The Journal of Experimental Education, 2006, 74(4), 311–327

Test Anxiety and the Immediate Feedback Assessment Technique

DAVID DiBATTISTA
Brock University

LEANNE GOSSE
Wilfrid Laurier University

ABSTRACT. The authors examined the relationship between the reactions of undergraduate students to using the Immediate Feedback Assessment Technique (IFAT), an answer form that provides immediate feedback on multiple-choice questions, for the first time on a major examination and their levels of test anxiety and trait anxiety. They also assessed whether students with higher levels of test anxiety and trait anxiety might be disadvantaged relative to other students by use of the IFAT in a testing situation. They found that preference of undergraduates (N = 185) for the IFAT was not related to test anxiety, nor did evidence indicate that the IFAT put students with higher levels of test anxiety at a disadvantage with respect to test performance. Using the IFAT did not generally increase test-related anxiety, and for a majority of students, immediate feedback actually reduced it. Nineteen percent of students felt that immediate feedback interfered with their test performance but would nevertheless still prefer to use the IFAT in future tests. Potential concerns that test-anxious students may either dislike the IFAT or be disadvantaged by its use appear unwarranted and should not deter instructors from adopting the IFAT.

Key words: feedback, Immediate Feedback Assessment Technique, multiple-choice testing, test anxiety

The authors thank the Brock University Centre for Teaching, Learning, and Educational Technology for a grant to support this research. Aspects of this research have been presented at the 2004 meetings of the Canadian Psychological Association and Eastern Psychological Association. The authors have no financial interest in the Immediate Feedback Assessment Technique.

Address correspondence to: David DiBattista, Department of Psychology, Brock University, St. Catharines, Ontario L2S 3A1, Canada. E-mail: david.dibattista@brocku.ca

ATTEMPTS TO PROVIDE IMMEDIATE FEEDBACK in the context of multiple-choice testing date back more than 75 years and have taken a variety of forms involving the use of mechanical devices (Pressey, 1950), latent image technology (Abplanalp, 1995), and computers (Wise, Plake, Pozehl, Barnes, & Luken, 1989). The Immediate Feedback Assessment Technique (IFAT) is a relatively new, commercially available answer form for multiple-choice testing that can be used conveniently with large classes.[1] On the IFAT form, there is a series of small boxes corresponding to the four alternatives for up to 50 multiple-choice questions, with each box covered by a waxy, opaque coating. The box associated with the correct alternative for a given question has a star in it, and the remaining boxes are blank. For each multiple-choice question, the student chooses the alternative believed to be correct and scratches the coating off the appropriate box. If the student's choice is correct, a star is revealed and the student goes on to the next item. If the student's choice is incorrect, the box is blank, and the student reconsiders the remaining alternatives and continues scratching boxes until the star is revealed. For each item, the student's final selection is always the correct answer. Students are awarded full credit for answering an item correctly on the first attempt, and at the instructor's discretion, they may earn progressively less credit for answering correctly on subsequent attempts, thus rewarding them for their proximate knowledge of the correct answer.

From a pedagogical perspective, an advantage of the IFAT is that it provides students with corrective immediate feedback for each test item. Although feedback interventions do not invariably lead to improvements in performance (Kluger & DeNisi, 1996, 1998), immediate feedback often has been shown to be more effective than delayed feedback in promoting learning in classroom situations (Kulik & Kulik, 1988; Sassenrath, Yonge, & Schrable, 1968). Furthermore, corrective feedback not only enhances learning in a variety of settings (Brosvic, Walker, Perry, Degnan, & Dihoff, 1997; Elliott, 1988; Phye & Bender, 1989; Taras, 2003) but is also more effective than feedback that provides information only about the correctness or incorrectness of responses (Bangert-Drowns, Kulik, Kulik, & Morgan, 1991; Kluger & DeNisi, 1996). Because the effects of feedback interventions on performance may be positive, neutral, or even negative, Kluger and DeNisi (1998) recommended that new techniques for providing feedback not be assumed to have positive effects but rather be investigated empirically. In keeping with this recommendation, Epstein and his colleagues have assessed the IFAT's contribution to learning in experiments in both the classroom (Dihoff, Brosvic, & Epstein, 2003; Epstein, Epstein, & Brosvic, 2001) and the laboratory (Epstein et al., 2002). They found that participants consistently learned

[1]Information about the IFAT, including details about cost and availability, can be obtained at http://www.epsteineducation.com or by contacting Michael L. Epstein, Department of Psychology, Rider University, Lawrenceville, NJ 18648.

more when they used the IFAT rather than a traditional answer form that provided no feedback, such as the more widely used, computer-scored, fill-in-the-blank Scantron form. In addition, the corrective immediate feedback provided by the IFAT also promoted learning more effectively than did similar feedback from computerized testing (Epstein et al., 2002). Thus, although both the IFAT and Scantron forms serve as effective assessment tools, only the IFAT contributes to students' learning.

