Journal of Information Systems Education

Volume 35 | Issue 4 Article 7

12-15-2024

Multiple-Choice Test Format and Student Test Anxiety: A Case Set in a Technical Analytics Class

Stephane E. Collignon
West Virginia University, stephane.collignon@mail.wvu.edu

Josey Chacko
Mount St. Mary's University, chacko@msmary.edu

Salman Nazir
West Virginia University, salman.nazir@mail.wvu.edu

Recommended Citation

Collignon, Stephane E.; Chacko, Josey; and Nazir, Salman (2024) "Multiple-Choice Test Format and Student Test Anxiety: A Case Set in a Technical Analytics Class," *Journal of Information Systems Education*: Vol. 35 : Iss. 4 , 467-480.
DOI: https://doi.org/10.62273/QDMI3914
Available at: https://aisel.aisnet.org/jise/vol35/iss4/7

Journal of Information Systems Education

Volume 35
Issue 4
Fall 2024

Multiple-Choice Test Format and Student Test Anxiety: A Case Set in a Technical Analytics Class

Stephane E. Collignon, Josey Chacko, and Salman Nazir

Recommended Citation: Collignon, S. E., Chacko, J., & Nazir, S. (2024). Multiple-Choice Test Format and Student Test Anxiety: A Case Set in a Technical Analytics Class. *Journal of Information Systems Education*, 35(4), 467-480. https://doi.org/10.62273/QDMI3914

Article Link: https://jise.org/Volume35/n4/JISE2024v35n4pp467-480.html

Received:	August 5, 2023
First Decision:	December 8, 2023
Accepted:	June 14, 2024
Published:	December 15, 2024

ISSN: 2574-3872 (Online) 1055-3096 (Print)

Journal of Information Systems Education, 35(4), 467-480, Fall 2024
https://doi.org/10.62273/QDMI3914

Multiple-Choice Test Format and Student Test Anxiety: A Case Set in a Technical Analytics Class

Stephane E. Collignon
John Chambers College of Business and Economics
West Virginia University
Morgantown, WV 26506, USA
stephane.collignon@mail.wvu.edu

Josey Chacko
Richard J. Bolte, Sr. School of Business
Mount St. Mary's University
Emmitsburg, MD 21727, USA
chacko@msmary.edu

Salman Nazir
John Chambers College of Business and Economics
West Virginia University
Morgantown, WV 26506, USA
salman.nazir@mail.wvu.edu

ABSTRACT

Most business schools require students to take at least one technical Management Information System (MIS) course. Due to the technical nature of the material, the course and the assessments tend to be anxiety inducing. With over three out of every five students in US colleges suffering from "overwhelming anxiety" in some form, we study whether or not the perception of test format congruence (i.e., ability to reward knowledge) leads to satisfaction with the test format and lower test anxiety. In this study, we also considered the impact risk-taking profiles have on satisfaction with the test format. Using data collected from our survey, we conducted an exploratory factor analysis on the measurement model and a confirmatory factor analysis on the structural model. We found that test congruence positively impacts satisfaction with the format, satisfaction impacts anxiety negatively, and risk profile does not seem to play a role. These findings contribute theoretically as we create an integrated framework grounded in different theoretical views. The findings also have practical implications as they allow instructors to see that aligning assessments to reward knowledge can help manage students' anxiety.

Keywords: Assessment, Student satisfaction, Information systems education, Large classes, Test anxiety

1. INTRODUCTION

Management Information System (MIS) programs often face the conundrum of teaching technical content to a mass business audience. In our study, the MIS department at a large land-grant university in the United States offers a required undergraduate business course—in a face-to-face format—which introduces students to the decision sciences aspect of business analytics. This course is technical and generally challenges our students, which, over the years, and across many student evaluations, was recognized to trigger frustrations. Due to the technical content in this course, as in standard MIS material, MIS educators are increasingly forced to recognize that it triggers higher anxiety amongst students (Zoller & Ben-Chaim, 1989), and in some cases, this anxiety leads to psychological distress and poor health. Since student anxiety is particularly acute during assessments, it would be beneficial to MIS educators to find new ways of reducing test anxiety.

Managing student anxiety is particularly important with over 60% of college students experiencing some form of "overwhelming anxiety" and many universities across the US, including ours, are elevating the resources supporting mental health (Wolverton, 2019). Naturally, one way MIS educators can contribute would be to monitor and predict stress (Kim et al., 2022) and identify some paths to reduce anxiety induced by tests. In our context, because the course is part of the required curriculum for all business majors, and there are multiple sections with large class sizes (100+), for purposes of efficiency, faculty use multiple-choice question (MCQ) tests. While many MCQ variations exist including negative marking

(not popular because punishment for guessing has a psychological disadvantage [Crocker & Algina, 1986; Lesage et al., 2013; Prieto & Delgado, 1999; Vanderoost et al., 2018]), the number right (NR) format (where students receive points for the correct answer and none for omission or incorrect responses) is dominant (Lesage et al., 2013). This popularity and reluctance to change format in some fields—like accounting and medicine—is partly due to the need to align with national certification practices. The field of MIS is less constrained. Thus, we changed the MCQ format used in our class, and we seized the opportunity to study whether perceptions surrounding the test format were linked with test anxiety. The MCQ format we use in our study, proposed by Collignon et al. (2020), is a validated variation on the standard Number Right (NR) format. This format enables instructors to gather more individualized information on student knowledge (Collignon et al., 2020), which is interesting in MIS because traditional MCQ formats do not provide sufficient information for individualized feedback on analytical MIS abilities (Kuechler & Simkin, 2003). Information gain is not the subject of our study, but the format change provided a stage for studying links between perceptions and test anxiety. This format capitalizes on the familiarity students have with the NR format but also introduces the "I don't know" option. As per Collignon et al. (2020), students indicate their lack of knowledge using the "I don't know" option to questions and receive a partial score of 0.35 (slightly above the threshold of the expected value of answering at random; based on best practices outlined by Budescu & Bar-Hillel, 1993 and Espinoza & Gardeazabal, 2010). Since scoring the "I don't know" option implies rewarding the acknowledgment of the lack of knowledge instead of letting students guess and gamble for points, we wondered if students' perceptions could vary and impact test anxiety. Since this format impacts two levers of control—knowledge, and gambling (i.e., risk-taking)—we wondered if these specific levers were linked to student satisfaction and test anxiety. Thus, without advocating for the use of any specific test format, we use the change of information technology (test format) as a context to observe the impact on concepts that are consequential to student learning. Specifically, we explore to what extent the levers of control (risk-taking and test congruence, i.e., rewarding knowledge) are related to satisfaction and anxiety. Our results show that the test format congruence is positively related to satisfaction, which is negatively related to anxiety. The student risk profile does not seem to be related to satisfaction with the test format. These results contribute both practically and theoretically to academic literature. Our framework integrates the path hypothesized based on different theories in a single model. By proceeding through exploratory factor analysis, our study also confirms simplified scales that are practical for shortening survey instruments. Practically, we are showing that instructors who change their MCQ design need to pay attention to congruence because they can impact student perceptions of the test format and reduce their anxiety.

The remainder of our paper is structured as follows. We first review the literature and then justify our hypotheses. After that, we present our results and discuss them before concluding.

