

## **Practice tests improve performance, increase engagement and protect from psychological distress**

**Stuart Marlin, Tori English, Lewis Morley, Tahlia O'Keefe-Quinn, Paige Whitfield**

School of Psychology, University of Newcastle, Australia.

---

### ***Abstract***

*The increasing prevalence of high levels of distress in university student populations has led academic and support staff to investigate options to help students cope with academic stress. Our research focused on investigating the benefit of early academic interventions for content engagement and feedback. In a 1<sup>st</sup> year psychology student sample of 547, we collected data on psychological measures (motivation and distress), practice test engagement and performance on assessment tasks. Assessment data from a baseline phase (practice tests were available) were compared to assessment data from an intervention (reward for undertaking practice tests). Our experiment also allowed an investigation into the type of benefit gained from practice tests engagement (content specific benefit vs general engagement effects). Results showed that undertaking practice tests ahead of assessment quizzes was associated with significantly higher assessment performance. Practice test uptake significantly increased when an incentive was in place resulting in much higher assessment scores for students. Students who reported high levels of distress on the DASS performed significantly lower on assessments. However, highly distressed students who undertook practice testing showed performance at the same level as non-distressed students.*

**Keywords:** *Practice; Testing; Stress; Engagement; Assessment.*

---

## **1. Introduction**

In educational settings, testing is most commonly used to evaluate a student's learning of core material and to assign students a grade for their course (Roediger, Putnam, & Smith, 2011). However, in addition to this summative assessment, formative testing has also been shown to have a significant impact on later assessment tasks (Sly, 2006). Some of the most common techniques for study involve rewriting notes, reviewing material covered in lectures or classes as part of a study group, rereading course notes and textbooks as well as completing practice questions or tests (Gurung, 2005). Extensive research has shown that testing, whether it be practice or the actual assessment, results in significant improvement of knowledge which is often superior to other methods of study (Carpenter, Pashler, and Vul, 2006; Cull, 2010; Glover, 1989; Sly, 2006). Roediger and Karpicke (2006) showed that practice tests are not superior to study alone if the assessment test occurs shortly after the study time (5 min). However, they found that knowledge appeared to significantly decay in the study alone condition with 50% retention at 2 days later and further decreased across a week. In comparison, knowledge also declined in the practice test condition, but even after 7 days was well above 50%. While uncued recall in short answer type questions has been generally shown to be the best type of testing (Carpenter, Pashler, and Vul, 2006; McDaniel, Anderson, Derbish & Morrisette, 2007), other research demonstrates the importance of timely feedback (Butler, Karpicke and Roediger, 2007; Butler, Karpicke and Roediger, 2008; Butler, Godbole & Marsh, 2013; Phelps, 2012). Moreover, a recent meta-analysis revealed that multiple-choice practice tests can improve learning more than short-answer practice tests (Adesope, Trevisan & Sundararajan, 2017). Despite currently being underutilised in higher education (Binks, 2017), mounting evidence suggests that multiple-choice practice testing, together with the meaningful feedback it provides (Gikandi, Morrow & Davis, 2011; Wojcikowski and Kirk, 2013), may be an effective formative learning strategy.

While many strategies may aid academic performance, self-testing (i.e. retrieval practice) has been highlighted as one of the most effective ways to engage with course material. Karpicke, Butler & Roediger (2009) found that although this strategy is highly effective, students report using it less frequently than other study habits. It is unclear from prior research whether this gap is present because students do not have access to practice assessments or if this method is not yet recognised as an effective learning strategy. Results from Hartwig and Dunlosky (2012) suggest that while students may engage in self-testing when preparing for exams, they view this as a tool to evaluate knowledge, rather than to aid learning.