A second advantage of the IFAT is that students not only readily accept its use, but they actually welcome it and would like to see its use expanded (DiBattista, Mitterer, & Gosse, 2004). More than 80% of undergraduates using the IFAT for the first time indicated that they would like to be able to use the IFAT in all of their courses that have multiple-choice tests, and 64% felt the IFAT to be fairer than the more commonly used machine-scored answer form they were accustomed to using. Moreover, although they were using the IFAT during an examination that counted for up to 30% of their course grade, 78% of students nevertheless indicated that the IFAT made the test feel something like a game. Students' acceptance of the IFAT has also been found to be essentially independent of both test performance and a variety of personal characteristics (DiBattista et al.). For example, students' preference for the IFAT was not found to be correlated with any of the following variables: overall test performance, performance on the multiple-choice portion of the test, age, number of courses previously taken, students' self-reported degree of preparedness for the test, and students' perceptions of the difficulty of the multiple-choice items. Overall, then, students' reactions to the IFAT are positive across students with a broad range of characteristics, suggesting that its use may contribute to the important but usually neglected goal of creating a more positive reaction to testing among students (McMorris, Boothroyd, & Pietrangelo, 1997).

Despite the acceptance of the IFAT, it must be kept in mind that certain subgroups of students may find use of the IFAT to be problematic. For example, certain aspects of the testing circumstances can put more highly anxious students at a disadvantage relative to their classmates (Hill & Wigfield, 1984). Ordinarily, students taking multiple-choice tests in the classroom obtain no feedback on their performance during the test, but the use of the IFAT changes the testing circumstances quite dramatically by providing students with immediate corrective feedback on an item-by-item basis. Of course, for almost all students, at least some of the feedback that they obtain while using the IFAT will be negative, and negative feedback tends to be most detrimental to more highly anxious individuals (Auerbach, 1973; Eysenck, 1982; Hill & Eaton, 1977). Students with higher levels of test anxiety may, therefore, be less accepting of the IFAT and perhaps even be at a disadvantage when using it instead of a more traditional multiple-choice answer form.

Although test anxiety was originally thought of as a unidimensional construct (Sarason, 1961), later theorists have conceptualized test anxiety as having two or

more dimensions. For example, Liebert and Morris (1967) and Spielberger, Gonzalez, Taylor, Algaze, and Anton (1978) have proposed that the two major dimensions of test anxiety are worry and emotionality, and Sarason (1984) has argued that test-irrelevant thinking and bodily symptoms are also important dimensions. Spielberger et al. have conceptualized test anxiety as being a situation-specific anxiety trait. According to this perspective, test-anxious students generally have higher levels of trait anxiety, tend to find examinations threatening because of their evaluative nature, and experience higher levels of state anxiety when taking tests (Spielberger & Vagg, 1995). It is not surprising that researchers have found test anxiety to be inversely related to students' performance in a wide variety of testing situations (Clark, Fox, & Schneider, 1998; Hembree, 1988; Musch & Bröder, 1999; Powers, 2001). Furthermore, a number of situational factors, such as time pressure (Plass & Hill, 1986), have been shown to exacerbate the anxiety associated with test situations and further detract from test performance.

Results of research on the effects of immediate feedback on anxiety during testing have not been consistent. Some researchers have found immediate feedback to be associated with decreases in anxiety (Arkin & Schumann, 1984; Morris & Fulmer, 1976; Rocklin & Thompson, 1985). On the other hand, researchers have also frequently observed increases in anxiety (Strang & Rust, 1973; Wise, Plake, Eastman, Boettcher, & Lukin, 1986; Wise et al., 1989), as well as reductions in test performance (Delgado & Prieto, 2003; Kluger & DeNisi, 1996; Wise et al., 1986). Wise et al. (1986) postulated that encountering failure during testing, and being informed of this failure via immediate feedback, may cause an increase in anxiety, which in turn contributes to impairment in performance. Factors such as overall test difficulty and the order of item difficulty (Wise et al., 1986) may influence the effects of immediate feedback on test-related anxiety, but factors influencing the nature of the relationship between immediate feedback and anxiety during testing have not been fully elucidated.

A number of studies have produced no evidence that students using the IFAT rather than a standard answer form either experience more anxiety during testing or answer fewer items correctly on initial attempts (Dihoff et al., 2003; Dihoff, Brosvic, Epstein, & Cook, 2004; Epstein et al., 2001, 2002). However, there has been no research to date on students' reactions to the IFAT as a function of their level of test anxiety. Test performance is typically inversely related to levels of test anxiety (Clark, Fox, & Schneider, 1998; Hembree, 1988; Musch & Bröder, 1999; Powers, 2001), which means that more highly test-anxious students will be likely to receive more frequent negative feedback when using the IFAT. Furthermore, negative feedback tends to have its greatest impact on both anxiety level (Auerbach, 1973) and performance (Eysenck, 1982; Hill & Eaton, 1977) among more highly anxious individuals. It is, therefore, reasonable to hypothesize that for students with higher levels of test anxiety, using the IFAT may exacerbate feelings of anxiety and impair test performance, which may in turn contribute to

reduced acceptance of the IFAT as a response technique. Examination of this hypothesis is particularly important because test anxiety is not uncommon, with a prevalence rate estimated to be in the range of 10% to 15% among postsecondary students (Hill & Wigfield, 1984). In the research presented here, we examined the relationship between the reactions of undergraduate students to using the IFAT for the first time on a major examination and their levels of test anxiety and trait anxiety. We also assessed whether students with higher levels of test anxiety and trait anxiety might be disadvantaged relative to other students by use of the IFAT in a testing situation.

Method

Participants

Participants were undergraduate students enrolled in an introductory level research design and statistics course taught by the first author and required for students majoring in psychology. None of the students had prior experience with the IFAT. Participation was voluntary, with participants being eligible to win a small monetary prize in a random drawing. All personal information remained entirely confidential. All procedures were reviewed and approved by the Brock University Research Ethics Board.

