Enhancing Electrical Science Learning Within A Novel Practice Testing Learning Framework

James Eustace
Department of Computing
National College of Ireland
Dublin, Ireland
james.eustace1@student.ncirl.ie

Pramod Pathak
Computing Department
National College of Ireland
Dublin, Ireland
pramod.pathak@ncirl.ie

***Abstract*—This paper presents a novel practice testing learning framework to enhance learning in electrical science. Several learning techniques are used by students to maximise their learning. Research has shown that most of those learning techniques are not effective. Retrieval practice also referred to as the 'testing effect', promotes practice testing as it is an effective technique. Accordingly, the testing effect phenomena and its potential application to electrical science education may be particularly crucial to student learning. However, research into practice testing has been limited to date in classroom contexts, and it is necessary to broaden the inquiry to enable instructors to implement practice testing as part of their teaching and learning strategy. This research seeks to address this gap by a study of n=161 apprentices on the National Electrical Apprenticeship programme in the Republic of Ireland, across multiple training centres, and how they benefited from e-assessment practice tests to enhance learning. For this purpose, a novel practice testing learning framework is applied supporting learning, item development, classification and mapping to learning outcomes. A One-Way ANOVA with the dependent variable the performance in the randomised unseen criterion test, with student test attempts as the factor found a significant improvement in performance between groups. A significant increase in performance was also evident for participants in the practice test topic not meeting the passing grade of the overall criterion test between groups with a mean performance of 72.35% following three or four practice tests compared to 52.94% with no practice tests in the practice test topic area. This finding suggests that participation in three or four practice tests enhances learning and improves self-regulation. The insights from this study will serve as guidelines for future educational research on practice testing towards enhancing learning in electrical science and other domains.**

Keywords— testing effect; electrical science; practice testing learning framework

I. INTRODUCTION

Learning and applying electrical science domain knowledge presents a significant challenge to new entrants to the electrical community of practice. Learning involves the acquisition and encoding of new information and learners apply different learning techniques. Tests are usually associated with measuring achievement. However, a growing body of research supports practice testing as an effective learning technique in itself, and the act of retrieving and applying knowledge in a test creates the 'testing effect' which enhances learning [1]–[3]. In this paper, Section A provides a brief overview of practice testing research and also emphasises potential boundary conditions may exist which are not yet defined. Section B discusses retrieval practice conditions and their impact on learning. Section C explores taxonomies and links these to theories and models for learning to build a framework supporting a practice testing approach while Section D describes the Practice Testing Learning Framework (PTLF) within an integrated pedagogical model that educators can apply to their teaching and learning strategies. Finally, Section E introduces self-regulatory processes and how these may be mapped be to the PTLF.

A. Practice Testing

Research has shown that many students struggle to regulate their learning, often using ineffective techniques such as rereading or highlighting [1]. Practice testing, also referred to as retrieval practice, benefits learning where the material is better remembered over time if the learning process includes tests on that material and has been shown to be highly effective across a range of conditions and criterion tasks and is supported by a large body of research [1]–[3]. The role of tests can be expanded in educational contexts with both direct and indirect effects on student performance [4]. The indirect effects of testing may encourage students to study and to distribute their effort throughout the course and adjust their studies based on feedback received. The primary focus of practice testing is on its direct effects and is not a result of additional exposure [3]. Studies of the testing effect phenomenon have identified the positive effect of retrieval during testing in consolidating learning compared to repeated study [5]–[7]. Testing effect size is typically determined by performance on the final criterion test with participants assigned to either a practice test group or a study only group.

Practice testing is a powerful learning technique, [1], [5], [8]–[11] in classroom studies where the final test material is identical to the practice test [12], [13], or modified or rephrased versions of the same test [14]. Much of the research in classroom contexts have tended to concentrate on test items requiring the retrieval of facts [13], [15], [16].

Much of the existing research on practice testing has tended to focus on recall with identical or similar test items in both practice and criterion tests. To test understanding and application of knowledge and skills test items should not come directly from assigned reading or course presentation [17] but should require learners to use knowledge or skill in new situations.

The application of knowledge to new situations, i.e. 'transfer' is a primary goal of education. A small number of recent studies have examined near transfer in classroom contexts such as middle school Science [18] and Introductory Biology [19]. In the McDaniel (2013) study, transfer was not supported by lower level or definitional questions but was supported by higher level items [19]. In another study,

978-1-5386-1174-6/18/$31.00 ©2018 IEEE

criterion test performance from MCQ testing using clickers in the classroom was enhanced when followed up with a short discussion [20]. However, in a study using authentic classroom materials, the testing effect was not evident in topically related questions [21]. This may indicate that the testing effect may have important boundary conditions not yet clearly defined.

B. Retrieval Practice Conditions

A significant amount of research has focused on manipulating retrieval practice conditions and comparing study only conditions to study test conditions. The benefits of testing with studying compared to study alone have been well established. Conditions that require more effort from the participant tend to result in enhanced learning performance. This section provides a brief review of selected practice conditions and their impact on learning.

Number of retrieval attempts – A condition of considerable interest is how much practice is enough and are there limited benefits to be gained from repeatedly taking practice tests? A gradual and significant improvement in criterion test performance in concept learning increasing from one and six practice attempts was observed [22] and taking three initial practice tests with a delay of 1 week to criterion test producing superior performance over restudy requiring transfer of knowledge in new contexts [23].