2. LITERATURE REVIEW

2.1 Testing and Student Test Anxiety

Test anxiety is a significant issue affecting institutions of higher learning for several reasons. First, about 20% of college students exhibit and are affected by high levels of test anxiety (Schouwenburg, 1999; von der Embse et al., 2018) and, increasingly, more and more students are experiencing general anxiety (Duraku, 2017), with more than 60% experiencing overwhelming anxiety (Wolverton, 2019). Second, studies indicate that test anxiety activates an increased risk for psychological distress like depression (Leadbeater et al., 2012; Zeidner, 1998, 2007), as well as poor health outcomes (Zeidner, 1998, 2007). Finally, test anxiety negatively affects student performance (Cassady & Johnson, 2002; Chapell et al., 2005; Ramirez & Beilock, 2011; Zeidner, 1998, 2007), and thus any assessment of a student's ability is muddled by test anxiety (Zeidner, 2007), and critically, because of poorer performance, students' options for advancement in life are curtailed (Zeidner, 1998). The impact of test anxiety on student performance is nuanced—there are facilitating aspects to test anxiety as well (Alpert & Haber, 1960). Hahn et al. (2017) explain this dual-natured phenomenon by linking it to the Yerkes-Dodson Curve (Yerkes & Dodson, 1908), i.e., performance improves with test anxiety (arousal) up to an optimal point and then reduces thereafter. In the end, too much anxiety is negative, and identifying the means to reduce it is of interest.

Test anxiety impacts students differently based on various factors and attributes. For example, test anxiety tends to be higher in freshmen than in seniors (Duraku, 2017; Sansgiry & Sail, 2006), higher in females (Cassady & Johnson, 2002; Zeidner, 1990; Zoller & Ben-Chaim, 1989), and higher in subject areas that are more technical/quantitative (Zoller & Ben-Chaim, 1989) even when the perception of difficulty is controlled for (Everson et al., 1993). Similarly, students' perceptions have an impact on test anxiety. For example, perceptions of course load and course difficulty were linked to increased test anxiety (Sansgiry & Sail, 2006; von der Embse et al., 2018), perceptions of test validity/fairness were linked negatively to test anxiety (Chu et al., 2014; Hamilton, 1994), and perceptions based on experience and past performance have an impact on anxiety (Hahn et al., 2017). In addition to all these factors, the format of the test instrument has an impact on anxiety.

2.2 Format of Test Instrument: The Importance of Fairness and Control

In the past, the literature on the impact of test instrument format and student test anxiety has been conflicted, with some studies indicating that there is no impact (Rowley, 1974). However, more recently, there is a preponderance of evidence indicating that the format of testing instruments does indeed have an impact on student test anxiety (Furnham et al., 2013; Guraya et al., 2018; Hahn et al., 2017; Meijer, 2001; Sparfeldt et al., 2013; Zeidner, 1990).

Research has clearly and consistently shown over the past several decades that students with high anxiety prefer some types of assessments over others (Furnham et al., 2013; Green, 1981; Struyven et al., 2005). This is important because when the testing instrument is the student's preferred format, test anxiety is lower (Zoller & Ben-Chaim,1989). Students prefer MCQs (Birenbaum, 2007; Zeidner, 1990) and specifically,

students with high anxiety prefer MCQs (Struyven et al., 2005). Students like the familiarity of MCQs (von de Embse et al., 2018) and view MCQs as less threatening (McDaniel et al., 1994) and less anxiety-inducing (Zeidner, 1987, 1990). Most essentially, students find MCQs fairer (Birenbaum, 2007) and students find assessments satisfactory when they are fair (Wygal et al., 2017).

In addition to the reasons outlined above, the student's perception of the level of control in a test has an impact on test anxiety (Shermis & Lombard, 1998). The level of control is a characteristic of the test we can act upon. As we used a format in which students could opt out of answering (i.e., checking "I don't know"), we could expect variation regarding perceptions of control in two ways. First, the change of format might attract students' attention to the congruence of the test format—i.e., whether the test format rewards knowledge (controlling performance through knowledge). Second, as students are offered the option to secure points rather than gamble, the change of format might raise their awareness about comfort with risk (control as knowing how to play the game).

2.3 Congruence

In general, in the education literature, congruence in educational assessments refers to the alignment between the material covered, the learning objectives, and the assessment used to assess the mastery of that material (Dick et al., 2008; Gagne et al., 2005; James, 2010). This definition looks at congruence from the instructor/instruction perspective. From the student/assessed perspective, congruence refers to how students perceive the fairness of the assessment instrument in relation to the course material (Wygal et al., 2017), or in other words, congruence is the alignment between the material learned and how well the test format assesses that knowledge or the lack of it. This is practically important for three reasons: One, congruence between material learned and assessment format leads to student satisfaction with the assessment format (Kincaid & Zemke, 2006). Second, if students are satisfied with the assessment instrument it has a positive impact on learning and academic achievement (Brown, 2011), and finally, perceptions by students on assessment congruence are linked negatively to text anxiety (Chu et al., 2014; Hamilton, 1994). This idea of congruence from the assessed perspective is important outside education as well. In the HR literature, Whiting et al. (2008) found that the more alignment there was between the appraisal instrument and employees' expectation of what needs to be appraised, the more employees were satisfied.

For the purposes of our study, we are specifically looking at assessment congruence from the perspective of the student—i.e., the alignment between the test format and the student's perception of fairness—that the test format assesses and indeed rewards knowledge. Past studies have established that a student's perception of the assessment instrument affects learning and academic achievement (Brown, 2011), notably via anxiety reduction (Zeidner, 1990). In the next section, we hypothesize the links between elements of control (congruence and risk profile), satisfaction, and anxiety.

3. HYPOTHESES DEVELOPMENT

We posit that satisfaction with the test format will be negatively correlated with anxiety for several reasons. First, if students are satisfied with the test format, they may perceive the situation to be of lower or manageable psychological demands and they may feel more empowered and in control (Elsass & Veiga, 1997). Studies that examined the relationship between job demands and physical and psychological strain suggest that as employees find themselves to be more satisfied and in control of their job demands, they may have lower levels of negative outcomes such as physical and psychological strain (Elsass & Veiga, 1997). Satisfaction with the test format may help students feel more optimistic about future outcomes and may enable them to believe that they have more control on the test environment, leading to a more hopeful outlook about future results. Such satisfaction may counteract anxiety related to circumstances that give less control and increase fear of unknown outcomes (Seligman, 2006).

When students are satisfied with the test format, they may perceive it as fair and unbiased. They may perceive that the test gives them a fair chance at doing well and correctly assesses their knowledge without unnecessarily creating difficulty for them and unjustly evaluating them. Feelings of fairness cause a reduced allostatic load due to the positive energy around the task, thus causing lower detrimental and physiological costs such as stress and anxiety (Galanakis et al., 2015; McEwen & Wingfield, 2003). This has the overall effect of reducing anxiety for students. On the contrary, a test that has a format that students are not satisfied with may increase their anxiety level.

Third, research suggests that when students are not satisfied with the testing format, they do not see the relevance of the format, causing them to feel more disengaged and disinterested (Jones & Egley, 2004). This lack of relevance affects the well-being of individuals and can lead to higher levels of stress and anxiety. This is based on self-determination theory, which posits that satisfaction of psychological needs such as relevance positively impacts well-being. Hence, if the test format fails to satisfy these motivations, test takers may feel disempowered which in turn may lead to higher levels of stress (Deci & Ryan, 2012). On the contrary, when students believe that the format is relevant and fair, they tend to be satisfied with the testing format, which lowers their anxiety (Chu et al., 2014). Based on these reasons, we hypothesize that satisfaction with the test format will be negatively associated with anxiety.

Hypothesis 1: Satisfaction with the test format will be negatively correlated with anxiety.