Although practice effects have been identified as having a direct impact on academic performance, anxiety and stress indirectly affect the information-processing components associated with learning tasks (Tobias, 1976). Tobias' model suggests that stress can interfere in pre-processing, during processing, and post-processing phases of learning, and can result in negative effects on student performance. Since almost all university students report

experiencing high levels of distress compared to 29% of the general population, the contribution of stress and anxiety is of great interest to educational psychology research (Bore, Pittolo, Kirby, Dluzewska, & Marlin, 2016; Stallman, 2010). Factors such as being female, reporting financial stress, and being a full-time student can predict higher distress, however, Stallman (2010) reported that the prevalence of stress is most evident in first year students and decreases across years of study. This can be attributed to a range of factors such as increased confidence and task familiarity as students engage in similar styles of content throughout their degree. Jackson, Kleitman and Aidman (2014) showed that practice improved performance and reduced stress in cognitive workload tasks. Putwain (2008) also found that task familiarity and workload can affect student stress levels. Students in his study were required to complete exams that were low, mid and high stakes. Students performed better as they completed more tasks suggesting that practice led to increased task familiarity which can aid in the reduction of stress and increased academic performance. Putwain also found that students performed better and reported the lowest stress around mid-stake exams, suggesting that there is an interaction between task familiarity, task weighting, and stress on performance.

The present experiment examined the relationship between practice test engagement and academic performance on assessments across the semester in a large first year course. We predicted that practice test engagement should improve assessment scores and that by providing an incentive to undertake practice tests (a second attempt at Quizzes 2 and 4 if students completed two or more practice tests), a larger number of students would undertake practice tests. Practice test uptake was expected to remain higher than the initial baseline (Quiz 1) as students would learn the value of practice for assessment outcomes. Within the incentive quiz conditions (Quizzes 2 and 4) we also manipulated the type of content within the practice tests which allowed for a direct investigation of the type of benefit of practice (content specific feedback vs general engagement with content). We predicted there would be both content specific benefits of practice as well as engagement effects. Finally we measured overall student distress using the short form Depression, Anxiety and Stress Scale (DASS-21: Lovibond & Lovibond, 1995) early in the semester and again towards the end of the semester. We predicted that students who reported higher levels of distress (stress and anxiety combined) would perform more poorly than those showing low levels of distress and that students with higher distress who engaged in practice would perform better than those with high stress who did not do practice tests.

## **2. Methods**

### **2.1. Participants**

The 547 participants (Male = 106, Female = 327, Undisclosed = 114) in this study were students recruited from the 2019 cohort of the first-year course “PSYC1010 Introduction to Psychology 1” at the University of Newcastle. Those who voluntarily consented to participate agreed to the data from their coursework and lab activities being analysed as part of the study.

### **2.2. Materials**

Each of the following measures were completed by participants as part of their coursework; a Demographic survey, the DASS-21 (Lovibond & Lovibond, 1995) and the Psychology Motivation Questionnaire II which was adapted from the Science Motivation Questionnaire II (Glynn, Brickman, Armstrong, & Taasobshirazi, 2011). A study habits self-report measure was completed prior to each assessment quiz. The questions for this measure focused on student engagement with course material, self-directed and peer assisted study habits, and the use of feedback given on practice tests when studying.

Students completed four multiple choice assessment quizzes over the semester with multiple choice practice tests (25 questions) available prior to each quiz. No incentives were in place for Modules 1 and 3. However, in Modules 2 and 4, students were given the opportunity to complete their assessment quiz a second time (with their highest mark recorded) if they completed at least two of the three available practice tests for that module. Prior to completing the practice tests in Modules 2 and 4, students were placed into one of two experimental groups that determined the type of practice questions they received. Group 1 received a higher number (12) of practice questions from content A and very few practice questions (4) from content B. For Group 2 these question weightings were reversed. Both groups received equal numbers of practice questions from content C (9).

### **2.3. Procedure**

Demographic data was collected in Week 3, distress data in Weeks 6 and 9 (with a 2 week break between Weeks 7 and 8) and motivation data in Week 8. Quiz 1 occurred online in Week 5, Quiz 2 was completed in Week 8, Quiz 3 in Week 10 and finally Quiz 4 was completed after Week 12. Practice tests and quizzes were of the same multiple choice format and examined the same concepts but were drawn from different pools of questions. Thus practice tests gave students feedback on concepts but those questions were not present in the assessment quiz. For Quizzes 2 and 4, the 2<sup>nd</sup> quiz attempt also had new questions which were not in either the practice test or the 1<sup>st</sup> quiz attempt. While Quizzes 2 and 4 had two attempts, only the 1<sup>st</sup> attempt was used in our analysis to compare performance across all quizzes.