Massed versus Spaced Practice - Spaced practice (also known as distributed practice) is a learning strategy, where learning is broken up into several short sessions over a period. Massed practice, on the other hand, involves repeated attempts taken directly after each other. One example of spaced practice comes from Carpenter and DeLosh (2005), in name learning, experiments 2 and 3, where they demonstrated that names repeatedly retrieved at spaced intervals were retained almost three times as well as names repeatedly retrieved in massed practice [24]. While much of the research has been in laboratory settings, there are some examples involving classroom settings such as with vocabulary learning [25] and US history facts [13].

Retention Interval (RI) - There are mixed results in relation to schedules of distributed practice, expanding, fixed or reducing schedules and that the optimal learning schedule may be conditional on the length of the retention interval. The retention interval is the time lag between the last retrieval practice and the criterion test. In one study, a contracting schedule was beneficial for retention intervals up to 7 days, but both fixed and expanding schedules were better for a long retention interval of 35 days [26].

Practice Test Performance – The number of practice test attempts should also be considered in terms of practice test performance resulting from successful retrieval relative to criterion test performance. While learners may attempt a practice test, effortful retrieval is required to enhance learning. Rawson and Dunlosky (2011) prescribe recalling concepts to an initial criterion of three correct recalls and then to relearn them three times at widely spaced intervals [27].

C. Learning Taxonomy

It is important to link theories of practice testing with theories and models of learning [2], and the Practice Testing Learning Framework provides a useful framework to describe the interactions of retrieval-based learning.

A number of studies have suggested that the benefits of the testing effect can be realised when items assess higher cognitive levels [19], [28]. The most well-known approach for classifying learning objectives and test items reflecting those objectives is Bloom’s taxonomy of the Cognitive Domain [29]. Findings from this research using higher level questions involving cognitive processes such as application, and analysis suggest that conceptual understanding improves compared to using lower level items focusing on recall. A revision of Bloom’s taxonomy placed renewed focus on the value of the original Handbook while incorporating new knowledge and thought into the framework [30] to include cognitive process and knowledge dimensions [31].

Webb’s Depth of Knowledge (DOK) Levels is an alternative taxonomy which proposes that depth-of-knowledge levels should be assigned to both learning objectives and assessment items to ensure their alignment. This alignment is a crucial factor to be considered for retrieval-based learning. According to Webb (2002), depth-of-knowledge levels are assigned to both objectives within standards and assessment items. This approach provides a critical component to ensure alignment between assessments and objectives [32]. The DOK model consists of four levels which include “recall and reproduction” at Level 1 and is defined as the recall of information such as a fact, definition, or a simple procedure, or performing a simple process. The term “simple” typically involves one step. A significant body of testing effect research has focused recalling identical items attempted before. Level 2 requires mental processing that goes beyond recalling or reproducing a response. The content knowledge or process involved at Level 2 is more complicated than at Level 1 and requires students to decide on how to approach a question or problem. Level 2 tasks require two or more steps to solve a problem. Strategic thinking at Level 3 requires reasoning, planning, using evidence, and is at a higher level of thinking than the previous two levels. The cognitive demands at Level 3 are complex and abstract. Extended thinking tasks at Level 4 have high cognitive demands and are very complex. Students are typically required to relate ideas within the content area or among content areas, make several connections and select or develop an approach on how the situation can be solved [32]. This taxonomy model supports alignment between the criterion and practice tests where the original Bloom's Taxonomy had limitations by using verbs to differentiate taxonomy levels with many of these verbs appear at multiple levels. These limitations are overcome by using the Cognitive Rigor (CR) Matrix [33] which combines the Revised Bloom's Taxonomy, Cognitive Process Dimension (CPD) with DOK Levels in a two-dimensional model. By supporting alignment between curriculum learning outcomes and test item development, Hess’s CR matrix supports retrieval-based learning within the Practice Testing Learning Framework.

Beyond the testing effect theories directly relating to practice testing several learning theories are relevant in the context of this paper. These include the Conversational Framework [34] and First Principles of Instruction [35].

D. Practice Testing Learning Framework (PTLF)

Practice testing is a strategy or technique for learning, which can be combined with other techniques and is considered here within an integrated pedagogical model. The authors propose that for the benefits of practice testing to be fully realised, practice testing must exist within a framework for learning. The PTLF [36] is an operationalisation of the Conversational Framework and where the e-assessment practice test environment is the "task practice environment" or teachers constructed environment [37].

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 [38]. This framework has been influential in the analysis of formative e-assessment within educational domains including higher and further education, and work-based learning [39], in the design of learning environments [40] and more recently in practice testing [36], [41].

An alternative framework known as First Principles of Instruction [35] adopts a problem-centred approach where learners are engaged in solving real-world problems. 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.

