of both).

Materials

The materials used in this study were the same as those used in previous work by Asher et al. (2025). We used a video from the YouTube channel CrashCourse, a creator of high-quality educational videos since 2006, about measures of central tendency that took 11 minutes and 22 seconds. We developed a total of 40 questions about the video. 20 questions were about information that was presented in the video (i.e., in the video at timestamp 1:36 the presenter says, "The mean, or average, takes the sum of all the numbers in a data set, and divides by the number of data points" and we created the corresponding question: "How do you calculate the mean of a dataset?"). The remaining 20 questions were generalization questions, questions that did not directly ask about information covered in the video, but could be responded based on that information. For example, based on the same statement from the presenter at timestamp 1:36, a generalization question would be: "Student A has 5 pencils, Student B has 2 pencils, Student C has 2 pencils. What is the mean number of pencils these students have?"

All questions were four-option multiple-choice. The 20 questions of each type were divided into two groups. The groups assigned as practice and final test questions were randomly selected for each participant. Thus, overall, there were 20 practice questions and 20 final test questions. We originally intended to counterbalance the order of the tests, but due to a coding error, participants who received practice were given the same test at both times. Correct response feedback was presented during practice, but not to the final questions.

All participants started with the same set of instructions telling them that they would receive instruction followed by a final test. The final test for all participants was about measures of central tendency (as described above). The instruction varied depending on the participant's condition. In the practice and combined conditions, practice questions were 4-item multiple choice, presented one at a time. The order of the questions and the options were randomized across participants. In the video condition, instruction did not include any practice questions and instead started directly with the video on Measures of Central Tendency. A timer tracked the total duration that participants spent on instruction (M = 13.32 minutes, SD = 6.72 minutes).

After completing the instruction, participants completed a trivia game with a series of trivia facts for five minutes, creating a delay between the instruction and posttest. The trivia questions were developed and normed by Tauber et al. (2013). After the trivia game, participants completed the final test. During the final test, participants answered each of the 20 multiple-choice questions one at a time in random order. No feedback was presented. At the end of the final test participants were thanked for their participation and sent back to Prolific for payment.

Predictions

We make the following predictions:

1. Before instruction, participants will predominantly prefer video-based instructional methods.
2. Participants whose preferences are honored will demonstrate higher test performance.
3. Practice-only instruction will be the most efficient, regardless of whether participants' preferences were accommodated, consistent with prior research.
4. Given that students' preferences are often misaligned with effective learning strategies, participants will maintain their initial instructional preferences even after instruction.

Results

Based on total time in the task, two outliers were identified, such as one participant who spent 42 minutes on the instruction portion, which was likely indicative of inactivity. To mitigate the influence of these outliers on the analysis, we winsorized the time data. This was done by setting a conservative cutoff for extreme values, defined as "Q3 + 3*IQR" (the third quartile plus three times the interquartile range). Any time values above this threshold were capped at the cutoff value, and this adjusted measure of time was used in the main analysis.

All analyses reported here were pre-registered at https://osf.io/wj7ym, unless otherwise noted.

Initial Preference for Practice-Based Instruction

Students' initial preferences for instructional methods to prepare for the test were not evenly distributed, $X^2(df = 2) = 24.66, p < .001$, Figure 1.

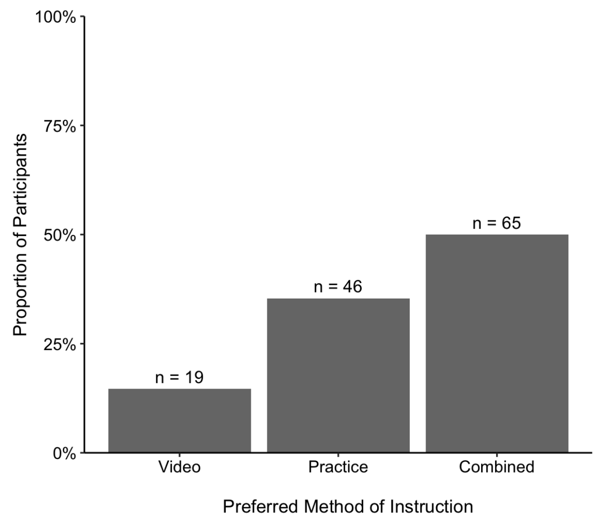

Figure 1: Initial Preferences for Instructional Method

To compare the frequencies of different preferences, we conducted post-hoc pairwise comparisons using Chi-square tests, correcting for multiple comparisons using the Bonferroni method, between each pair of instructional method categories. In comparison to video, participants preferred practice (p < .001) and combined (p = .002) instruction. There was no significant difference in preferences for combined versus practice (p = .214).