Of the 215 students enrolled in the course at the time of the test, 185 used the IFAT on the test and also completed at least two of the three questionnaires, and we present the data for these students here. Ninety percent (n = 167) of the participants were female, an accurate reflection of the actual composition of the class. Students ranged in age from 19 to 43 years (M = 20.9, SD = 2.7).

Materials

The trait anxiety scale was taken from the International Personality Item Pool (IPIP; 2001), which provides a variety of scales that are freely available for commercial and scientific use. Respondents use a 5-point scale (1 = *very inaccurate* to 5 = *very accurate*) to rate themselves on 10 statements, such as "I worry about things" and "I am relaxed most of the time" (negatively keyed). The IPIP trait anxiety scale has an acceptable 2-week test–retest reliability of 0.91 (DiBattista, 2003). Goldberg (1999) reported that this scale has Cronbach's alpha coefficient equal to .83 and a correlation, corrected for unreliability of the scales, of .90 with the Revised NEO Personality Inventory (NEO-PI-R) anxiety scale (Costa & McCrae, 1992).

The Revised Test Anxiety Scale (RTAS) was originally developed by combining the Test Anxiety Inventory (Spielberger et al., 1978) and the Reactions to Tests scale (Sarason, 1984), and retaining 18 items having an alpha coefficient

equal to .88 (Benson, Moulin-Julian, Schwarzer, Seipp, & El-Zahhar, 1992). Respondents use a 4-point scale (1 = *almost never* to 4 = *almost always*) to rate themselves on statements such as "I worry a great deal before taking important tests" and "During tests, I find myself thinking of things unrelated to the material being tested." Scores on the RTAS can range from 18 to 72. Exploratory and confirmatory factor analyses on multinational samples revealed that the RTAS has the same four-factor structure as the Reactions to Tests scale, namely, tension, worry, test-irrelevant thinking, and bodily symptoms (Benson et al.).

After using the IFAT for the first time, participants completed a questionnaire designed to elicit their reactions to the IFAT. Most items used a 5-point Likert scale (1 = *strongly disagree*, 3 = *neither disagree nor agree*, 5 = *strongly agree*). Five items (Table 1) related to students' general acceptance of the IFAT, and six items related specifically to anxiety in the context of the test situation (Table 2). Cronbach's alpha coefficients were .71 and .81, respectively, for the two sets of items. Two close-ended questions asked students to select from separate lists the aspects of the IFAT that they liked the most and the least; alternatives presented in the lists were derived from written comments provided by several hundred students who had previously used the IFAT for in-class tests. At the beginning of the term, all students in the course indicated that they had previously used Scantron answer forms for multiple-choice testing on one or more occasions in other postsecondary courses, and for this reason the Scantron form served as the primary point of comparison in some of the questionnaire items.

Procedure

Over a period of 4 weeks, students completed during their regular class time a series of paper-and-pencil questionnaires, including those mentioned above; certain aspects of the questionnaires were unrelated to this research study and will not be discussed further. After students had completed the final in-class questionnaire, the format and use of the IFAT were described in detail during class time, and students were informed that they would be using it later in the course. Care was taken to describe the IFAT in a neutral fashion so that students' reactions to the IFAT would not be biased by the instructor's comments. For example, students were shown an IFAT form and told how to use it, but they were not told of research indicating that students prefer using IFAT (DiBattista et al., 2004).

Two weeks later, the students used the IFAT on a 3-hr test that was administered at the half-way point of the course and counted for 31% of the course grade. The 37 four-option multiple-choice items counted for 37% of the marks on the test, with 1.0, .25, .1, and 0 marks being awarded for correct answers given on the first, second, third, and fourth attempt, respectively. The multiple-choice items were presented in a random order at the end of the test, and all students completed the multiple-choice portion of the test after completing the nonmulti-

TABLE 1. Responses to Individual Questionnaire Items Pertaining to Acceptance of the Immediate Feedback Assessment Technique (IFAT)

Question	Disagree (%)	Agree (%)	*M*	*SD*	*t*[a]	*d*
1. I would like it if I could use the IFAT form in all of my courses that have multiple-choice tests.	9	77	4.16	1.09	13.2***	1.06
2. Getting immediate feedback on the multiple-choice items that interfered with my performance on the test.	53	19	2.45	1.17	–5.87***	–.47
3. I think that the IFAT is fairer than other types of multiple-choice response techniques (e.g., Scantron).	8	59	3.79	1.07	9.14***	.74
4. I like the fact that the IFAT form allows me to get part marks on multiple-choice questions.	1	95	4.72	.58	37.0***	2.97
5. Even if the IFAT did not allow me to get part marks, I would still like to be able to use it on multiple-choice tests.	34	53	3.27	1.46	2.26*	.18

Note. $n = 154$. The response scale for each item was 1 = *disagree strongly*, 3 = *neither agree nor disagree*, 5 = *agree strongly*.
[a]One-sample *t* tests compared the mean score for each item with a test value of 3, which was the value assigned to the response *neither agree nor disagree*. All tests were two-tailed and had degrees of freedom = 153. Effect size(d) = (mean – 3)/*SD*.
*$p < .05$. **$p < .01$. ***$p < .001$.

ple-choice portion, which consisted of statistical problems, short answers, and essays. Immediately after completing the test, students completed the questionnaire asking for their reactions to the IFAT.