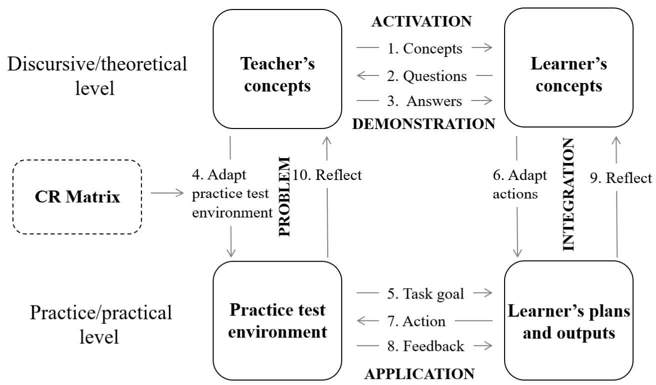

Fig. 1. Practice Testing Learning Framework (PTLF)

E. Self-regulation

Meta-cognition is defined as the awareness of and knowledge about one's thinking [42] and is a complex construct which involves cognitive knowledge and cognitive regulation. Cognitive knowledge is made up of multiple types such as declarative, procedural and conditional knowledge as well as different types of cognitive regulation planning, monitoring or regulating [43]. The ability of students to self-regulate or monitor their learning is essential to effectively their guide study and learning effort. Self-regulation is the self-directed process by which learners transform their mental abilities into academic skills. Learning is an activity that students do for themselves in a proactive way rather than as a covert event that happens to them in reaction to teaching [42].

The structure of self-regulatory processes occurs in three cyclical phases; the forethought phase; performance phase and self-reflection phase. The forethought phase describes processes and beliefs that occur before efforts to learn; the performance phase to processes that occur during learning, and finally the self-reflection phase to processes that occur after the learning effort [42] and mapped be to the PTLF.

Test expectancy influences students' perceptions of what constitutes learning and may impact how students prepare for tests and monitor learning [19], [44]. While practice testing provides feedback on learning and may improve meta-cognitive awareness for students [45], [46], they may not be aware of its benefits [4], [5], [47]. Students are generally under confident about testing [21] while relying on less effective techniques such as rereading or highlighting [1]. However, while student metacognitive beliefs can change when they experience the benefits of testing and guided with support [48] further research is required as to whether testing enhances metacognitive effects and whether these effects are long-lasting and independent of the specific information being learned [21], [49].

II. METHODS

This study investigated how apprentices engaged with practice tests during their training and how this engagement enhanced their performance. Two research questions guided this study: (1) Does e-assessment practice testing within the Practice Testing Learning Framework promote learning in Electrical Science? (2) What are the optimum parameters for this learning transfer?

A. Course and Materials

The course selected for this study was Electrical, an apprenticeship for Electrical apprentices enrolled in a 4-year national programme. On successful completion of the Electrical apprenticeship, the apprentice is awarded a Quality and Qualifications Ireland (QQI) Level 6 Advanced Certificate Craft - Electrical. This is the only route to becoming an electrician in Ireland. This 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 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 Motrol Control, (5) Fundamentals of Alternative Electrical Energy Sources, (6) Team Leadership and (7) Communications.

The study focused on a unit within the Electrical Science Module. 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 corectly 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.

Online Practice Tests were developed and made available to participants via Moodle for Ohms Law, The Basic Circuit. This topic is assessed by 10 items in the Criterion test which is typically administered around week 12 of the course.

The learning outcomes associated with Ohms Law/The Basic Circuit are illustrated in Table 1. The key learning points for the first learning outcome on the identification of graphical symbols include; the cell, battery, resistors, ammeter, voltmeter, ohmmeter, cables crossing, and cables joined; components of a simple electric circuit and the concepts of the open and short circuit. The key learning points for the second learning outcome include; the structure of the atom, nucleus and shells of orbiting electrons; definitions and charges or electron, proton and neutron; laws of charges' balanced and unbalanced atoms; conductors and insulators; current flow as movement of free electrons; EMF as a cause of current flow; sources of EMF; conventional current flow versus actual current flow; resistance as opposition to current flow; definition of Coulomb and definition of Ohm's law. The key learning points for the third learning outcome include heating, magnetic and chemical effects of electric current. The key learning points for the final learning outcome include; Ohm's Law calculations: transposition of formulae; addition, subtraction, multiplication and division of indices; units and symbols of current, voltage and resistance and charge.

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 required depth of knowledge. The first three learning outcomes require understanding at level 1 and 2 while the fourth learning outcome requires application at level 2. The practice tests were integrated into the learning environment as described within the PTLF. The electrical science concepts were delivered using traditional pedagogy with questions and feedback at the discursive/theoretical level between the instructor and students in classroom sessions. The practice test environment was configured with practice tests which presented the task goal to learners, requiring action and receiving feedback on completion.

TABLE 1 Learning outcomes supported by Practice Tests

Unit 1	Learning Outcome(s)	CPD*	DOK**
Ohms Law/The Basic Circuit	1.1 Identify graphical symbols associated with the basic circuit	2	2
	1.2 State the units associated with basic electrical quantities	2	1
	1.3 State the three main effects that electric current has upon the basic circuit	2	1
	1.4 Calculate circuit values using Ohm's Law	3	2

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

A survey of participants was conducted which explored the perceived effectiveness of study methods employed by participants; participant engagement with practice tests; perceived difficulty of topics in the Criterion test and perceived benefits of practice testing.

B. Participants

The participants in the study were n=161 Electrical apprentices on Phase 2 of their national Electrical apprenticeship programme between May and October 2016. The apprentices were assigned to classes in Education and Training Board (ETB) Training Centre's by SOLAS, the statutory authority with responsibility for apprenticeship in Ireland. The assignment of apprentices to classes is based on apprentice registration number which is allocated at the beginning of an apprenticeship. Typical class size, 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 to all apprentices.

C. Procedures

Ohms Law/The Basic Circuit – three versions of the practice test were developed, with three attempts allowed at each. Versions 2 and 3 became available once the previous version of the test was completed. Each practice test version consisted of 15 MCQ's, with a minimum forced delay of 1 day between attempts for that version and a 20-minute time limit allowed per 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 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. Participation in the study was optional.

