•

著者	SAINOO-FULLER Sebastian
journal or publication title	The Journal of Nagasaki University of Foreign Studies
number	12
page range	161-178
year	2008-12-30
URL	http://id.nii.ac.jp/1165/00000184/

Assessment regimes considered: the effect stress has on memory recall, performance in online and paper testing.

Sebastian SAINOO-FULLER

概　要

この研究は、特に学習者のテスト行為とその考え方に注目して、高等教育レベルの学生の経験的知識へ、オンライン評価素材が及ぼす影響について考察したものである。考察の結果、総合的評価においてはオンラインと筆記に測定可能な違いが見られないことを示唆する。実際には、オンラインでテストを受験する学生は、テストへの脅威感がより低いことが分かった。形成的評価を考慮すれば、講義の評価テスト前にオンラインで練習問題をしておけば、若干利点があると思われる。学生はテストのフォーマットだけでなく、事前に示されたテストの内容に慣れるのである。本稿は、オンラインテストを導入すれば、学生がその科目のテストの準備をするのに役立つことを提案し、また、安全なWebベースのテストは、学生の自信を高め、授業時間を増やすことを通じて、学部生の教育に一助となることを明らかにするものである。

Abstract:

This study considered the effects of online assessment materials on tertiary level students' experiences focusing specifically on learners' testing behaviors and beliefs. The results suggest that there is no significantly measurable difference in when comparing online to paper-pencil summative assessments. In fact, students taking tests online reported lower levels of perceived test threat. Regarding formative assessment, it is suggested that a small benefit for using online practice tests prior to graded course exams. The students become familiar with the prescribed content of the test as well as the testing format. The results support the integration of online practice tests to help students prepare for course exams and also reveal that secure web-based testing can aid undergraduate instruction through improved student confidence and increased classroom time.

Introduction

The application of the internet to supply students with access to course materials has become standard practice for instructing undergrad-uates (Duchastel, 1996). Online materials often include links to a syllabus, an outline of prescribed topics, materials, and links to monitored chat rooms (Wheeler, 2000). However, recent developments with user-friendly web-based assessment packages such as Mozilla and secure Internet testing have encouraged many educators to design online quizzes, tests and assignments. Although there is great enthusiasm among educators regarding the opportunities for online application of both formative and summative assessment materials, there is little evidence considering the impact of web-based assessment practices on student performance (Buchanan, 1998; 2000). Furthermore, the impact of online testing on students' attitudes and anxieties is an under-considered topic. This investigation assessed undergraduate students' experiences within the context of a course utilizing online

assessments. In particular, two primary research questions were examined: (1)What differences are there in students' perceptions and performances for graded tests based on the format of delivery (online vs. paper test)?; and (2) How are undergraduate students' experiences uniquely influenced by the provision of online formative assessments (practice quizzes)?

Perhaps the most extensive body of research that has explored the experience of learners in various testing conditions comes from the test anxiety literature, which has detailed a variety of conditions and criteria that tend to positively or negatively influence academic test performance. One generality in this body of research is that understanding students' experiences with tests is facilitated when viewing the entire learning and testing process as a recursive cycle. Three phases are included in the L-T Cycle: test preparation (forethought), test performance, and test reflection (Schutz & Davis, 2000; Zeidner, 1998). Students with high levels of cognitive test anxiety and other negative test perceptions have difficulty operating in all three of these phases (Cassady, 2004b). The conclusion from this line of research has been that the beliefs and behaviors students maintain during each of these phases directly influence performance. The current study targeted students' experiences in the test preparation and performance phases, and used the established framework of the L-T Cycle to investigate theoretical benefits and drawbacks related to online testing.

Test Preparation

In the test preparation phase, students with high levels of cognitive test anxiety tend to procrastinate, worry over potential failure, utilize ineffective study strategies, and demonstrate insufficient cognitive processing skills to gain effective conceptual understanding for the content (Cassady, 2004b; Culler & Holohan, 1980; Hembree, 1988; Wittmaier, 1972). There is evidence that students with test anxiety develop these patterns due to deficient abilities in effectively predicting and encoding prescribed content (Cassady, 2004a; Naveh-Benjamin et al., 1987), with some research pointing directly to the articulatory processing loop, which controls verbal processing in working memory (Ikeda, Iwanaga, & Seiwa, 1996). These pervasive processing failures have been explained through skill deficit models, where the students simply have not developed the necessary strategies to encode, organize, and store the materials at hand (e.g., Naveh-Benjamin et al., 1987). Training the learner to employ effective strategies for test preparation should alleviate such a skill deficit, and consequently promote higher test performance for students who have a history of test anxiety and test failure. The L-T Cycle framework predicts that once a student gains an effective study strategy for encoding and storing core content, the traditional deleterious effects of test anxiety will be less dramatic because the student will recognize the content is accessible and the self-deprecating ruminations and coping strategies such as procrastination and task avoidance will be less readily activated (Cassady, 2004b). Another proposition for helping learners overcome the effects of cognitive test anxiety is to reduce the perceived threat of an evaluative event. For example, Cassady (2004a) found that under conditions where there was no external evaluative pressure (i.e., ungraded tests of memory in a laboratory setting), the influence of test anxiety on performance was significantly lower than in conditions of high external evaluative pressure (college entrance exams). This pattern of results indicates that when the evaluative stress is removed, the processing deficits are attenuated, supporting the proposition that the test anxious learner has the basic cognitive skills to encode, organize, and store core content.

This study was designed to extend the laboratory-based finding with contrived materials to a realistic educational setting by providing ungraded practice tests as a test preparation strategy available to learners in educational psychology courses.

Test Performance

The classic view of test anxiety has been focused on the test performance phase, where learners fail to perform well due to task interference. This interference can take many forms, including: (a) sudden, inexplicable loss of previously mastered information at the time of testing (Covington & Omelich, 1987); (b) interfering self-deprecating ruminations (Sarason, 1986); (c) distracting thoughts of failure brought on by feelings of threat to self imposed by the test (Cassady, 2002; 2004b; Schwarzer & Jerusalem, 1992); or (d) physiological reactions that impair stable cognitive action (e.g., headache, perspiration, heart palpitation; Sarason, 1986). These distracters during the testing event naturally reduce the ability of the learner to effectively locate and use relevant information stored in long-term memory.

