# Retrieval Practice, Enhancing Learning in Electrical Science

James Eustace and Pramod Pathak
*School of Computing, National College of Ireland, IFSC, Dublin, Ireland*

Keywords: Testing Effect, Retrieval Practice, Practice Testing Learning Framework, Electrical Science, Problem Solving.

Abstract: After an initial learning period taking a practice test improves long-term retention compared with not taking a practice test. This testing effect finding has significant relevance for education, however, integrating retrieval practice effectively into teaching and learning activities presents challenges to educators in classroom situations. This paper extends previous research applying the Practice Testing Learning Framework (PTLF) to support teaching and learning using practice testing in the classroom with materials that range in complexity from understanding to problem-solving in electrical science. Findings from this study show the number of practice tests completed and overall engagement with practice tests had a significant impact on criterion test performance in the topics enhancing learning where practice tests were available and more effective than other techniques employed. The testing effect was evident with materials involving problem-solving and the authors recommend the PTLF to integrate retrieval practice into teaching and learning activities.

## 1 INTRODUCTION

Retrieval practice has interested researchers for over a hundred years (Abbott, 1909) and is supported by a considerable body of research (Roediger and Karpicke, 2006; Roediger *et al.*, 2011; Roediger, Putnam and Smith, 2011) and with increased interest in the last decade (Karpicke, 2017). Studies on the effects of retrieval practice on students' learning of course material are limited and how to integrate retrieval practice effectively into teaching and learning activities has not been addressed and presents challenges to educators particularly where problem-solving, and transfer is required. This paper addresses this gap with the Practice Testing Learning Framework (PTLF) and its application using three topics in Electric Science in the national Electrical Apprenticeship Programme in Ireland.

Students struggle to regulate their learning often using ineffective techniques (Dunlosky *et al.*, 2013), techniques which they are familiar with and have used before. Apprentices are no different and have difficulty with theory examinations in the Electrical Apprenticeship at Phase 2. Apprentices range in age from 16, the minimum age of employment as an Electrical apprentice to more mature learners. This paper investigates whether retrieval practice enhances learning in electrical science in topics requiring application and problem-solving within the PTLF, and is the treatment more effective than other techniques employed by apprentices?

Section 2 provides a review of recent research of retrieval practice in classroom contexts and an introduction to the PTLF is outlined in Section 3, while Section 4 describes methods for the quasi-experimental design adopted for this study. Section 5 presenting the results followed by a discussion in Section 6 and the conclusions are presented in Section 7.

## 2 RETRIEVAL PRACTICE IN CLASSROOM CONTEXTS

Retrieval practice research in classroom contexts has been limited with studies emerging using computerised quizzing in electrical science (Eustace and Pathak, 2018), introductory statistics (Eustace and Pathak, 2017), mathematics for computing (Eustace *et al.*, 2015), engineering (Butler *et al.*, 2014), Spanish language learning (Lindsey *et al.*, 2014) and introductory psychology (Pennebaker *et al.*, 2013).

Paper-based practice tests in introductory psychology (Batsell *et al.*, 2017) and the use of clicker systems in middle school science (Roediger *et al.*, 2011; McDaniel *et al.*, 2013), educational psychology (Mayer *et al.*, 2009) are also evident in a



Eustace, J. and Pathak, P.
Retrieval Practice, Enhancing Learning in Electrical Science.
DOI: 10.5220/0007674102620270
In *Proceedings of the 11th International Conference on Computer Supported Education (CSEDU 2019)*, pages 262-270
ISBN: 978-989-758-367-4

number of classroom studies. The next section will look at computerised quizzing in classroom contexts.

### 2.1 Computerised Quizzing

The use of computerised or online quizzes allows practice tests to be delivered at scale over time providing learners with feedback. One study employed the OpenStax Tutor system using a within-subjects experimental design. Improvements were observed in learning in the STEM classroom with retrieval practice items requiring the application of a concept by combining three principles from cognitive science, (1) repeated retrieval practice, (2) distribution over time and (3) feedback. The experiment utilised a single factor, intervention versus standard practice, within groups with random assignment to groups, n=40. The experiment was conducted within the classroom and unlike laboratory experiments with high levels of control which may not transfer to the classroom, learners were not restricted from other activities so as not to impact on the learning process. The small to medium effect sizes observed from the 'noisy' classroom reflect the improved learning within the classroom (Butler *et al.*, 2014).

Another study administered the quizzes using a course management system in an introductory biology course. This study investigated the effect of exam-question level on fostering student conceptual understanding using low-level and high-level quizzes and exams. The findings of the study recommend that assessments should be designed at higher levels of Bloom's taxonomy to assess the desired learning outcomes which in turn helps students to direct their learning leading to deeper conceptual understanding. The problem remains however that many science instructors fail to design assessments that assess the required cognitive process and often test only content knowledge (Jensen *et al.*, 2014).

