Student Anxiety, Acceptance, and Experience Using the Immediate Feedback – Assessment Technique®

Kirstin Kuchler
University of Minnesota - Twin Cities, kuchl001@umn.edu

Lizbeth Finestack
University of Minnesota - Twin Cities, finestac@umn.edu

Follow this and additional works at: https://ir.library.illinoisstate.edu/tlcsd

Part of the Curriculum and Instruction Commons, and the Higher Education and Teaching Commons

Student Anxiety, Acceptance, and Experience Using the Immediate Feedback – Assessment Technique®

Cover Page Footnote

We have no known conflicts of interest to disclose. Correspondence concerning this article should be addressed to Kirstin Kuchler, M.A., CCC-SLP, University of Minnesota-Twin Cities, Department of Speech-Language-Hearing Sciences, 164 Pillsbury Drive SE, Minneapolis, MN 55455. Email: kuchl001@umn.edu

Multiple-choice tests are a common assessment technique used across all levels of education. One format for multiple-choice assessments is the Immediate Feedback Assessment Technique® (IF-AT®; DiBattista et al., 2004; DiBattista & Gosse, 2006; Merrel et al., 2015), also called answer-until-correct (Clark et al., 1998; Muniz & Menendez, 2011). The IF-AT® is a paper form that requires test-takers to scratch off their answer choice which reveals a star, indicating the answer is correct, or a blank space, indicating the answer is incorrect. Computer versions of this technique are also used. When test-takers receive feedback that they answered the question incorrectly, they have the option to keep choosing responses until the question is answered correctly. Scoring systems can vary, but with the IF-AT® subsequent attempts typically receive fewer points than the first attempt. This allows test-takers to earn more points than they would have with only one opportunity to respond (DiBattista & Gosse, 2006; DiBattista et al., 2004). The method also rewards points to students who have some knowledge but may not answer correctly on their first attempt (Epstein et al., 2001). Additionally, educators may use the IF-AT® as a practice test for students and teachers to identify potential weaknesses and help students improve their knowledge or as a post-test in which they still have the opportunity to learn while demonstrating their knowledge.

Researchers have examined the use of the IF-AT® as a way to systematically improve the learning process (Farland et al., 2015), to improve test reliability and validity of test results (Muniz & Menendez, 2011), to use as a motivating scoring system (DiBattista et al., 2009; Muniz & Menendez, 2011), and to examine the effect on a variety of emotions during test taking (Attali & Powers, 2010; Daniels & Gierl, 2017; DiBattista & Gosse, 2004; DiBattista et al., 2009; Green, 1981; Lee et al., 2012; Martlett & Watson, 1968; Rocklin & Thompson, 1985; Vanderoost et al., 2018). In the current study, we examined the use of the IF-AT® in an undergraduate course with focus on self-reported test anxiety, acceptance of, and experience using the format.

The non-traditional answer-until-correct testing method has multiple names and can be used in multiple formats. The IF-AT® is a specific package of testing materials created and patented by Epstein and Epstein (Epstein et al., 2001). Here we use the term answer-until-correct unless the IF-AT® was the specific tool used in the study described.

Answer-until-correct Testing Format. There are several potential advantages to using the answer-until-correct format relative to traditional multiple-choice formats. One advantage is that it can capture different levels of student knowledge. Dichotomous scoring, in which a single response for a test item is scored as either correct or incorrect, does not make the distinction between (a) the test-takers who answer correctly because they know the answer and (b) the test-takers who answer correctly because they guessed correctly (Epstein et al., 2001.). Furthermore, Epstein et al. (2001) assert that traditional multiple-choice tests do not reward students who are able to narrow the answers down to two. Conversely, the answer-until-correct format captures students' partial knowledge by rewarding test-takers who can narrow down the answer by giving some points to successive responses. A potential disadvantage of the answer-until-correct format is that it may reward students who guess by giving some points to those who try all but one answer. Three studies have examined various scoring systems when using the answer-until-correct format (i.e., Dihoff et al., 2004; Muniz & Menendez, 2011; Vanderoost et al., 2018).

Muniz and Menendez (2011) modeled the answer-until-correct format using eight scoring conditions on tests with seven different test items (i.e., 10, 20, 30, 40, 50, 75, 100). This simulation created 56 conditions that could be evaluated to determine the condition in which reliability of the scores and validity of the inferences could be maximized from the viewpoint of Classical Test Theory. The scoring sequences that resulted in higher validity and reliability than dichotomous testing were those for which answer tries 1 through 5 were awarded 1, 0.5, 0.25, 0.12, and 0 or 1, 0.25, 0.12, 0.06, and 0 points, respectively. The differences between these point systems and dichotomous scoring was not significantly different, leading the authors to conclude that the dichotomous scoring system may be sufficient for classroom testing, but more nuanced responses could increase the validity and reliability of more general knowledge assessments.

To evaluate the merits of the answer-until-correct format for assessing partial knowledge, Merrel et al. (2015) examined answers chosen on the second try after initially selecting the incorrect response. The researchers assessed 8,775 responses that were incorrectly answered on the first try. There were five answer options students could choose from for each question, thus there was a 25% chance of getting the answer correct on the second try. Merrel et al. found that the second choice was correct 44.9% of the time. Researchers employed a *t*-test to determine the difference between the actual number of correct second-try responses and random chance correct second-try responses and found that students were using "some level of discernment to determine which of their remaining possibilities were the correct choice" (Merrel et al., 2015, p. 52). These findings indicate that answer-until-correct may be an appropriate and empirically supported format to allow students to receive credit for partial knowledge.

Extending this work, Vanderoost et al. (2018) compared two scoring methods to determine which method would reduce grade inflation and allow students to express partial knowledge without giving credit for guessing. In this study, there were five answer choices to each multiple-choice question. The researchers compared negative marking and elimination testing with adapted scoring. Negative marking is a method in which the test taker either responds once to the question or leaves the question blank if the answer is unknown. Researchers awarded a correct response 1 point, an incorrect response -0.25 points, and a question left blank 0 points. This scoring was designed to reflect if the test taker has full knowledge, misconception, or no knowledge, respectively. Alternatively, elimination testing with adapted scoring allows students to respond to the question with one response indicating full knowledge or indicate which of the five responses could or could not be the answer, revealing partial knowledge and partial misconceptions. Study results indicated no significant differences for average exam scores between scoring types (elimination testing and negative marking), but that elimination testing allowed students to reveal their partial knowledge on a topic. Elimination testing with adapted scoring allowed researchers to better understand when students were guessing (had doubt in their response) and when they were narrowing down an answer (had partial knowledge). Answer-until-correct is a simplified form of both test formats that allows the evaluation of partial knowledge.

Another potential advantage of the answer-until-correct multiple-choice testing format is that it can be used as a learning tool. In a study by Dihoff et al. (2004), three groups of students received six practice tests prior to in-class examinations. The control group received scantron traditional multiple-choice practice tests for all six tests. One experimental group received all IF-AT® practice tests. The other experimental group received three scantron traditional multiple-choice practice

tests (S) and three IF-AT® (I) practice tests in one of the following four combinations: SSSIII, IIISSS, ISISIS, or SISISI. The researchers found that there was no difference between the amount of self-reported study time and time on practice tests in relation to the scores on the in-class exams or the final exam. There were statistically significant results for the test-takers' scores that revealed that the experimental group with six IF-AT® practice tests had the best scores on the in-class examinations. Those with only scantrons had the lowest scores on the in-class examinations. When questions on the final exam were similar to those on the practice tests, the results suggested that the instant feedback provided by the IF-AT® form helped to improve students' course knowledge.

