

Practice Tests: Do They Work? Can They Work Better?

Anastasios Karakolidis, Michael O'Leary & Darina Scully

Centre for Assessment Research Policy and Practice in Education (CARPE)

Dublin City University

Introduction

Within the diverse field of certification and licensure, millions of tests are administered each year by non-governmental bodies and state agencies. A perusal of relevant websites suggests that most of the major certification and licensure tests have associated sample items, practice tests and/or preparatory tests (*prep tests*) either in hard copy or online available to potential candidates. All three are described in the standards of the National Commission for Certifying Agencies (NCCA, 2016). Sample items are offered simply to give potential test candidates an idea of what items on the test may look like. Practice tests are made up of numerous sample items and may, in some instances, be constructed to match a test plan. *Prep tests* can be distinguished from practice tests on the basis that they are always constructed to match test specifications closely, and are used to provide candidates with a predicted score for the real test. While sample items are usually provided at no cost and *prep tests* have to be purchased, practice tests can either be offered for free (often by the organisation responsible for setting or administering the test¹) or charged for by the entities that developed them (the certification/licencing authority² or other organisations³). These distinctions are worth noting; however, for simplicity, the general term 'practice tests' is used in this paper to refer to various types of practice and preparation materials.

¹ see, for example, <http://www.licensingcertification.org/for-candidates/practice-test/>

² e.g. the National Board of Medical Examiners
(http://www.nbme.org/PDF/NSAS/NSAS_Program_Information_Guide.pdf)

³ e.g. *Measure Up's* Microsoft Practice Tests for IT certification
(<http://www.measureup.com/Microsoft-Practice-Tests-C318.aspx>)

The Value of Practice Tests

The costs of practice tests varies considerably from relatively inexpensive (\$5 each) to \$100 or more⁴. Even within the same organisations it is noted that some practice tests are offered for free while others have to be paid for⁵. As far as the authors are aware, research on the sales of credentialing/licensure practice tests in the US and/or worldwide has never been conducted so the exact value in dollars of this market is unknown at this time. A complication in evaluating the extent of the market for practice tests is that some organisations may build the cost of developing and maintaining practice tests into the fee charged to candidates for the actual tests. That said, the certification/licensure testing market globally is likely to be valued in the billions of dollars given the very large numbers of those seeking certification or a license to practice⁶ and the fact that fees for most tests cost at least \$50. According to an article on the American Psychological Association (APA) website, the licensure fees to practice psychology can range from \$500 to \$1000 and cites one applicant as estimating “he spent nearly \$3,000 on licensing, including \$1,200 on EPPP study materials and a prep course”⁷. If we make the very conservative assumption that around one in four candidates spend between \$10 to \$25 (circa 10% of the cost of the lower priced tests) to take practice tests, then it is safe to say that many hundreds of millions of dollars are spent on this activity each year. So are practice tests worth the investment of time and money from the candidates’ point of view? This paper presents the key findings arising from a literature review on the efficacy of practice tests. On the basis of these findings, a series of recommendations for best practice regarding the development and provision of practice tests to test are made, and potential directions for future research outlined.

⁴ See, for example, <https://nationalcounselingexam.com/SubscriptionInfo>

⁵ See, for example, <https://www.socialworktestprep.com/free-practice-tests>

⁶ One company alone delivers an average of 7-10 million tests per year according to information provided on its website- see <https://www.prometric.com/en-us/about-prometric/pages/prometric-advantage-overview.aspx>