A Web-based flash-card tutoring system, the Colorado Optimized Language Tutor (COLT) was used for Spanish foreign-language instruction to create a personalised review system. The system required students to type the Spanish translations after presenting them with vocabulary words and short sentences in English after which corrective feedback was provided. The study found that personalized review enhanced performance by 16.5% over current educational practice (massed study) with a 10.0% improvement over a one-size-fits-all strategy for spaced study (Lindsey *et al.*, 2014).

The TOWER (Texas Online World of Educational Research), an online teaching and learning platform that provides student feedback on their performance as they learn the material was used in another study. The first 10-minutes of each class were devoted to an 8-item daily quiz with seven of the questions on previous material and a personalised question the student had answered incorrectly on a previous quiz. For the duration of the study, students took 26, 8-item multiple-choice quizzes at the beginning of every class via their own devices. No final or other exams were administered. Their performance was based on an analysis of students' overall grade based on quizzes and writing assignments and comparison with classes in previous years (Pennebaker *et al.*, 2013).

## 3 PRACTICE TESTING LEARNING FRAMEWORK

The studies to date employing e-assessment have not positioned practice tests within an overall learning model or framework. The PTLF (Eustace and Pathak, 2018) illustrated in Figure 1, addresses this gap and is an operationalisation of the Conversational Framework, where the e-assessment practice test environment is the "task practice environment" or teachers constructed environment (Laurillard, 2002). Within the PTLF, learning activities occur between two levels; the discursive/theoretical level and the practice/practical level and these activities reflect the learning process (Laurillard, 2009).

First Principles of Instruction (Merrill, 2002) adopts a problem-centred approach where learners are engaged in solving real-world problems and is integrated within the PTLF in Figure 1. Activation occurs when existing knowledge is leveraged by the teacher as a foundation for new knowledge and this new knowledge is demonstrated to the learner. The teacher adapts the practice test environment to present the problem. The learner applies new knowledge with engagement in the task practice environment. The learner adapts their practice to the problem using existing knowledge and through their action and subsequent feedback integrates new knowledge through reflection.

The Cognitive Rigor (CR) Matrix (Hess *et al.*, 2009) combines the Revised Bloom's Taxonomy, Cognitive Process Dimension (CPD) with DOK Levels (Webb, 2002) in a two-dimensional model. By supporting alignment between learning outcomes and test item development, Hess's CR matrix supports retrieval-based learning within the Practice Testing Learning Framework.

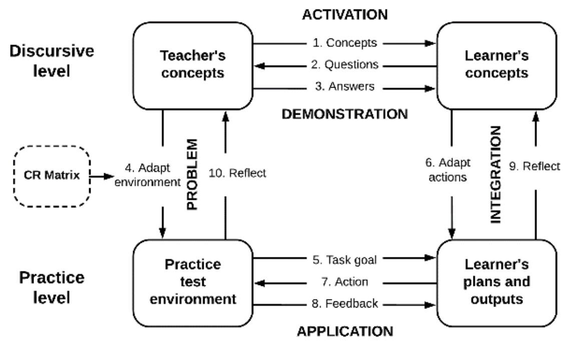


Figure 1: Practice Testing Learning Framework (PTLF) adapted from (Laurillard, 2002).

### 3.1 Testing Effect Theories

Theories proposed to account for the testing effect include the amount of exposure, elaborative retrieval and transfer appropriate processing and are briefly described here for context.

The amount of exposure hypothesis (Slamecka and Katsaiti, 1988) argue that the testing effect is due to the over-learning of the items through repeated exposure and successful recalling. However the testing effect cannot be fully explained by the amount of exposure alone as while additional studying can produce better retention in the short term, it is surpassed by practice testing which produces better long-term retention on tested material (Roediger and Karpicke, 2006; Roediger *et al.*, 2011) and on non-tested material (Chan, 2010; Little and Bjork, 2014).

The elaborative retrieval hypothesis (Carpenter, 2009; McDaniel *et al.*, 2011) advances the idea that the process of retrieval modifies memory and increases the probability of future successful retrieval. In support of the retrieval effort hypothesis, findings indicate that as retrieval difficulty during practice tests increase, subsequent criterion test performance also increases (Pyc and Rawson, 2009; Smith and Karpicke, 2013; Vaughn, Rawson and Pyc, 2013). For a testing effect to occur, students must be able to answer questions successfully. If the practice test difficulty is such that no items are recalled or if the correct answers to the non-recalled items are not given for the test topic, then minimal, none or even a negative test effect may occur.