Contemporary views of test anxiety have demonstrated additional problems in the performance phase for those test-anxious students with poor study skills (e.g., Naveh-Benjamin et al., 1987). These students face additional difficulty because the encoding and storage processes in the test preparation phase have been adversely affected as well, significantly reducing the probability of competent performance under pressure.

To reduce the impact of test anxiety and related test perceptions on test performance, the use of practice tests in an instructional program can serve two purposes: (a) provide ungraded testing experiences that serve as effective test preparation activities and (b) provide non-threatening practice exams that build student confidence through repeated attempts and presumed success with realistic testing materials. In this study, online presentation of practice tests was used as a simplified means to make practice tests consistently and readily available to students.

Online Formative and Summative Assessment

There is a limited research base on the use of online tools to deliver formative and summative assessments. However, the research base on traditional testing formats is relevant and provides insight into the experiences of learners. To frame the theoretical framework for this study, we present the literature demonstrating that (a) formative assess-ments can serve as effective test preparation events, (b) providing multiple formative assessments can influence learners' test perceptions, and (c) migrating traditional multiple-choice tests to an online testing protocol provides no universal performance or perception variances.

Impact of Formative Assessment on Learning and Achievement

The decision to use formative assessment in instruction is typically motivated by an attempt to provide the instructor with an accurate estimation of student ability at a particular point in the course, or to provide the students with an assessment task similar in nature to the summative test (Buchanan, 1998). This allows the student to identify strengths and weaknesses and to better prepare for the "real" exam. One of the great advantages of online test programs is the ability to deliver practice tests that serve as formative assessment tools for the students. Practice tests have been shown to increase students' final outcome performance by roughly twelve percent (Bocij & Greasley,

1999; also see Carrier & Pashler, 1992; Dempster, 1997; Glover, 1989; McDaniel, Kowitz, & Dunay, 1989). Delivering practice tests online may provide an additional benefit to the student by allowing her or him to complete the test conveniently without the environmental distractions that are common during in-class practice tests. Because different conceptualizations for "practice test" or "practice quiz" are common, there are dramatically different educational, cognitive, and theoretical implications when employing the different strategies of practice testing; thus, operationalization is key. In this discussion, unless otherwise noted, practice quizzes and formative assessments refer to assessment tools that are completed by students prior to a summative (graded) assessment. These practice tests are similar to summative assessments in format and difficulty level, but do not impact the students' course grade and are comprised of a different set of items. The utility of formative assessment is partly reliant upon the manner through which the feedback is provided to the learner. The most desirable feedback approach appears to be immediate post-performance reporting, which provides feedback directly after the entire quiz or test has been completed (King & Behnke, 1999). This method takes advantage of a primary benefit of computer-assisted assessment by supplying timely feedback (Clariana, Ross, & Morrison, 1991; Jongekrijg & Russell, 1999), while avoiding the problem of inducing anxiety or distraction that can arise when providing performance indicators directly after each item (Wise, Plake, Eastman, Boettcher, & Luken, 1986; Wise, Plake, Pozehl, Barnes, & Lukin, 1989). The anxiety induced in item-by-item feedback has been shown to hamper performance through motivational processes such as learned helplessness or externalized attributions of control over performance (Boggiano & Ruble, 1986).

Formative Assessment and Students' Perceptions of Tests

The benefits of repeated formative assessment for students are likely to rest in their perceptions of test preparedness for the summative measure. Bandura (1986) proposed repeated exposure to successful testing experiences for students with high anxiety would promote self-efficacy for later tests. The use of formative assessments (where no evaluative pressure is imposed) as practice for tests is likely to increase the probability that students will have a positive experience in the testing event with respect to anxiety. In these formative assessment experiences, perceived threat, self-awareness, cognitive test anxiety, and emotionality should all be lower than in standard summative assessment sessions (Kurosawa & Harackiewicz, 1995; Schwarzer & Jerusalem, 1992). With the suppression of these affective detractors, the student is more likely to be able to benefit from self-regulatory processes in the practice testing session, leading to higher performance, growth, and subsequent success (Bandura, 1986; Schutz & Davis, 2000).

Online Summative Assessment

Summative assessment in an online environment differs in form and function from the formative assessment process. Not only are the summative assessments graded, but the methods through which students access and respond to the tests usually differ. The summative assessment process requires high levels of control and security in the testing process to ensure reliability and validity in scores, attention to technical problems that may arise during the testing session, and assurance that the online nature of the testing process itself has no impact on actual performance. An additional concern that is often raised by instructors considering online summative assessment is that online testing will induce heightened levels of anxiety over the test, leading to performance levels that underestimate true ability.

The advantages for providing course tests online can include flexibility in delivering tests to students and efficiency in scoring, depending upon the method of delivery chosen by the instructor. With the online delivery of tests, students are not necessarily bound by the traditional artificial academic scheduling constraints. Specifically, (a) they can complete exams at different times of the day to fit their convenience; (b) they can potentially complete the tests in different locations if the test is not a required "closed-book" exam; and (c) unless there is an explicit reason for a time limit, students can take as long as needed to complete the exam. In a similar line, an additional benefit that can be gained through online summative assessment is that additional class time may be gained in traditional on-campus courses. That is, rather than taking a class period to have the students complete the course exam, the instructor can use the class period for instruction.

In perhaps the most complete examination of online summative assessment to date, Bocij & Greasley (1999) reported that students claimed online testing was superior because they were less distracted with the process of handwriting their responses, which helped them maintain focus on the test items and were less panicked. The lower levels of panic were impacted in part by the fact that online tests took less time to complete. Students in Bocij & Greasley's (1999) work reported the tests were fair, unbiased, and "less threatening than conventional examinations" (p. 14). Finally, the authors reported that performance gains were noted in the online testing conditions, but these effects were not present for the high ability students who appeared to be unaffected by test delivery format.