III. Data Analysis and Results

A. Number of Retrieval Practice Attempts

The number of practice tests undertaken had a significant impact on the participant's performance, with n=161, in the Ohms Law/The Basic Circuit topic (Fig. 2). Just over 27%, n=44, did not avail of any of the practice tests provided by the study, with 117 taking practice tests. The no practice test group had a mean score of 74%. The most significant cohort, n=49, attempted of 3 to 4 practice tests, with a mean performance of 83%. A drop in mean performance was observed for participants with nine attempts which was the maximum number of attempts allowed. A One-Way ANOVA with the dependent variable, the performance in the Ohms Law/The Basic Circuit topic in the criterion test, with student engagement with the practice tests as the factor found a significant improvement in performance with $p=.012$, $F=3.035$ between groups.

Of the 161 participants, 118 successfully passed the overall criterion test with 43 referred (not meeting the standard) on the first attempt. For these 118 participants, the

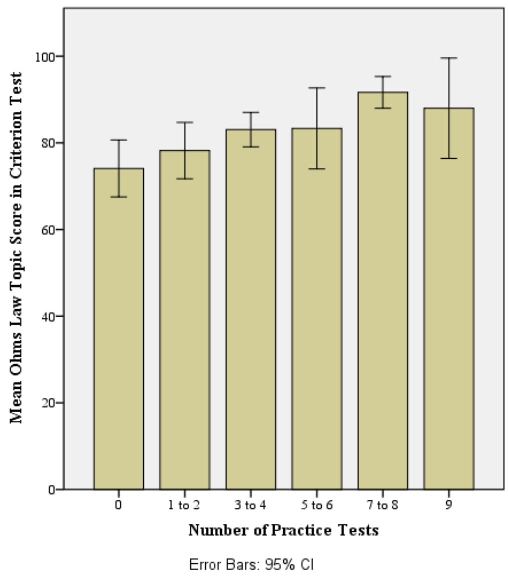

Fig. 2. Mean topic result with number of practice tests for all participants

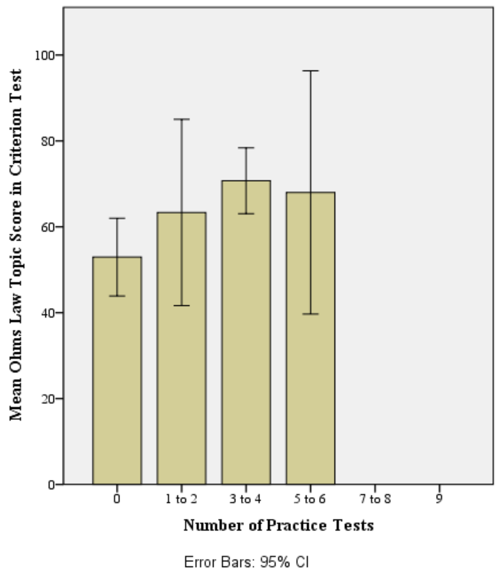

Fig. 3. Mean topic result with number of practice tests for participants that did not pass the overall criterion test

performance overall was high (TABLE 2), with some improvement with the increased number of practice tests. A One-Way ANOVA with the dependent variable, the performance in the Ohms Law/The Basic Circuit topic in the criterion test for the 118 participants, with student engagement with the practice tests as the factor did not find a significant improvement in performance with p=.229, F=1.232 between groups.

TABLE 2 Performance of 118 participants in Ohms Law/The Basic Circuit that passed the overall criterion test

Practice Test Attempts	N	Mean	Std. Dev	Std. Error	95% Confidence Interval for Mean Lower Bound	Upper Bound	Min	Max
0	27	87.4074	10.22538	1.96788	83.3624	91.4524	70	100
1 to 2	22	82.2727	13.42770	2.86279	76.3192	88.2262	60	100
3 to 4	35	88.0000	10.79215	1.82421	84.2928	91.7072	60	100
5 to 6	13	89.2308	13.82120	3.83331	80.8787	97.5828	50	100
7 to 8	11	91.8182	6.03023	1.81818	87.7670	95.8693	80	100
9	10	88.0000	16.19328	5.12076	76.4160	99.5840	60	100
Total	118	87.2881	11.81488	1.08765	85.1341	89.4422	50	100

Of the 43 participants not meeting the criterion test standard on the first attempt (Fig. 3), a One-Way ANOVA with the dependent variable, the performance in the Ohms Law/The Basic Circuit topic, with student engagement with the practice tests as the factor found a significant improvement in performance with p=.038, F=2.824 between groups. When the data is reviewed in terms of no practice testing or with practice testing for this cohort, with 17 not taking the practice tests with a mean of 52.94% and 26 taking practice tests with a mean of 69.23%, a significant improvement in performance is observed with p=.004, F=9.124 between groups.

B. *Massed versus Spaced Practice*

Of the n=117 participants who engaged with the practice tests, n=91 passed the criterion test (TABLE 3), with 26 not passing (TABLE 4).

TABLE 3 Massed/Spaced Practice - Criterion Test Pass

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean Lower Bound	Upper Bound	Min	Max
Massed	36	83.3333	12.42118	2.07020	79.1306	87.5361	60	100
Spaced	55	89.8182	11.62547	1.56758	86.6754	92.9610	50	100
Total	91	87.2527	12.29868	1.28925	84.6914	89.8141	50	100

A One-Way ANOVA with the dependent variable, the performance in the Ohms Law/The Basic Circuit topic in the criterion test, with massed or spaced practice as the factor found a significant improvement in performance for spaced practice with p=.013, F=6.413 between groups with a medium effect size, Cohen's $d = 0.54$.