Congruence is the perception that the test format tests knowledge conveyed in class. When students believe that the test format is appropriate for testing the knowledge imparted to them in class, rewards knowledge attained in class, and is geared to truly reward people who know the correct answer, they tend to believe that the test format is congruent. Kincaid and Zemke (2006) found that assessment congruence is related to satisfaction with the testing format. That is, as students perceive that the assessment tool is appropriate for testing the knowledge that they have acquired in class and correctly rewards their knowledge, they tend to feel more satisfied with the testing format. In a study on employee performance evaluation, it was found that employees are more satisfied with their evaluation process if the instrument is congruent and evaluates the right points (Whiting et al., 2008). Specifically, when employees perceived a congruence between their ideal appraisal system and the current appraisal system

used by their employer, they indicated a higher level of satisfaction with the appraisal system. We also find evidence in information systems literature that when information representation matches the skill and the type of problem to be solved, users may be better able to achieve their tasks (Vessey & Galletta, 1991). This is based on the Cognitive Fit theory which suggests that when the cognitive demands of a task match the cognitive abilities and the mental model of individuals, there is a greater chance that it may enhance satisfaction and help achieve better performance outcomes.

Research has also reported that when the test format aligns with students' expectations and values, they may feel more satisfied with the test format. This is based on the Expectancy-Value theory, which posits that individuals are more motivated and satisfied when they perceive a task as valuable and expect to have the capability to do well on the task (Wigfield & Eccles, 2000). These studies indicate that congruence of the test format may be positively related to satisfaction with the test format. Hence, we propose:

Hypothesis 2: Congruence will be positively correlated with test format satisfaction.

A risk personality is one where an individual is willing to engage in behaviors that may have uncertain or potentially negative outcomes. With True/False MCQs offering an "I don't know" option in the medical field, Anderson (2004) found that risk personality impacts performance but did not investigate if satisfaction with the test format is impacted. Our study investigates that point. Some test takers may thrive in environments where they can take calculated risks and engage in challenging tasks. They may perceive a greater potential for reward in such environments. Therefore, when provided with opportunities where risk-taking behavior is rewarded, they may feel more satisfied with the testing format. This is in line with literature that suggests that people with risk-taking personalities tend to seek out challenging tasks as they may perceive them to be more stimulating and rewarding (Zuckerman, 2007). Since an environment that allows risk takers to engage in challenging tasks aligns better with their preferences, risk takers' satisfaction may be impacted positively when a test format rewards gambling and negatively when it rewards "playing safe."

Individuals with risk personalities may also enjoy the challenging environment of the test if they think that they perform better in challenging situations. This may lead to higher satisfaction with a testing format that allows the test takers to compete with their peers and allows them to show their mastery of concepts. Such students may thrive in challenging environments since these environments tend to allow these students to excel and achieve higher performance. This phenomenon is also observed in corporate settings where CEOs who are overconfident and enjoy risky and challenging environments are especially enthusiastic in such hard settings (Hirshleifer et al., 2012).

Unlike typical risk situations, in this exam, the test format introduces a new variable where students, even those who seek risk, must consider what is best for themselves even if the decision results in selecting the "I don't know" option. Expected Utility Theory (EUT) suggests that when faced with risky decisions, individuals assess a course of action that maximizes the outcome in their favor (Schoemaker, 1982). As per EUT, if we assume students are fully rational, taking risks (as in guessing the answer) should only happen when students can guess with an expected value above the "I don't know" reward (i.e., 0.35 in our case). Superior expected values can happen for True/False questions (value of 0.5 even without knowledge), or if students can make educated guesses that result in higher expected values (e.g., if they know for sure that two out of four alternative answers are incorrect). Only in those cases would taking risk make sense. There is a countermanding thought process where students with risk personalities have to consider the best outcome even if it does not satisfy the risk-seeking personality. Since the test format allows students to not take the risk of guessing and enables them to simply indicate that they do not know the answer, people with risky personalities might feel less stimulated and less satisfied by the test format. Such individuals may also feel that this test format prevents them from outshining their peers as it allows students who don't know to gain by simply selecting the "I don't know" option and getting some credit for it. As the test format levels the playing field and rewards any students who choose the "I don't know" option, the risk-taking personalities may become dissatisfied with the format since everyone can get some credit for choosing the "I don't know" option. In such scenarios, the relative benefit of taking risks becomes less attractive. In their study, Collignon et al. (2020) show that students in general seem close to fully rational when using the same test format we use for our study. However, Collignon et al. (2020) also observed marginal irrational behaviors. Few students may wrongly believe that they can take educated guesses and do better than the "I don't know" option. Such students may perceive the inclusion of the "I don't know" option as a format that substantially reduces their advantage in test scores as other students can attain credit without taking any risk. This can notably happen if students sometimes miscalculate their expected value (Wygal et al., 2017). Such a format may, therefore, prevent risk personalities from the opportunity to thrive as the challenge is taken out of the situation. This may have a negative effect on satisfaction with the test format for these risk personalities. Hence, we propose:

Hypothesis 3: Risk personality will be negatively associated with test format satisfaction.

4. METHODOLOGY

"Education research tends to use the classroom as the laboratory; ideas are developed and tested by asking students about their perceived learning and/or satisfaction with a specific pedagogy" (Apostolou et al., 2013, p. 108).

In the undergraduate core business curriculum of a land grant university, a decision science class offered by the MIS department is required. Given the size of our classes, it was chosen to test students with multiple-choice questions. As the regular format of MCQ allows students to choose an option at random, it is difficult to know what students know accurately. To remedy this issue, a validated test format was used where partial points were offered when students admitted they don't know (Collignon et al., 2020). We seized the opportunity offered by the change of format to study the relationship between test format satisfaction and test anxiety. The change of format was announced and explained in class on several occasions before the test. The day of the test, the change of format was explained again with two practice questions. Consequently, when we constructed our survey, we were sure

that students knew what "this test format" or "this type of test" referred to when they were prompted to answer questions (please see prompt in Appendix A). It was also announced that grades distribution would not change and that selecting "not knowing" would not lead to passing grades. The sole purpose was to gain information that would benefit feedback provided to students.

4.1 Survey Instrument Development

We reviewed the literature for survey items. When items were not readily available, we adapted scales from related literature. That adaptation work was performed with the help of two experts who are not co-authors (one with a Ph.D. in education, the other with a Ph.D. in MIS). The final instrument is provided in Appendix A. When reviewing the literature for pre-existing instrument items regarding satisfaction and fairness, we came across the works of Whiting et al. (2008) and Ling and Libby (2010). In the work of Whiting et al. (2008), they mix items such as "My current performance appraisal system is fair" and "I am satisfied with my current performance appraisal system." Additionally, prior to this study, our personal experience is that students often respond to questions regarding test satisfaction through perceptions of fairness. This oblique answer can indicate some conceptual blurring between test fairness and satisfaction. Because of this blurring and the strong relationship of both concepts perceptible in our literature review, we chose to survey students both on their satisfaction (OS1 and OS2) and their perception of test format fairness (FA1 to FA3) without expecting students to part between concepts. Nevertheless, via the exploratory factor analysis (EFA) presented in the analysis section, we left the possibility of being contradicted by the data, if items pertaining to satisfaction did not load together with items pertaining to fairness.

Regarding the Anxiety scale, finding items was a more straightforward task, and our adapted questions were inspired by Sarason (1984) and Cassady and Johnson (2002). Based on these sources and prior findings that support within educational testing environments the idea that stress and anxiety go together (Schouwenburg, 1999; Zeidner, 1998), our instrument for test anxiety contains vocabulary pertaining directly to test anxiety (AX1 and AX2) but also to stress and nervousness (ST1 to ST4). Although we anticipated these items would load together, the EFA was again there to possibly contradict us.