Current Investigation

As mentioned above, this investigation addressed two research questions. The first was a comparison of the effect of delivering course exams online versus in class on paper. This portion of the study involved examining the affective experiences of one instructor's students. The students were enrolled in the same course, separated by one year. The only evaluative difference existing between the two courses was the method of delivering the course exams. For the first group of students, all tests were delivered in class on paper. For the second group, all tests were delivered online in a computer-based testing laboratory staffed by testing proctors who ensured the security of the testing process and corrected any technical issues that arose. Students' levels of cognitive test anxiety, emotionality, and perceived threat of tests were compared to determine if there were differential perceptions of tests for students experiencing the two alternate methods of test delivery. These data were intended to examine the extent to which online testing leads to heightened levels of fear, anxiety, or worry over tests. The hypothesis underlying this question was that the method of presentation would have no meaningful detrimental impact for the students in any of these variables.

The second part of the study examined the relationships among the use of online formative assessments, student performance, and test perceptions. For both groups of students, online practice tests were made available as a test preparation option for only the third exam. It was expected that the students using online formative assessment tests (as practice) would have higher rates of performance on subsequent summative assessment measures. Due to the differential patterns of behavior and performance traditionally noted in students with test anxiety based in part on study strategies (Naveh-Benjamin, McKeachie, & Lin, 1987), no a priori predictions regarding the relationship between online formative assessment and test perceptions were reasonable.

Method

Undergraduate students in introductory educational courses were the participants in this investigation. Participants were drawn from intact classes of students enrolled in the same university during the six months of testing. Eighty-four undergraduate students participated in the in-class testing group in the fall of 2006. In the in-class testing group there were 74 females and 10 males, which was representative of the population in the elementary education program that the courses served. Ninety-two participants were included in the online testing group in the fall of 2006. There were 24 males and 68 females in the online testing condition. The participants in the study were all volunteers; participation in the study served as one of many options to complete a course requirement on professional research.

Test Anxiety

Test anxiety research has repeatedly validated the existence of two interrelated factors commonly referred to as worry and emotionality (Hembree, 1988). Although over two decades of research has confirmed the presence of both factors, there is clear evidence that the cognitive factor has the most direct negative impact on test performance (Deffenbacher, 1980; Sarason, 1986). The term "cognitive test anxiety" refers to the wide variety of thoughts and beliefs that can impair performance either during a learner's attempts to prepare for or take an examination (Cassady, 2004b). These cognitive barriers include (a) comparing self-performance to peers, (b) considering the consequences of failure, (c) low levels of confidence in performance, (d) excessive worry over evaluation, (e) feeling unprepared for tests, or (f) limitations in retrieval cues utilization (Deffenbacher, 1980; Geen, 1980; Hembree, 1988; Morris, Davis, & Hutchings, 1981; Sarason, 1986). The Cognitive Test Anxiety scale (Cassady & Johnson, 2002) is a 27-item instrument focused on only the cognitive domain of test anxiety. Students respond to the items on this instrument using a four-point Likert-type scale ("Not at all typical of me," "Only somewhat typical of me," "Quite typical of me," "Very typical of me"). Previous research with this instrument has demonstrated high internal consistency (alpha >.90) as well as construct stability as measured by test-retest consistency at three administration periods (beginning, middle, end of academic semester, r's 0.88 to 0.93) (Cassady, 2001b). To measure cognitive test anxiety, the Cognitive Test Anxiety scale was completed by all students no more than 2 days prior to the taking of the third examination. The timing of the test administration was determined by prior investigations with similar samples (Cassady, 2004b) that demonstrated students had sufficient experience with the course testing procedures to have an adequate understanding of the specific test conditions and procedures for the given course. The second factor of test anxiety is known as emotionality (Liebert & Morris, 1967). This factor is the individual's subjective awareness of heightened autonomic arousal during examinations (Schwarzer, 1984). To measure the emotionality component of test anxiety, the Bodily Symptoms subscale of Sarason's (1984) Reactions to Tests was administered. This 10-item scale addresses students' self-perceived physiological reactions during tests (e.g., sweating, increased heart rate, headache). The students responded to the items using the same response scale as the Cognitive Test Anxiety scale.

Perceived Test Threat

The Perceived Threat of Tests is an 18-item self-report instrument that focuses on the perception of the upcoming test as threatening, either due to general difficulty of course content or personal barriers to success on the test (Cassady,

2004b). Participants respond to a four-point Likert-type scale, with responses ranging from strongly disagree to strongly agree. Select items are reverse-coded such that high values on the Perceived Threat of Tests instrument reveal high levels of perceived threat.

Test Preparation Strategies

An 8-item study skills survey was also used in this investigation to gather self-report information on the students' study habits and strategies using the same response options as in the Cognitive Test Anxiety scale (Cassady, 2004b). The items assessed students' chosen study activities as well as their perceived ability with test preparation strategies (e.g., reading comprehension and task focus). A combined score for the study skills items represents an overall study efficacy rating from the student, with a high score indicating they rate themselves highly on positive test preparation activities.

Use of the online practice tests was also coded as an indicator of individuals' test preparation activities. For the paper-based testing group, students self-reported the use of the practice tests in response to a dichotomous (yes-no) query after the third exam. Advances in available online courseware in the fall of 2000 enabled tracking of individual users for the online testing group. Thus, for that group only, actual number of times each participant accessed practice tests was available. Because the paper-based testing group data were self-reported and did not meet the assumption of interval data, the main analyses exploring the impact of online practice tests were conducted on data collected only from the online testing group.