In an exploratory analysis, we further investigated if learners' preferences were related to their self-efficacy beliefs. We hypothesized that participants with higher baseline statistics self-efficacy would be more likely to select practice-only because they would feel more equipped to answer questions independently without the instructional support of a video. To test this hypothesis, we coded the instructional methods: video as 0, combined as 0, and practice as 1. We then conducted a logistic regression to assess if baseline self-efficacy for statistical ability predicted whether participants selected the practice-only condition.

Contrary to our prediction, baseline statistics self-efficacy was not related to a preference for practice questions. To investigate if this null result occurred because of flaws with our one-item measure of self-efficacy, we correlated baseline self-efficacy for statistics and arts with final test performance. Under the assumption that if we are successfully measuring baseline self-efficacy, statistics self-efficacy should be positively related to final test performance, whereas arts self-efficacy should not. As predicted, the correlation between baseline statistics self-efficacy and final test scores was moderate and statistically significant ($r = 0.35$, $t(128) = 4.28$, $p < .001$). In contrast, the correlation between baseline arts self-efficacy and final test performance was negative and not statistically significant ($r = -0.05$, $t(128) = -0.59$, $p = .558$). Thus, our measure of self-efficacy appeared to be a valid measure of students' self-perceived ability in statistics, even though it was not a significant predictor of instructional method preference.

Regulation condition and assigned instructional method were not significant predictors of participants' statistics self-efficacy after the final test (*p*s > .05).

No Differences in Test Performance by Instruction

We used a linear mixed effects model to analyze final test performance. Specifically, we regressed each students' performance on knowledge and application questions on the following predictors: regulation condition (coded as 0 for preferences honored and 1 for preferences dishonored), assigned instructional method (dummy coded with combined as the reference group), and baseline statistics self-efficacy (standardized). We also included an interaction term between regulation condition and instructional method to explore whether the effect of preference status on learning outcomes varied depending on the instructional method. In addition, we included a contrast to differentiate between application and knowledge subscores on the final test (coded as -0.5 for knowledge and 0.5 for application), along with a by-participant random slope for the application vs. knowledge contrast to account for individual variability in performance. We preregistered that we would use practice test scores as another predictor, but this would require excluding all participants in the video condition, so it was not used.

Compared to those who received combined instruction, participants who completed practice-based instruction ($b = 0.27$, $t(123) = 0.94$, $p = .351$) or video instruction ($b = -0.33$, $t(123) = -1.40$, $p = .165$) did not significantly differ in performance. There were no significant differences in final test performance based on the contrast between application and knowledge subscores ($b = -0.13$, $t(129) = -1.67$, $p = .098$) or regulation condition ($b = 0.39$, $t(123) = 1.75$, $p = .083$).

The interaction between regulation condition (honor vs. dishonor) and instructional method was significant for the video condition compared to the combined condition ($b = -0.71$, $t(123) = -2.04$, $p = .043$). Participants who were dishonored and assigned to the video condition performed better than those who were honored, but the pattern was reversed for the combined condition, Figure 2. That is, while being assigned videos when activities or combined was chosen resulted in better learning than deciding to learn with videos, being assigned to videos or practice when a student chose combined resulted in worse learning. One possibility is that students with higher knowledge chose effective activities and benefited even from less effective methods (videos). Thus, it is possible that calibration is correlated with prior knowledge, such that students with less knowledge may have preferred less efficient approaches like videos. In contrast, participants with more knowledge were more likely to choose practice and perform well regardless of their assigned instructional method.

Finally, as mentioned before, higher baseline statistics self-efficacy was significantly associated with better performance on the final test ($b = 0.29$, $t(123) = 4.12$, $p < .001$).