Despite these advantages, there are some potential disadvantages to using the answer-until-correct format. One disadvantage is that test-takers may experience increased test anxiety if during the exam they find themselves performing poorly or not as well as they had expected.

Assessment Feedback and Test Anxiety. Several studies have evaluated student anxiety associated with the answer-until-correct format. For example, Clark et al. (1998) compared the relationships between anxiety and performance on test questions that varied in the feedback provided: dichotomous scoring multiple-choice with no feedback, dichotomous scoring multiple-choice with immediate feedback, and answer-until-correct multiple-choice. For the dichotomous scoring with no feedback format, students chose their answer and received no feedback about its correctness. The dichotomous scoring with feedback format was a dichotomous-scoring, multiple-choice format in which students received feedback about the correctness of their choice immediately but were not able to try again. The answer-until-correct format allowed students to answer until they selected the correct choice. Prior to testing, students completed a survey about general anxiety, which allowed the researchers to place students in high anxiety or low anxiety groups. After testing, participants reported on their anxiety during testing. Study results indicated that there was not a significant relationship between the assessment feedback students received and anxiety reported during testing. There was also not a statistically significant relationship between test performance and type of feedback. However, there was a significant relationship between pre-test anxiety and anxiety reported during testing indicating that students who reported high levels of anxiety pre-test experienced higher levels of test anxiety regardless of the testing format.

In another study, Rocklin and Thompson (1985) examined the relationship between student anxiety and test performance when using the IF-AT® with tests varying in difficulty. The researchers surveyed the students prior to test taking to determine if they typically experienced low, medium, or high anxiety. Then, students experienced a hard test in two conditions (i.e., with and without immediate feedback) and an easy test in the same two conditions. Researchers found that hard tests yielded a lower mean proportion correct for students who had low, medium, and high anxiety regardless of whether they received immediate feedback. Performance on the easy test was more variable with students who reported low anxiety receiving higher scores when they received feedback than without. Similar to the findings of the Clark et al. (1998) study, it seemed that test difficulty impacted perceived anxiety more than the question format and that the IF-AT® did not lead to additional test anxiety.

IF-AT® and Acceptance. Studies of the IF-AT® indicate that its use is widely accepted by test-takers because of the immediate feedback received, opportunity to learn while testing, opportunity

to receive partial credit, and ability to know the test score upon test completion (Clark et al., 1998; DiBattista et al., 2004). For example, DiBattista et al. (2004) studied psychology students' acceptance of the IF-AT®. Of the 234 students who took the test, 177 students responded to the optional survey about their experience using the IF-AT® testing method. All questions had statistically significant positive responses, indicating students liked using the IF-AT® format. Results were not correlated with personal characteristics (e.g., age, gender) which indicates that the acceptance of the IF-AT® format spans a wide range of students.

In another study examining students' testing format preferences, Clark et al. (1998) found that test-takers preferred answer-until-correct tests over two other options. The 73 participants in a self-paced psychology course were randomly assigned a test feedback condition for the first two units and were given the opportunity to choose their test feedback condition for the five subsequent units. On the first opportunity to switch feedback condition, of those assigned to the answer-until-correct condition, only 13% chose to switch formats. In contrast, 84% of students who chose to switch from the other two conditions chose the answer-until-correct format. On the final test of the semester, 84% of students chose to switch from their assigned test to answer-until-correct format and 72% of all students chose answer-until-correct test condition. This clearly indicates a majority of students preferred the answer-until-correct test condition.

Current Study. The purpose of the current study was to further examine undergraduate college students' acceptance of the IF-AT® for course exams and anxiety related to the IF-AT® format. We also examined changes in reported IF-AT® experiences between the first two course exams and compared reported acceptance and anxiety after the two exams. Additionally, we evaluated the relationship between exam scores and the levels of self-reported anxiety and acceptance. The following research questions guided our study.

1)   What levels of *anxiety* do undergraduate students report experiencing when test taking using the IF-AT®? Do students' levels of *anxiety* change from Exam 1 to Exam 2 when using the IF-AT®?

2)   What levels of *acceptance* do undergraduate students report when test taking using the IF-AT®? Do students' levels of *acceptance* change from Exam 1 to Exam 2 when using the IF-AT®?

3)   What levels of *experience* do undergraduate students report when test taking using the IF-AT®? Do students report changes in their *experience* after one use of the IF-AT® format in a course?

4)   What is the relationship between students' objective examination scores and reported students' levels of anxiety, acceptance, and experience?

Based on the findings of Clark et al. (1998) and Rocklin and Thompson (1985), we hypothesized that students would experience some anxiety as is natural during test taking tasks, but that anxiety levels would not change across exams. Clark et al. (1998) and DiBattista et al. (2004) findings led us to hypothesize that students would generally accept the IF-AT® testing format and that level of acceptance would not change across exams. Furthermore, we expected that students would report improved experiences when test taking using the IF-AT® based on results in Dihoff et al. (2004). Finally, research by Muniz and Menendez (2011), Vanderoost et al. (2018), and Merrel et al.

(2015) influenced our hypothesis that the relationship between objective examination scores and students' levels of anxiety, acceptance, and experience would not be greater than chance.

Method

Participants. This study was deemed exempt from requiring oversight from the university institutional review board for human subjects research. Participants were drawn from the same 3000-level undergraduate course, Language Acquisition and Science, offered from 2011 through 2019. The exam content changed minimally over this time period. The six classes included in the present study ranged in size from 17 to 80 students. Anonymity was imperative so that students did not receive the impression that their participation was tied to their grade; therefore, gender, age, and ethnicity demographics were not collected. The same professor taught the class each semester. A total of 388 students were given the option to complete a survey regarding their opinion of the IF-AT® after completing their exams. After Exam 1, 318 students (81.9% response rate) completed the survey; after Exam 2, 152 students (39.2% response rate) completed the survey. No incentive was given to students to participate. Participation was on a completely volunteer basis, which could explain the large decrease of participation from completing the Exam 1 survey to the Exam 2 survey.