We conducted statistical analyses using the SPSS for Windows statistical package, version 11.0.1 (SPSS Inc., Chicago, Illinois). Effect sizes were computed following the recommendations of Olejnik and Algina (2000).

Results

The mean (± *SD*) test grade was 69.8 ± 14.8. The mean percentage of marks earned on the multiple-choice portion of the test was 68.3 ± 12.7, with students

TABLE 2. Responses to Individual Questionnaire Items Pertaining to the Immediate Feedback Assessment Technique (IFAT) and Anxiety

Question	Disagree (%)	Agree (%)	*M*	*SD*	*t*[a]	*d*
1. Using the IFAT made me feel less anxious than I otherwise would have while doing the multiple-choice items.	36	35	3.05	1.25	.45	.04
2. Using the IFAT made me feel more anxious than I otherwise would have while doing the multiple-choice items.	38	30	2.82	1.20	−1.81	−.15
3. Whenever I got a multiple-choice item correct on the first try, I could feel myself becoming less anxious.	13	71	3.97	1.18	10.2***	.82
4. Whenever I got a multiple-choice item wrong on the first try, I could feel myself becoming more anxious.	12	70	3.90	1.11	10.1***	.81
5. Getting immediate feedback on multiple-choice items made me feel less anxious than I otherwise would have.	20	54	3.49	1.16	5.20***	.42
6. Getting immediate feedback on multiple-choice items made me feel more anxious than I otherwise would have.	45	31	2.74	1.29	−2.50*	−.20

Note. $n = 154$. The response scale for each item was 1 = *disagree strongly*, 3 = *neither agree nor disagree*, 5 = *agree strongly*.

[a]One-sample *t* tests compared the mean score for each item with a test value of 3, which was the value assigned to the response *neither agree nor disagree*. All tests were two-tailed and had degrees of freedom = 153. Effect size(*d*) = (mean − 3)/*SD*.

*$p < .05$. **$p < .01$. ***$p < .001$.

earning 62.5 ± 14.3% and 5.9 ± 2.1%, respectively, on their initial and subsequent attempts on the multiple-choice items. Thus, as in a previous study in which the same grading scheme was used, students' scores were about 6 percentage points higher than they would have been if partial credit had not been available (DiBattista et al., 2004).

As Table 1 indicates, students had very positive attitudes toward the IFAT. More than three quarters of students agreed that they would like to be able to use the IFAT in all of their courses that have multiple-choice tests; only 9% disagreed. In addition, only 19% reported that getting immediate feedback on multiple-choice items interfered with their performance on the test, whereas 59% of the students found the IFAT to be fairer than other types of multiple-choice response techniques. It is not surprising that almost all students liked being able to earn partial credit on multiple-choice questions, with more than half the students nevertheless saying that they would still like to be able to use the IFAT even if a partial-credit grading scheme were not used.

When asked to indicate the one aspect of the IFAT that they liked the most, students selected being able to earn partial credit (26%), learning the right answer to every question (20%), getting a second chance to respond to a missed item (12%), knowing their grade on the multiple-choice portion of the test (11%), and gaining confidence when getting the right answer (10%). The aspects that students liked the least were that it is too easy to scratch on the wrong line of the IFAT form (30%), they lose confidence when they get the wrong answer (30%), and it takes more time to answer questions when using the IFAT instead of the more traditional response formats (12%).

Table 3 shows details regarding RTAS and trait anxiety scores. The full-scale RTAS and the trait anxiety scores demonstrated acceptable levels of internal consistency, as measured by Cronbach's alpha coefficient. Furthermore, three of the four RTAS subscales also showed acceptable internal consistency, although the alpha coefficient value for the bodily symptoms subscale proved to be low (Kline, 1986).

The mean RTAS score in the present sample was quite similar to means reported by Benson et al. (1992). However, although women typically have higher levels of test anxiety than men (Hembree, 1988), we detected no statistically sig-

TABLE 3. Trait Anxiety and Revised Test Anxiety Scale (RTAS) Scores

			RTAS subscales			
Measure	Trait anxiety (n = 175)	RTAS (n = 185)	Tension	Worry	Test-irrelevant thinking	Bodily symptoms
M	31.60	37.45	14.22	12.78	5.58	4.86
SD	6.95	8.68	3.58	3.99	2.06	1.73
SEM	1.11	1.04	0.57	0.60	0.35	0.67
Range	12–47	20–61	6–20	6–23	3–11	3–12
α	.84	.88	.84	.85	.83	.61

nificant sex difference in RTAS scores here (women: 37.69 ± 9.08; men: 35.26 ± 7.93), $t(183) = 1.09$, $p > .05$. On the other hand, as is often found to be the case (Ben-Zur & Zeidner, 1988), the trait anxiety scores of women (32.31 ± 6.57) were significantly higher than those of men (25.00 ± 7.11), $t(173) = 4.32$, $p < .001$; effect size = mean difference/total standard deviation = .95.

Correlations among various measures of test performance and RTAS and test anxiety scores are shown in Table 4. As expected, RTAS scores were negatively correlated with all measures of test performance. The correlations between RTAS scores and performance on the multiple-choice and nonmultiple-choice portions of the test were not statistically significantly different, $t(182) = -1.82$, $p > .05$; therefore, test anxiety was not more strongly related to performance on one or the other component of the test.