The most longstanding hypothesis looks at the testing effect from transfer-appropriate processing (TAP) viewpoint where memory performance is dependent on the overlap between the encoding process and the retrieval process (Morris *et al.*, 1977). It is the student engagement with similar operations or processes during testing that results in enhanced performance compared with items not tested or only restudied. This theory has been broadly considered from the perspective of matching item formats from practice test to criterion test and in situations where the criterion test differs from the initial test or requires transfer, then TAP predicts a reduction in the size of the testing effect (Rohrer *et al.*, 2010). However, an alternative perspective of TAP found a testing effect when matching cognitive processes between initial and criterion tests using high-level items in new contexts enhanced performance in both high and low-level conditions (Jensen *et al.*, 2014).

This study does not align particularly to any one of these theories in isolation but does support exposure to appropriate materials and that elaborative retrieval invoking the cognitive process at the required depth of knowledge can enhance learning.

## 4 METHODS

Conducting research in the classroom is not without challenge particularly where the treatment impacts on classroom activities, summative assessment outcomes and overall course performance. The methods adopted in this study minimised disruption to apprentices and were implemented for three topics. The previous study involving one topic, had two limitations which are addressed in this study by increasing the number of treatment topics and comparing within group performance of treatment and non-treatment topics. The methods and procedures adopted reflect these concerns with all apprentices enrolled in the practice test treatment with three topics using a within-group design, with participant performance being measured and compared against performance in the non-treatment topics. While the difficulty of topics does vary, the topics selected for treatment reflected broadly the non-treatment topics involving application and problem-solving. Learner ability, motivation and other learning opportunities are uncontrolled factors however by comparing participant performance within the topics in the ‘noisy’ classroom should address these validity concerns.

### 4.1 Course and Materials

As with the previous study (Eustace and Pathak, 2018), this was an apprenticeship for Electrical apprentices enrolled in a 4-year national programme.

The apprenticeship consists of four on-the-job phases with an approved employer and three off-the-job phases in an educational organisation. The study was conducted during Phase 2, which is delivered in the Education and Training Board (ETB) training centres over 22 weeks. The course in Phase 2 consists of seven modules of learning: (1) Electrical Science, (2) Installation Techniques 1, (3) Installation Techniques 2, (4) Panel Wiring and Motor Control, (5) Fundamentals of Alternative Electrical Energy Sources, (6) Team Leadership and (7) Communications.

Table 1: Learning outcomes supported by Practice Tests.

| Unit | Learning Outcome(s) | CPD* | DOK** |
|---|---|---|---|
| Ohms Law/The basic circuit | Identify graphical symbols associated with the basic circuit | 2 | 2 |
| | State the units associated with basic electrical quantities | 2 | 1 |
| | State the three main effects that electric current has upon the basic circuit | 2 | 1 |
| | Calculate circuit values using Ohm's Law | 3 | 2 |
| Resistance network measurement | Identify the differences between series, parallel and series/ parallel resistive circuits using a multimeter | 2 | 2 |
| | Calculate the total resistance, voltage and current of series, parallel and series/parallel resistive circuits using the relevant formulae and a multimeter | 3 | 2 |
| | Identify the differences between series, parallel combinations of cells in relation to the voltage and current outputs using the relevant formulae and a multimeter | 3 | 2 |
| | Explain resistivity and list the factors which affect it | 2 | 1 |
| Cables and cable terminations | Describe the construction, sizes and applications of PVC and PVC/PVC cables (up to 16mm$^2$) and of flexible cords (up to 2.5mm$^2$) | 2 | 2 |
| | Apply proper safety, care, handling and storage techniques to tools | 2 | 2 |

*CPD* 1 = Remember, 2 = Understand, 3 = Apply.*
*DOK** 1 = Recall and Reproduction, 2 = Skills and Concepts*.

The practice test development approach employed Hess's CR matrix to map learning outcomes to support item development and classification (Table 1). Test items were designed to reflect the cognitive process dimension with the required depth of knowledge. The treatment topics included questions requiring application and problem solving with the practice tests integrated into the learning environment as described in the PTLF.

## 4.2 Participants

The participants in the study were n=164 Electrical apprentices on Phase 2 of their national Electrical apprenticeship programme in 2016. The apprentices were assigned to classes in Education and Training Board (ETB) Training Centre's by SOLAS, the coordinating provider for the Electrical Apprenticeship. The assignment of apprentices to classes is based on apprentice registration number which is allocated at the beginning of an apprenticeship with participants being drawn for locations nationally to make up each class. The typical apprenticeship class size at Phase 2 is small relative to other courses with n=14. All participants are enrolled in the Apprenticeship Moodle Learning Management System following registration, which provides apprentices assess to course material and resources. The practice tests were provided as an optional course resource, and in contrast to laboratory studies where the learning process is highly controlled, participants manage their own learning regarding taking the practice tests. The concerns around the validity of results and findings are discussed further in Section 6.