Procedures. The IF-AT® test form used for the two exams was ordered from Epstein Enterprises (2020, August 13). The form has 50 rows of 4 possible answers per row. There is a fifth column at the end of each row that allows the examiner to enter the score received on each question (i.e., 3, 2, 1, 0). When the students scratched off their answer choice, a star was revealed indicating that it was the correct answer, or a blank space indicating that it was the incorrect answer. Prior to Exam 1, the instructor introduced the IF-AT® form to students. The instructor first asked students to raise their hand if they had used the form before. Many, but less than half, of the students indicated that they had used the form in another class. Next, the instructor presented a PowerPoint slide with a picture of an IF-AT® form similar to the one the students would use on their exams. Then, the instructor distributed cut-up IF-AT® forms. Each segment had approximately five answer rows available. The instructor presented sample test questions for students to practice responding to using the IF-AT® form. Because the instructor randomly distributed the cut-ups, the IF-AT® answer rows did not match the questions. The instructor prompted students to scratch the response that matched the answer they thought was correct, but to keep scratching the form until they found the star that indicated they had selected the "correct" answer, which would be indicated by a star in the answer box. The instructor encouraged the students to try different tools for scratching off their responses (e.g., coin, credit card, retracted pen) to determine which worked best for them. This allowed the students the opportunity to familiarize themselves with the form and the amount of pressure needed to scratch the answer box. The instructor also encouraged students to scratch off the entire response box as the star was not always in the exact center of the box. The instructor advised students to complete the exam as they typically would: they should circle their responses on the paper exam they were given and then transfer their responses to the IF-AT® form. The instructor also advised students to use a piece of paper to underline the row to which they were responding to ensure that they were marking the correct row on the IF-AT® form. The instructor told the students that there would be extra rows available at the bottom of their exams where they could practice scratching answers on the day of the exam if they would like.

On each exam, there were 30 questions on the multiple-choice section; therefore, 20 rows on the IF-AT® form were unused by the students. Each question included four choices and was worth a maximum of 3 points. Students received 3 points if they answered the question correctly on their first attempt, 2 points if they answered correctly on their second attempt, 1 point if they answered correctly on their third attempt, and no points if they had to scratch off all of the options. After the 30 multiple-choice questions there were short-answer, matching, and diagram-labeling questions worth a total of 10 points. The entire exam was worth 100 points. Each exam covered material presented since the last exam (i.e., was not cumulative). At the end of Exam 1 and Exam 2, the instructor provided students with an optional survey to complete regarding their anxiety, acceptance, and experience using the IF-AT® format. The Exam 2 survey included many of the same questions as Exam 1, but four questions were replaced with questions addressing students' comfortability using the IF-AT® form a second time and changes in preparedness and study habits for the second exam.

Optional Survey. We created a survey, based on multiple sources (DiBattista et al., 2004; DiBattista & Gosse, 2006; DiBattista et al., 2009;), to gather information on undergraduates' attitudes regarding use of the IF-AT®. The optional survey asked students to respond to 15 questions using a 5-point Likert scale that ranged from "disagree strongly" (1) to "agree strongly" (5). There were three domains of questions: Anxiety, Acceptance, and Experience. The Exam 1 survey (see Appendix A) investigated students' self-reported anxiety and acceptance using the IF-AT® testing format. The Exam 2 survey (see Appendix B) investigated students' self-reported anxiety, acceptance, and experience using the IF-AT® testing format. Four questions were replaced on the Exam 2 survey for a total of 19 unique questions on the survey over the course of the two exams.

Measures. Three questions were included on the Exam 1 and Exam 2 surveys to create the Anxiety domain. Six questions were included on the Exam 1 survey and five questions were included on the Exam 2 survey to create the Acceptance domain. Four questions were included on the Exam 2 survey to create the Experience domain. Remaining survey questions were used as reliability measures. Table 1 includes each question on the Exam 1 and Exam 2 surveys. The questions were numbered by the order they were displayed to the students and categorized by the domains.

Response Reliability. The Exam 1 survey had three pairs of questions used to determine reliability of student responses (Questions 1 and 3, Questions 4 and 13, and Questions 6 and 14). For example, Questions 1 and 3 asked students to indicate how they felt about receiving or not receiving a star: "Whenever I scratched a box and found the star, I felt as if I was being rewarded for my efforts." vs. "Whenever I scratched a box and did not find the star, I felt as if I was being punished." If a student responded with a 5 on the first question, it is expected that the student would respond with a 1 on the second question. We recoded negatively worded questions (e.g., "Whenever I scratched a box and did not find the star, I felt as if I was being punished.") to the same scale as the positively worded questions and then we subtracted the scores from each other

Table 1

Survey Questions by Domain and Exam Number

Question	Exam 1	Exam 2
Anxiety		
(1) Whenever I scratched a box and found the star, I felt as if I was being rewarded for my efforts.[3]	x	x
(13) Whenever I got a multiple-choice item correct on the first try, I could feel myself becoming less anxious.[4]	x	x
(14) Using the IF-AT® made me feel less anxious than I otherwise would have while doing the multiple-choice items.[6]	x	x
Acceptance	x	x
(7) I would like it if I could use the IF-AT® in all of my courses that have multiple-choice tests.	x	x
(9) I think that the IF-AT® is fairer than an ordinary response form for multiple-choice tests.	x	
(10) I like the fact that the IF-AT® lets me know the right answer to every question.	x	x
(11) I like the fact that the IF-AT® form allows me to get part marks on multiple-choice questions.	x	x
(12) The grading scheme that was used for this test was fair.		x
(15) The IF-AT® allowed me to learn from my mistakes.		x
Experience		x
(3) I felt more comfortable using the IF-AT® for Exam 2 than Exam 1.		x
(5) I changed my study strategy for Exam 2, relative to my strategy for Exam 1.	x	
(8) I felt more prepared for Exam 2 than Exam 1.	x	
(10) I made fewer careless mistakes on Exam 2, relative to Exam 1.		
Questions Used for Reliability	x	
(3) Whenever I scratched a box and did not find the star, I felt as if I was being punished.	x	
(4) Whenever I got a multiple-choice item wrong on the first try, I could feel myself becoming more anxious.		
(6) Using the IF-AT® made me feel more anxious than I otherwise would have while doing the multiple-choice items.	x	x
Questions Not Used in Analyses	x	x
(2) I do not like the fact that the IF-AT® does not let me go back and change my answers the way that an ordinary response form does.	x	x
(5) Whenever I scratched a box and did not find the star, I felt somewhat distracted and found it harder to concentrate on the test.		
(8) Because I was using the IF-AT® rather than an ordinary response form, this test took me longer than it otherwise would have.		

Note. Questions are numbered to indicate the order in which they appeared to the participants. Superscripts indicate the matching reliability question.

to determine the reliability of responses. Complete reliability was not expected for all responses; therefore, responses were included in the final statistical analysis when reliability fell within 2 points. For example, if a student responded with a 5 on the "positive" question they could respond with 3, 2, or 1 on the "negative" question and still be considered a reliable respondent and included in the final analysis. Unreliable participant responses were removed from analyses. Thus, after reliability measures were applied to the data, there were 226 reliable respondents to the Exam 1 survey and 101 reliable respondents to the Exam 2 survey. We did not include responses to Questions 3, 4, and 6 for our study analyses because they were used for reliability purposes; we did not want to use them twice to inflate scoring. To analyze the results consistently across domains we chose to use the positively worded questions, which occurred more frequently throughout the survey.