Procedures

Students in the in-class testing group took four tests during the semester, including one comprehensive examination. The first three tests of the semester are the focus of this investigation, given the unique nature of final examinations regarding content coverage and student preparation (see Cassady & Johnson, 2002 for detail). The three tests were each completed during 75-minute class sessions in the regular course meeting room. The instructor was present for the exam administration. The tests were multiple-choice exams ranging in length from 32 to 36 items, with an average difficulty index (the percentage of test takers correctly answering the item) of 0.76. Two days prior to taking the third exam, students in the study completed the self-report instruments. This contrived timing of data collection was intended to provide sufficient situational anxiety to capture heightened rates of perceived threat and emotionality (Cassady, 2004b). Logistic and ethical concerns prevented completing the scales on the day of testing. Logistically, there was no reliable time for the students to all complete the items directly prior to the test and maintain sufficient time to complete the exam items. Ethically, it is conceivable that completing the cognitive test anxiety scale or perceived test threat measure would induce additional anxiety that could have a detrimental impact on performance if taking the test immediately thereafter. The students in the online testing sample also took four exams, including one comprehensive examination. The tests differed slightly in content due to differences between the courses. However, the tests were also multiple choice tests of similar length with an average difficulty index of 0.74. The students in this sample took all exams in a secured computer-based testing laboratory at their convenience, determining at which point during a 7-day period they would complete the exam. Tests were proctored by a laboratory assistant, who logged students onto the proper test and ensured the security of the testing session. The computer-based testing laboratory was accessible during the weekends, and until midnight every day for student use. Students in this sample completed

the test anxiety and perceived test threat instruments no more than two days prior to taking the test (completing the surveys online, with date stamping to ensure the appropriate time lapse).

Online Formative Assessments

For both semesters, online practice tests1 were made available to students after the second exam, as an additional test preparation option. The practice tests were announced in class as well as through the online course management system. All practice tests were created to provide related (but not identical) items for student preparation for the course exams. There were four practice tests offered to the students, with each test providing no less than 10 items targeting one of the chapters covered in the third course exam. Starting four weeks prior to the third exam, students had freedom to access the practice tests at any time, as many times as desired.

Results

The results are organized to present the analyses centering on the two primary questions. First, is there a meaningful difference between the paper-based and online-testing groups in test perceptions and performance? Secondly, what unique contribution to student performance does using online practice tests provide when simultaneously accounting for prior performance and test perceptions?

Online vs. In-class Summative Assessment

Given Bocij & Greasley's (1999) finding that performance gains observed in computer-based testing conditions did not occur for the higher-ability students, the participants in this study were split into three groups based on performance on the first two exams (which occurred prior to collection of any data for this study). Using the students' mean performance levels on the first two exams, quartile splits were established. The top 25% were considered the high-scoring group, the bottom 25% were the low-scoring group, and the middle 50% were the average-scoring group. Using this contrived grouping system, a 3 5 2 multivariate analysis of variance was conducted, examining the main effects and interaction of the independent variables: prior performance (high, average, low) and assessment format (paper, online) on the dependent measures cognitive test anxiety, emotionality, perceived test threat, study skills, and quiz usage. The results of the MANOVA revealed significant main effects for both prior performance, $F(10, 294) = 4.08$, $p < .001$, $\eta^2 = .12$, and assessment format, $F(5, 146) = 18.48$, $p < .001$, $\eta^2 = .39$. The interaction effect was not statistically significant, $F(10, 294) = 1.25$, $p = .26$, $\eta^2 = .04$. The absence of a significant interaction does not confirm the finding by Bocij and Greasley (1999) demonstrating differential benefits for online testing for the high and low ability students.

Prior Test Performance Effects

Follow-up between-subjects analyses of variance revealed several statistically significant effects. For simplicity, only significant effects are presented. For the main effect of prior test performance, a statistically significant difference was observed for the following dependent variables: (a) cognitive test anxiety, $F(2, 150) = 10.90$, $p < .001$, $\eta^2 = .13$; (b) perceived test threat, $F(2, 150) = 7.14$, $p < .001$, $\eta^2 = .08$; and (c) quiz use, $F(2, 150) = 4.38$, $p < .02$, η 2 = 06. Examination of the means in Table 1 illustrate the effects of Scheffe's post-hoc analyses (all p's <.05) which

demonstrated that (a) low-scoring students held significantly higher levels of cognitive test anxiety than both the average- and high-scoring students; (b) low-scoring students held higher levels of perceived test threat than the high-scoring students; and (c) more students in the high-scoring group reported using the practice tests than students in the average-score group. Note that although the differences are all statistically significant, the effect sizes are weak.

Table 1: Means and Standard Deviations on Test Perception and Preparation Measures: Assessment Format and Prior Performance

Prior Test Performance	Low	Average	High
Paper-Based Testing			
n=17 n=30 n=18			
Cognitive Test Anxiety[a]	80.41 (12.04)	70.10 (16.26)	65.72 (15.69)
Emotionality[b]	17.65(4.83)	16.97(6.12)	17.39(5.28)
Perceived Test Threat[c]	56.53 (5.35)	53.20 (7.18)	2.72 (6.52)
Study Skills[d]	17.65(5.18)	18.87(5.18)	20.83 (5.22)
Quiz use[e]	.65(.49)	.40 (.49)	.44 (.51)
Online Testing			
n=24 n=44 n=23			
Cognitive Test Anxiety	4.33 (16.73)	1.23 (13.16)	58.70 (13.26)
Emotionality	18.00 (7.46)	18.11 (7.00)	15.74 (5.57)
Perceived Test Threat	48.29(5.17)	46.41 (4.29)	42.48(6.04)
Study Skills Scale	20.50 (6.33)	21.14 (5.02)	22.04 (3.77)
Quiz use	.63 (.49)	.43 (.50)	.87(.34)

Notes: Possible score range is 27 to 108. Possible score range is 10 to 40. Possible score range is 18-72. Possible score range is to 32. Quiz use is determined by a dummy-code of 0 = "no" and 1 = "yes." Higher scores indicate a greater percentage of the group using the quizzes.

Testing Format Effects

Between-subjects analyses for the main effect of testing format revealed significant differences for (a) perceived test threat, $F(1, 150) = 76.68$, $p < .001$, $\eta^2 = .34$ and (b) self-reported study skills, $F(2, 150) = 5.90$, $p < .02$, $\eta^2 = .04$. The means displayed in Table 1 reveal that students in the online testing group had meaningfully lower levels of perceived test threat. The results also demonstrate that the weak effect size for self-reported study skills favored the online testing group.

A separate univariate analysis of covariance was run to examine the effect of online testing on Test 3 performance, using the average performance level on Test 1 and Test 2 as the covariate. The results revealed no significant difference based on the format of the test administration, $F(1, 172) = .07$, $p = .79$, $\eta^2 = .00$.