⁷ <http://www.apa.org/gradpsych/2004/01/get-licensed.aspxit>

The Scope of this Research

For the purposes of this research brief, a search for literature relating to practice tests was conducted using the academic search engines *Scopus*, *Dublin City University Library Database*, and *Google Scholar*. Potentially relevant articles were identified, and citation searches for these articles were also carried out in an effort to identify additional studies of interest. Approximately 25-30 academic papers were reviewed, 18 of which were considered key and are cited in this final report. Most of the research used was published within the past ten years, although some relevant studies from the 1990s are included. The quality of the academic literature was mixed. Many of the studies were small in scale (sample size was generally between 30 and 80 participants; although two studies had samples of > 100), and some had methodological flaws (*e.g.* non-experimental design, allowing participants to choose whether or not to take a practice test). The vast majority of the research was undertaken in undergraduate educational contexts in the US. Earlier studies employed simple word recall tests, and although this is an effective way of testing hypotheses, these findings may not necessarily translate to the recall of more complex material. More recent studies; however, replicated some of the initial findings using MCQ knowledge tests of higher level academic content (*e.g.* research methods, psychology, geography, neuroscience), which more closely approximate the structure and complexity of credentialing/licensure tests. In addition to the literature search, the websites of some certification and licensure examination boards were consulted with the aim of getting an understanding of typical test preparation practices in these milieu.

Key Findings

1. *Practice tests can be beneficial to prospective test candidates*
 - Completion of a practice test is likely to improve a candidate's performance in the actual certification test. There is a wealth of evidence to support this. The benefits of practice tests are in part due to the *direct testing effect*, whereby simply engaging in the process of information recall improves future recall (Roediger & Karpicke, 2006), but can also be explained by more *indirect* effects. Specifically, completing a practice test can increase learners' awareness of their current state of knowledge,

which in turn improves the quality of encoding during subsequent restudy sessions, and thus eventual test performance (*e.g.* Little & McDaniel, 2015; Soderstrom & Bjork, 2014). This phenomenon is known as *test-potentiated learning*. Recent empirical investigations employing the ‘think-aloud’ methodology have begun to uncover some of the mechanisms underlying test-potentiated learning. Fernandez and Jamet (2016), for example, found that students who have been exposed to practice tests make greater use of self-regulated learning processes, such as self-questioning and knowledge elaboration during restudy sessions.

- Some research (*e.g.* Balch, 1998; Prior & Lorch, 2012) suggests that students need to thoroughly engage with practice tests in order to realize their benefits. That is, a reasonable level of processing needs to occur. In a simulated last-minute test preparation scenario; however, Pilotti et al. (2009) demonstrated that simply reviewing practice test questions along with their answers (an activity fostering reasonably low level processing) can improve future test performance as much as engaging in active retrieval practice i.e. mental activity to recall information.
- A good deal of research suggests that practice tests will be most beneficial when the both the test itself and the conditions in which it is taken closely mirror those of the actual test (Roedinger & McDermott, 1993), for example, in quiet conditions, and under the time restrictions. Perlman (2003) also suggested that practice tests taken in this way can also help test-takers adopt strategies for dealing with test-taking anxiety.

2. *Practice tests can have some benefits even when feedback and/or the rationale for correct answers are not provided*

- In some of the studies discussed above (*e.g.* Soderstrom & Bjork, 2014; Little & McDaniel, 2015), participants simply completed practice tests. They did not receive any feedback on the correctness of their answers, or on their performance overall; nonetheless, positive effects on future test performance were still observed. As Glenberg et al. (1987) pointed out; individuals obtain *self-generated feedback* from the experience of retrieval success or failure during a practice test. In many cases, it seems that this self-generated feedback is sufficient to facilitate test candidates’

metacognitive judgements of their learning, and improve their future study.

- Despite the fact that practice tests have been shown to have benefits in the absence of feedback, it is still preferable if they at least inform candidates of the correct answers. This requires little cost or effort on the part of the test developers and, as Fernandez and Jamet (2016) pointed out, it ensures accurate feedback is obtained. The provision of corrective feedback also circumvents the problem of *negative suggestibility* in multiple-choice practice tests, whereby previously-seen distractors are more likely to be mistaken for correct answers in later testing situations (Butler & Roediger, 2008; Roediger et al., 2011).
- The provision of more detailed feedback during the taking of practice tests, such as rationales or explanations for the correct answers, may provide further benefits to potential test candidates. Fakcharoenphol, Potter and Stelzer (2011), for example, found that providing explanatory solutions in practice exams yielded nearly twice the gain relative to providing only the answer. However, this was in the context of problem-solving/analysis level items on an undergraduate physics exam, and may not translate to simpler recall or comprehension level items.