Data Analysis. A teaching assistant entered students' responses to each survey question into a database. We compiled responses across courses for analyses and completed descriptive and statistical analyses using RStudio (RStudio Team, 2019). To answer Research Questions 1, 2, and 3, related to the impact of students' use of the IF-AT® on their anxiety, acceptance, and experience, we examined survey data for each of the three domains of interest. First, we analyzed the data descriptively. Then we completed McNemar tests, a type of Chi-square analysis for related samples, to test for differences in responses between Exams 1 and 2 (*McNemar Test Calculator,* 2020). There were 101 matched samples included in these analyses (participants who completed both Exams 1 and 2, identified based on student names). We collapsed the five Likert-like scale responses (i.e., agree strongly, agree, neutral, disagree, and disagree strongly) into two groups for statistical analyses (agree strongly/agree and neutral/disagree/disagree strongly), but not for descriptive analyses. We calculated Cohen's g effect sizes for the Chi-square analyses by calculating the greater of $b/(b+c)$ or $c/(b+c)$ which equals P; g is P – 0.5 (Mangiafico, 2016). In a 2 x 2 table, a and d represent concordant cells; and b and c represent discordant cells. We interpreted values of 0.05 – 0.14, 0.15 – 0.24, and > 0.25 as small, medium, and large effect sizes, respectively (Cohen, 1988).

For Research Question 4, we used ordinal logistic regression to predict responses on the survey based on exam score using RStudio (RStudio Team, 2019). Ordinal logistic regression is used when researchers are working to predict a dependent variable that is ordinal in nature or does not have a consistent measurement between ratings on the scale (i.e., agree strongly to disagree strongly; Sullivan & Artino, 2013). Many times, Likert responses are clustered at the extremes which could be problematic for parametric measures because they assume normal distribution. Sullivan and Artino (2013) found support for using parametric measures as a robust way of analyzing Likert scale responses, which supported our use of the ordinal logistic regression analyses. Ordinal logistic regression was used to assess ordinal measures therefore, there is no need to collapse the variables into smaller groups for these analyses. Ordinal logistic regression is reported in 95% confidence intervals (CI) which are a range of the lower limit and upper limit of means. CIs are significant when they do not include zero in the range.

Results

Anxiety. Figure 1 displays the levels of anxiety undergraduate students reported experiencing when using the IF-AT® for Exam 1 and Exam 2. On the figure, responses to the right of the

horizontal line are either *agree* or *agree strongly.* For Exam 1, most students (86%, $n = 195$) *agreed* or *agreed strongly* that they felt rewarded when they received a star, and 82% ($n = 183$) *agreed* or *agreed strongly* that they became less anxious when they got the right answer on the first try. Only 39% of students ($n = 87$) *agreed* or *agreed strongly* that they felt less anxious than they would have on traditional multiple-choice tests. Students felt similarly on Exam 2 such that 86% ($n = 86$) *agreed* or *agreed strongly* that they felt rewarded when they received a star, 74% ($n = 74$) *agreed* or *agreed strongly* that they became less anxious when they got the right answer

Figure 1. *Participant Responses for the Anxiety Domain on Exam 1 and Exam 2*

Anxiety Domain

Exam 1 Responses

1. Whenever I scratched a box and found the star, I felt as if I was being rewarded for my efforts.

1% — 2% — 11% 33% 54%

13. Whenever I got a multiple choice item correct on the first try, I could feel myself becoming less anxious.

0% — 2% 16% 35% 47%

14. Using the IFAT made me feel less anxious than I otherwise would have while doing the multiple-choice items.

5% 22% 34% 23% 16%

Exam 2 Responses

1. Whenever I scratched a box and found the star, I felt as if I was being rewarded for my efforts.

1% — 3% — 11% 43% 43%

13. Whenever I got a multiple choice item correct on the first try, I could feel myself becoming less anxious.

4% — 4% 19% 37% 37%

14. Using the IFAT made me feel less anxious than I otherwise would have while doing the multiple-choice items.

7% 23% 33% 19% 19%

■ Disagree Strongly ■ Disagree ■ Neutral ■ Agree ■ Agree Strongly

on the first try, and 38% ($n = 38$) of students *agreed* or *agreed strongly* that they felt less anxious than they would have on traditional multiple-choice tests. Taken together, these responses suggest that, although students felt positive about certain aspects of the IF-AT®, they generally reported they were not less anxious than they otherwise would have been when taking multiple-choice tests.

Next, we examined if students' levels of anxiety changed from Exam 1 to Exam 2 when using the IF-AT® based on three separate questions. Consistent with the descriptive analyses, there was little change in the level of anxiety students felt over the course of taking two exams as none of the

McNemar analyses were significant (all $ps > .12$), although the analysis for Question 13 yielded a medium effect size (see Table 2). This suggests that fewer students felt less anxious when they answered a question correctly on Exam 2.

Figure 2. *Participant Responses for the Acceptance Domain on Exam 1 and Exam 2*

Acceptance Domain

Exam 1 Responses

7. I would like it if I could use the IF-AT in all my courses that have multiple-choice tests.
9. I think that the IF-AT is fairer than an ordinary response form for multiple-choice tests.
10. I like the fact that the IF-AT lets me know the right answer to every question.
11. I like the fact that the IF-AT form allows me to get part marks on multiple-choice questions.
12. The grading scheme that was used for this test was fair.
15. The IF-AT allowed me to learn from my mistakes.

Exam 2 Responses

7. I would like it if I could use the IF-AT in all my courses that have multiple-choice tests.
9. I think that the IF-AT is fairer than an ordinary response form for multiple-choice tests.
11. I like the fact that the IF-AT form allows me to get part marks on multiple-choice questions.
12. The grading scheme that was used for this test was fair.
15. The IF-AT allowed me to learn from my mistakes.