The Role of Practice Testing in the L-T Cycle

The first indirect test on the efficacy of online practice tests was through student self report. For both semesters, a subset of the participants provided ratings of the usefulness of the online practice tests by responding to the statement, "I found the online quizzes to be helpful in preparation for the exam." Only six of the 64 students who responded to this

Likert-scaled item disagreed with the statement (41 "agree"; 17 "strongly agree"). Chi-square analyses revealed no differential rates of endorsing the statement based on method of summative assessment, $X2\ (3, N = 64) = 2.64$, $p > .05$. Only the online summative assessment group provided data regarding the total number of uses for the practice quizzes (recall that the paper assessment group provided only nominal data indicating use or no-use). Therefore, the remaining analyses focusing on the influence of practice testing on the L-T Cycle are restricted to the online summative assessment group. This has the additional benefit of eliminating the effect of having differing formats for the practice (online) and summative (paper) assessments.

The data presented in Table 2 demonstrate a complex relationship among the various constructs of perceived test threat, cognitive test anxiety, performance, and study strategies. The addition of the online practice quizzes for only the third course exam provided a unique context for students' test preparation that had not been available in previous exams. Initial ANOVA-based analyses revealed no consistent pattern of impact for the online practice quizzes on outcomes for the third exam, when using prior test performance as a covariate. However, it is clear from earlier analyses that those students who are likely to use the quizzes differ from those who are not, presenting a condition that cannot be easily interpreted through standard ANOVA. Given the complexity of the relationships among these variables in the L-T Cycle, more detailed examination with structured equation modeling was employed to investigate the unique influence of practice tests on perceptions and performance.

Table 2: Intercorrelation Matrix for the Online Testing Group (n = 91)

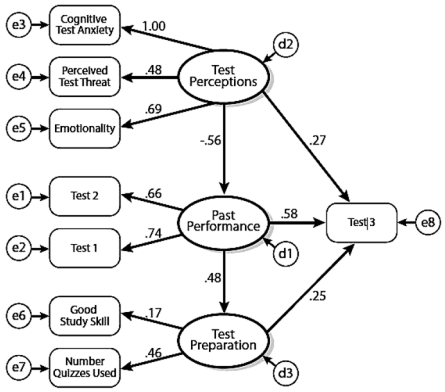
	1	2	3	4	5	6
1. Exam 1 Performance						
2. Exam 2 Performance	.52**					
3. Exam 3 Performance	.38**	.32**				
4. Cognitive Test Anxiety	-.40**	-.40**	-.12			
5. Emotionality	-.10	-.22	-.11	.69**		
6. Perceived Test Threat	-.43**	-.36**	-.15	-.48**	.30**	
7. Number of Practice Quizzes Used	.16	.19	.25*	-.07	.02	-.03
8. Study Skills and Habits.11	.09	.14	-.07	-.03	-.31	.01

Notes: *p<.01 **p<.001

We created two viable models based on the extant research involving test perceptions, preparation, and performance. Both structural equation models proposed that three latent variables provided direct effects on performance on the third exam. These three variables (Test Perceptions, Past Performance, and Test Preparation) also were modeled to influence one another, which led to the primary difference between the two presented models. Model A (Figure 1) rests on the proposition that Test Perceptions is primarily a stable entity that has influence over upcoming and past test performances. This proposition rests on the assumption that perceptions of tests develop over time and are likely to maintain stability across one academic semester, as has been supported in earlier work with these materials (Cassady, 2001a). Perceptions of tests were also hypothesized to influence Test Preparation indirectly through Past Performance, and have indirect influence on test performance through the other two latent variables. Past Performance was hypothesized to be related directly to Test Preparation and current test performance (also influencing

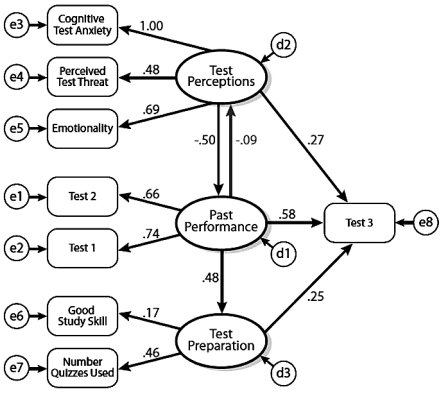
current performance indirectly through test preparation). The path linking Past Performance to Test Preparation is consistent with the L-T Cycle framework. In that model, during the test reflection phase, attributions accounting for success of failure in previous testing situations dictate the types of preparation strategies that are selected. Furthermore, those attributions are connected to the learner's perceptions of tests in general (see Cassady, 2004b).

Figure 1: Model A



Model B (Figure 2) differed by including an additional path leading from prior test performance to test perceptions. The notion is that past performances contribute to the overall level and orientation of beliefs about tests, recognizing a bi-directional relationship between test perceptions and performances in the past. This relationship is particularly compelling in a condition such as the current study, where the Past Performance variable is composed entirely of tests from the same course as the outcome variable (i.e., Test 3)

Figure 2: Model B



As demonstrated in Figures 1 and 2 and Table 3, with the exception of the addition of the path from Past Performances to Test Perceptions that appears only in Model B, the estimates for the paths are identical for the two models. Most effect sizes (path coefficients) were moderate to low. Past Performance had a greater direct effect on scores on Test 3 than did either Test Perceptions or Test Preparation. Test Perceptions had a moderate effect on Past Performance as did Past Performance on Test Preparation. The indirect effect of Test Perceptions through Past Performances on Test Preparation was small. Small indirect effects on the Test 3 scores were also noted for Test Perceptions, as modeled through both Past Performance and Test Preparation.