TABLE 4 Massed/Spaced Practice - Criterion Test Referral

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean Lower Bound	Upper Bound	Min	Max
Massed	10	71.0000	14.49138	4.58258	60.6335	81.3665	50	100
Spaced	16	68.1250	18.69715	4.67429	58.1620	78.0880	40	100
Total	26	69.2308	16.95242	3.32464	62.3835	76.0780	40	100

There was no significant finding between massed and spaced practice for participants who did not meet the criterion test standard (TABLE 4). Participants who massed their practice performed marginally better (71%) than those that spaced their practice (68%) with a small effect size Cohen's *d* = 0.17.

C. Retention Interval (RI)

For participants that passed the criterion test, 25% (n=23) had a RI=0 days, same day as criterion test, 37% (n=34), had a RI=1 to 3 days, 15% (n=14) had a RI=4 to 7 days, and 9% (n=9) had a RI=8 to 14 days. The remaining 14% of participants had a RI ranging from 15 to 21 days (n=4), 22 to 28 days (n=4) and 29+ days (n=3). A One-Way ANOVA did not find a significant difference between groups. A small effect was observed with Cohens *d* effect size = 0.27 for a RI=8 to 14 days compared to RI=0 days.

For participants that did not pass the criterion test, 42% (n=11) had a RI=0 days, same day as criterion test, 34% (n=9), had a RI=1 to 3 days, 12% (n=3) had a RI=4 to 7 days and 12% (n=3) had an RI=8 to 14 days. While a One-Way ANOVA did not find a significant difference with p=.08, F=2.568 between groups, a large effect was observed with Cohens *d* = 0.82 for a RI=1 to 3 days compared to a RI=0 days.

D. Practice Test Performance

Overall performance in the practice tests had a significance impact in the criterion test result. Participants who had a maximum score in their practice tests of less than 69%, n=15, had a mean performance in the criterion test topic of 70 while participants who scored 70% or greater, n=102, had a mean performance of 85.2 in the criterion test topic. A One-Way ANOVA with the dependent variable, the performance in the Ohms Law/The Basic Circuit topic in the criterion test, with student performance on the practice tests as the factor found a significant improvement in performance with p=.000, F=14.264 between groups. A large effect was observed with Cohens *d* = 0.90.

For participants that passed the criterion test, who had a maximum score in their practice tests of less than 69%, n=7, had a mean performance in the criterion test topic of 77.14 while participants who scored 70% or greater, n=84, had a mean performance of 88.09 in the criterion test topic. A One-Way ANOVA with the dependent variable, the performance in the Ohms Law/The Basic Circuit topic in the criterion test, with student performance on the practice tests as the factor found a significant improvement in performance with p=.023, F=5.373 between groups. A medium effect was observed with Cohens *d* = 0.70.

For participants that did not pass the criterion test, who had a maximum score in their practice tests of less than 69%, n=8, had a mean performance in the criterion test topic of 63.75 while participants who scored 70% or greater, n=18, had a mean performance of 71.67 in the criterion test topic. A One-Way ANOVA found no significant improvement in performance with p=.281, F=1.218 between groups. A small effect was observed with Cohens *d* = 0.45.

E. Results of Survey

There were 40 respondents to the online survey of which 36 reported they participated in practice tests and 4 did not. Two respondents stated they did not know about the practice tests while the other two were aware but did not feel they were beneficial enough from them to participate.

Effectiveness of Study Techniques – Dunlosky et. al identified a number of techniques with varying degrees of effectiveness for learning [1]. These techniques were used as the basis for asking the respondents to rate how effective they thought the study methods they used were. The results are illustrated in Fig 5. Practice testing was rated highest with a weighted average of 3.67 followed closely by highlighting/underlining (3.61), with rereading notes (3.22), distributed practice (3.19) and summarizing (3.17). The high rating of known ineffective techniques such as highlighting, and rereading is consistent with existing research with learners perceiving these to be effective.

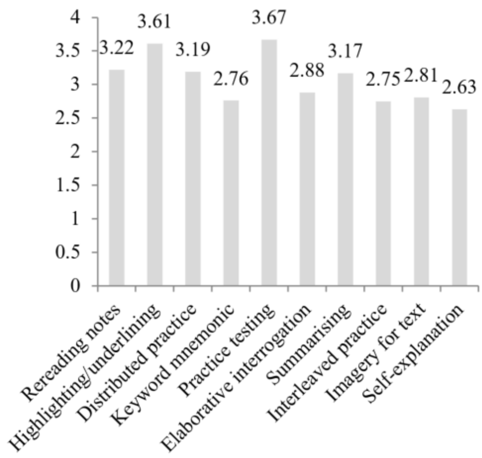

Fig. 4. Percieved effectiveness of study techniques

Difficulty of topics – Respondents, when asked how difficult the topics they studied were, found Power and Energy the most difficult with a weighted average of 2.8, Ohms Law/The Basic Circuit and Resistance Network Measurement following closely behind (2.63). 95% of respondents found the practice tests to be either helpful or very helpful.