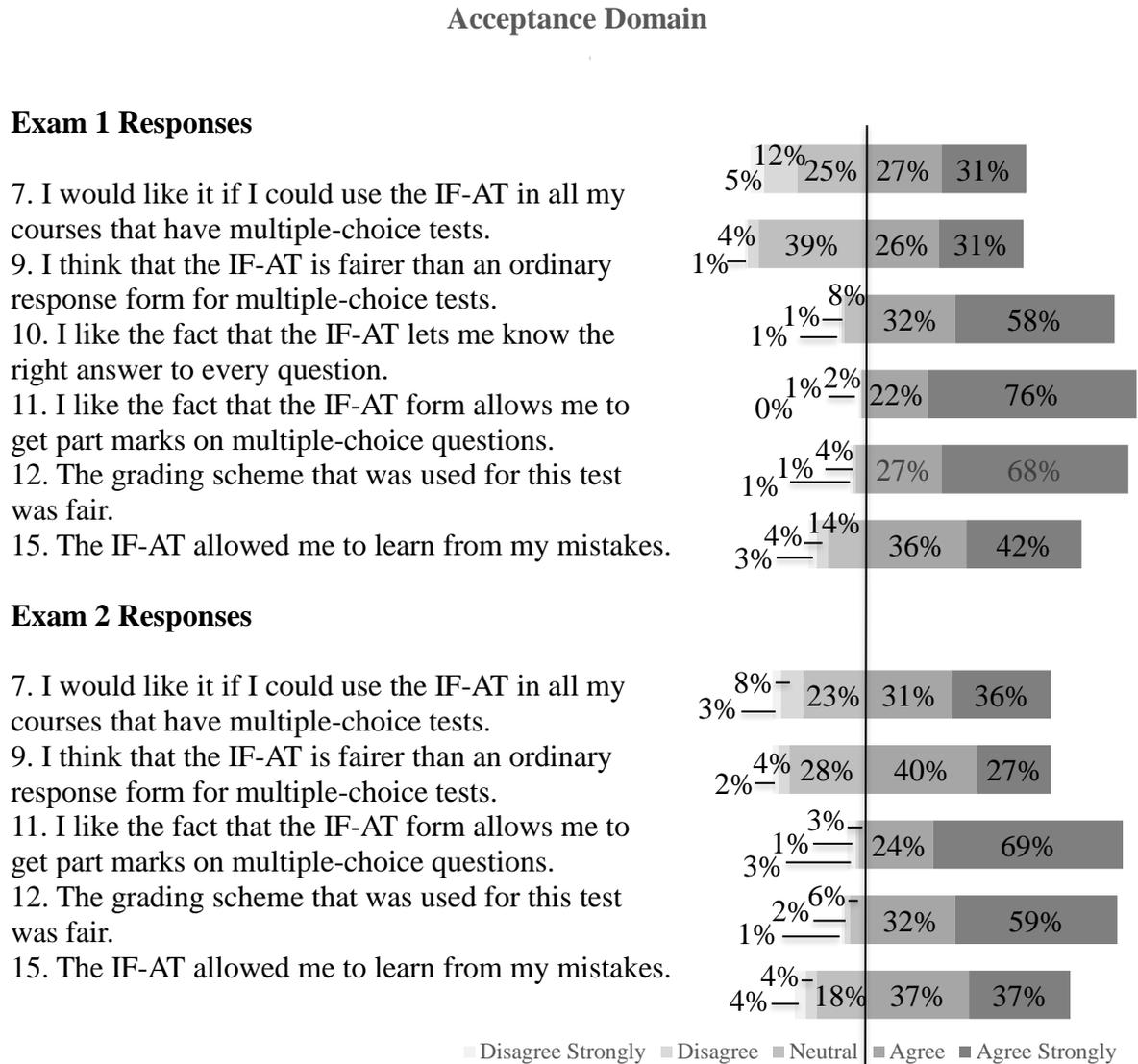Legend: ■ Disagree Strongly ■ Disagree ■ Neutral ■ Agree ■ Agree Strongly

Table 2. *McNemar Analyses Comparing Responses from Exam 1 and Exam 2 Items on Anxiety and Acceptance Domains*

	Exam 2 Responses		
Exam 1 Responses	Agree Strongly / Agree	Disagree Strongly / Disagree / Neutral	χ^2 (effect size)
Anxiety Domain			
1. Being Rewarded			
Agree Strongly / Agree	81	7	0.083
Disagree Strongly / Disagree / Neutral	5	8	(0.08)
13. Becoming Less Anxious			
Agree Strongly / Agree	65	18	2.37
Disagree Strongly / Disagree / Neutral	9	9	(0.17)
14. Less Anxious Than Traditional Tests			
Agree Strongly / Agree	24	14	0
Disagree Strongly / Disagree / Neutral	14	49	(0)
Acceptance Domain			
7. Use In All Courses			
Agree Strongly / Agree	56	8	0.21
Disagree Strongly / Disagree / Neutral	11	26	(-0.08)
9. IF-AT® is Fairer			
Agree Strongly / Agree	49	8	3.12
Disagree Strongly / Disagree / Neutral	18	26	(-0.19)
11. Like Part Marks			
Agree Strongly / Agree	93	5	1.5
Disagree Strongly / Disagree / Neutral	1	2	(0.33)
12. Grading Scheme Is Fair			
Agree Strongly / Agree	91	3	1.33
Disagree Strongly / Disagree / Neutral	0	6	(0.5)
15. Learn From My Mistakes			
Agree Strongly / Agree	72	13	2.72
Disagree Strongly / Disagree / Neutral	5	11	(0.22)

Note: McNemar test was conducted on a total of 101 matched pairs of participants for all questions.
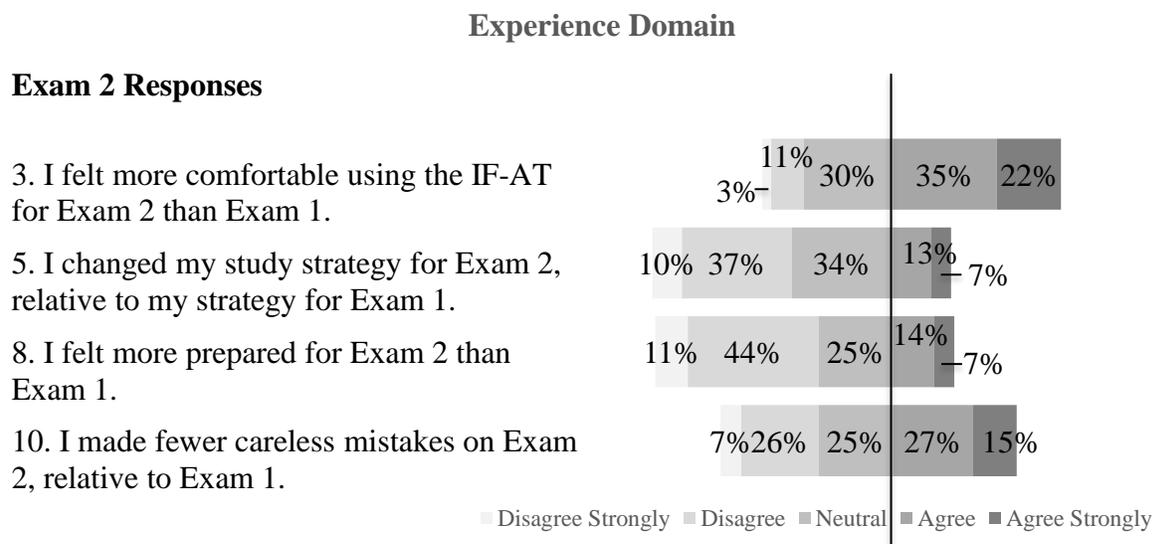
Acceptance. Figure 2 displays the level of acceptance students reported experiencing when test taking using the IF-AT® for Exam 1 and Exam 2. Of the six questions on Exam 1, there were three questions in particular for which the overwhelming majority of students *agreed and agreed*

strongly: (a) Question 10: 90% of students (*n* = 203) liked that the format told them the correct answer, (b) Question 11: 97% of students (*n* = 220) liked getting part marks, and (c) Question 12 - 94% of students (*n* = 215) liked the grading scheme for the IF-AT®. For Exam 2, students were similarly accepting of receiving part marks (Question 11: 93%, *n* = 94) and of the grading scheme (Question 12: 91%, *n* = 91) as on Exam 1. Additionally, for both Exams 1 and 2, most students marked *agree* or *agree strongly* that they liked that they were allowed to learn from their mistakes (Question 15: Exam 1: 78%, *n* = 177; Exam 2: 74%, *n* = 74).