There was a statistically significant direct relationship between RTAS scores and trait anxiety scores, $r(173) = .47$, $p < .001$, a finding that has been commonly reported (Hembree, 1988). However, despite this substantial relationship, the data in Table 4 indicate that trait anxiety was not significantly correlated with any measure of test performance. Furthermore, the relationships among RTAS scores and the various measures of performance all remained significant and were not appreciably diminished when we controlled for the effect of trait anxiety by means of partial correlation.

Students' preference for the IFAT (Table 1, Item 1) was not statistically significantly related to age, gender, number of courses previously taken, the extent to which students felt prepared for the test, the perception of the difficulty of the multiple-choice items, the number of points that students earned by responding correctly to multiple-choice items after their first attempt or to any measure of test performance ($p > .05$ in all cases). Furthermore, preference for the IFAT was

TABLE 4. Correlations Between Measures of Test Performance and Revised Test Anxiety Scale (RTAS) and Trait Anxiety Scores

Test portion	RTAS *r*	RTAS-partial[a] *r*	Trait anxiety *r*
All	−.20**	−.19*	−.02
Multiple-choice	−.26***	−.20**	−.14
Nonmultiple-choice	−.15*	−.16*	.04

Note. Degrees of freedom for statistical tests: RTAS = 183; RTAS-partial = 172; trait anxiety = 173.

[a]RTAS-partial is the partial correlations between RTAS and test performance variables, controlling for the effect of trait anxiety.

*$p < .05$. **$p < .01$. ***$p < .001$.

not significantly related to either test anxiety or trait anxiety. Specifically, students' scores on the RTAS and the trait anxiety scale had correlations of only –.02 and –.07, respectively, with preference for the IFAT (Table 1, Item 1).

The extent to which students felt that immediate feedback interfered with their test performance (Table 1, Item 2) was not statistically significantly associated with RTAS or trait anxiety scores, or with any measure of test performance ($p > .05$ in all cases). However, this variable was inversely related to preference for the IFAT, $r(152) = -.33$, $p < .001$. Nevertheless, even for students ($n = 29$) who reported that immediate feedback interfered with their performance, the mean IFAT-preference score (3.52 ± 1.33) value was significantly greater than 3, the midpoint value on the 5-point disagree–agree scale, $t(28) = 2.10$, $p < .05$; effect size = (mean – 3)/*SD* = .39. Furthermore, only 24% (7/29) of these students said that they would not prefer to use the IFAT in all of their courses that use multiple-choice tests. Thus, even students who felt immediate feedback interfered with their test performance demonstrated a clear preference for the IFAT.

As the data in Table 2 indicate, using the IFAT did not have a consistent effect on students' self-reported anxiety during the test. However, more than one third of students felt that using the IFAT actually reduced their anxiety while doing the multiple-choice items, and fewer than one third felt that it increased their anxiety. It is not surprising that more than two thirds of students found that answering a multiple-choice item correctly on the first attempt reduced their anxiety, whereas answering it incorrectly increased their anxiety. However, when asked specifically about how receiving immediate feedback on multiple-choice items influenced their anxiety, more than half of students said it reduced their anxiety, whereas fewer than one third said it increased their anxiety.

Table 5 shows that responses to the anxiety-related questionnaire items were not statistically significantly associated with either RTAS or trait anxiety scores, although they were correlated with preference for the IFAT (Table 1, Item 1). Students' preference for the IFAT was lower when using the IFAT, getting immediate feedback on multiple-choice items, or answering an item incorrectly on the first attempt increased their anxiety, and preference for the IFAT was higher under the opposite conditions. Also, as mentioned earlier, students' perceptions that immediate feedback interfered with their test performance (Table 1, Item 2) were inversely related to their preference for the IFAT ($r = -.33$). Furthermore, students' perceptions about the negative effect of immediate feedback on performance were also significantly related to their reports that their anxiety increased from using the IFAT, $r(152) = .53$, $p < .001$, getting immediate feedback on multiple-choice items, $r(152) = .46$, $p < .001$, and answering an item incorrectly on the first attempt, $r(152) = .35$, $p < .001$. When these anxiety-related measures were controlled for by means of partial correlation, the correlation between preference for the IFAT and students' perceptions that immediate feedback interfered with their test performance disappeared, $r(152) = -.03$, $p > .05$.

TABLE 5. Correlations Between the Anxiety-Related Questionnaire Items and Preference for the Immediate Feedback Assessment Technique (IFAT), Revised Test Anxiety Scale (RTAS) Score, and Trait Anxiety Score

Question	Preference for IFAT[a] *r*	RTAS score *r*	Trait anxiety *r*
1. Using the IFAT made me feel less anxious than I otherwise would have while doing the multiple-choice items.	.48***	−.04	−.06
2. Using the IFAT made me feel more anxious than I otherwise would have while doing the multiple-choice items.	−.48***	.00	.06
3. Whenever I got a multiple-choice item correct on the first try, I could feel myself becoming less anxious.	.19*	.12	−.02
4. Whenever I got a multiple-choice item wrong on the first try, I could feel myself becoming more anxious.	−.21**	.04	.08
5. Getting immediate feedback on multiple-choice items made me feel less anxious than I otherwise would have.	.47***	.12	.02
6. Getting immediate feedback on multiple-choice items made me feel more anxious than I otherwise would have.	−.40***	.02	.04

Note. $n = 154$. The response scale for each item was 1 = *disagree strongly*, 3 = *neither agree nor disagree*, 5 = *agree strongly*.
[a]Preference for the IFAT is the score for Item 1 of Table 1. Degrees of freedom for statistical tests: preference for IFAT = 152; RTAS = 152; trait anxiety = 145.
*$p < .05$. **$p < .01$. ***$p < .001$.