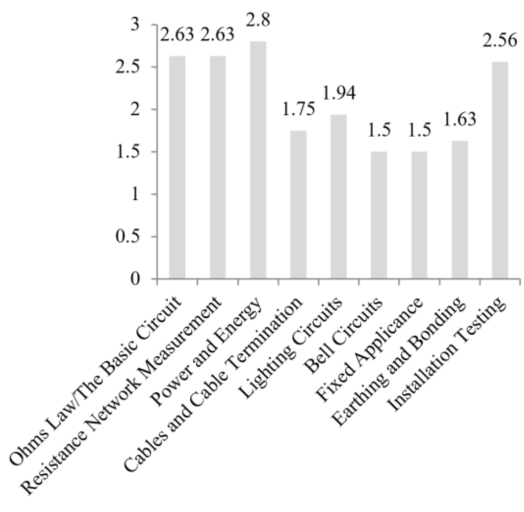

Fig. 5. Topic difficulty as reported by respondents

Impact of Practice Testing - When asked how the practice tests affected their preparation for the criterion test three main themes emerged; (1) Identifying gaps in knowledge, (2) Better understanding of the topic and (3) Preparation for type and format of questions.

IV. Discussion

This paper investigated e-assessment practice testing within a Practice Testing Learning Framework (PTLF) in electrical science and how participants employed the practice tests to support their learning. A discussion follows on the implications of retrieval practice in electrical engineering education, the impact of parameters on learning transfer and the contribution of the PTLF to practice testing research.

A. Practice Testing Learning Framework

While a substantial body of research supports the benefits of retrieval practice there are challenges to converting that research to practice. The PTLF addresses many of these challenges by combining the Conversational Framework [37] with a practice testing development approach using the CR matrix taxonomy model. The PTLF has now been deployed in undergraduate courses, including mathematics for computing [36] and statistics [41] and in apprenticeship in electrical science. For practice testing to be implemented effectively within classroom contexts, educationalists should consider applying a test development approach which supports alignment between practice test items and course learning outcomes which is better achieved by employing the CR matrix compared to using a single taxonomy model such as Blooms taxonomy.

B. Parameters

This study explored several parameters and their impact on learning transfer.

Number of retrieval practice attempts – The number of retrieval practice attempts had a significant impact on performance in Ohms Law/The Basic Circuit on the criterion test. The largest cohort, n=49, availed of 3 to 4 practice tests, with a mean performance of 83%. A drop in mean performance was observed for participants with nine attempts which was the maximum number of attempts allowed. As the practice tests were provided as an optional resource, just over 27%, n=44, did not avail of any of the practice tests. This group had a mean score of 74%. This cohort can be divided into two groups, those that passed the criterion test n=27 and those that did not pass n=17. Results from the survey indicated that members of this cohort felt the practice tests were not beneficial or were unaware of the practice tests.

Overall, the findings indicate 3 to 4 practice tests as an optimum number of practice tests for all learners, learners with higher levels of self-regulation with a mean score = 88% and learners with lower levels of self-regulation with a mean score of 70.71%.

Spaced versus Massed Practice - For participants who demonstrated higher levels of self-regulation in passing the criterion test a significant improvement in performance and a medium effect size, Cohen's d = 0.54, was observed when the retrieval practice attempts were spaced. This finding is consistent with spaced practice research. An opposite effect was observed for participants with lower levels of self-regulation with those massing their practice performing marginally better that those that spaced their practice with a Cohen's d = 0.17.

Retention Interval (RI) - For participants that passed the criterion test there was no significant difference between groups with small effect observed with Cohens $d = 0.27$ for a RI=8 to 14 days compared to a RI=0 days. For participants that did not pass the criterion test, a large effect was observed with Cohens $d = 0.82$ for a RI=1 to 3 days compared to a RI=0 days. These findings suggest that taking a practice test on the day of the criterion test is not as effective as taking the last retrieval practice several days before hand.

Practice Test Performance – Overall performance in the practice tests had a significance impact in the criterion test topic result with a large effect was observed with Cohens $d = 0.90$. This finding indicates that by maximizing performance during practice tests benefits learning and is a strong indicator of future performance in criterion test.

For participants that did not pass the overall criterion test a small effect was observed with Cohens $d = 0.45$. While this cohort did not engage in as many tests as more regulated participants, their performance in the practice tests tended to reflect their performance in the criterion test topic.

V. Conclusion and future work

In conclusion, and reflecting on the research questions, retrieval practice within the PTLF enhances performance in electrical science and benefits learning. The PTLF can be adapted by educators to their own teaching practice. The application of the PTLF in mathematics, statistics and electrical science demonstrates its use in STEM domains in traditional and flipped classrooms.

Determining optimum parameters that can be generalized presents challenges with many variables. From this study, taking 3 to 4 practice tests, maximizing performance when taking the tests, spacing the practice tests over time and leaving a retention interval of a number of days prior to the criterion test is recommended for electrical science.

Future work will explore the parameters further with different cohorts and in new electric science topics within the PTLF. The impact of the PTLF on learner self-regulation and the learner experience, their perceived and realized benefits will also be examined. The potential future application of the PTLF to distance learning and MOOC's and how this work may be generalized to engineering and STEM subjects will make a significant contribution to the body of research.

References

[1] J. Dunlosky, K. A. Rawson, E. J. Marsh, M. J. Nathan, and D. T. Willingham, "Improving Students' Learning With Effective Learning Techniques: Promising Directions From Cognitive and Educational Psychology," Psychol. Sci. Public Interes., vol. 14, no. 1, pp. 4–58, Jan. 2013.