Students were more evenly split on acceptance of other aspects of the IF-AT® format. On Exam 1, 58% of students (*n* = 131) *agreed* and *agreed strongly* that they would prefer to use the IF-AT® in all their courses (Question 7). Additionally, 56% of students (*n* = 128) *agreed or agreed strongly* that the IF-AT® format is fairer than other multiple-choice tests (Question 9). Following Exam 2, student ratings were even more favorable of the IF-AT®, as 67% of students (*n* = 67) *agreed* and *agreed strongly* that they both would prefer to use the IF-AT® in all their courses (Question 7) and that the IF-AT® is a fairer format (Question 9).

The McNemar analyses comparing responses related to acceptance for Exam 1 and Exam 2 indicated no significant differences for all responses (all *ps* >.08). However, the effect sizes for many of the questions were medium or large (see Table 2). The effect size for Question 9 was medium, suggesting that more students indicated that the grading scheme was fairer after Exam 1 than after Exam 2. The effect size for Question 15 was also medium, which may suggest that fewer students felt that they learned from their mistakes after Exam 2 compared to Exam 1. There was a large effect size associated with the analysis for Question 11, indicating that most students like receiving partial credit for questions on both exams. There was also a large effect size for Question 12 with almost all students indicating that the grading scheme was fair for both Exams 1 and 2.

Figure 3. *Participant Responses for the Experience Domain on Exam 2*

Experience Domain

Exam 2 Responses

3. I felt more comfortable using the IF-AT for Exam 2 than Exam 1. — 3%, 11%, 30%, 35%, 22%

5. I changed my study strategy for Exam 2, relative to my strategy for Exam 1. — 10%, 37%, 34%, 13%, 7%

8. I felt more prepared for Exam 2 than Exam 1. — 11%, 44%, 25%, 14%, 7%

10. I made fewer careless mistakes on Exam 2, relative to Exam 1. — 7%, 26%, 25%, 27%, 15%

Disagree Strongly ▪ Disagree ▪ Neutral ▪ Agree ▪ Agree Strongly

Experience with IF-AT®. We evaluated students' experiences using the IF-AT® format over the course of two exams, with three questions on Exam 2 (see Figure 3). For Question 3, most students (57%, $n = 57$) reported that they *agreed or agreed strongly* that they were more comfortable using the IF-AT® on Exam 2, relative to Exam 1. Additionally, for Question 10, 42% of students ($n = 42$) *agreed or agreed strongly* that they made fewer careless mistakes on the Exam 2 compared to Exam 1. For Questions 5 and 8, most students *disagreed or disagreed strongly* with the statements indicating that they changed their study habits for Exam 2 (47%, $n = 47$) and felt more prepared for Exam 2 (55%, $n = 55$), respectively.

Relationship between Objective Exam Scores and Three Domains. The purpose of this research question was to evaluate the extent to which there was a relationship between objective exams scores and each of the three domains. To do this, we first calculated the mean scores for both Exam 1 and Exam 2. On Exam 1, students received a mean score of 91.52% (SD: 6.13; min-max: 71% - 100%). On Exam 2, the mean score was 88.81% (SD: 7.31; min-max: 61.75% -100%). Using these exams scores, we will present the results of the ordinal regression for each of the domains.

We fit models that averaged the results of the individual questions in the Anxiety and Acceptance Domains. There was not a statistically significant relationship between the Anxiety Domain and scores on Exam 1 (95% CI [-0.02, 0.05]) or Exam 2 (95% CI [-0.04, 0.04]). Neither was there a statistically significant relationship between the Acceptance Domain and scores on Exam 1 (95% CI [-0.04, 0.04]) or Exam 2 (95% CI [-0.03, 0.05]). Finally, there was not a significant relationship between the Experience Domain and scores on Exam 2 (95% CI [-0.06, 0.02]).

While none of the domains with questions collapsed were statistically significant, several individual questions were significant (i.e., Questions 3, 10, 11, and 12).Question 12 of the Acceptance Domain was statistically significant for both Exam 1 (95% CI [0.01, 0.10]) and Exam 2 (95% CI [0.02, 0.12]), indicating that the students who judged the grading scheme as fair were more likely to have earned higher grades on the exams. On Exam 2, Question 11 was significant (95% CI [0.02, 0.14]), indicating that students who were more likely to agree that they liked receiving partial credit tended to earn a higher grade on Exam 2. Within the Experience Domain two individual questions were statistically significant: Question 3 (95% CI [0.01, 0.11]) and Question 10 (95% CI [0.01, 0.10]), indicating that students who received higher grades on Exam 2, were more likely to agree that they felt more comfortable using the IF-AT® and made fewer mistakes on Exam 1 than Exam 2, respectively.

Discussion

The purpose of this study was to determine students' anxiety, acceptance, and experience related to using the IF-AT® as a method of multiple-choice testing during an undergraduate language acquisition and science course. Additionally, we examined if survey responses changed from Exam 1 to Exam 2 to evaluate consistency and generalizability of responses. We also analyzed the relationship between survey responses and exam grades.

It should be noted that when interpreting the results of the survey, some of the survey items were written in a one directional format (e.g., I felt more prepared for Exam 2 than Exam 1). These

statements could not be interpreted as students feeling the opposite of the statement if they disagree. For example, Question 14, in which there is a high neutral or disagree response, stated, "Using the IF-AT® made me feel less anxious than I otherwise would have while doing multiple-choice tests." For the students who disagreed or responded neutrally, the conclusion that they felt more anxious using the IF-AT® could not be made. Thus, due to this limitation, we were cautious in our interpretations. Future research should consider including survey questions in which a broader interpretation of students' positive and negative feelings is evaluated and any potential for acquiescent bias is avoided (Bandalos, 2018).

Another note of caution for interpretation of the results is that there were lower rates of responses after Exam 2 than Exam 1. Offering an incentive could have reduced the gap of participation. Reviewed literature incentivized participation with a $10 gift certificate (Dibattista et al., 2004), or eligibility to "win a small monetary prize in a random drawing win a small monetary prize in a random drawing" (Dibattista & Gosse, 2006, p. 315).

Anxiety. Research Question 1 examined levels of anxiety, which is often associated with test taking. The majority of students did not report experiencing less anxiety by using the IF-AT® compared to other tests. This finding is consistent with those of Clark et al. (1998) and Rocklin and Thompson (1985). Further, students agreed at similar rates after Exams 1 and 2 that they felt like they were being rewarded when they found a star and they felt less anxious using the IF-AT® than traditional tests. The opinion of students that did change from Exam 1 to Exam 2 was that fewer students agreed that getting the answer right on the first try reduced their anxiety. Overall, these findings support our prediction that students feel anxiety while test taking. However, our prediction that student anxiety levels would remain the same across tests was not supported.

Test anxiety levels vary among test takers because of the general level of anxiety they experience (Clark et al., 1998) or testing difficulty (Rocklin & Thompson, 1985) rather than testing format. Studies that previously evaluated the IF-AT® found that students liked that the IF-AT® gave them immediate feedback, learning opportunities, partial credit, and knowledge of test score upon completion not because it decreased their test anxiety. While it is difficult to quantify test difficulty in the current study, previous research indicates that test difficulty impacts perceived anxiety more than the question format and that the use of IF-AT® does not lead to additional test anxiety (Clark et al., 1998; Rocklin & Thompson, 1985). Because the current research did not gather information on general anxiety, we cannot say definitively whether student testing anxiety was related to students' personal general anxiety or the use of the IF-AT®. This is a limitation of the current study.