## 4.3 Procedures

As in the previous study (Eustace and Pathak, 2018), the materials developed for the study consisted of MCQ test items, assembled into a test bank. A Moodle Learning Management System was used to deploy the practice tests. The criterion test is the national T1 Theory Test used in the Apprenticeship Programme. It consists of 75 items, four option MCQ's with one correct option. The criterion test is unseen to participants and delivered by a different assessment management system, not linked to the practice test item bank. Apprentices must successfully answer at least 52 of the items correctly to pass the criterion test.

The criterion test topics include Ohms Law/The basic circuit; Resistance network measurement; Power and energy; Cables and cable termination; lighting circuits; Bell circuits; Fixed appliance and socket Circuits; Earthing and bonding and Installation testing.

Practice tests were developed and made available to participants via Moodle for three topics, Ohms Law/The basic circuit, Resistance network measurement and Cables and cable terminations.

Each practice test consisted of 15 MCQ's, with a minimum forced delay of 1 day between attempts and a 20-minute time limit for each test. Feedback was deferred, apprentices were required to select an answer to each question and then submit the test before the test is graded, or feedback is given. Feedback is shown immediately after the attempt showing whether correct, and the marks received. Participants attempted practice tests in their own time which were available for the duration of the course.

Participants were informed by email and Moodle message that the practice tests consisted of 15 multiple choice questions and once they started, they had 20 minutes to complete the test. Participants were also informed they would have to wait 1 day between attempts at the same version and that the practice test results were not included in the course result calculation. The remaining six topics were used as a control as no practice tests were provided. All topics were assessed in the criterion test which is typically administered around week 12 of the course. The additional 'noisy' activities of apprentices and instructors were not controlled, i.e. participants may have undertaken self-testing, taken additional instructor-led paper-based tests or applied preferred study techniques.

# 5 RESULTS

## 5.1 Number of Practice Tests

The number of practice tests taken by apprentices had a significant impact on criterion test performance in all treatment topics, and presented in Figures: 2, 3 and 4 respectively.

Table 2 reflects reduced participation after the initial Ohms Law practice tests with higher numbers not taking the practice tests in resistance and cables. Cohen's *d* is used in this study and provides a measure of effect size, with 0.2, 0.5, and 0.8 reflecting small, medium, and large effect sizes, respectively (Cohen, 1988).

A One-Way ANOVA with the dependent variable, the performance in the Ohms Law/The basic circuit topic in the criterion test, with the number of practice tests as the factor found a significant improvement in performance with p = .006, F = 3.765 between groups. A small effect size with Cohen's *d* = 0.48 was observed between no practice tests and 3 to 4 practice tests and a large effect size observed between no practice tests and 7 to 9 practice tests with Cohen's *d* = 0.89.

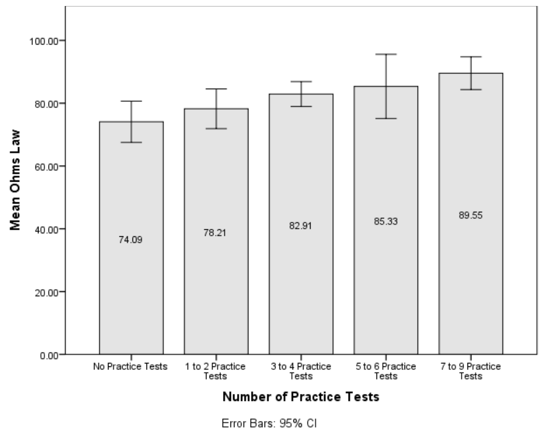


Figure 2: Performance in Ohms Law.

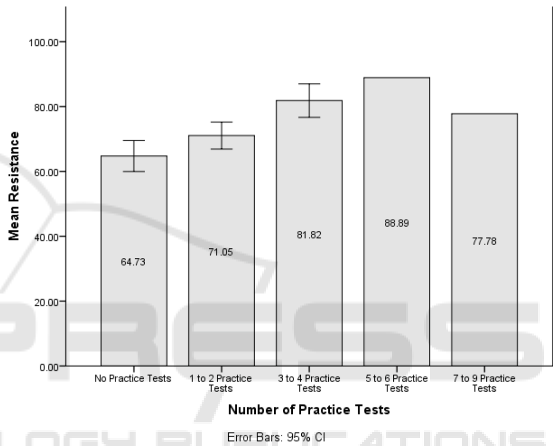


Figure 3: Performance in Resistance.

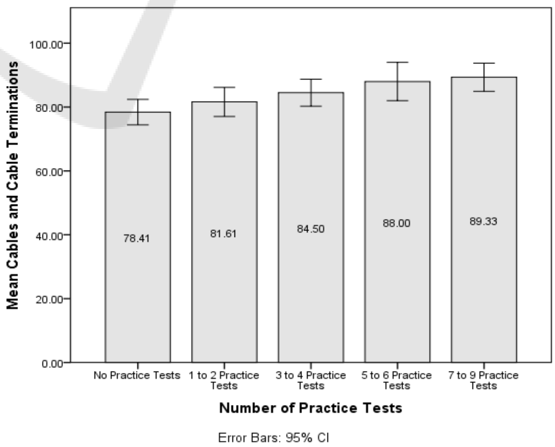


Figure 4: Performance in Cables.