[2] J. D. Karpicke, Retrieval-Based Learning : A Decade of Progress, Third Edit. Elsevier, 2017.

[3] H. L. I. Roediger and J. D. Karpicke, "The Power of Testing Memory Basic Research and Implications for Educational Practice.," Perspect. Psychol. Sci., vol. 1, no. 3, pp. 181–210, Sep. 2006.

[4] H. L. I. Roediger, A. L. Putnam, and M. A. Smith, "Ten benefits of testing and their applications to educational practice.," in The

psychology of learning and motivation (Vol 55): Cognition in education., vol. 55, J. P. Mestre and B. H. Ross, Eds. San Diego, CA US: Elsevier Academic Press, 2011, pp. 1–36.

[5] J. D. Karpicke and H. L. I. Roediger, “The critical importance of retrieval for learning.,” Science (80-.)., vol. 319, no. 5865, pp. 966–968, Feb. 2008.

[6] S. K. Carpenter, “Cue strength as a moderator of the testing effect: The benefits of elaborative retrieval.,” J. Exp. Psychol. Learn. Mem. Cogn., vol. 35, no. 6, pp. 1563–1569, Nov. 2009.

[7] H. L. I. Roediger and A. C. Butler, “The critical role of retrieval practice in long-term retention.,” Trends Cogn. Sci., vol. 15, no. 1, pp. 20–27, Jan. 2011.

[8] K. Rawson and J. Dunlosky, “When Is Practice Testing Most Effective for Improving the Durability and Efficiency of Student Learning?,” Educ. Psychol. Rev., vol. 24, no. 3, pp. 419–435, Sep. 2012.

[9] M. Smith and J. Karpicke, “Retrieval practice with short-answer, multiple-choice, and hybrid tests,” Memory, no. September, pp. 37–41, Sep. 2013.

[10] H. L. I. Roediger and J. D. Karpicke, “Test-Enhanced Learning: Taking Memory Tests Improves Long-Term Retention.,” Psychol. Sci., vol. 17, no. 3, pp. 249–255, Mar. 2006.

[11] J. D. Karpicke and J. R. Blunt, “Retrieval Practice Produces More Learning than Elaborative Studying with Concept Mapping.,” Science (80-.)., vol. 331, no. 6018, pp. 772–775, Feb. 2011.

[12] M. A. McDaniel, P. K. Agarwal, B. J. Huelser, K. B. McDermott, and H. L. I. Roediger, “Test-Enhanced Learning in a Middle School Science Classroom: The Effects of Quiz Frequency and Placement.,” J. Educ. Psychol., vol. 103, no. 2, pp. 399–414, May 2011.

[13] S. K. Carpenter, H. Pashler, and N. J. Cepeda, “Using tests to enhance 8th grade students’ retention of US history facts,” Appl. Cogn. …, vol. 23, no. 6, pp. 760–771, Sep. 2009.

[14] H. L. I. Roediger, P. K. Agarwal, M. A. McDaniel, and K. B. McDermott, “Test-Enhanced Learning in the Classroom: Long-Term Improvements from Quizzing,” J. Exp. Psychol. Appl., vol. 17, no. 4, pp. 382–395, Dec. 2011.

[15] H. L. I. Roediger and E. J. Marsh, “The Positive and Negative Consequences of Multiple-Choice Testing.,” J. Exp. Psychol. Learn. Mem. Cogn., vol. 31, no. 5, pp. 1155–1159, Sep. 2005.

[16] K. B. McDermott, P. K. Agarwal, L. D’Antonio, H. L. I. Roediger, and M. a McDaniel, “Both multiple-choice and short-answer quizzes enhance later exam performance in middle and high school classes.,” J. Exp. Psychol. Appl., vol. 20, no. 1, pp. 3–21, Mar. 2014.

[17] T. M. Haladyna, Developing and Validating Multiple-Choice Test Items, vol. 3. New Jersey: Lawrence Erlbaum Associates, 2004.

[18] M. A. McDaniel, R. C. Thomas, P. K. Agarwal, K. B. McDermott, and H. L. I. Roediger, “Quizzing in Middle-School Science: Successful Transfer Performance on Classroom Exams,” Appl. Cogn. Psychol., vol. 27, no. 3, pp. 360–372, May 2013.

[19] J. L. Jensen, M. A. McDaniel, S. M. Woodard, and T. A. Kummer, “Teaching to the test…or testing to teach: Exams requiring higher order thinking skills encourage greater conceptual understanding.,” Educ. Psychol. Rev., Jan. 2014.

[20] R. E. Mayer et al., “Using Technology-Based Methods to Foster Learning in Large Lecture Classes: Evidence for the Pedagogic Value of Clickers,” 2009.

[21] C. L. Wooldridge, J. M. Bugg, M. A. McDaniel, and Y. Liu, “The Testing Effect with Authentic Educational Materials: A Cautionary Note,” J. Appl. Res. Mem. Cogn., vol. 3, no. 3, pp. 214–221, Jul. 2014.

[22] C. Wiklund-Hörnqvist BertNyberg, Lars, “Strengthening concept learning by repeated testing.,” Scand. J. Psychol., vol. 55, no. 1, pp. 10–16, Feb. 2014.

[23] A. C. Butler, “Repeated testing produces superior transfer of learning relative to repeated studying.,” J. Exp. Psychol. Learn. Mem. Cogn., vol. 36, no. 5, pp. 1118–1133, Sep. 2010.