As congruence was a notion borrowed from another stream of literature, we did not find survey items that were ready to use. However, Whiting et al. (2008) list 47 items pertaining to the congruence of a job performance appraisal system. We replaced the idea of job task vs. appraisal system alignment with knowledge vs. test format alignment and created our items that way. With the help of our expert panel, three items resulted from that process (CO1 to CO3).

The literature about measuring risk taking recommends mixing questions on different topics (DOSPERT scale). As recommended by Coppola (2014), we used a short DOSPERT scale in our survey to measure risk taking. One question asked directly about self-perception of risk-taking (RP1) and two others asked about risk-taking in recreational and financial situations (RP2 and RP3).

4.2 Data Collection and Sample

The students took their second mid-term exam with the changed format. Our survey was opened immediately after the exam and stayed open for four days. No incentive was given to take the survey. As per IRB guidelines, the results of the survey were anonymous. 289 students took the class, and of these, 277 answered the survey.

When collecting data on dependent and independent variables simultaneously in a survey, common method bias can occur (Podsakoff et al., 2003). We applied several techniques to minimize that phenomenon. We ensured the students that their answers were strictly anonymous and, since the survey was online, we randomized the questions (Podsakoff et al., 2003). We also added an attention check question (see Appendix A) to make sure that students were answering while having read the questions. A total of 245 responses are usable (after discarding partial responses, failures to answer the attention check, etc.).

	Ethnicity	College Year
• Average age: 20.8 • Male: 154 (63%) • Female: 91 (37%) • Self-determined: 0 • Average reported usage of "I don't know" (out of 25 questions): 2.06 (std dev 1.7)	• White/Caucasian: 180 (73%) • Black/African Am.: 7 (3%) • Asian: 40 (16%) • Pacific Islander: 2 (1%) • Latino: 5 (2%) • Native American Indian: 0 • Middle Eastern: 7 (3%) • Other: 4 (2%)	• Freshman: 1 (<1%) • Sophomore: 70 (29%) • Junior: 143 (58%) • Senior: 29 (12%) • Other: 3 (1%)

Table 1. Sample Descriptive Statistics

The course in which our study was set requires having taken another quantitative class and serves as a prerequisite for other classes. Thus, our sample's average age and number of years in college (shown in Table 1) make sense. Similarly, in the year the study was conducted, our college recorded that about 60% of its students were male, corresponding to our gender distribution.

5. ANALYSIS AND RESULTS

Analyses were performed using SPSS and AMOS 28.

5.1 Measurement Model Testing

An exploratory factor analysis helps determine the number of constructs measured in a survey and what items are constitutive of these constructs. We ensured that our data could be submitted to an EFA by first looking at multicollinearity issues in the item correlation matrix. The analysis revealed potential multicollinearity issues between items of the Anxiety scale (correlations above .8, as per Field, 2005). By removing items AX2, ST3, and ST4, no correlation above .8 was left in the correlation matrix, and the determinant of the item correlation matrix was above 0.00001—the threshold to submit the matrix to a factor analysis (Field, 2005).

A first EFA was performed with Promax rotation because we expected some correlations between factors (other rotations,

such as, Oblimin or Varimax were performed with minor discrepancies). The factor analysis yielded a pattern matrix with all items loading properly on the four intended constructs. However, Item CO2 of the Congruence scale showed low communalities with other items, and although it loaded with other Congruence items in the pattern matrix, its loading coefficient was low. We tested the internal reliability of our instrument's scales by calculating Cronbach's alpha. Alpha for all scales but one was above the recommended threshold of 0.7 (Nunnally, 1978). With its three items, the Congruence scale's alpha was 0.658, which is generally considered too low. Thus, we concluded that item CO2 was triggering internal reliability issues in the scale and removed it from the model. With items CO1 and CO3 only, Cronbach's alpha for Congruence is 0.740. Hence, internal reliability is good for all scales as shown in Appendix C.

A second EFA was run with the remaining items of our survey and all indicators complied with expected thresholds. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy is at a meritorious level - 0.831 (Kaiser & Rice, 1974). Bartlett's test of sphericity is statistically significant. The communalities for all items are above 0.5, and all items load only on one factor with loadings significant at .0001 and above 0.4, providing evidence of unidimensionality (Anderson & Gerbing, 1988; Hair et al., 2006).

We found evidence of convergent validity because all items load on their factor with loadings above 0.5 (Fornell & Bookstein, 1982). Composite Reliabilities (CR) are calculated for factors with more than two items and are above or near the expected 0.7 (Netemeyer et al., 1990; Nunnally & Bernstein, 1994).

Regarding discriminant validity, items do not load across factors and the square root of average variance extracted (AVE) for each construct is greater than correlations with other factors (Barclay et al., 1995). Also, the heterotrait-monotrait ratios (HTMT) are all below 0.85 (Kline, 2011).

Common method bias was tested with Harman's one-factor test (Podsakoff et al., 2003). One factor explains only 35% of the variance. In the factor correlation matrix, factors do not display correlations above 0.9, which means there is no sign of bias (Bagozzi et al., 1991). The second to smallest positive value in the item correlation matrix is also a good conservative proxy for the correlation between studied variables and a hidden common-method bias variable (Lindell & Whitney, 2001; Malhotra et al., 2006). Looking at the item correlation matrix, the second lowest statistically significant coefficient is 0.245 (with a p-value below 0.05). Squared, this coefficient shows that a maximum of 6% of the variance in this study could be explained through common method bias. We found this acceptable.

Finally, following the two-step procedure recommended by Anderson and Gerbing (1988), before testing the structural model, we tested our model's fit by performing a Confirmatory Factor Analysis (CFA) in AMOS. The model indicators of fit are very good (Byrne, 2013; Hu & Bentler, 1999): PCMIN/DF = 1.412 < 3, Comparative Fit Index (CFI) = 0.986 > 0.95, Tucker-Lewis Index (TLI) = 0.980 > 0.95, RMSEA = 0.41 < 0.5, a close fit with PCLOSE at 0.749 > 0.05.

5.2 Structural Model Testing

The second step of the two-step procedure is to test the structural equation model. This model also provided a good fit (PCMIN/DF = 1.563, CFI = 0.979, TLI = 0.972, RMSEA = 0.48 with PCLOSE at 0.547). The results of our model testing are shown in Figure 1 below.

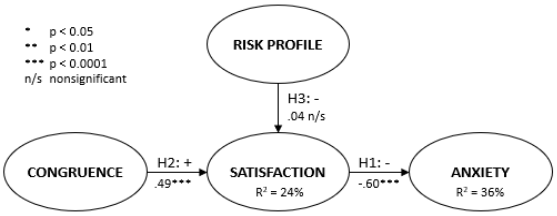


Figure 1. Structural Model Results

	Hypothesis	Result
H1	Satisfaction with the test format will be negatively correlated with anxiety.	Supported
H2	Congruence will be positively correlated with test format satisfaction.	Supported
H3	Risk personality will be negatively associated with test format satisfaction.	Not Supported

Table 2. Support for the Hypotheses

6. DISCUSSION

Information Systems classes and analytical courses tend to be challenging for the typical business student because of the technical nature of the material (May & Dhillon, 2009). This technical material is linked to higher test anxiety for students (Zoller & Ben-Chaim, 1989). In order to test the acquisition by students of that technical material, MCQ tests are the dominant mode of assessment. This dominance lies on grounds internal and external to MCQ tests. Internally, MCQ tests are characterized by efficiency, objectivity, validity, and reliability (Collignon et al., 2020; Lesage et al.; 2013; Stevens et al., 2023). Externally, several underlying pressures explain the widespread use of MCQ, including large class sizes (Collignon et al., 2020), the need for continuous improvement (Palocsay et al., 2020), the pressure on faculty to link assessment to learning objectives (Stevens et al., 2023), as well as the growing need to reach classes using some type of online format (Stevens et al., 2023).