A One-Way ANOVA with the dependent variable, the performance in the Resistance network measurement topic in the criterion test, with the number of practice tests as the factor found a significant improvement in performance with p = .003, F = 4.243 between groups, A large effect size with Cohen's *d* = 1.05 was observed between no practice tests and 3 to 4 practice tests.

A One-Way ANOVA with the dependent variable, the performance in the Cables and cable terminations topic in the criterion test, with the number of practice tests as the factor found a significant improvement in performance with p = .014, F = 3.209 between groups. A small effect size with Cohen's *d* = 0.42 was observed between no practice tests and 3 to 4 practice tests and a large effect size observed with a Cohen's *d* = 0.87 between no practice tests and 7 to 9 practice tests.

Table 2: Number of participants taking the practice tests.

| Attempts | Ohms Law | Resistance | Cables |
|---|---|---|---|
| 0 | 44 | 69 | 63 |
| 1 to 2 | 28 | 71 | 31 |
| 3 to 4 | 55 | 22 | 40 |
| 5 to 6 | 15 | 1 | 15 |
| 7 to 9 | 22 | 1 | 15 |

## 5.2 Overall Participant Engagement

To evaluate the performance of apprentices relative to their engagement with the practice tests the results of the four groups are presented in Table 3. Within the study 42 did not take any practice tests, 38 completed some practice tests, 63 completed all practice tests at least once, and 21 completed all practice tests three or more times. A One-Way ANOVA with the dependent variables, the performance in the topics in the criterion test, with completion of the practice tests as the factor found a significant improvement in performance between groups for all three with p = .004, F = 4.680 for Ohms Law, p = .004, F = 4.566 for Resistance and p = .003, and F = 4.752 for Cables.

Table 3: Mean score in criterion test topics.

| Number of Practice Tests | N | Ohms Law | Resistance | Cables |
|---|---|---|---|---|
| No tests completed | 42 | 73.57 | 64.29 | 77.38 |
| Completed some tests | 38 | 80.53 | 68.71 | 79.47 |
| All tests at least once | 63 | 83.02 | 70.55 | 85.87 |
| All tests three+ times | 21 | 89.52 | 82.01 | 87.14 |
| | 164 | 80.85 | 69.99 | 82.38 |

Not completing any practice tests and completing all practice tests three or more times for Ohms Law had a medium effect size with Cohen's *d* = 0.62. Resistance had a large effect size with Cohen's *d* = 1.04 and Cables had a medium effect size with Cohen's *d* = 0.69.

## 5.3 Overall Criterion Test Performance

The overall criterion test topic performance was reviewed to determine the relative difficulty of topics for apprentices and is presented in Figure 5.

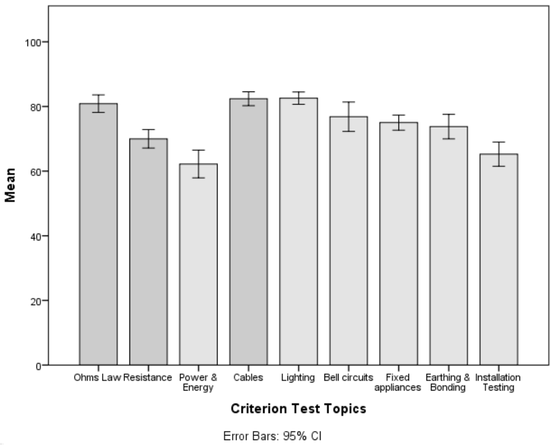


Figure 5: Overall criterion test topic performance.

The mean score in each topic of the criterion test reflects the relative difficulty with Ohms Law = 80.85, Resistance = 69.9864, Power and energy = 62.2, Cables = 82.38, Lighting = 82.56, Bell circuits = 76.83, Fixed appliances = 75, Earthing and bonding = 73.78 and Installation testing = 65.24. The three most difficult topics based on the results are Power and energy, Installation testing and Resistance, in that order.

## 5.4 Within-group Results

The statistics from a paired sample T-Test (Table 4) of apprentices that completed at least one or more attempts of each of the three topics where practice tests were available, and their performance in the remaining topics found a mean of 81.42 with a standard deviation of 11.21 while topics without practice tests for the same group had a mean of 73.32 with a standard deviation of 13.83.