Acceptance. Research Question 2 asked whether students would accept the IF-AT® as a testing format. More than 55% of students *agreed* or *agreed strongly* with all the questions about acceptance on the surveys after both exams. The questions asking about the grading scheme and getting part marks drew more than 90% acceptance on both exams and on Exam 1 over 90% of students also liked that the correct answer to every question was revealed. One student commented: "I really like walking out of the test having a good idea of how I scored." Moreover, nearly two-thirds of students completing the Exam 2 survey *agreed* or *agreed strongly* that they would want to use the IF-AT® in other courses, and that it was more fair than other multiple-choice tests. These descriptive statistics confirm our hypothesis and reveal that the majority of students accepted the

IF-AT®, though responses in favor of the format were not unanimous. Our conclusion is supported by the findings of Clark et al. (1998) that confirmed 72% of participants chose the IF-AT® format when given the option to choose which testing format they used and the findings of DiBattista et al. (2004) that indicated students with a wide range of demographics responded positively with statistical significance. This acceptance rate warrants strong consideration by instructors for more prevalent IF-AT® use in college classrooms particularly because anxiety is present regardless of testing format.

For both the Anxiety and Acceptance Domains, overall responses were consistent as the results of the McNemar Tests indicated that there were not statistically significant changes in responses from Exam 1 to Exam 2. When analyzing the questions individually, there was a notable change from Exam 1 responses to Exam 2 responses for Question 13 in the Anxiety Domain, wherein there was an 8% drop in students agreeing that they felt themselves becoming less anxious when they found a star on the first try. Most of those responses moved to disagreement rather than neutral response. We hypothesize that this change could be because the novelty of the format wore off or fewer students who agreed with the statement self-selected to complete the survey on Exam 2. The discrepancy between the sample sizes on the Exam 1 survey and Exam 2 survey is a limitation of this study, though the level of acceptance the students' displayed cannot be overlooked.

Experience. Research Question 3 asked about student experience using the IF-AT®. Student experience on the current study is difficult to quantify. Students were split on the ways their experiences changed from using the IF-AT® on Exam 1 to Exam 2. More than half reported feeling more comfortable using the testing format, but only 42% felt they made fewer careless mistakes on Exam 2. Most students felt neutrally or disagreed that their study habits changed and that they were more prepared for Exam 2. Due to the one directional nature of Question 8 (students' preparedness), we cannot make the conclusion that students felt more prepared for Exam 1. Nevertheless, one limitation of this study is that data was not collected to understand why students did not feel more prepared for Exam 2. Future studies could examine if students need to be instructed how to change their study habits to best use the IF-AT® format and to further understand what preparation students needed for Exam 2.

Relationship between Domains and Exam Scores. Question 4 asked about the relationship between the exam scores and the responses in each of the survey domains. While many students felt favorably toward the IF-AT®, statistical testing did not find significant relationships between aggregate domains (Anxiety, Acceptance, or Experience) and the students' exam scores. None of the domains were statistically significant for predicting exam grades. This could indicate that students, as a whole, evaluated their anxiety and acceptance of the IF-AT® independently of their exam results. This conclusion is bolstered by the condition of the students knowing their exam results on 90% of the test before completing the survey, which is supported by finding that 90% of students agreed or agreed strongly with the Exam 1 survey statement, "I like the fact that the IF-AT® lets me know the right answer to every question." The current study results were similar to results found by DiBattista and Gosse (2006); some students are more anxious with the IF-AT® testing format, but this is not the case for most students.

Four individual questions (two in Acceptance Domain and two in Experience Domain), however, revealed a statistically significant relationship between exam scores and responses on the survey.

The results, which were reported in confidence intervals, were so close to overlapping with zero that they may hold little practical significance.

Test difficulty was not measured in this study and therefore could be considered a limitation. It is difficult to know whether Exam 2 was more difficult than Exam 1, if students felt less prepared to be assessed on the content of Exam 2, or if something changed in the teaching of content over the years of the study or the student characteristics over the years.

Considerations for the Classroom. Communication Sciences and Disorders instructors should consider using the IF-AT® for several reasons established in the current study and from previous literature. Courses that lend themselves to assessments in a multiple-choice format could benefit from the learning opportunities that the IF-AT® offers. Findings from our study as well as previous research, indicate that students will experience anxiety regardless of testing format (Clark et al., 1998); therefore, the benefits of the IF-AT® format may outweigh these costs. Students can earn additional points and show their partial knowledge of the subject when using the IF-AT® (Epstein Educational Enterprises, 2019; Merrel et al., 2015). Clark et al. (1998) found that when students were offered a choice of testing format 13% of students chose to switch away from using answer-until-correct, but 84% of students chose to switch away from another format to using answer-until-correct. While our study found that over 58% of students would use the IF-AT® in all their courses, fewer than 18% of students disagreed or disagreed strongly with this statement. Although not all students accepted the IF-AT® testing format, the benefits should be considered if the course content is appropriate for multiple-choice testing.

Disclosures

We have no known conflicts of interest to disclose.

Author Note.

Correspondence concerning this article should be addressed to Kirstin Kuchler, M.A., CCC-SLP, University of Minnesota-Twin Cities, Department of Speech-Language-Hearing Sciences, 164 Pillsbury Drive SE, Minneapolis, MN 55455. Email: kuchl001@umn.edu

References

Attali, Y., & Powers, D. (2010). Immediate feedback and opportunity to revise answers to open-ended questions. *Educational and Psychological Measurement*, *70*(1), 22–35. https://doi.org/10.1177/0013164409332231

Bandalos, D.L. (2018). *Measurement theory and applications for the social sciences*. Guilford Press.

Clark, J., Fox, P., & Schneider, H. (1998). Feedback, test anxiety and performance in a college course. *Psychological Reports, 82,* 203-208. https://doi.org/10.2466/pr0.1998.82.1.203

Cohen, J. (1988). Set Correlation and Contingency Tables. Applied Psychological Measurement, 12(4), 425–434. https://doi.org/10.1177/014662168801200410

Daniels, L.M. & Gierl, M.J. (2017). The impact of immediate test score reporting on university students' achievement emotion in the context of computer-based multiple-choice tests. *Learning and Instruction, 52*, p. 27-35. https://doi.org/10.1016/j.learninstruc.2017.04.001

DiBattista, D., Mitterer, J.O., & Gosse, L. (2004). Acceptance by undergraduates of the immediate feedback assessment technique for multiple-choice testing. *Teaching in Higher Education*, *9*(1), 17-28. https://doi.org/10.1080/1356251032000155803

DiBattista, D. & Gosse, L. (2006). Test anxiety and the immediate feedback assessment technique. *The Journal of Experimental Education*, *74*(4), 311-327. https://doi.org/10.3200/JEXE.74.4.311-328

DiBattista, D., Gosse, L., Sinnige-Egger, J., Candale, B., & Sargeson, K. (2009). Grading scheme, test difficult, and the immediate feedback assessment technique. *The Journal of Experimental Education*, *77*(4), 311-336. https://doi.org/10.3200/JEXE.77.4.311-338.