Among MCQ tests, there are many formats. The use of the "number right" (positive marking) format (the correct answer gets a point, and the others get zero) dominates in the USA. In this specific MIS technical context, we used the version proposed by Collignon et al. (2020) and seized the opportunity to study variations around satisfaction with the MCQ test format to observe how anxiety could be managed. Our study does contribute as we can see that there is variance associated with test format satisfaction, that it can be acted upon, and that it matters if we want to help reduce test anxiety.

Our study has theoretical and practical implications. Our model confirms and integrates several theoretical grounds (cognitive fit, self-determination theory). Put together these theoretical grounds show there is a path between making a test

congruent with the knowledge to be tested, satisfying students' expectations, and reducing anxiety. The high statistical significance of path coefficients ($p<0.0001$) and the percentage of variation explained in our constructs (24% for satisfaction and 36% for anxiety) tend to strongly support our model. Doing so, we bring support to recent literature, which posits that test format matters when it comes to reducing test anxiety.

The change of format was maybe suppressing some perception that gambling could be rewarded. This is why we included a risk profile in the study. We hypothesized that risk personality would be negatively related to satisfaction because risk-friendly students may see fewer opportunities to guess in the new format. Fewer opportunities to guess means an environment where risk-takers may feel less comfortable and less capable of outperforming others. We offer to interpret the lack of significant results for that hypothesis in two ways depending on whether we abide by the assumption of student rationality. Rational students should still see that, in some cases, their expected value of guessing is above the "I don't know" reward, and they should still proceed with guessing (e.g., true/false questions, and questions where they can eliminate some alternative answers). Although the proportion of opportunities may vary depending on the test format and the value of the "I don't know" compensation, students may not be sensitive to the proportion of guessing opportunities as long as they exist. If so, it makes sense that the level of risk profile does not correlate with satisfaction. On the other hand, parting from the assumption of rationality, we thought students may overestimate the edge they lose and be dissatisfied. But maybe we underestimated the appeal of safety to the risk-averse. Anderson (2004) and Collignon et al. (2020) observed students choosing the safe option while guessing had a greater expected value. This phenomenon, in addition to potentially overestimating the edge guessing can procure, may create a disconnect between risk profile and satisfaction. Adopting the instructor's perspective, it is comforting to see that students were interested in the congruence of the test and being tested on their knowledge rather than on their capability to play a game of chance with their academic success at stake. Beyond this comforting thought, the finding is interesting as we may omit that aspect in future studies and instructors can feel comfortable modifying the format and suppressing gambling opportunities.

Practically, paying attention to the test format is not fortuitous. Instructors often automatically resort to the most common format of MCQ whereas other formats can help provide more tools to provide individual feedback to students. Additionally, even after having chosen a format, instructors can be flexible and can reduce test-induced anxiety by adopting subtle modifications. To do so, they can investigate their student evaluations of instruction to see comments that pertain to the format of the test. Our study shows that instructors satisfying student expectations (while keeping the test congruent), should help students experience less anxiety.

Our contribution to research does not stop at the theoretical framework level. There are also some technical aspects of our study that contribute to the academic study of test formats. Notably, in the development of our scales, we chose to sometimes include items that pertain to subtly different concepts but that we expected to act as one. The EFA supported our approach. First, the central concept of satisfaction was developed integrating the notion of being satisfied with the fairness of the format. The EFA supported this design as the items pertaining to finding the test format satisfactory or liking the test format were loaded together with items surveying if students found the test format fair. This came as a confirmation of anecdotal observations that students tend to strongly associate fair and satisfactory when it comes to test format.

Another scale that was designed by mixing items from slightly different concepts is the anxiety scale. Our items use vocabulary pertaining to anxiety, stress, and nervousness. As per Schouwenburg (1999) we mixed the scale, but by going through an EFA, we opened the door to items not loading together. As expected, the items loaded together—so much so that we encountered problems of multicollinearity that we resolved by suppressing some items. This experience with the satisfaction and anxiety scales can help further researchers in justifying mixed and short scales, which is valuable when designing surveys. Indeed, shorter scales help reduce fatigue when surveying people, which in turn supports reliability and power.

We are aware that there are many limitations to our study. We focused entirely on whether the students felt that the format of the test rewarded knowledge (congruence) and whether risk profiles were impacted by their expectations. Other aspects of the test format certainly affect satisfaction with the format and anxiety, such as length, number of multiple choices per question, and the threshold of the reward for not answering. By focusing on congruence, we studied one aspect of student control over test performance (knowledge), but other mechanisms giving more control to students could help in reducing anxiety too.

Nowadays, many classes are offered in hybrid formats, while our study sets our test only within a face-to-face environment. Because of that restriction, we wonder if the same test format has the same effect on satisfaction and anxiety online. MIS is a field that opens to many careers where workers can work remotely. It might be interesting to further study the interaction of anxiety reduction mechanisms while students are isolated from others.

Additionally, because our study is embedded in an MIS context that general business students and even students enrolled in IT-related majors find hard and stressful (Sharma et al., 2020), we assumed that our context is conducive to anxiety. In the literature, some studies have suggested that test anxiety can have positive effects before reaching a climax and then being debilitating. Our study focused on the debilitating aspect of anxiety, and we geared toward finding ways of reducing anxiety without really knowing if we had passed the threshold where anxiety really became debilitating. It is likely that that threshold is an individual characteristic and difficult to find at the class level.

Finally, because the study is a picture in one instant, we could see how congruence, satisfaction, and anxiety relate but did not observe if anxiety was indeed reduced in the long term when implementing the updated test format. Further research, structured longitudinally, could help with such endeavors. Also, with a longitudinal approach, we could study if students can exercise more control over anxiety and performance by banking on answering "I don't know" and making sure they receive improved individualized feedback.

7. CONCLUSION

Our paper integrates several theoretical grounds in one model that shows that the test format congruence (i.e., ensuring it affords testing knowledge conveyed in class) impacts positively student satisfaction with the test. In turn satisfaction with the test format reduces test anxiety. These findings matter because educational institutions have seen student mental health degrade over the past decade and many institutions are looking for ways to help students. Our study shows that even when using tools such as MCQs, format congruence matters when managing student anxiety. Of course, our study was implemented at one point in time, and further research is needed to evaluate the effects of format congruence over time as students become more familiar with less traditional MCQ formats.

8. REFERENCES

Alpert, R., & Haber, R. N. (1960). Anxiety in Academic Achievement Situations. *Journal of Abnormal and Social Psychology*, 61(2), 207-215. https://doi.org/10.1037/h0045464

Anderson, J. (2004). Medical Teacher 25th Anniversary Series Multiple-Choice Questions Revisited. *Medical Teacher*, 26(2), 110-113. https://doi.org/10.1080/0142159042000196141

Anderson, J. C., & Gerbing, D. W. (1988). Structural Equation Modeling in Practice: A Review and Recommended Two-Step Approach, *Psychological Bulletin*, 103(3), 411-423. https://doi.org/10.1037//0033-2909.103.3.411

Apostolou, B., Dorminey, J. W., Hassell, J. M., & Watson, S. F. (2013). Accounting Education Literature Review (2010-2012). *Journal of Accounting Education*, 31(2), 107-161. https://doi.org/10.1016/j.jaccedu.2013.03.001

Bagozzi, R. P., Yi, Y., & Phillips, L. W. (1991). Assessing Construct Validity in Organizational Research, *Administrative Science Quarterly*, 36(3), 421-458. https://doi.org/10.2307/2393203

Barclay, D., Higgins, C. A., & Thompson, R. L. (1995). The Partial Least Squares Approach to Causal Modeling, Personal Computer Adoption and Use as an Illustration. *Technology Studies*, 2(2), 285-309.