Table 4: Paired sample T-Test descriptive statistics of the group that participated in all practice tests at least once.

| | Mean | N | Std. Dev | Std. Error Mean |
|---|---|---|---|---|
| Topics with PT | 81.42 | 84 | 11.21 | 1.22 |
| Topics without PT | 73.32 | 84 | 13.83 | 1.51 |

The results from the paired sample T-Test found a statistically significant difference between the mean of topics with practice tests and mean of topics without practice tests within the group with a Sig. (2-Tailed) value $p < .001$, $t = 5.960$.

The statistics from a paired sample T-Test (Table 5) of apprentices that did not participate in any practice tests found a mean of 71.75 with a standard deviation of 16.19 for the treatment topics while non-treatment topics for the same group had a mean of 70.99 with a standard deviation of 13.42.

Table 5: Paired sample T-Test descriptive statistics of the group that did not participate in any practice tests.

| | Mean | N | Std. Dev | Std. Error Mean |
|---|---|---|---|---|
| Topics with PT | 71.75 | 42 | 16.19 | 2.50 |
| Topics without PT | 70.99 | 42 | 13.42 | 2.07 |

The results from the paired sample T-Test found no statistically significant difference between the mean of topics with practice tests and mean of topics without practice tests within the group that did not participate in practice tests with a Sig. (2-Tailed) value $p = .635$, $t = .478$.

### 5.5 Survey - No Practice Test Group

A survey post-experiment was conducted of the 42 apprentices that did not complete any practice tests with 11 respondents (26.19%). Two themes emerged; lack of awareness of the practice tests as 55% of respondents claimed they didn't know about them and 45% preferred to use their own revision techniques.

## 6 DISCUSSION

A discussion follows on the limitations of the experimental design, participant engagement in practice tests, the within-group results and reflection on earlier findings.

### 6.1 Validity of Experimental Design

Typically with practice testing in laboratory conditions, students are presented with materials to be learned and then randomly assigned to groups where one group completes a practice test or sequence of practice tests, and the other group studies the material again. Learner performance is evaluated in a criterion test or delayed criterion test and results evaluated between groups. It could be argued that the participants that availed of practice tests in this study were more motivated or higher performing and that other uncontrolled factors such as self-testing influenced the enhanced performance. While these factors are difficult to control in the classroom, the finding of significance with the paired sample T-Test found that learning was enhanced for practice test participants that engaged in the practice tests in those treatment topics but did not enhance their learning in the non-treatment topics. Interestingly, the results of the paired sample T-Test of the group that did not participate in any practice tests found no significant difference between their performance in treatment and non-treatment topics. These findings would suggest that the practice testing treatment condition had a significant impact on learning and that the within-group design is valid for the classroom experiment.

### 6.2 Engagement in Practice Tests

The number of practice tests taken by participants had a significant impact on criterion test performance in all treatment topics. These findings are consistent with research involving Ohms Law with 3 to 4 practice tests enhancing the learning process (Eustace and Pathak, 2018). Participants found resistance network measurement a more difficult topic than the other treatment topics, and this is reflected in the no practice test group not meeting the minimum standard compared to the other treatment topics. Participants did not take as many attempts at the resistance practice test with only 2 participants taking 5 or more practice tests. 69 did not avail of any practice tests of which 42 scored less than 70% in the resistance topic on the criterion test, and 23 did not pass the criterion test overall suggesting these learners had difficulty during the learning process. The increase in the number of learners not availing of the resistance and cables practice tests compared to the ohms law practice tests may be attributed to novelty, as Ohms Law presented first or increased workload as learners were exposed to more material which may have taken their attention. Overall, taking all practice tests at least once enhanced learning, and taking 3 or more practice tests had the greatest effect.

### 6.3 Effect Size

The largest effect size was observed for the resistance topic, the more difficult of the three treatment topics. Of the group (42) that did not complete any practice tests the mean performance in the criterion test was 73.57, 64.29, 77.38 for Ohms Law, resistance and cables respectively. The impact of the practice tests was most pronounced in the Resistance topic with a large effect size, Cohen's $d = 1.05$ between no practice tests and 3 to 4 practice tests. This topic had a greater number of items requiring application and problem solving and the testing effect is more pronounced for these items. In Ohms Law, a large effect size was observed between no practice tests and 7 to 9 practice tests with Cohen's $d = 0.89$. Medium effect sizes were observed in Ohms Law and cables for 3 to 4 practice tests compared to no practice tests.

### 6.4 Practice Test Participation

For those that did not avail of the practice tests, accessing the Moodle system and difficulty with passwords contributed to the lack of awareness with others preferring to use alternative study techniques. Several techniques have been identified with varying degrees of effectiveness (Dunlosky *et al.*, 2013), and less effective techniques such as highlighting and rereading continue to be popular with apprentices (Eustace and Pathak, 2018). There was no significant difference between treatment and non-treatment topics for this group that did not avail of the practice tests.