It appears then that students' perception that immediate feedback interferes with their test performance may be mediated by increased anxiety associated with using the IFAT. However, further examination of the available data did not provide any indication of why certain students are more likely than others to find use of the IFAT more anxiety producing. We had postulated that students who had higher RTAS scores and lower multiple-choice scores might find the IFAT to

be anxiety provoking, feel that immediate feedback interfered with their performance, and have a lower preference for the IFAT. However, linear regression analyses with RTAS scores and scores on the multiple-choice portion of the test as predictor variables did not produce statistically significant results in predicting any of these variables. For example, with preference for the IFAT (Table 1, Item 1) as the criterion variable and RTAS and multiple-choice scores as predictor variables, the value of R was only .02, $F(2, 151) = .05$, $p > .05$.

Discussion

As in previous research (DiBattista et al., 2004; Epstein & Brosvic, 2002), students had favorable responses to the IFAT, with more than three quarters indicating that they would like to use the IFAT in all of their courses having multiple-choice tests. It is not surprising that almost all students liked having the opportunity to obtain partial credit on multiple-choice items. However, because more than half of the students would like to use the IFAT even if partial credit were not available, the availability of partial credit does not by itself account for their acceptance of the IFAT. For example, students may also prefer the IFAT because they believe that it is fairer than other types of multiple-choice response techniques, lets them learn more, and makes the test feel a bit like a game (DiBattista et al.). Students reacted less favorably to other aspects of the IFAT. For example, many students reported that they lost confidence when they answered incorrectly, and many found that the layout of the IFAT form makes it too easy to scratch on the wrong line, a technical problem that might be resolved by redesigning the form.

Our findings agree with previous research that has consistently shown test anxiety to be inversely related to test performance in a variety of settings (Clark et al., 1998; Hembree, 1988; Musch & Bröder, 1999; Powers, 2001). We found inverse associations among RTAS scores and various measures of test performance, including overall test scores and scores on the multiple-choice and non-multiple-choice portions of the test. In contrast, trait anxiety was not associated with any of these measures of test performance, and, furthermore, the associations between RTAS scores and test performance were not statistically significantly diminished when trait anxiety was statistically controlled. These findings are consistent with the notion that test anxiety is related to, yet distinct from, trait anxiety (Spielberger et al., 1978).

Despite the inverse relationship between test anxiety and test performance, students' preference for the IFAT was not found to be related to test anxiety, any measure of test performance, or any of a number of demographic variables. Furthermore, there was no indication that even those students with higher levels of test anxiety and poorer performance on the multiple-choice portion of the test had a reduced preference for IFAT.

In the context of academic testing, immediate feedback has been associated with both increases (Strang & Rust, 1973; Wise et al., 1986, 1989) and decreases in state anxiety (Arkin & Schumann, 1984; Morris & Fulmer, 1976; Rocklin & Thompson, 1985). Like other researchers (Dihoff et al., 2003, 2004), we did not find that the IFAT increased test-related anxiety for students in general. Moreover, the immediate feedback provided by the IFAT actually reduced the test-related anxiety of a majority of students, despite the fact that the IFAT was being used for the first time and was novel, which might be expected to heighten students' anxiety (Zuckerman, 1976). Some of the reduction in anxiety associated with the IFAT seems to be attributable to the fact that anxiety is reduced when students respond correctly on the initial attempt, and for most students, there will be many more correct than incorrect responses. A second anxiety-reducing factor may be that the IFAT makes the multiple-choice portion of the exam feel like a game (DiBattista et al., 2004), which may serve to de-emphasize the evaluative aspects of the testing situation. In general, reductions in test-related anxiety resulting from immediate feedback were associated with a higher preference for the IFAT, and vice versa. Moreover, there was no indication that even those students with higher levels of test anxiety and poorer multiple-choice performance found the use of the IFAT to be particularly anxiety provoking.

A minority of students (19%) indicated that getting immediate feedback interfered with their test performance. It is not surprising that this belief was inversely related to preference for the IFAT, although it is important to note that there was no evidence that the belief was, in fact, associated with an actual impairment of test performance. The students' perception that immediate feedback interferes with their test performance may be mediated by increased state anxiety associated with using the IFAT. However, it is at present unclear why the IFAT increases anxiety more in some students than in others, although the effect does not seem to be related to either test anxiety or trait anxiety. In addition, there was no indication that even those students with either higher levels of test anxiety or poorer performance on the multiple-choice portion of the test, or both, were particularly likely to feel that immediate feedback interfered with their performance.

Kluger and DeNisi (1998) have suggested that feedback during testing may sometimes cause attention to be directed from the task at hand to the self, thus interfering with performance. Further research may show whether students who feel that the immediate feedback provided by the IFAT interferes with their test performance tend to have either personality or cognitive traits, or both, that make them especially prone to this type of detrimental attentional shift. Because negative immediate feedback has been found to have particularly detrimental effects on task performance among individuals with an internal attributional style (Mikulincer, 1988), a high need for structure (Mikulincer, Yinon, & Kabili, 1991), and a proneness to cognitive interference (Mikulincer, 1989), these three variables may be of particular interest in determining why some stu-

dents believe that the immediate feedback provided by the IFAT interferes with their test performance.