Birenbaum, M. (2007). Assessment and Instruction Preferences and Their Relationship With Test Anxiety and Learning Strategies. *Higher Education*, 53(6), 749-768. https://doi.org/10.1007/s10734-005-4843-4

Brown, G. T. L. (2011). Self-Regulation of Assessment Beliefs and Attitudes: A Review of the Students' Conceptions of Assessment Inventory. *Educational Psychology*, 31(6), 731-748. https://doi.org/10.1080/01443410.2011.599836

Budescu, D., & Bar-Hillel, M. (1993). To Guess or Not to Guess: A Decision-Theoretic View of Formula Scoring. *Journal of Educational Measurement*, 30(4), 277-291. https://doi.org/10.1111/j.1745-3984.1993.tb00427.x

Byrne, B. M. (2013). *Structural Equation Modeling with AMOS: Basic Concepts, Applications, and Programming*. (2nd ed.). New York, NY: Taylor and Francis Group. https://doi.org/10.4324/9781410600219

Cassady, J. C., & Johnson, R. E. (2002). Cognitive Test Anxiety and Academic Performance. *Contemporary Educational Psychology*, 27(2), 270-295. https://doi.org/10.1006/ceps.2001.1094

Chapell, M. S., Blanding, Z. B., Silverstein, M. E., Takahashi, M., Newman, B., Gubi, A., & McCann, N. (2005). Test Anxiety and Academic Performance in Undergraduate and Graduate Students. *Journal of Educational Psychology*, 97(2), 268-274. https://doi.org/10.1037/0022-0663.97.2.268

Chu, M.-W., Guo, Q., & Leighton, J. P. (2014). Students' Interpersonal Trust and Attitudes Towards Standardised Tests: Exploring Affective Variables Related to Student Assessment. *Assessment in Education: Principles, Policy & Practice*, 21(2), 167-192. https://doi.org/10.1080/0969594X.2013.844094

Collignon, S. E., Chacko, J., & Wydick M. M., (2020). An Alternative Multiple-Choice Question Format to Guide Feedback Using Student Self-Assessment of Knowledge. *Decision Sciences Journal of Innovative Education*, 18(3), 456-480. https://doi.org/10.1111/dsji.12213

Coppola, M. (2014). Eliciting Risk-Preferences in Socio-Economic Surveys: How Do Different Measures Perform? *The Journal of Socio-Economics*, 48, 1-10. https://doi.org/10.1016/j.socec.2013.08.010

Crocker, L., & Algina, J. (1986). *Introduction to Classical and Modern Test Theory*. New York: Holt, Rinehart & Winston.

Deci, E. L., & Ryan, R. M. (2012). Motivation, Personality, and Development Within Embedded Social Contexts: An Overview of Self-Determination Theory. In *The Oxford Handbook of Human Motivation* (pp. 85-107). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195399820.013.0006

Dick, W., Carey, L., & Carey, J. O. (2008). *The Systematic Design of Instruction* (7th ed.). Boston, MA, USA: Allyn & Bacon.

Duraku, Z. H. (2017). Factors Influencing Test Anxiety Among University Students. *The European Journal of Social & Behavioural Sciences*, 18(1), 69-78. https://doi.org/10.15405/ejsbs.206

Elsass, P. M., & Veiga, J. F. (1997). Job Control and Job Strain: A Test of Three Models. *Journal of Occupational Health Psychology*, 2(3), 195-211. https://doi.org/10.1037/1076-8998.2.3.195

Espinoza, M. P., & Gardeazabal, J. (2010). Optimal Correction for Guessing in Multiple-Choice Tests. *Journal of Math Psychology*, 54(5), 415-425. https://doi.org/10.1016/j.jmp.2010.06.001

Everson, H. T., Tobias, S., Hartman, H., & Gourgey, A. (1993). Test Anxiety and the Curriculum: The Subject Matters. *Anxiety, Stress, and Coping*, 6(1), 1-8. https://doi.org/10.1080/10615809308249528

Field, A. P. (2005). *Discovering Statistics Using SPSS (2nd Edition)*. London: Sage.

Fornell, C., & Bookstein, F. (1982). Two Structural Equation Models: Lisrel and PLS Applied to Consumer Exit-Voice Theory. *Journal of Marketing Research*, 19(4), 440-452. https://doi.org/10.1177/002224378201900406

Furnham, A., Nuygards, S., & Chamorro-Premuzic, T. (2013). Personality, Assessment Methods and Academic Performance. *Instructional Science*, 41(5), 975-987. https://doi.org/10.1007/s11251-012-9259-9

Gagne, R. M., Wager, W. W., Golas, K. C., Keller, J. M., & Russell, J. D. (2005). *Principles of Instructional Design* (5th ed.). Cengage Learning. https://doi.org/10.1002/pfi.4140440211

Galanakis, M., Bithava, I., Emmanouil, C., Lali, P., Symeonidi, A., & Darviri, C. (2015). Evidence for the Inter/Intra-Relationship Between the Sense of Fairness at Workplace, Distress, and Health Outcome: A Systematic Review. *Psychology*, 6, 2081-2090. https://doi.org/10.4236/psych.2015.616203

Green, K. E. (1981). Test Anxiety Level and Test Format Preference. *Psychological Reports*, 48(2), 537-538. https://doi.org/10.2466/pr0.1981.48.2.537

Guraya, S. Y., Guraya, S. S., Habib, F., AlQuiliti, K. W., & Khoshhal, K. I. (2018). Medical Students' Perception Of Test Anxiety Triggered by Different Assessment Modalities. *Medical Teacher*, 40(sup1), S49-S55. https://doi.org/10.1080/0142159X.2018.1465178

Hahn, H., Kropp, P., Kirschstein, T., Rücker, G., & Müller-Hilke, B. (2017). Test Anxiety in Medical School Is Unrelated to Academic Performance but Correlates With an Effort/Reward Imbalance. *PLOS ONE*, 12(2), e0171220. https://doi.org/10.1371/journal.pone.0171220

Hamilton, L. S. (1994). *An Investigation of Students' Affective Responses to Alternative Assessment Formats*. [Washington, D.C.]: Distributed by ERIC Clearinghouse, https://eric.ed.gov/?id=ED376203

Hair, J., Black, W., Babin, B., Anderson, R., & Tatham, R. (2006). *Multivariate Data Analysis* (6th ed.). Upper Saddle River, NJ: Pearson Prentice Hall.