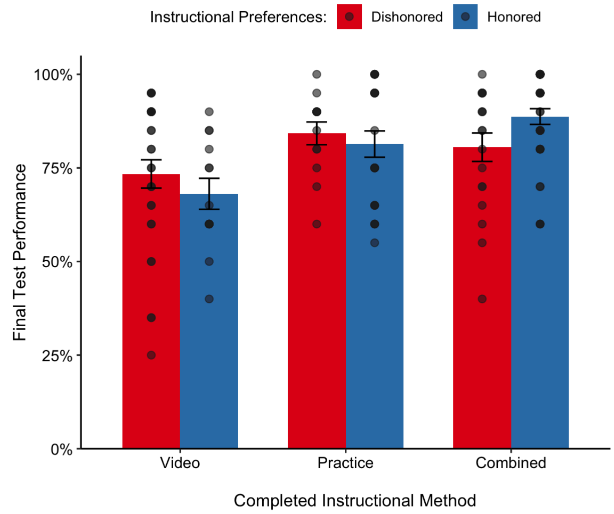

Figure 2: Final Test Performance by Assigned Instruction and Regulation Condition

Practice was the Most Efficient Instruction

To investigate to which degree combining practice and video is an efficient learning approach, we compared how long it took participants to complete instruction. We preregistered a measure of efficiency that divided final test performance by total instructional time. We wanted to separate the two measures, so efficiency is solely measured by instructional time. However, our conclusions do not change, regardless of the efficiency calculation.

We conducted a regression analysis predicting instructional time. We included the following predictor variables: regulation condition (coded as 0 for preferences honored and 1 for preferences dishonored), assigned instructional method (dummy coded with combined as the reference group), and baseline statistics self-efficacy (standardized). We included an interaction term between regulation condition and instructional method. Final test score was preregistered as a predictor, but, because the test happens after the instruction, we removed it.

Participants in the video condition spent significantly less time on instruction compared to those in the combined condition (an average of 7.84 minutes or a 40.83% time decrease; $b = -1.06$, $t(123) = -5.62$, $p < .001$). Furthermore, participants in the practice condition spent significantly less time on instruction compared to the combined condition (an average of 12.6 minutes or a 65.63% time decrease; $b = -1.75$, $t(123) = -7.67$, $p < .001$), Figure 3. None of the other predictors significantly predicted time spent in instruction.

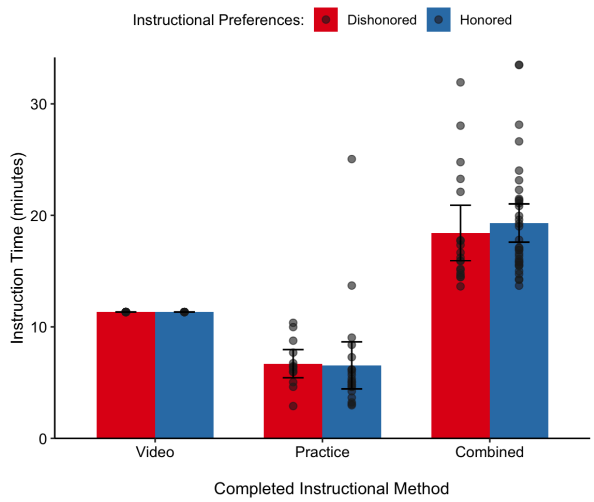

Figure 3: Efficiency of by Assigned Instruction and Regulation Condition

Preferences Shifted to Combined Instruction

We hypothesized that participants' preferences for instructional methods after the final test would be consistent with their initial preferences. We preregistered an analysis to focus on changes in preference (shift vs. no shift) based on honor and dishonor conditions. There was no significant difference between the conditions ($p > .05$), so we decided to explore the post-instruction distribution of preferences. We conducted a Chi-square test for selected instructional methods after the final test among participants. The results of the Chi-square test were significant, $X^2(df = 2) = 98.28$, $p < .001$, Figure 4.