A limitation of this study is that only a small minority of the participants were men. However, this limitation is mitigated by the fact that there were no statistically significant differences between men and women in test anxiety, nor were there significant differences between the responses of male and female students to any of the questionnaire items in Tables 1 and 2 (detailed results not presented here). A second limitation of this study, like almost all studies that have involved the IFAT (e.g., DiBattista et al., 2004; Dihoff et al., 2003, 2004; Epstein et al., 2001, 2002), is that the research participants were undergraduate students. Our decision to use undergraduates was primarily based on the need to have a large sample of students who would be taking an important multiple-choice test in a real-life setting, thus providing the type of environmental context that is particularly appropriate for studying test anxiety. Because multiple-choice tests are so widely used in undergraduate settings, the choice of undergraduates was appropriate for this study. Nevertheless, multiple-choice testing is also widely used outside of the educational system, such as in military and industrial settings. In the future, it will be desirable to study the use of the IFAT in environments involving multiple-choice testing in which users are not traditional students.

We had suspected that students with high levels of test anxiety might be disadvantaged by using the IFAT and like using it less than would other students. However, the evidence presented here does not indicate that the IFAT discriminates against students based on their level of test anxiety, nor is preference for the IFAT related to the level of test anxiety. Furthermore, there is no evidence that even those students who are high in test anxiety and perform poorly on the multiple-choice portion of the test are either particularly disadvantaged by the use of the IFAT or inclined to dislike it. Although negative findings must always be interpreted with caution, these results have practical implications for the use of the IFAT. Instructors who suspect that the IFAT discriminates against students with higher levels of test anxiety may choose not to use it. However, the findings reported here indicate that such concerns are likely to be unwarranted and, therefore, should not deter instructors from adopting the IFAT. Nevertheless, it may be helpful to allow students to become accustomed to using the IFAT form before they are required to use it in a multiple-choice test that counts for course credit. Prior familiarization will reduce the novelty of the IFAT in the first testing situation and perhaps allay the anxiety that use of the IFAT may induce in a minority of students (Zuckerman, 1976).

REFERENCES

Abplanalp, P. (1995). Application of latent image technology to the multiple-choice test format. *Optometric Education, 20,* 118–123.

Arkin, R. M., & Schumann, D. W. (1984). Effects of corrective testing: An extension. *Journal of Educational Psychology, 76,* 835–843.

Auerbach, S. M. (1973). Effects of orienting instructions, feedback information and trait-anxiety level on state-anxiety. *Psychological Reports, 33,* 779–786.

Bangert-Drowns, R. L., Kulik, C. C., Kulik, J. A., & Morgan, M. T. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research, 61,* 213–238.

Benson, J., Moulin-Julian, M., Schwarzer, C., Seipp, B., & El-Zahhar, N. (1992). Cross-validation of a Revised Test Anxiety Scale using multi-national samples. In K. A. Hagtvet & T. B. Johnsen (Eds.), *Advances in anxiety research* (Vol. 7, pp. 62–83). Amsterdam: Swets & Zeitlinger.

Ben-Zur, H., & Zeidner, M. (1988). Sex differences in anxiety, curiosity, and anger: A cross-cultural study. *Sex Roles, 19,* 335–347.

Brosvic, G. M., Walker, M. A., Perry, N., Degnan, S., & Dihoff, R. E. (1997). Illusion decrement as a function of duration of inspection and figure type. *Perceptual and Motor Skills, 84,* 779–783.

Clark, J. W., Fox, P. A., & Schneider, H. G. (1998). Feedback, test anxiety and performance in a college course. *Psychological Reports, 82,* 203–208.

Costa, P. T., Jr., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (Neo-PI-R) and NEO Five-Factor Inventory (NEO-FFI): Professional manual.* Odessa, FL: Psychological Assessment Resources.

Delgado, A. R., & Prieto, G. (2003). The effect of item feedback on multiple-choice test responses. *British Journal of Psychology, 94,* 73–85.

DiBattista, D. (2003) *Test-retest reliability of the IPI Trait Anxiety Scale.* Unpublished raw data.

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,* 17–28.

Dihoff, R. E., Brosvic, G. M., & Epstein, M. L. (2003). The role of feedback during academic testing: The delay retention effect revisited. *The Psychological Record, 53,* 533–548.

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.

Elliott, D. (1988). The influence of visual target and limb information on manual aiming. *Canadian Journal of Psychology, 42,* 57–68.

Epstein, M. L., & Brosvic, G. M. (2002). Students prefer the Immediate Feedback Assessment Technique. *Psychological Reports, 90,* 1136–1138.

Epstein, M. L., Epstein, B. B., & Brosvic, G. M. (2001). Immediate feedback during academic testing. *Psychological Reports, 88,* 889–894.

Epstein, M. L., Lazarus, A. D., Calvano, T. B., Matthews, K. A., Hendel, R. A., Epstein, B. B., et al. (2002). Immediate feedback assessment technique promotes learning and corrects inaccurate first responses. *The Psychological Record, 52,* 187–201.

Eysenck, M. W. (1982). *Attention and arousal: Cognition and performance.* Berlin: Springer-Verlag.