Hirshleifer, D., Low, A., & Teoh, S. H. (2012). Are Overconfident CEOs Better Innovators? *The Journal of Finance*, 67(4), 1457-1498. https://doi.org/10.1111/j.1540-6261.2012.01753.x

Hu, L., & Bentler, P. M. (1999). Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1-55. https://doi.org/10.1080/10705519909540118

James, M. (2010). Educational Assessment: Overview. In Peterson, P., Baker, E., & McGaw, B. (Eds), *International Encyclopedia of Education* (3rd ed., vol. 3, pp. 161-171). Oxford: Elsevier. https://doi.org/10.1016/B978-0-08-044894-7.01713-9

Jones, B. D., & Egley, R. J. (2004). Voices From the Frontlines: Teachers' Perceptions of High-Stakes Testing. *Education Policy Analysis Archives*, 12(39), n39. https://doi.org/10.14507/epaa.v12n39.2004

Kaiser, H. F., & Rice, J. (1974). Little Jiffy, Mark IV. *Educational and Psychological Measurement*, 34(1), 111-117. https://doi.org/10.1177/001316447403400115

Kincaid, C., & Zemke, D. M. V. (2006). Perceptions of Cheating: An Exploratory Study. *Journal of Hospitality & Tourism Education*, 18(1), 47-55. https://doi.org/10.1080/10963758.2006.10696849

Kim, T., Kim, H., Lee, H. Y., Goh, H., Abdigapporov, S., Jeong, M., Cho, H., Han, K., Noh, Y., Lee, S. J., & Hong, H. (2022, April). Prediction for Retrospection: Integrating Algorithmic Stress Prediction Into Personal Informatics Systems for College Students' Mental Health. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (pp. 1-20). https://doi.org/10.1145/3491102.3517701

Kline, R. B. (2011). *Principles and Practice of Structural Equation Modeling* (3rd ed.). Guilford Press.

Kuechler, W. L. & Simkin, M. G. (2003). How Well Do Multiple Choice Tests Evaluate Student Understanding in Computer Programming Classes? *Journal of Information Systems Education*, 14(4), 389-400.

Leadbeater, B., Thompson, K., & Gruppuso, V. (2012). Co-Occurring Trajectories of Symptoms of Anxiety, Depression, and Oppositional Defiance From Adolescence to Young Adulthood. *Journal of Clinical Child & Adolescent Psychology*, 41(6), 719-730. https://doi.org/10.1080/15374416.2012.694608

Lesage, E., Valcke, M. & Sabbe, E. (2013). Scoring Methods for Multiple Choice Assessment in Higher Education—Is It Still a Matter of Number Right Scoring or Negative Marking? *Studies in Educational Evaluation*, 39(3), 118-193. https://doi.org/10.1016/j.stueduc.2013.07.001

Lindell, M. K., & Whitney, D. J. (2001). Accounting for Common Method Variance in Cross-Sectional Research Designs. *Journal of Applied Psychology*, 86(1), 114-121. https://doi.org/10.1037/0021-9010.86.1.114

Ling, C., & Libby, T. (2010). Writing Mini-Cases: An Active Learning Assignment. *Issues In Accounting Education*, 25(2), 245-265. https://doi.org/10.2308/iace.2010.25.2.245

Malhotra, N. K., Kim, S. S., & Patil, A. (2006). Common Method Variance in IS Research: A Comparison Of Alternative Approaches and a Reanalysis of Past Research. *Management Science*, 52(12), 1865-1883. https://doi.org/10.1287/mnsc.1060.0597

May, J., & Dhillon, G. (2009). Interpreting Beyond Syntactics: A Semiotic Learning Model for Computer Programming Languages. *Journal of Information Systems Education*, 20(4), 431-438.

McDaniel, M. A., Blischak, D. M., & Challis, B. (1994). The Effects of Test Expectancy on Processing and Memory of Prose. *Contemporary Educational Psychology*, 19(2), 230-248. https://doi.org/10.1006/ceps.1994.1019

McEwen, B. S., & Wingfield, J. C. (2003). The Concept of Allostasis in Biology and Biomedicine. *Hormones and Behavior*, 43(1), 2-15. https://doi.org/10.1016/S0018-506X(02)00024-7

Meijer, J. (2001). Learning Potential and Anxious Tendency: Test Anxiety as a Bias Factor in Educational Testing. *Anxiety, Stress, & Coping*, 14(3), 337-362. https://doi.org/10.1080/10615800108248361

Netemeyer, R. G., Johnston, M. W., & Burton, S. (1990). Analysis of Role Conflict and Role Ambiguity in a Structural Equations Framework. *Journal of Applied Psychology*, 75(2), 148-157. https://doi.org/10.1037/0021-9010.75.2.148

Nunnally, J. C. (1978). *Psychometric Theory* (2nd ed.). New York: McGraw-Hill.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric Theory* (3rd ed.). New York, NY: McGraw-Hill.

Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional Manipulation Checks: Detecting Satisficing to Increase Statistical Power. *Journal of Experimental Social Psychology*, 45(4), 867-872. https://doi.org/10.1016/j.jesp.2009.03.009

Palocsay, S. W., Stevens, S. P., & Novoa, L. J. (2020). STRATA: A Spreadsheet Tool for Multidimensional Analysis of Operations Research/Management Science Assessment Test Data. *INFORMS Transactions on Education*, 21(1), 18-33. https://doi.org/10.1287/ited.2019.0229

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies. *Journal of Applied Psychology*, 88(5), 879-903. https://doi.org/10.1037/0021-9010.88.5.879

Prieto, G., & Delgado, A. R. (1999). The Effect of Instructions on Multiple-choice Test Scores. *European Journal of Psychological Assessment*, 15(2), 143-150. https://doi.org/10.1027//1015-5759.15.2.143

Ramirez, G., & Beilock, S. L. (2011). Writing About Testing Worries Boosts Exam Performance in the Classroom. *Science*, 331(6014), 211-213. https://doi.org/10.1126/science.1199427

Rowley, G. L. (1974). Which Examinees Are Most Favoured by the Use of Multiple Choice Tests? *Journal of Educational Measurement*, 11(1), 15-23. https://doi.org/10.1111/j.1745-3984.1974.tb00966.x

Sansgiry, S. S., & Sail, K. (2006). Effect of Students' Perceptions of Course Load on Test Anxiety. *American Journal of Pharmaceutical Education*, 70(2), 26. https://doi.org/10.1016/S0002-9459(24)07668-X

Sarason, I. G. (1984). Stress, Anxiety, and Cognitive Interference: Reactions to Tests. *Journal of Personality and Social Psychology*, 46(4), 929-938. https://doi.org/10.1037/0022-3514.46.4.929

Schoemaker, P. J. (1982). The Expected Utility Model: It's Variants, Purposes, Evidence and Limitations. *Journal of Economic Literature*, 20(2), 529-563.

Schouwenburg, H. C. (1999). Test Anxiety Unveiled. Test Anxiety: The State of the Art. Moshe Zeidner. Plenum, New York, 1998. 440+xxi pages. *European Journal of Personality*, 13(4), 327-329. https://doi.org/10.1002/(SICI)1099-0984(199907/08)13:4%3C327::AID-PER345%3E3.0.CO;2-G

Seligman, M. E. (2006). *Learned Optimism: How to Change Your Mind and Your Life.* New York, NY: Vintage.

Sharma, M., Biros, D., Ayyalasomayajula, S., & Dalal, N. (2020). Teaching Tip: Teaching Programming to the Post-Millennial Generation: Pedagogic Considerations for an IS Course. *Journal of Information Systems Education*, 31(2), 96-105.