### 3. Results

Students uptake of practice tests dramatically increased between the first baseline condition (Module 1 where practice tests were encouraged) and the first incentive condition (Module 2 where a 2<sup>nd</sup> quiz attempt was given to those who did multiple practice tests). As shown in Figure 1, practice test uptake increased from 59% to 91% between Modules 1 and 2 and remained above 70% for the duration of the course. There was a smaller increase from Module 3 (73%) to Module 4 (82%). Higher practice rates were also associated with increased quiz performance (see Figure 1b) with significant improvements between Quiz 1 and Quiz 2 ( $t(517) = -8.281, p < .001$ ) and between Quiz 3 and Quiz 4 ( $t(491) = -10.630, p < .001$ ). What is perhaps more telling is the relationship between the number of practice tests completed (0 - 3) and quiz performance.

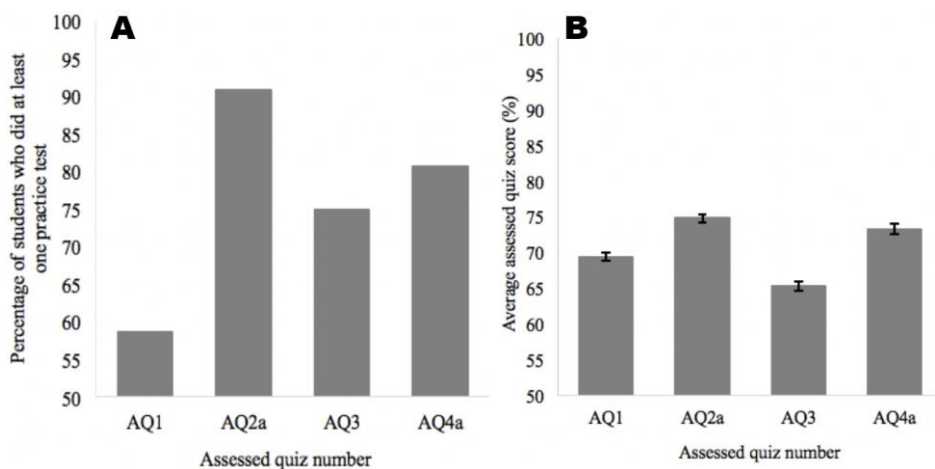


Figure 1. The percentage of participants doing at least one practice test for each module (A) and average quiz performance (B). Error bars represent standard error.

Practice test uptake was grouped into 3 categories (None (0), some (1 or 2) and All (all 3 available practice tests)). Figure 2a shows a consistent pattern of performance on quizzes relative to their practice test completion with significant differences for each quiz when comparing None vs Some, None vs All and Some Vs All. Thus practice tests were associated with improved quiz performance and incentives to increase practice test uptake were successful for Modules 2 and 4 and resulted in overall better performance. Moreover, students who reported that they completed practice tests but did not review their feedback performed no better than students who did no practice tests. Consistent with this finding was a strong relationship between overall grade and motivation score ( $r = .289$ ). Interestingly, all analyses of content specific effects for Modules 2 and 4 showed nonsignificant effects. Analysis of the relationship between distress and performance showed students with low

distress scored an average of 5 points higher on quizzes than students who reported high distress. Significant differences in quiz performance for distress level were found in Quizzes 1, 3 and 4 but not for Quiz 2 where over 90% of students did the practice. We also found the significant reduction in performance due to distress was only present in those students who had low engagement with practice tests ( $t(67) = 2.15, p = .035$ ) but was not present in highly distressed students who also had high practice test engagement (see Figure 2b).