## 7 CONCLUSIONS

This paper investigated e-assessment practice testing within the PTLF in electrical science and the impact of practice tests to enhance learning as measured in the criterion test topic performance. Findings from this study show the number of practice tests completed and overall engagement with practice tests had a significant impact on criterion test performance in the topics where practice tests were available. Assessment items should be designed to the cognitive process dimension and depth of knowledge to assess the intended learning outcomes which in turn helps participants to direct their learning leading to deeper conceptual understanding. The testing effect was evident with materials involving problem-solving and the authors recommend the use the PTLF to integrate retrieval practice into teaching and learning activities. The practice tests were more effective that the other techniques employed by apprentices. Future work will extend this research further to include all topics assessed in the criterion test for electrical science and examine the learner experience and the experience of instructors applying the PTLF.

## REFERENCES


Abbott, E. E. (1909) ‘On the analysis of the factor of recall in the learning process.’, *The Psychological Review: Monograph Supplements*, 11(1), pp. 159–177. doi: 10.1037/h0093018.

Batsell, W. R. *et al.* (2017) ‘Ecological Validity of the Testing Effect’, *Teaching of Psychology*, 44(1), pp. 18–23. doi: 10.1177/0098628316677492.

Butler, A. C. *et al.* (2014) ‘Integrating Cognitive Science and Technology Improves Learning in a STEM Classroom’, *Educational Psychology Review VO - 26*. Springer, (2), p. 331. Available at: http://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,cookie,shib&db=edsjsr&AN=edsjsr.43549798&site=eds-live&scope=site&custid=ncirlib.

Carpenter, S. K. (2009) ‘Cue strength as a moderator of the testing effect: The benefits of elaborative retrieval.’, *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Carpenter, Shana K., Department of Psychology, Iowa State University, W112 Lagomarcino Hall, Ames, IA, US, 50011-3180, shacarp@iastate.edu: American Psychological Association, 35(6), pp. 1563–1569. Available at: 10.1037/a0017021.

Chan, J. C. K. (2010) ‘Long-term effects of testing on the recall of nontested materials.’, *Memory*. Psychology Press (UK), 18(1), pp. 49–57. Available at: 10.1080/09658210903405737.

Cohen, J. (1988) *Statistical power analysis for the behavioural sciences*. 2nd edn. New York: LAWRENCE ERLBAUM ASSOCIATES.

Dunlosky, J. *et al.* (2013) ‘Improving Students’ Learning With Effective Learning Techniques: Promising Directions From Cognitive and Educational Psychology’, *Psychological Science in the Public Interest*, 14(1), pp. 4–58. doi: 10.1177/1529100612453266.

Eustace, J., Bradford, M. and Pathak, P. (2015) ‘A Practice Testing Learning Framework to Enhance Transfer in Mathematics’, in Muntean, C. and Hofmann, M. (eds) *The 14th Information Technology & Telecommunications Conference*, pp. 88–95.

Eustace, J. and Pathak, P. (2017) ‘Enhancing Statistics Learning within a Practice Testing Learning Framework’, in *ICERI2017*. Seville, Spain, pp. 1128–1136.

Eustace, J. and Pathak, P. (2018) ‘Enhancing Electrical Science Learning Within A Novel Practice Testing Learning Framework’, in *Frontiers in Education Conference (FIE)*,

Hess, K. K. *et al.* (2009) *Cognitive Rigor: Blending the Strengths of Bloom’s Taxonomy and Webb’s Depth of Knowledge to Enhance Classroom-Level Processes, Online Submission*. Online Submission.

Jensen, J. L. *et al.* (2014) ‘Teaching to the test…or testing to teach: Exams requiring higher order thinking skills encourage greater conceptual understanding.’, *Educational Psychology Review*. Germany: Springer. Available at: 10.1007/s10648-013-9248-9.

Karpicke, J. D. (2017) *Retrieval-Based Learning : A Decade of Progress*. Third Edit, *Learning and Memory: A Comprehensive Reference*. Third Edit. Elsevier. doi: 10.1016/B978-0-12-805159-7.02023-4.

Laurillard, D. (2002) *Rethinking university teaching: A conversational framework for the effective use of learning technologies*, *Rethinking university teaching A conversational framework for the effective use of learning technologies*. Available at: http://www.worldcat.org/isbn/0415256798.

Laurillard, D. (2009) 'The pedagogical challenges to collaborative technologies.', *International Journal of Computer-Supported Collaborative Learning*. Springer Science & Business Media B.V., 4(1), pp. 5–20. Available at: 10.1007/s11412-008-9056-2.

Lindsey, R. V. *et al.* (2014) 'Improving Students' Long-Term Knowledge Retention Through Personalized Review', *Psychological Science*, 25(3), pp. 639–647. doi: 10.1177/0956797613504302.

Little, J. L. and Bjork, E. L. (2014) 'Optimizing multiple-choice tests as tools for learning.', *Memory & Cognition*. Germany: Springer. Available at: http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2014-34514-001&site=ehost-live.