3. *It remains unclear whether practice tests can help candidates to predict their performance on future tests*

- Many psychological studies of memory suggest that we tend to exhibit a *foresight bias*, i.e. we are overconfident when predicting our likely recall capabilities (e.g. Koriat & Bjork, 2005). Given that practice tests increase learners' awareness of what they do and do not know, it seems logical to assume that practice tests may reduce the foresight bias, and help test-takers to form more accurate predictions of their eventual performance. Research examining this issue; however, has returned mixed results.
- One of the first studies examining the role of practice tests in facilitating learners' abilities to predict their eventual performance (Maki & Serra, 1992), revealed that

those who completed practice tests failed to produce more accurate predictions than those who did not. This was also supported by Bol and Hacker (2001), who found that not only did practice tests fail to facilitate test-takers' predictions of their eventual performance, they were associated with a *decrease* in predictive accuracy. More recent research has yielded directly contradictory findings: specifically, that those who have been exposed to practice tests can predict their future test scores with greater accuracy than those who have not (*e.g.*, Fernandez & Jamet, 2016; Little & McDaniel, 2015; Soderstrom & Bjork, 2014; Pilotti, Chodorow & Petrov, 2009).

- The reason for these equivocal findings is not clear. Bol (2001) argued that if test-takers expect that items in the final test will be identical to those used in the practice test, they may familiarize themselves with these items only, or focus only on the content tested by these particular items. This may lead to an erroneous sense of confidence about their likely performance in the actual test. Alternatively, if test-takers experience a great deal of retrieval failure during a practice test, they may predict a poor performance in the eventual test. However, if retrieval failure leads to more effective subsequent study, it makes sense that these predictions could be inaccurate, in that eventual test performance will be much higher than what was predicted after taking the practice test. Ultimately, the accuracy of test-takers' predictions of their eventual performance may depend on the amount of time available for re-study between the practice test and the actual test.

Recommendations

Based on the research literature consulted thus far, the following recommendations are made:

1. *Certification/Licensure organisations should continue to develop and offer practice tests to prospective test candidates*
- The provision of practice tests is good practice, because it allows potential test-takers to gain familiarity with the question format, the testing environment and the overall exam experience. There is substantial evidence to suggest that engaging with

these practice tests can help potential test candidates to identify areas of strength and weakness, and improve their future study strategies.

2. *Practice tests should include basic corrective feedback*

- Test-takers should be provided with the correct responses to each question, as this can consolidate the benefits of the practice test, and avoids the potential of negative suggestibility. More detailed feedback, and links to relevant study materials *etc.* would also be helpful, but this is not *necessary*, as practice tests can provide benefits to test candidates without including this information.

3. *Practice tests should be as similar to the actual tests as possible*

- It may be obvious, but nonetheless important, to state that practice tests should measure the same construct in the same way as the actual test (without the need to be predictive). The practice test should include questions from the same item bank as the actual test but the sets of items used should not overlap as this will undermine the validity of the final test. That said, it is worth bearing in mind that candidates may benefit from the inclusion of non-identical practice items due to indirect testing effects and test-potentiated learning (*i.e.* they may adopt better study strategies, which in turn can improve their performance on different items measuring similar content).
- Potential test candidates should be encouraged to take practice tests in a similar environment to that in which they will take the eventual test.
- If a pass/fail status is to be provided at the end of the practice test, it is extremely important that the practice test contains the same number of items as the official test, and that the content is represented in the same proportions, such that the resultant score is as reliable as that of the official test.