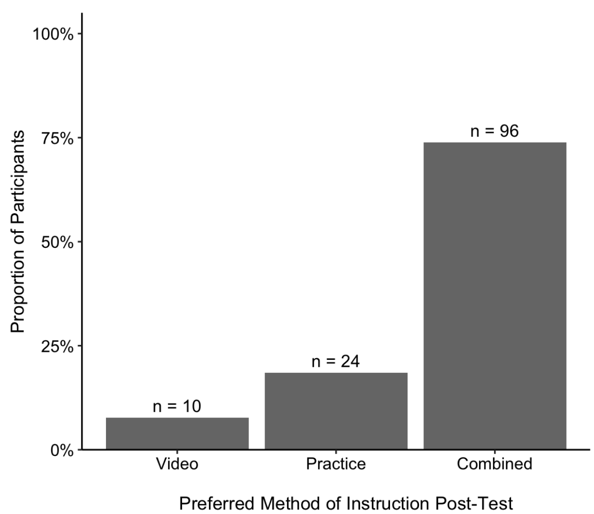

Figure 4: Preferences for Instructional Method After Instruction and Final Test

We conducted post-hoc pairwise comparisons to further investigate differences in preferences for instructional methods. These comparisons were adjusted for multiple testing using the Bonferroni correction method. All three comparisons were significant ($p < .05$). After the post test, the majority of participants selected the combined condition of video and practice as their preferred method of instruction.

Regardless of honored or dishonored assignment to instructional method, there were no significant predictors of participants' preferences for instructional methods after the final test.

Discussion

Scientific evidence and practical understanding indicate that combining practice with lecture enhances learning (Walker et al., 2008). More recently, practice only, without lecture, was shown to decrease instructional time without sacrificing performance. In Asher et al. (2025), eliminating traditional lectures and instructing exclusively through practice with feedback allowed students to learn at least 15% faster. However, students reported preferring a combination of lecture and practice or only the lecture. The practice-only approach also undermined interest in the content among students with low initial confidence. This evidence, along with decades of research on the importance of giving students choice in how they learn (Ryan & Deci, 2000; Patall et al., 2008), suggests that there may be benefits for honoring a students choice to learn from a lecture, even if lecture was not the most efficient method. In this study, we aimed to answer whether practice is really enough to learn and if it matters if students are in control of choosing their instructional methods.

One of our key findings is a clear initial preference for practice compared to video-only instruction. Based on findings from prior studies where participants were asked, after instruction, which type of instruction they would prefer (e.g., Asher et al. 2025; Deslauriers et al., 2019), we initially hypothesized that participants would select video as their preferred instructional method to prepare for the test. However, when asked at the start of the experiment, 50% of participants chose the combination of a video and practice, and another 35.38% of participants chose the practice-only condition.

Contrary to our predictions and prior work, students might have the correct initial intuitions about how to study but change their minds after instruction. These findings challenge the belief that students' learning decisions are generally miscalibrated (Dunlosky & Rawson, 2012), and instead suggest a more nuanced model. Students might have a correct understanding of learning processes at the start, but then use erroneous information from recent experience to update their instructional decisions, leading to less optimal learning. Our findings also suggest that future work should compare study decisions before and after the instructional phase, and current work only asking students for their intuitions after learning might introduce bias.

After instruction, 73.85% of participants selected the combined condition. These post-test preferences align with patterns observed in previous research where students showed a tendency to favor more comprehensive instructional methods (Carvalho et al., 2017). The observed shift among our participants reflects an important insight into how students approach learning and test preparation. Our results suggest that while many students initially make correct decisions about how they study, they may misjudge the effectiveness of their choices after experiencing instruction. Specifically, students may feel compelled to "do everything" to ensure that they are fully prepared, especially after only receiving one of the instructional methods (either video or practice). It is also possible that, after self-assessing, participants believe their final test performance could have been improved by using multiple instructional methods to learn more about the material.

A preference for the combined approach relates to the common psychological tendency of "more is better," such that students perceive multiple forms of instruction as a more complete or robust way to master material, even if it takes more time. Our findings may reflect a cognitive bias that quantity is a proxy for quality (Kleinberg & Marsh, 2023). After taking a test, students may believe that utilizing all available resources can further increase their chances of success, even if they initially chose the practice-only condition.

Despite the post-test preference for combined instruction, our findings show that participants with practice-only instruction performed just as well, in less time, compared to those who received the combination of video and practice. In this study, practice with feedback was the most efficient instructional method, which is consistent with previous literature on the testing effect (Carvalho et al., 2022).

Notably, students' preferences for instructional methods did not significantly affect the performance or instructional time for practice-only. It is possible that background knowledge of central tendency was helpful for participants to benefit from practice. Additionally, the match between the practice and test allowed participants who received practice-based instruction to solve application questions on the test by memorizing answers (Asher et al., 2025).