Mayer, R. E. *et al.* (2009) 'Using Technology-Based Methods to Foster Learning in Large Lecture Classes: Evidence for the Pedagogic Value of Clickers'. Available at: http://www.bergen.edu/Portals/0/Docs/Clickers/using technology based methods.pdf.

McDaniel, M. A. *et al.* (2011) 'Test-Enhanced Learning in a Middle School Science Classroom: The Effects of Quiz Frequency and Placement.', *Journal of Educational Psychology*, 103(2), pp. 399–414. Available at: 10.1037/a0021782.

McDaniel, M. A. *et al.* (2013) 'Quizzing in Middle-School Science: Successful Transfer Performance on Classroom Exams', *Applied Cognitive Psychology*, 27(3), pp. 360–372. doi: 10.1002/acp.2914.

Merrill, M. D. (2002) 'First principles of instruction', *Educational Technology Research and Development*, pp. 43–59. doi: 10.1007/BF02505024.

Morris, C. D., Bransford, J. and Franks, J. J. (1977) 'Levels of processing versus transfer appropriate processing.', *Journal of Verbal Learning & Verbal Behavior*. Netherlands: Elsevier Science, 16(5), pp. 519–533. Available at: 10.1016/S0022-5371(77)80016-9.

Pennebaker, J. W., Gosling, S. D. and Ferrell, J. D. (2013) 'Daily online testing in large classes: Boosting college performance while reducing achievement gaps', *PLoS ONE*, 8(11). doi: 10.1371/journal.pone.0079774.

Pyc, M. A. and Rawson, K. A. (2009) 'Testing the retrieval effort hypothesis: Does greater difficulty correctly recalling information lead to higher levels of memory?', *Journal of Memory and Language*. Pyc, Mary A., Kent State University, Department of Psychology, P.O. Box 5190, Kent, OH, US, 44242-0001: Elsevier Science, 60(4), pp. 437–447. Available at: http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2009-04548-002&site=ehost-live.

Roediger, H. L. I. *et al.* (2011) 'Test-Enhanced Learning in the Classroom: Long-Term Improvements from Quizzing', *Journal of Experimental Psychology: Applied*. Journal of Experimental Psychology: Applied, 17(4), pp. 382–395. Available at: http://dx.doi.org/10.1037/a0026252.

Roediger, H. L. I. and Karpicke, J. D. (2006) 'Test-Enhanced Learning: Taking Memory Tests Improves Long-Term Retention.', *Psychological Science*. Roediger, Henry L., III, Department of Psychology, Washington University Campus Box 1125, One Brookings Dr., St. Louis, MO, US, 63130, roediger@artsci.wustl.edu: Blackwell Publishing, 17(3), pp. 249–255. Available at: 10.1111/j.1467-9280.2006.01693.x.

Roediger, H. L. I., Putnam, A. L. and Smith, M. A. (2011) 'Ten benefits of testing and their applications to educational practice.', in Mestre, J. P. and Ross, B. H. (eds) *The psychology of learning and motivation (Vol 55): Cognition in education.* San Diego, CA US: Elsevier Academic Press (The psychology of learning and motivation; Vol 55; 0079-7421 (Print);), pp. 1–36. Available at: 10.1016/B978-0-12-387691-1.00001-6.

Rohrer, D., Taylor, K. and Sholar, B. (2010) 'Tests enhance the transfer of learning.', *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Rohrer, Doug, Department of Psychology, University of South Florida, PCD4118G, Tampa, FL, US, 33620, drohrer@cas.usf.edu: American Psychological Association, 36(1), pp. 233–239. Available at: 10.1037/a0017678.

Slamecka, N. J. and Katsaiti, L. T. (1988) 'Normal forgetting of verbal lists as a function of prior testing.', *Journal of Experimental Psychology: Learning, Memory, and Cognition*. US: American Psychological Association, 14(4), pp. 716–727. Available at: 10.1037/0278-7393.14.4.716.

Smith, M. and Karpicke, J. D. (2013) 'Retrieval practice with short-answer, multiple-choice, and hybrid tests', *Memory*, (September), pp. 37–41. doi: 10.1080/09658211.2013.831454.

Vaughn, K. E., Rawson, K. a and Pyc, M. a (2013) 'Repeated retrieval practice and item difficulty: does criterion learning eliminate item difficulty effects?', *Psychonomic bulletin & review*, 20(6), pp. 1239–45. doi: 10.3758/s13423-013-0434-z.

Webb, N. (2002) 'An analysis of the alignment between mathematics standards and assessments for three states', *annual meeting of the American Educational …*. Available at: http://addingvalue.wceruw.org/Related Bibliography/Articles/Webb three states.pdf (Accessed: 10 May 2014).