Table 3: Model Comparison Data

	Model A	Model B
Direct Effects		
Test Perceptions – Test 3	.27	.27
Test Perceptions – Past Performance	-.56	-.50
Test Preparation – Test 3	.25	.25
Past Performance – Test Perception	—	-.09
Past Performance – Test Preparation	.48	.48
Past Performance – Test 3	.58	.58
Indirect Effects		
Test Perception – Test Preparation	-.27	-.24
Total Effects		
Test Perception – Test 3	-.117	-.075
Past Performance – Test 3	.700	.707
Test Preparation – Test 3	.248	.248
Fit Statistics		
χ 2(18)	30.40	30.40
p	.03	.03
χ 2/df (ratio)	1.69	1.69
TLI	.88	.88
CFI	.92	.92
PCFI	.59	.59
RMSEA	.09	.09
AIC	66.40	.40

Following established criteria for model comparisons (Gridley, 2002) the fit statistics for the two models are considered identical (Table 3). The addition of a path from Past Performances to Test Perceptions in addition to the one from Test Perceptions to Past Performances does not significantly modify the statistical explanations available in the models. Therefore, there are no differences between the models in their ability to fit the data. While parsimony would suggest adopting Model A, Model B provides a more theoretically tenable solution given the acknowledgement of the influence of past performances on the formation of test perceptions. In essence, Model B illustrates that although Test Perceptions and Past Performance exert influence upon one another, the downward path in both models is dominant.

The intriguing finding with the models in this study highlight the potential impact of the online practice quizzes. The direct effect of Test Perceptions to Test 3 performance and Past Performance confirm prior results demonstrating an overall impact of test perceptions, specifically cognitive test anxiety, on test performance levels. However, in the unique testing situation under investigation in this study, that is a testing condition accompanied by online practice quizzes, examination of the total effects indicated that the standard negative influence of Test Perceptions was no longer prevalent.

Discussion

The rapid growth of using the Internet to deliver course materials, including assessment measures, has opened a new branch of research in effective instructional practice (Wheeler, 2000). However, to date there has been limited information examining the learning benefits gained through systematic use of these online instructional tools (Buchanan, 1998; 2000). Structured around the established framework of the L-T Cycle and the broad base of research on the impact of testing conditions on students with test anxiety, this study begins to answer fundamental questions regarding the utility of online testing practices, and has documented specific benefits of providing both formative and summative assessments online.

Online Summative Assessment

The results provide no support that online testing will induce additional anxiety or impact performance levels. However, it is important to recognize these results should not be over generalized to all undergraduate students; all participants in this study were involved in courses that required frequent use of the Internet to access course materials and information. This systematic access to technology tools and materials likely facilitated any adjustment students needed to make to use online evaluative materials. It is improbable that students with lower levels of online experience would have similar comfort levels, and the level of emotionality and anxiety may be expected to rise for students without systematic exposure to computer-based instructional processes (Cassady, 2001a).

The only meaningful difference reported by students in the two testing conditions was the heightened level of perceived threat reported by students taking tests on paper. We propose this outcome was mostly influenced by the lack of personal control over the testing events (Boggiano & Ruble, 1986; Butler, 2003). Given the flexibility afforded by the secure computer-based testing laboratories, the online testing group was permitted to complete each test over the course of an entire week, including evenings and weekends. This led to anecdotal reports from the students that they enjoyed being able to take tests on “light” days. This ability to schedule the tests seemed to allow the students to reduce the level of contextual stress by strategically placing their testing times in convenient time slots. For the students taking tests during assigned times, there was no ability to choose what day would work best with their schedules. These students frequently reported they had several other assignments or tests during the same day or week that the test was given. As many students have reported, “everything is due at the same time.” Thus, while the students reported great satisfaction in their level of choice in testing, this benefit of online assessment resulted in a confound in these analyses; it is impossible with the current data to determine that the reduced test threat in the online condition is not simply due to the ability to choose testing time. However, even as a negative effect, this condition of flexible timing for testing is more easily achieved in online testing given logistic concerns.

The data suggest that providing tests online in a secure, proctored computer-based testing laboratory may not simply provide a reasonable alternative method for gathering summative assessment data from students, but may actually be a preferable method. In addition to lower levels of perceived test threat and the obvious benefits of ease in scoring or test delivery, online testing can also provide increased instructional time. In this case, the gains in instructional time were a by-product of delivering the tests outside of the confines of class meeting rooms and sessions. The use of online testing produced approximately 4.5 additional hours of instructional time, as compared to in-class testing. This additional time was gained by replacing three 75-minute class periods formerly reserved for

testing (total time = 3.75 hours) as well as an additional 15 minutes per test for returning corrected tests and providing the correct answers, which was administered automatically through the online testing module (conservative estimate; total time = 4.5 hours).

The only noted barriers to effective assessment in an online environment are the standard logistical concerns. First, as more instructors become proficient with online testing, labs become stressed to meet the need for testing. This institutional barrier warrants considerable attention due to the expense associated with creating and maintaining additional testing laboratories that can be monitored. Second, some students struggled with responding on screen rather than on paper. In particular, some students found it hard to keep track of items they had skipped over to come back to later. The standard solution to this barrier has been to suggest that all students bring blank paper to work with during the test period. Recent advancements in online testing programs have also helped to alleviate this problem by providing reminders to test takers when an item has been left unanswered before closing the testing session. Third, students in the online testing condition were not able to ask questions of the instructor during the assessment period. Losing the ability to clarify questions with the instructor prior to responding is a barrier highlighted by a few students who describe question-asking during the test as a coping behavior they periodically employ during testing. Finally, testing security is a constant concern in online testing. Use of secure testing facilities an software solutions that can randomize pre-selected equivalent content items help combat these concerns. Just as instructors have to be conscientious in overcoming the “fraternity test file” from previous semesters with paper-based testing, instructors using online assessments need to monitor the test conditions to preserve the integrity of assessment.

Online Formative Assessment

Previous studies have discussed the availability of online formative assessment tools (Buchanan 1998; 2000), however no data have been available demonstrating the overall impact on students’ performances or perceptions of testing events. Students overwhelmingly reported that they found the online formative assessment tools (practice quizzes/tests) to be useful in preparation for the exam. Although student perceptions of utility are important in determining the impact of practice tests on the L-T Cycle, particularly when taking the impact of cognitive test anxiety and perceived threat into account (Cassady, 2004b), the contribution of this study comes from the results generated in this exploration of the relationships among test perceptions, test preparation, and prior performance variables.