Our results suggest that having control over the choice of instructional method may not influence learning outcomes. These results contradict some prior work that self-regulated learning typically leads to better outcomes, particularly when students are well-calibrated in their decisions (Gureckis & Markant, 2012; Murphy et al., 2004). Thus, future work should explore the role of choice in tasks that require generalization of knowledge and skills.

Although lack of control over instruction resulted in equal learning outcomes in our study, we must also consider possible negative motivational effects. Practice-based methods may be detrimental for less-confident students who become discouraged by failure or negative feedback (Asher et al., 2025). In our analyses, instructional methods did not significantly predict post-test statistics self-efficacy. Nonetheless, future research should focus on the interaction between instructional methods and more comprehensive motivational measures, such as confidence and interest.

The majority of participants choosing the combined instruction, after instruction, may not only reflect the feeling that more is better but also a way to address motivational needs. In particular, the video component may provide necessary support for less confident students, helping them feel more prepared and capable before engaging in the practice questions. A combination of video and practice could be crucial for sustaining interest, particularly for students who feel overwhelmed by immediate practice testing without prior knowledge (Cooper et al., 2018).

The shift in instructional preferences also highlights a key challenge: how to convince students that practice alone can provide the necessary preparation for success. People often believe that more hours equals more success, but this approach can lead to inefficiency and wasted time. By changing the narrative to "work smarter, not harder," we can help optimize students' time and efforts.

Future research should focus on strategies to guide students in selecting and committing to efficient practice-based approaches. Although our study did not find direct benefits of self-regulation, existing evidence suggests that self-regulated learning leads to more effective and sustained learning outcomes. One potential area is integrating feedback that enhances students' confidence to foster calibrated self-regulation. Targeted feedback that conveys high expectations has been used to build trust in classrooms (Yeager et al., 2014), thus it may also be used to encourage students to make informed, autonomous decisions that align with the most effective study strategies.

Our findings contribute to the growing body of evidence on the efficiency and effectiveness of practice-based instruction. Importantly, instructional methods that prioritize practice could be particularly beneficial in high-pressure environments where there is limited time to cover complex material, such as introductory STEM courses. These courses are often characterized by heavy workloads and intense time constraints, which place a significant burden on students. Instructional methods that maximize learning efficiency could play a crucial role in helping students keep up with the pace of the course.

To ensure that practice-only instruction feels sufficient, not just before instruction but also after a test, future research should explore ways to reinforce the effectiveness of practice as a primary instructional method. It may be beneficial to highlight the time savings from practice-only instruction. For example, leveraging social comparison of the time and performance outcomes among peers who are engaging in combined methods may further demonstrate efficiency. As students see improved performance and saved time, they may feel more confident that practice alone is highly effective for learning. Ultimately, by strategically encouraging students to engage in practice with feedback, we can mitigate the feeling of needing to do more to prepare for a test and reduce instructional time.

Acknowledgements

This research was supported by the National Science Foundation under Grant No. 2418656 to Paulo F. Carvalho.

References

Asher, M. W., Sana, F., Koedinger, K. R., & Carvalho, P. F. (2025). Practice with feedback versus lecture: Consequences for learning, efficiency, and motivation. *Journal of Applied Research in Memory and Cognition*. Advance online publication. https://doi.org/10.1037/mac0000205

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-Regulated Learning: Beliefs, Techniques, and Illusions. *Annual Review of Psychology*, *64*(1), 417–444. https://doi.org/10.1146/annurev-psych-113011-143823

Capone, R. (2022). Blended learning and student-centered active learning environment: A case study with STEM undergraduate students. *Canadian Journal of Science, Mathematics and Technology Education*, *22*(1), 210-236.