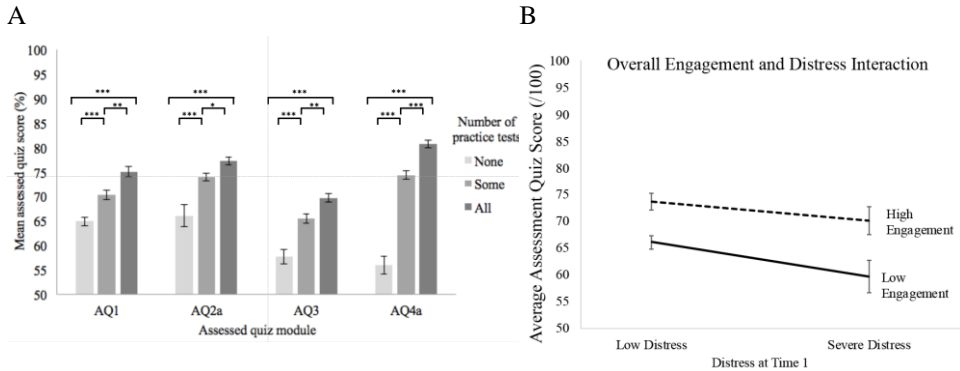


Figure 2. Average quiz score as a function of engagement in practice tests (A) and as a function of engagement for high vs low distressed students (B). Error bars represent standard error. \*  $p > .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$ .

#### 4. Discussion

The present study aimed to investigate the relationship between engagement in practice tests and performance on assessment quizzes. We also examined the role practice tests could play for those students who display high levels of distress. Our results show a clear benefit of engagement with practice tests for student learning. In our baseline conditions, the practice test uptake was lower and there was a corresponding reduction in quiz scores. When we provided a strong incentive to do the practice tests, uptake increased with a corresponding increase in performance on assessment. The effects of practice test engagement also relate to how many practice tests are done, with the most benefit coming from three practice tests. Students who completed no practice tests for Module 1 and continued to do so for the remaining modules showed no improvement in scores across the semester, whereas those who began doing practice in Module 2 or 3 or 4 showed improvements from when engagement began. One key aspect of this project was a clear distinction between practice tests, which were for formative feedback only, and assessment quizzes". Students undertake practice tests with less stress than an assessment quiz and use them for self evaluation of learning, which is essential for them to see their value (Karpicke, Butler & Roediger, 2009). Interestingly, our results showed the benefits of practice were not specifically related to the content for which they were given practice questions. Students who had proportionately

much higher numbers of questions on one type of content did not perform better on that content than they did for content where they received relatively few questions. This suggests that students are using practice tests as general performance feedback which may increase other types of review, as supported by the research of Carpenter (2012). An additional benefit observed with practice tests may be a general improvement in scores as they gain more experience with the test format (Meir, 2017; Snooks, 2005). This general reduction in performance stress that is observed across all students was especially relevant to students who present with an ongoing higher level of distress. Our results show that highly distressed students performed significantly poorer in assessments than students with low levels of distress. More importantly, providing these highly distressed students with practice tests allowed them to perform as well as those with low levels of distress.

## References

- Adesope, O. O., Trevisan, D. A., & Sundararajan, N. (2017). Rethinking the use of tests: A meta-analysis of practice testing. *Review of Educational Research*, 87(3), 659-701. doi:10.3102/0034654316689306
- Bore, M., Pittolo, C., Kirby, D., Dluzewska, T., Marlin, S. (2016) Predictors of psychological distress and well-being in a sample of Australian undergraduate students. *Higher Education Research and Development*. 35 (5), pp. 869-880
- Binks, S. (2017). Testing enhances learning: A review of the literature. *Journal of Professional Nursing*. doi: 10.1016/j.profnurs.2017.08.008
- Butler, A. C., Godbole, N., & Marsh, E. J. (2013). Explanation feedback is better than correct answer feedback for promoting transfer of learning. *Journal of Educational Psychology*, 105(2), 290-298. doi:10.1037/a0031026
- Butler, A. C., Karpicke, J. D., & Roediger, H. L. (2008). Correcting a metacognitive error: feedback increases retention of low-confidence correct responses. *Journal of Experimental Psychology, Learning & Memory Cognition*, 34(4), 918-928. doi:10.1037/0278-7393.34.4.918
- Butler, Andrew C., Karpicke, J. D., & Roediger, Henry L., III. (2007). The effect of type and timing of feedback on learning from multiple-choice tests. *Journal of Experimental Psychology: Applied*, 13(4), 273-281.
- Carpenter, S. (2012). Testing enhances the transfer of learning. *Current Directions in Psychological Science*, 21(5), 279-283.
- Carpenter, S. K., Pashler, H., & Vul, E. (2006). What types of learning are enhanced by a cued recall test? *Psychonomic Bulletin & Review*, 13(5), 826-830. doi: 10.3758/BF03194004
- Gikandi, J. W., Morrow, D., & Davis, N. E. (2011). Online formative assessment in higher education: A review of the literature. *Computers & Education*, 57(4), 2333-2351. doi:10.1016/j.compedu.2011.06.004