The small but positive impact of practice test use on subsequent course examination performance provides preliminary evidence that online practice tests can serve as an effective test preparation strategy. The data in this study support the pattern of results predicted by the testing phenomenon (Glover, 1989), where the completion of a realistic testing event can promote performance on subsequent assessment tasks. In addition, the similarity between the formative and summative assessment tools in function, difficulty, and format likely facilitated the transfer of content information or contextual cues from the practice setting to the final performance session, which should aid recall of the target information (McDaniel et al., 1989; Roediger & Guynn, 1996).

The formative assessment generator used in this study also provided the pedagogically desirable method of immediate post-test feedback (King & Behnke, 1999; Wise et al., 1989). The feedback process is accomplished through a separate pop-up browser window. This allows the user to simultaneously view the corrective feedback and the original question, promoting the user’s ability to modify existing cognitive structures and retrieval cues.

With respect to the L-T cycle, the addition of online quizzes to learners' test preparation strategies provided a unique structured study tool that helped to alleviate the overall effect of Test Perceptions on Test 3 performance. In repeated studies of cognitive test anxiety and performance, there has been a stable and definite trend documenting a significant negative relationship for students from undergraduate populations (Cassady, 2004a; 2004b; Cassady & Johnson, 2002; Cassady et al., 2004). This trend was repeated in this sample as well for the first two course examinations, for which there were no practice tests available. However, as shown in Table 2, there was no significant correlation between Test 3 performance and cognitive test anxiety or perceived test threat. Indeed, only prior test performances and the use of the practice tests were significantly related to Test 3 performance. As illustrated in Figure 2 (Model B), although Test Perceptions continue to have influence on the overall model, the influence in this unique condition appears to be in driving the learner toward a more useful study strategy (practice tests) that nullifies the standard effects of test perception. It is essential to stress that the benefits seen for those students using the formative assessment quizzes were not likely a mere consequence of delivery method. We predict that all benefits observed in this study would be replicated with paper-pencil practice tests, provided they matched the actual tests in format and difficulty level. The unique contributions provided by the QuizEditorJS software used in this study rest in the primary benefits afforded through computerized delivery of assessment: greater student access, flexibility, ease of constructing the assessment tools, and immediate formative feedback (Bransford, Brown, & Cocking, 1999; Buchanan, 2000; Dempster & Perkins, 1993). Allowing students to freely access practice tests and receive immediate corrective feedback provides personal control over test preparation. This method of delivery also has benefits over the standard in-class short quiz approach in that students can repeatedly access a variety of different practice tests.

Limitations and Future Directions

Naturally, the conduct of research with samples of convenience in naturally occurring educational settings provides multiple threats to external validity that are key to vary in replication studies in order to confirm the effects are not situation-specific. The primary limitation in this study is the small sample size, particularly in the online testing sample upon which the bulk of the formative assessment data analyses (i.e., SEM)

are based. The small sample size harms the power for extrapolation, which naturally affects significance testing, but more importantly provides concern for the stability of the two models. Additional participants in the present study would have enabled more detailed analyses of the contributing factors leading to the positive effects associated with the practice quizzes. In particular, we are interested in exploring which students are most likely to access the quizzes and what role success or failure on initial attempts with practice quizzes has on repeated attempts. The presence of confounded variables also needs to be controlled in future investigations. First, the individual's control over the timing of the test administration is likely to influence the perceived level of cognitive test anxiety and perceived test threat. To address this concern, providing the on-paper group with the option to take the test at any point in a given time frame would control the confounding variable.

The second confound in the study is that all practice tests were provided online. Does presentation format of the practice quizzes matter? Most textbook publishers provide student study guides for core undergraduate course textbooks that include practice test items. Would the same benefits be granted with use of these materials? The limitations to this study preclude a definitive answer, however we propose that the presentation format likely does

matter. Specifically, the issue of importance is a positive match in presentation format between the formative and summative assessments. It is a well-established effect that memory performance is improved in conditions where retrieval cues sparked in the testing condition are more consistent with the cues available during encoding (Roediger & Guynn, 1996; Tulving & Thompson, 1973), or provide more specific "diagnostic" information that facilitates reconstruction of the target content (Nairne, 2002a; 2002b).

A third confounding condition that could be controlled in future investigations is related to the comparison of the online and paper-based testing conditions. In the study, the paper-based class received fewer instructional periods given their in-class testing requirement. It is possible that the effects in this study are influenced by the different amount of instructional time.

References

Bandura, A. (1986). Social foundations of thought and action: A social-cognitive theory. Englewood Cliffs, NJ: Prentice-Hall.

Bocij, P. & Greasley, A. (1999). Can computer-based testing achieve quality and efficiency in assessment? International Journal of Educational Technology, 1(1), 17 pages.

Boggiano, A. K. & Ruble, D. N. (1986). Children's responses to evaluative feedback. In R. Schwarzer (Ed.) Self-related cognitions in anxiety and motivation (pp. 195–228). Hillsdale, NJ: LEA.

Bransford, J. D., Brown, A. L., & Cocking, R. R. How people learn: Brain, mind, experience, and school. Washington, DC: National Academy Press.

Buchanan, T. (1998). Using the World Wide Web for formative assessment. Journal of Educational Technology Systems, 27(1), 71–79.

Buchanan, T. (2000). Potential of the Internet for personality research. In M.H. Birnbaum (Ed.) Psychological experiments on the Internet. San Diego: Academic Press.

Butler, D. L. (2003). The impact of computer-based testing on student attitudes and behavior. The Technology Source, January/February.

Carrier, M. & Pashler, H. (1992). The influence of retrieval on retention. Memory & Cognition, 20, 633–642.

Cassady, J. C. (2001a). Integrating technology instruction in pre-professional training programs. Trainer's Forum, 19(3), 1–2; 8–10.

Cassady, J. C. (2001b). The stability of undergraduate students' cognitive test anxiety levels. Practical Assessment, Research & Evaluation, 7(20).

Cassady, J. C. (2004a). The impact of cognitive test anxiety on text comprehension and recall in the absence of salient evaluative pressure. Applied Cognitive Psychology, 18(3), 311–325.

Cassady, J. C. (2004b). The influence of cognitive test anxiety across the learning teaching cycle. Learning and Instruction, 14(6), 569–592.

Cassady, J. C. & Johnson, R. E. (2002). Cognitive test anxiety and academic performance. Contemporary Educational Psychology, 27, 270–295.