Carvalho, P. F., McLaughlin, E. A., & Koedinger, K. R. (2017). *Is there an explicit learning bias? Students beliefs, behaviors and learning outcomes*. In *CogSci.*

Carvalho, P. F., McLaughlin, E. A., & Koedinger, K. R. (2022). Varied practice testing is associated with better learning outcomes in self-regulated online learning. *Journal of Educational Psychology, 114*(8), 1723–1742. https://doi.org/10.1037/edu0000754

Cooper, K. M., Downing, V. R., & Brownell, S. E. (2018). The influence of active learning practices on student anxiety in large-enrollment college science classrooms. *International Journal of STEM Education*, *5*(1), 23. https://doi.org/10.1186/s40594-018-0123-6

Deslauriers, L., McCarty, L. S., Miller, K., Callaghan, K., & Kestin, G. (2019). Measuring actual learning versus feeling of learning in response to being actively engaged in the classroom. *Proceedings of the National Academy of Sciences*, *116*(39), 19251–19257. https://doi.org/10.1073/pnas.1821936116

Dunlosky, J., & Rawson, K. A. (2012). Overconfidence produces underachievement: Inaccurate self evaluations undermine students' learning and retention. *Learning and Instruction*, *22*(4), 271–280. https://doi.org/10.1016/j.learninstruc.2011.08.003

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, *111*(23), 8410–8415. https://doi.org/10.1073/pnas.1319030111

Gureckis, T. M., & Markant, D. B. (2012). Self-Directed Learning: A Cognitive and Computational Perspective. *Perspectives on Psychological Science*, *7*(5), 464–481. https://doi.org/10.1177/1745691612454304

Hartwig, M. K., & Dunlosky, J. (2012). Study strategies of college students: Are self-testing and scheduling related to achievement? *Psychonomic Bulletin & Review, 19*(1), 126–134. https://doi.org/10.3758/s13423-011-0181-y

Kleinberg, S., & Marsh, J. K. (2023). Less is more: Information needs, information wants, and what makes causal models useful. *Cognitive Research: Principles and Implications*, *8*(1), 57. https://doi.org/10.1186/s41235-023-00509-7

Murphy, R. J., Gray, S. A., Straja, S. R., & Bogert, M. C. (2004). Student Learning Preferences and Teaching Implications. *Journal of Dental Education*, *68*(8), 859–866. https://doi.org/10.1002/j.0022-0337.2004.68.8.tb03835.x

Patall, E. A., Cooper, H., & Robinson, J. C. (2008). The effects of choice on intrinsic motivation and related outcomes: A meta-analysis of research findings. *Psychological Bulletin, 134*(2), 270–300. https://doi.org/10.1037/0033-2909.134.2.270

Roediger, H. L., & Karpicke, J. D. (2006). The Power of Testing Memory: Basic Research and Implications for Educational Practice. *Perspectives on Psychological Science*, *1*(3), 181–210. https://doi.org/10.1111/j.1745-6916.2006.00012.x

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68–78. https://doi.org/10.1037/0003-066X.55.1.68

Schunk, D. H. (1990). Goal Setting and Self-Efficacy During Self-Regulated Learning. Educational Psychologist, 25(1), 71–86. https://doi.org/10.1207/s15326985ep2501_6

Tauber, S. K., Dunlosky, J., Rawson, K. A., Rhodes, M. G., & Sitzman, D. M. (2013). General knowledge norms: Updated and expanded from the Nelson and Narens (1980) norms. *Behavior Research Methods, 45*(4), 1115–1143. https://doi.org/10.3758/s13428-012-0307-9

Tullis, J. G., Fiechter, J. L., & Benjamin, A. S. (2018). The efficacy of learners' testing choices. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(4), 540–552. https://doi.org/10.1037/xlm0000473

Walker, J. D., Cotner, S. H., Baepler, P. M., & Decker, M. D. (2008). A Delicate Balance: Integrating Active Learning into a Large Lecture Course. *CBE—Life Sciences Education*, *7*(4), 361–367. https://doi.org/10.1187/cbe.08-02-0004

Yeager, D. S., Purdie-Vaughns, V., Garcia, J., Apfel, N., Brzustoski, P., Master, A., Hessert, W. T., Williams, M. E., & Cohen, G. L. (2014). Breaking the cycle of mistrust: Wise interventions to provide critical feedback across the racial divide. *Journal of Experimental Psychology: General*, *143*(2), 804–824. https://doi.org/10.1037/a0033906

Zimmerman, B. J. (1990). Self-Regulated Learning and Academic Achievement: An Overview. *Educational Psychologist*, *25*(1), 3–17. https://doi.org/10.1207/s15326985ep2501_2