- Glynn, S. M., Brickman, P., Armstrong, N., & Taasoobshirazi, G. (2011). Science motivation questionnaire II: Validation with science majors and non-science majors. *Journal of Research in Science Teaching* 48(10), 1159-1176. doi: 10.1002/tea.20442
- Hartwig, M. K., & Dunlosky, J. (2012). Study strategies of college students: Are self-testing and scheduling related to achievement? *Psychonomic Bulletin and Review*, 19(1), 126–134. doi:10.3758/s13423-011-0181-y
- Jackson, S. A., Kleitman, S., & Aidman, E. (2014). Low cognitive load and reduced arousal impede practice effects on executive functioning, metacognitive confidence and decision making. *PLoS ONE*, 9(12), 1–29. doi:10.1371/journal.pone.0115689
- Karpicke, J. D., Butler, A. C., & Roediger, H. L. (2009). Metacognitive strategies in student learning: Do students practise retrieval when they study on their own? *Memory*, 17(4), 471–479. doi:10.1080/09658210802647009
- McDaniel, M. A., Anderson, J. L., Derbish, M. H., & Morrisette, N. (2007). Testing the testing effect in the classroom. *European Journal of Cognitive Psychology*, 19(4–5), 494–513. doi:10.1080/09541440701326154
- Meier, B. Optimizing learning in undergraduate psychology students: the impact of advance quizzing, review, and classroom attendance. *Cogn. Research* 2, 39 (2017). doi:10.1186/s41235-017-0075-2
- Phelps, R. P. (2012). The effect of testing on student achievement, 1910-2010. *International Journal of Testing*, 12(1), 21-43. doi: 10.1080/ 15305058. 2011. 602920
- Putwain, D. (2008). Do examinations stakes moderate the test anxiety-examination performance relationship? *Educational Psychology*, 28(2), 109–118. doi:10.1080/01443410701452264
- Roediger, H. L., Putnam, A. L., & Smith, M. A. (2011). Chapter one – Ten benefits of testing and their applications to educational practice. *Psychology of Learning and Motivation*, 55, 1-36. doi: 10.1016/B978-0-12-387691-1.00001-6
- Sly, L. (2006). Practice tests as formative assessment improve student performance on computer-managed learning assessments. *Assessment & Evaluation in Higher Education*, 24(3), 339-343. doi: 10.1080/0260293990240307
- Snooks, M. (2005). Using practice tests on a regular basis to improve student learning. *New Directions for Teaching and Learning*.100 109-113.
- Stallman, H. M. (2010). Psychological distress in university students: A comparison with general population data. *Australian Psychologist*, 45(4), 249–257. doi: 10.1080/00050067.2010.482109
- Tobias, S. (1976). Achievement treatment interactions. *Review of Educational Research*, 46(1), 61–74. doi: 10.3102/00346543046001061
- Wojcikowski, K., & Kirk, L. (2013). Immediate detailed feedback to test-enhanced learning: An effective online educational tool. *Medical Teacher*, 35(11), 915-919. doi: 10.3109/0142159X.2013.826793