Cassady, J. C., Mohammed, A., & Mathieu, L. (2004). Cross-cultural differences in test anxiety: Women in Kuwait and the United States. Journal of Cross-Cultural Psychology, 35(6), 715–718.

Clariana, R. B., Ross, S. M., & Morrison, G. R. (1991). The effects of different feedback strategies using computer-administered multiple-choice questions as instruction. Educational Training, Research, and Development, 39, 5–17.

Covington, M. V. & Omelich, C. L. (1987). “I knew it cold before the exam”: A test of the anxiety-blockage hypothesis. Journal of Educational Psychology, 79, 393–400.

Culler, R. E. & Holohan, C. J. (1980). Test anxiety and academic performance: The effects of study-related behaviors. Journal of Educational Psychology, 72, 16–26.

Deffenbacher, J. L. (1980). Worry and emotionality in test anxiety. In I. G. Sarason, (Ed.) Test anxiety: Theory, research, and applications (pp. 111–124). Hillsdale, NJ: Lawrence Erlbaum.

Dempster, F. N. (1997). Using tests to promote classroom learning. In R. F. Dillon (Ed.) Handbook of testing (pp. 332–346). Westport, CT: Greenwood Press.

Dempster, F. N. & Perkins, P. G. (1993). Revitalizing classroom assessment: Using tests to promote learning. Journal of Instructional Psychology, 20, 197–203.

Duchastel, P. (1996). A Web-based model for university instruction. Journal of Educational Technology Systems, 25, 221–228.

Geen, R. G. (1980). Test anxiety and cue utilization. In I.G. Sarason (Ed.) Test anxiety: Theory, research, and applications (pp. 43–62). Hillsdale, NJ: LEA.

Glover, J. A. (1989). The “testing” phenomenon: Not gone but nearly forgotten. Journal of Education Psychology, 81, 392–399.

Gridley, B. E. (2002b). In search of an elegant solution: Reanalysis of Plucker, Callahan, and Tomchin, with respects to Pyryt and Plucker. Gifted Child Quarterly, 46, 224–234.

Hembree, R. (1988). Correlates, causes, and treatment of test anxiety. Review of Educational Research, 58, 47–77.

Jongekrijg, T. & Russell, J. D. (1999). Alternative techniques for providing feedback to students and trainees: A literature review with guidelines. Educational Technology, 39(6), 54–58.

Ikeda, M., Iwanga, M., & Seiwa, H. (1996). Test anxiety and working memory system. Perceptual and Motor Skills, 82, 1223–1231.

King, P. E. & Behnke, R. R. (1999). Technology-based instructional feedback intervention. Educational Technology, 39(5), 43–49.

Kurosawa, K. & Harackiewicz, J. M. (1995). Test anxiety, self-awareness, and cognitive interference: A process analysis. Journal of Personality, 63, 931–951.

Liebert, R. M. & Morris, L. W. (1967). Cognitive and emotional components of test anxiety: A distinction and some initial data. Psychological Reports, 20, 975–978.

McDaniel, M. A., Kowitz, M. D., & Dunay, P. K. (1989). Altering memory through recall: The effects of cue-guided retrieval processing. Memory & Cognition, 17, 423–434.

Morris, L. W., Davis, M. A., & Hutchings, C. H. (1981). Cognitive and emotional components of anxiety: Literature review and a revised worry-emotionality scale. Journal of Educational Psychology, 73, 541–555.

Nairne, J. S. (2002a). The myth of the encoding-retrieval match. Memory, 10, 389–395.

Nairne, J. S. (2002b). Remembering over the short-term: The case against the standard model. Annual Review of Psychology, 53, 53–81.

Naveh-Benjamin, M., McKeachie, W. J., & Lin, Y. (1987). Two types of test-anxious students: Support for an information processing model. Journal of Educational Psychology, 79, 131–136.

Roediger, H. L. & Guynn, M. J. (1996). Retrieval processes. In E. C. Carterette & M. P. Friedman (Series Eds.) & E. L. Bjork & R. A. Bjork (Vol. Eds.) Handbook of Perception and Cognition (2nd Ed.). Memory. San Diego, CA: Academic Press.

Sarason, I. G. (1984). Stress, anxiety, and cognitive interference: Reactions to Tests. Journal of Personality and Social Psychology, 46, 929–938.

Sarason, I. G. (1986). Test anxiety, worry, and cognitive interference. In R. Schwarzer (Ed.) Self-related cognitions in anxiety and motivation (pp. 19–34). Hillsdale, NJ: LEA.

Schutz, P. A. & Davis, H. A. (2000). Emotions and self-regulation during test taking. Educational Psychologist, 35, 243–256.

Schwarzer, R. (1984). Worry and emotionality as separate components in test anxiety. International Review of Applied Psychology, 33, 205–220.

Schwarzer, R. & Jerusalem, M. (1992). Advances in anxiety theory: A cognitive process approach. In K. A. Hagtvet & T. B. Johnsen (Eds.) Advances in test anxiety research (Vol. 7, pp. 2–31). Lisse, The Netherlands: Swetts & Zeitlinger.

Tulving, E. & Thomson, D. (1973). Encoding specificity and retrieval processes in episodic memory. Psychology Review, 80, 352–373.

Wheeler, S. (2000). Instructional design in distance education through telematics. Quarterly Review of Distance Education, 1(1), 31–44.

Wise, S. L., Plake, B. S., Eastman, L. A., Boettcher, L. L., & Luken, M. E. (1986). The effects of item feedback and examinee control on test performance and anxiety in a computer-administered test. Computers in Human Behavior, 2, 21–29.

Wise, S. L., Plake, B. S., Pozehl, B. J., Barnes, L. B., & Luken, M. E. (1989). Providing item feedback in computer-based tests: Effects of initial success and failure. Educational and Psychological Measurement, 49, 479–486.

Wittmaier, B. C. (1972). Test anxiety and study habits. The Journal of Educational Research, 65, 352–354.

Zeidner, M. (1998). Test anxiety: The state of the art. New York: Plenum Press.

fuller@tc.nagasaki-gaigo.ac.jp