Goldberg, L. R. (1999). A broad-bandwidth, public-domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality psychology in Europe* (Vol. 7, pp. 7–28). Tilburg, The Netherlands: Tilburg University Press.

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

Hill, K. T., & Eaton, W. O. (1977). The interaction of test anxiety and success-failure experiences in determining children's arithmetic performance. *Developmental Psychology, 13,* 205–211.

Hill, K. T., & Wigfield, A. (1984). Test anxiety: A major educational problem and what can be done about it. *The Elementary School Journal, 85,* 105–126.

International Personality Item Pool. (2001). *A scientific collaboratory for the development of advanced measures of personality traits and other individual differences.* Retrieved September 2002 from http://ipip.ori.org/

Kline, P. (1986). *A handbook of test construction.* London: Methuen.

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback intervention on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119,* 254–284.

Kluger, A. N., & DeNisi, A. (1998). Feedback interventions: Toward the understanding of a double-edged sword. *Current Directions in Psychological Science, 7,* 67–72.

Kulik, J. A., & Kulik, C. C. (1988). Timing of feedback and verbal learning. *Review of Educational*

Research, 58, 79–97.
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.
McMorris, R. F., Boothroyd, R. A., & Pietrangelo, D. J. (1997). Humor in educational testing: A review and discussion. *Applied Measurement in Education, 10,* 269–297.
Mikulincer, M. (1988). Reactance and helplessness following exposure to unsolvable problems: The effects of attributional style. *Journal of Personality and Social Psychology, 54,* 679–686.
Mikulincer, M. (1989). Cognitive interference and learned helplessness: The effects of off-task cognitions on performance following unsolvable problems. *Journal of Personality and Social Psychology, 57,* 129–135.
Mikulincer, M., Yinon, A., & Kabili, D. (1991). Epistemic needs and learned helplessness. *European Journal of Personality, 5,* 249–258.
Morris, L. W., & Fulmer, R. S. (1976). Test anxiety (worry and emotionality) changes during academic testing as a function of feedback and test importance. *Journal of Educational Psychology, 68,* 817–824.
Musch, J., & Bröder, A. (1999). Test anxiety versus academic skills: A comparison of two alternative models for predicting performance in a statistics exam. *British Journal of Educational Psychology, 69,* 105–116.
Olejnik, S., & Algina, J. (2000). Measures of effect size for comparative studies: Applications, interpretations, and limitations. *Contemporary Educational Psychology, 25,* 241–286.
Phye, G. D., & Bender, T. (1989). Feedback complexity and practice: Response pattern analysis in retention and transfer. *Contemporary Educational Psychology, 14,* 97–110.
Plass, J. A., & Hill, K. T. (1986). Children's achievement strategies and test performance: The role of time pressure, evaluation anxiety, and sex. *Developmental Psychology, 22,* 31–36.
Powers, D. E. (2001). Test anxiety and test performance: Comparing paper-based and computer-adaptive versions of the Graduate Record Examination. *Journal of Educational Computing Research, 24,* 249–273.
Pressey, S. L. (1950). Development and appraisal of devices providing immediate automatic scoring of objective tests and concomitant self-instruction. *The Journal of Psychology, 29,* 417–447.
Rocklin, T., & Thompson, J. M. (1985). Interactive effects of test anxiety, test difficulty, and feedback. *Journal of Educational Psychology, 77,* 368–372.
Sarason, I. G. (1961). Test anxiety and the intellectual performance of college students. *Journal of Educational Psychology, 52,* 210–206.
Sarason, I. G. (1984). Stress, anxiety, and cognitive interference: Reactions to tests. *Journal of Personality and Social Psychology, 46,* 929–938.
Sassenrath, J. M., Yonge, G. D., & Schrable, K. (1968). Immediate and delayed feedback on examinations and immediate and delayed retention. *California Journal of Educational Research, 19,* 226–231.
Spielberger, C. D., Gonzalez, H. P., Taylor, C. J., Algaze, B., & Anton, W. D. (1978). Examination stress and test anxiety. In C. D. Spielberger & I. G. Sarason (Eds.), *Stress and anxiety* (Vol. 5, pp. 167–191). New York: Wiley.
Spielberger, C. D., & Vagg, P. R. (1995). Test anxiety: A transactional process model. In C. D. Spielberger, & P. R. Vagg, *Test anxiety: Theory, assessment, and treatment* (pp. 3–14). Washington, DC: Taylor & Francis.
Strang, H. R., & Rust, J. O. (1973). The effects of immediate knowledge of results and task definition on multiple-choice answering. *The Journal of Experimental Education, 42,* 77–80.
Taras, M. (2003). To feedback or not to feedback in student self-assessment. *Assessment & Evaluation in Higher Education, 28,* 549–565.
Wise, S. L., Plake, B. S., Eastman, L. A., Boettcher, L. L., & Lukin, M. E. (1986). The effects of item feedback and examine 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, L. E. (1989). Providing item feedback in computer-based tests: Effects of initial success and failure. *Educational and Psychological Measurement, 49,* 479–486.
Zuckerman, M. (1976). Sensation seeking and anxiety, traits and states, as determinants of behavior in novel situations. In I. G. Sarason & C. D. Spielberger (Eds.), *Stress and anxiety* (Vol. 3, pp. 141–170). Oxford, UK: Hemisphere.