4. *Non-governmental bodies and state agencies should take into account ethical issues surrounding the development & provision of practice materials*

- Without evidence, it is important that practice tests are not promoted as tools that will allow candidates to predict their performance in the actual exams, because this is not always the case. Rather, they should be presented as learning aids that may *help improve eventual performance* by alerting potential candidates to their current level of knowledge and motivating them to adopt more effective study strategies. Test candidates should be encouraged to take practice tests relatively early on in their studies, such that sufficient time is available for restudy upon completion of the practice test.
- Given the beneficial impact that practice tests seem to have on test-takers' eventual performance, it has been argued (*e.g.* Lohman, 2013) that this type of material should be equally available to all. The implication here may be that, ethically speaking, certification/licensing entities should provide practice tests to all prospective candidates free of charge. It is acknowledged; however, that the development of practice tests requires additional financial input (*e.g.* the development of equivalent items that can never be used on the actual tests) and, as a consequence, a small, non-restrictive fee (circa \$5) may need to be considered. Alternatively, the registration fee for the actual exam could be slightly increased, with a free practice test included in this price. It may also make economic sense that a higher charge for one set of practice tests could be used to offset the cost of another *e.g.* in situations where a judgement can be made about the ability of different cohorts of test takers to pay a charge .
- The limitations placed by the National Commission for Certifying Agencies (NCCA) on test sponsors being involved in the development of study aides involvement should be borne in mind. The NCCA's *Standard 3* makes it clear that test preparation materials should always be seen as optional and that testing organisations must never give the impression that the purchase of its preparation materials is the only or preferred means for preparing adequately for the certification examination . In

some cases, the NCCA recommends that test development staff who work on the actual exam should not be involved in developing practice materials for that exam.

- With all this in mind it is worth noting many organizations in the field of certification and licensure testing provide a handful of sample items (either with or without correct answers) on their websites, with some also offering more extensive practice materials. These take the form of 'prep tests' that can be used to generate a score/indicate likely performance on the actual test, and some also offer detailed feedback and rationales for answers. There is usually a charge for more extensive materials⁸, but there are also instances where detailed practice tests can be accessed for free.⁹

Potential Research Studies

It would be meaningful, and relatively straightforward for many organisations involved in certification and licensure testing, to conduct research studies examining the effectiveness of the practice tests currently offered by them. This could be implemented by simply asking test candidates (as part of a post-test survey on the day of actual testing) to report whether or not they have taken a practice test; and if relevant, how many practice tests they have taken, approximately how long before the actual test these tests were taken, and to what extent the conditions in which they took the practice tests reflected 'real' test conditions. The relationship between each of these variables and performance on the actual exams could then be investigated. Given the cross-sectional nature of this proposed study, causal relationships could not be implied; nonetheless, this could provide valuable information about the relationship between candidates' preparation behaviours and engagement with practice tests, and their eventual test scores. Alternatively (or additionally), qualitative data concerning test candidates' attitudes towards practice tests (*e.g.* if they are helpful, if they are relevant to the actual exams *etc.*) could be collected.

⁸ NCBTMB offers access to > 3,000 practice questions and a personalized report analysing strengths and weakness for \$29.97 a month or \$49.97 for 3 months.

⁹ MTEL offers detailed Practice Tests with Multiple Choice Question Analyses, free of charge.

Non-governmental bodies and state agencies that currently offer full practice exams that provide a valid score, could gather data that would make it possible to conduct a study to examine the extent to which the practice exam can be used to predict (within error bounds) first-time pass status on the actual exam. This information could, potentially, be very useful to prospective candidates. If the practice test was to successfully predict the likelihood of passing the actual exam, then it could be used with a degree of accuracy in assessing test readiness. Previous such studies, undertaken for the National Board for Certification in Occupational Therapy practice exam (Avi-Itzhak, 2015) and the Mock Board Dental Hygiene Exam (Dadian, Geurink, Olney & Littlefield, 2002), revealed overall successful prediction rates of 80% and 75% respectively.