Dihoff, R.E., Brosvic, G.M., Epstein, M.L., & Cook, M.J. (2004). Provision of feedback during preparation for academic testing: Learning is enhanced by immediate but not delayed feedback. *The Psychological Record*, *54*, 207-231.

Epstein Educational Enterprises. (2020, August 13). *IF-AT: Immediate feedback-assessment technique* http://www.epsteineducation.com/home/default.aspx

Epstein, M. L., Epstein, B. B., & Brosvic, G. M. (2001). Immediate feedback during academic testing. *Psychological reports*, *88*(3 Pt 1), 889–894. https://doi.org/10.2466/pr0.2001.88.3.889

Farland, M.Z., Barlow, P.B., Lancaster, T.I., & Franks, A.S. (2015). Comparison of answer-until-correct and full-credit assessments in a team-based learning course. *American Journal of Pharmaceutical Education*, *79*(2), 21. http://doi.org/10.5688/ajpe79221

Green, K. (1981). Item-response changes on multiple-choice tests as a function of test anxiety. *The Journal of Experimental Education*, *49*(4), 225-228. https://doi-org.ezp3.lib.umn.edu/10.1080/00220973.1981.11011788

Lee, S., Su, H., & Lee, S. (2012). Effects of computer-based immediate feedback on foreign language listening and test-associated anxiety. *Perceptual and Motor Skills, 114*(3), 995-1006. https://doi.org/10.2466/28.11.21.PMS.114.3.995-1006.

Mangiafico, S.S. (2016). *Summary and Analysis of Extension Program Evaluation in R, version 1.18.1*.Rutgers Cooperative Extension. rcompanion.org/handbook/

Martlett, N.L. & Watson, D. (1968). Test anxiety and immediate or delayed feedback in a test-like avoidance task. *Journal of Personality and Social Psychology, 8(2), 200-203.*

McNemar Test Calculator, (2020, August 22). McNemar's test to analyze a matched case-control study. https://www.graphpad.com/quickcalcs/McNemar1.cfm.

Merrel, J.D., Cirillo, P.F., Schwartz, P.M., & Webb, J.A. (2015). Multiple-choice testing using Immediate Feedback—Assessment Technique (IF AT®) forms: Second-chance guessing vs. second-chance learning? *Higher Education Studies*, *5*(5). https://doi.org/10.5539/hes.v5n5p50

Muniz, J. & Menendez, F. (2011). The answer-until-correct item format revisited. *Methodology*, *7*(3), 103-110. https://doi.org/10.1027/1614-2241/a000028

RStudio Team (2019). RStudio: Integrated Development for R. RStudio, PBC. http://www.rstudio.com/.

Rocklin, T. & Thompson, J.M. (1985). Interactive effects of test anxiety, test difficulty, and feedback. *Journal of Educational Psychology, 77*(3)*, 368-372,* https://doi.org/10.1037/0022-0663.77.3.368

Sullivan, G.M., & Artino, A.R. (2013). Analyzing and interpreting data from Likert-type scales. *Journal of Graduate Medical Education, 5*(4)*, 541-54*2. https://doi.org/10.4300/jgme-5-4-18

Vanderoost, J., Janssen, R., Eggermont, J., Callens, R., & De Laet, T. (2018). Elimination testing with adapted scoring reduces guessing and anxiety in multiple-choice assessments, but does not increase grade average in comparison with negative marking. *Plos One*, *13*(10). https://doi.org/10.1371/journal.pone.0203931

Appendix A

OPTIONAL IF-AT® SURVEY Exam 1 (No Credit)

Survey Question	1 Disagree Strongly	3 Neither Agree Nor Disagree			5 Agree Strongly
1. Whenever I scratched a box and found the star, I felt as if I was being rewarded for my efforts.	1	2	3	4	5
2. I do not like the fact that the IF-AT® does not let me go back and change my answers the way that an ordinary response form does.	1	2	3	4	5
3. Whenever I scratched a box and did not find the star, I felt as if I was being punished.	1	2	3	4	5
4. Whenever I got a multiple-choice item wrong on the first try, I could feel myself becoming more anxious.	1	2	3	4	5
5. Whenever I scratched a box and did not find the star, I felt somewhat distracted and found it harder to concentrate on the test.	1	2	3	4	5
6. Using the IF-AT® made me feel more anxious than I otherwise would have while doing the multiple-choice items.	1	2	3	4	5
7. I would like it if I could use the IF-AT® in all of my courses that have multiple-choice tests.	1	2	3	4	5
8. Because I was using the IF-AT® rather than an ordinary response form, this test took me longer than it otherwise would have.	1	2	3	4	5
9. I think that the IF-AT® is fairer than an ordinary response form for multiple-choice tests.	1	2	3	4	5
10. I like the fact that the IF-AT® lets me know the right answer to every question.	1	2	3	4	5
11. I like the fact that the IF-AT® form allows me to get part marks on multiple-choice questions.	1	2	3	4	5
12. The grading scheme that was used for this test was fair.	1	2	3	4	5
13. Whenever I got a multiple-choice item correct on the first try, I could feel myself becoming less anxious.	1	2	3	4	5
14. Using the IF-AT® made me feel less anxious than I otherwise would have while doing the multiple-choice items.	1	2	3	4	5
15. The IF-AT® allowed me to learn from my mistakes.	1	2	3	4	5

Other IF-AT® Comments:

_____

Appendix B

OPTIONAL IF-AT® SURVEY Exam 2 (No Credit)

Survey Question	1 Disagree Strongly	3 Neither Agree Nor Disagree			5 Agree Strongly
1. Whenever I scratched a box and found the star, I felt as if I was being rewarded for my efforts.	1	2	3	4	5
2. I do not like the fact that the IF-AT® does not let me go back and change my answers the way that an ordinary response form does.	1	2	3	4	5
3. I felt more comfortable using the IF-AT® for Exam 2 than exam 1.	1	2	3	4	5
4. Whenever I got a multiple-choice item wrong on the first try, I could feel myself becoming more anxious.	1	2	3	4	5
5. I changed my study strategy for Exam 2, relative to my strategy for exam 1.	1	2	3	4	5
6. Using the IF-AT® made me feel more anxious than I otherwise would have while doing the multiple-choice items.	1	2	3	4	5
7. I would like it if I could use the IF-AT® in all of my courses that have multiple-choice tests.	1	2	3	4	5
8. I felt more prepared for Exam 2 than Exam 1.	1	2	3	4	5
9. I think that the IF-AT® is fairer than an ordinary response form for multiple-choice tests.	1	2	3	4	5
10. I made fewer careless mistakes on Exam 2, relative to Exam 1.	1	2	3	4	5
11. I like the fact that the IF-AT® form allows me to get part marks on multiple-choice questions.	1	2	3	4	5
12. The grading scheme that was used for this test was fair.	1	2	3	4	5
13. Whenever I got a multiple-choice item correct on the first try, I could feel myself becoming less anxious.	1	2	3	4	5
14. Using the IF-AT® made me feel less anxious than I otherwise would have while doing the multiple-choice items.	1	2	3	4	5
15. The IF-AT® allowed me to learn from my mistakes.	1	2	3	4	5

Other IF-AT® Comments:

_____

_