[24] S. K. Carpenter and E. L. DeLosh, “Application of the testing and spacing effects to name learning.,” Appl. Cogn. Psychol., vol. 19, no. 5, pp. 619–636, Jul. 2005.

[25] H. S. Sobel, N. J. Cepeda, and I. V Kapler, “Spacing effects in real-world classroom vocabulary learning.,” Appl. Cogn. Psychol., vol. 25, no. 5, pp. 763–767, Sep. 2011.

[26] C. E. Küpper-Tetzel, I. V Kapler, and M. Wiseheart, “Contracting, equal, and expanding learning schedules: The optimal distribution of learning sessions depends on retention interval.,” Mem. Cognit., vol. 42, no. 5, pp. 729–41, 2014.

[27] K. A. Rawson and J. Dunlosky, “Optimizing schedules of retrieval practice for durable and efficient learning: How much is enough?,” J. Exp. Psychol. Gen., vol. 140, no. 3, pp. 283–302, Aug. 2011.

[28] E. J. Marsh, H. L. I. Roediger, R. a Bjork, and E. L. Bjork, “The memorial consequences of multiple-choice testing,” Psychon. Bull. Rev., vol. 14, no. 2, pp. 194–199, Apr. 2007.

[29] B. S. Bloom, Taxonomy of Educational Objectives, vol. 16. 1956.

[30] L. W. Anderson and D. R. Krathwohl, A taxonomy for learning, teaching, and assessing: A revision of Bloom’s taxonomy of educational objectives, vol. Complete e. 2001.

[31] V. A. Jideani and I. A. Jideani, “Alignment of Assessment Objectives with Instructional Objectives Using Revised Bloom’s Taxonomy-The Case for Food Science and Technology Education.,” J. Food Sci. Educ., vol. 11, no. 3, pp. 34–42, Jul. 2012.

[32] N. Webb, “Depth-of-knowledge levels for four content areas,” Lang. Arts, 2002.

[33] K. K. Hess, B. S. Jones, D. Carlock, and J. R. Walkup, “Cognitive Rigor: Blending the Strengths of Bloom’s Taxonomy and Webb’s Depth of Knowledge to Enhance Classroom-Level Processes,” Online Submission, Mar. 2009.

[34] D. Laurillard, Rethinking University Teaching, vol. 2nd. 2002.

[35] M. D. Merrill, “First principles of instruction,” Educational Technology Research and Development, vol. 50, no. 3. pp. 43–59, 2002.

[36] J. Eustace, M. Bradford, and P. Pathak, “A Practice Testing Learning Framework to Enhance Transfer in Mathematics,” in The 14th Information Technology & Telecommunications Conference, 2015, pp. 88–95.

[37] D. Laurillard, Rethinking university teaching: A conversational framework for the effective use of learning technologies, vol. 2nd. 2002.

[38] D. Laurillard, “The pedagogical challenges to collaborative technologies.,” Int. J. Comput. Collab. Learn., vol. 4, no. 1, pp. 5–20, Mar. 2009.

[39] N. Pachler, C. Daly, Y. Mor, and H. Mellar, “Formative e-assessment: Practitioner cases,” Learn. Digit. Worlds Sel. Contrib. from CAL 09 Conf., vol. 54, no. 3, pp. 715–721, 2010.

[40] M. Neo, K. T.-K. Neo, and S. T. Lim, “Designing a web-based multimedia learning environment with laurillard’s conversational framework: an investigation on instructional relationships,” TOJET Turkish Online J. Educ. Technol., vol. 12, no. 3, pp. 39–50, 2013.

[41] J. Eustace and P. Pathak, “Enhancing Statistics Learning within a Practice Testing Learning Framework,” in ICERI2017, 2017, pp. 1128–1136.

[42] B. J. Zimmerman, “Becoming a Self-Regulated Learner: An Overview,” Theory Pract., vol. 41, no. 2, p. 64, 2002.

[43] E. R. Lai, “Metacognition : A Literature Review Research Report,” no. April, pp. 1–49, 2011.

[44] K. Thiede, J. Wiley, and T. Griffin, “Test expectancy affects metacomprehension accuracy,” Br. J. Educ. …, vol. 81, no. 2, pp. 264–273, Jun. 2011.

[45] P. K. Agarwal, L. D’Antonio, H. L. I. Roediger, K. B. McDermott, and M. A. McDaniel, “Classroom-based programs of retrieval practice reduce middle school and high school students’ test anxiety,” J. Appl. Res. Mem. Cogn., Jul. 2014.

[46] M. A. Smith, H. L. I. Roediger, and J. D. Karpicke, “Covert retrieval practice benefits retention as much as overt retrieval practice.,” J. Exp. Psychol. Learn. Mem. Cogn., vol. 39, no. 6, pp. 1712–1725, Nov. 2013.

[47] C. I. Johnson and R. E. Mayer, “A testing effect with multimedia learning.,” J. Educ. Psychol., vol. 101, no. 3, pp. 621–629, Aug. 2009.

[48] J. Tullis, J. Finley, and A. Benjamin, “Metacognition of the testing effect: Guiding learners to predict the benefits of retrieval.,” Mem. Cognit., vol. 41, no. 3, pp. 429–442, Apr. 2013.

[49] S. K. Carpenter, “Testing Enhances the Transfer of Learning,” Curr. Dir. Psychol. Sci., vol. 21, no. 5, pp. 279–283, Oct. 2012.