The studies suggested here, if undertaken, could be used by an organisation to enhance its approach to the provision of practice tests and to justify its current and/or future investment in them. In addition, for some non-governmental bodies and state agencies, these studies, having the potential to be larger in scale than any other undertaken heretofore, would contribute greatly to the research literature in the field.

References

- Avi-Itzhak, T. (2015). Ability of the National Board for Certification in Occupational Therapy Practice Test to Estimate the Probability of First-Time Pass Status on the National Certification Exam. *The American Journal of Occupational Therapy*, 69(2).
- Balch, W.R. (1998) Practice versus review exams and final exam performance. *Teaching of Psychology*, 25, 181-185.
- Bol, L. & Hacker, D. (2001). A Comparison of the Effects of Practice Tests and Traditional Review on Performance and Calibration. *The Journal of Experimental Education*, 69(2), 133-151.
- Butler, A.C. & Roediger, H.L. (2008). Feedback enhances the positive effects and reduces the negative effects of multiple-choice testing. *Memory and Cognition*, 36(3), 604-616.
- Dadian, T., Geurink, K, Olney, C. & Littlefield, J. (2002). The effectiveness of a mock board experience in coaching students for the Dental Hygiene National Board Examination. *Journal of Dental Education*, 66, 643-648.
- Fernandez, J. & Jamet, E. (2016). Extending the testing effect to self-regulated learning. *Metacognition and Learning*. DOI: 10.1007/s11409-016-9163-9.
- Fakcharoenphol, W., Potter, E. and Stelzer, T. (2011). What students learn when studying physics practice exam problems. *Physics Education Research*, 7 (1).
- Glenberg, A.M., Sanocki, T., Epstein, W. & Morris, C. (1987). Enhancing calibration of comprehension. *Journal of Experimental Psychology: General*, 116(2), 119.
- Koriat, A. & Bjork, R.A. (2005). Illusions of competence in monitoring one's knowledge during study. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 31, 187-194.
- Little, J.L. & McDaniel, M.A. (2015). Metamemory monitoring and control following retrieval practice for text. *Memory and Cognition*, 43(85), 85-98.
- Lohman, D. (2013) *Ability tests, the internet and practice tests: A recipe for invalidity*. Unpublished manuscript, College of Education, University of Iowa.
- Maki, R. & Serra, M. (1992). Role of Practice Tests in the Accuracy of Test Predictions on Text Material. *Journal of Educational Psychology*, 84(20), 200-210.
- National Commission for Certifying Agencies (NCCA)(2016). *NCCA standards for the accreditation of certification programs*. Available at:
<http://www.credentialingexcellence.org/p/cm/ld/fid=66>

- Perlman, C.L. (2003). Practice Tests and Study Guides: Do They Help? Are They Ethical? What is Ethical Test Preparation Practice? In: *Measuring Up: Assessment Issues for Teachers, Counselors, and Administrators* (387-396).
- Pilotti, M., Chodorow, M. & Petrov, R. (2009). The Usefulness of Retrieval Practice and Review-Only Practice for Answering Conceptually Related Test Questions. *The Journal of General Psychology*, 132(2), 179-204.
- Prior, T.J. & Lorch, M. (2012). A clerical error provides a unique opportunity to study the effect of previously seen exam questions on exam results. *New Directions*, 1 (8) .
- 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.
- Roediger, H.L., Putnam, A.L, & Smith, M.A. (2011). Ten Benefits of Testing and Their Applications to Educational Practice. In: J. Mestre & B. Ross (Eds.), *The Psychology of Learning and Motivation: Cognition in Education*, (Vol. 55, pp. 1-35). Oxford: Elsevier.
- Soderstrom, N. C. & Bjork, R.A. (2014). Testing facilitates the regulation of subsequent study time. *Journal of Memory and Language*, 73, 99-115.