Shermis, M. D., & Lombard, D. (1998). Effects Of Computer-Based Test Administrations on Test Anxiety and Performance. *Computers in Human Behavior*, 14(1), 111-123. https://doi.org/10.1016/S0747-5632(97)00035-6

Sparfeldt, J. R., Rost, D. H., Baumeister, U. M., & Christ, O. (2013). Test Anxiety in Written and Oral Examinations. *Learning and Individual Differences*, 24, 198-203. https://doi.org/10.1016/j.lindif.2012.12.010

Stevens, S. P., Palocsay, S. W., & Novoa, L. J. (2023). Practical Guidance for Writing Multiple-Choice Test Questions in Introductory Analytics Courses. *INFORMS Transactions on Education*, 24(1), 51-69. https://doi.org/10.1287/ited.2022.0274

Struyven, K., Dochy, F., & Janssens, S. (2005). Students' Perceptions About Evaluation and Assessment in Higher Education: A Review. *Assessment & Evaluation in Higher Education*, 30(4), 325-341. https://doi.org/10.1080/02602930500099102

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), e0203931. https://doi.org/10.1371/journal.pone.0203931

Vessey, I., & Galletta, D. F. (1991). Cognitive Fit: An Empirical Study of Information Acquisition. *Information Systems Research*, 2(1), 63-84. https://doi.org/10.1287/isre.2.1.63

von der Embse, N., Jester, D., Roy, D., & Post, J. (2018). Test Anxiety Effects, Predictors, and Correlates: A 30-Year Meta-Analytic Review. *Journal of Affective Disorders*, 227, 483-493. https://doi.org/10.1016/j.jad.2017.11.048

Wigfield, A., & Eccles, J. S. (2000). Expectancy-Value Theory of Achievement Motivation. *Contemporary Educational Psychology*, 25(1), 68-81. https://doi.org/10.1006/ceps.1999.1015

Whiting, H. J., Kline, T. J., & Sulsky, L. M. (2008). The Performance Appraisal Congruency Scale: An Assessment of Person-environment Fit. *International Journal of Productivity and Performance Management*, 57(3), 223-236. https://doi.org/10.1108/17410400810857239

Wolverton, B. (2019, February 21). As Students Struggle With Stress and Depression, Colleges Act as Counselors. *The New York Times*. https://www.nytimes.com/2019/02/21/education/learning/mental-health-counseling-on-campus.html

Wygal, D. E., Stout, D. E., & Cunningham, B. M. (2017). Shining Additional Light on Effective Teaching Best Practices in Accounting: Self-Reflective Insights from Cook Prize Winners. *Issues In Accounting Education*, 32(3), 17-31. https://doi.org/10.2308/iace-51743

Yerkes, R. M., & Dodson, J. D. (1908). The Relation of Strength of Stimulus to Rapidity of Habit-Formation. *Journal of Comparative Neurology and Psychology*, 18(5), 459-482. https://doi.org/10.1002/cne.920180503

Zeidner, M. (1987). Essay Versus Multiple-Choice Type Classroom Exams: The Student's Perspective. *The Journal of Educational Research*, 80(6), 352-358. https://doi.org/10.1080/00220671.1987.10885782

Zeidner, M. (1990). College Students' Reactions Towards Key Facets of Classroom Testing. *Assessment & Evaluation in Higher Education*, 15(2), 151-169. https://doi.org/10.1080/0260293900150206

Zeidner, M. (1998). *Test Anxiety: The State of The Art.* New York, NY: Kluwer Academic/Plenum Publishers.

Zeidner, M. (2005). *Test Anxiety: The State of The Art.* New York, NY: Kluwer Academic/Plenum Publishers.

Zeidner, M. (2007). Test Anxiety in Educational Contexts: Concepts, Findings, and Future Directions. In P. A. Schutz & R. Pekrun (Eds.), *Emotion in Education* (pp. 165-184). Elsevier Academic Press. https://doi.org/10.1016/B978-012372545-5/50011-3

Zoller, U., & Ben-Chaim, D. (1989). Interaction Between Examination Type, Anxiety State, and Academic Achievement in College Science; An Action-Oriented

Research. *Journal of Research in Science Teaching*, 26(1), 65-77. https://doi.org/10.1002/tea.3660260107

Zuckerman, M. (2007). *Sensation Seeking and Risky Behavior*. American Psychological Association. https://doi.org/10.1037/11555-000

AUTHOR BIOGRAPHIES

Stéphane E. Collignon is an associate professor with the department of MIS at West Virginia University. He received his Ph.D. from Virginia Tech. Collignon's research interests are information systems issues in logistics, privacy management, and education. Collignon works with quantitative methodologies but is also an expert in qualitative research (Grounded Theory). He teaches information systems, data management, decisions sciences and business intelligence. Collignon worked for 7 years in R&D for two French distribution companies. His research notably appeared in *Information & Management*, *Journal of Business Logistics*, *European Journal of Operational Research*, *Decision Sciences Journal of Innovative Education*.

Josey Chacko is an associate professor in the Richard J. Bolte Sr. School of Business at Mount St. Mary's University. Prior to earning his Ph.D. at Virginia Tech, he worked in manufacturing for close to 4 years. He is actively involved with the Information Systems for Crisis Response and Management (ISCRAM) group. Dr. Chacko's research has been published in various journals including *Decision Support Systems*.

Salman Nazir is an associate professor of MIS at West Virginia University. He received his Ph.D. from McGill University, Montreal, Canada. He enjoys studying agile development teams using quantitative and qualitative methods. He also studies agility at the organizational level. His research has been published in the *Journal of Strategic Information Systems*, *European Journal of Information Systems*, *Journal of the Association of Information Systems*, *Information & Management*, and *Information Technology & Management*.

APPENDICES

Appendix A. Survey Instrument

The students were prompted to indicate their level of agreement on a Likert scale from 1 to 5 (strongly disagree to strongly agree) as follows:

KEEPING IN MIND THE NEW TYPE OF MULTIPLE-CHOICE TEST IN WHICH YOU CAN CIRCLE ONE ANSWER OR OPT FOR THE 0.35pt OPTION (MID-TERM II TYPE)

Please indicate your extent of agreement to the following statements on the scale shown below:
[Likert scale from 1 to 5, strongly disagree to strongly agree].

Construct	Item	Item	Mean	Std Dev
Satisfaction	OS1	I think this format of testing is a satisfactory way to test people	3.48	0.818
	OS2	I like this type of test for assessing students	3.17	0.954
	FA1	I think this type of test is fair	3.63	0.833
	FA2	I believe this way of testing students is fair	3.58	0.834
	FA3	In my opinion, this format of test is a fair way of evaluating students	3.47	0.871
Anxiety	ST1	I feel this format of testing is stressful for students	3.05	1.025
	ST2	This type of test causes me stress	3.01	1.129
	ST3	This format of test makes me nervous	3.04	1.141
	ST4	I feel nervous when I take this type of test	3.07	1.123
	AX1	This format of test makes me anxious	3.00	1.107
	AX2	I feel anxious when I take this type of test	3.07	1.146
Congruence	CO1	I think this format of test rewards people who know the answers to the questions	3.51	1.035
	CO2	I think only students who know the class material perform well with this type of test	3.38	0.949
	CO3	In my opinion, people who truly know the answer are better rewarded with this format of test	3.33	1.076
Risk Profile	RP1	I perceive myself as a risk taker	3.21	0.984
	RP2	In games of chance I play for high stakes	3.07	1.038
	RP3 (rev)	In general, if I were to invest I would prefer to invest in stock with minimal risk and I am willing to accept the associated lower return	2.85	0.964

Appendix B. Pattern Cross Loading Matrix

	Factor			
	1	2	3	4
RP1	.045	.026	**.751**	-.008
RP2	.028	.013	**.904**	-.052
RP3(rev)	-.098	-.064	**.568**	.068
CO1	-.012	.004	-.009	**.790**
CO3	.047	.007	.015	**.730**
FA1	**.734**	-.024	.031	.039
FA2	**.892**	.076	-.034	-.045
FA3	**.860**	.106	.031	.044
OS1	**.612**	-.168	.032	.103
OS2	**.708**	-.169	-.085	-.092
AX1	-.004	**.865**	-.020	.009
ST1	-.009	**.833**	.022	.006
ST2	-.003	**.898**	-.026	-.008

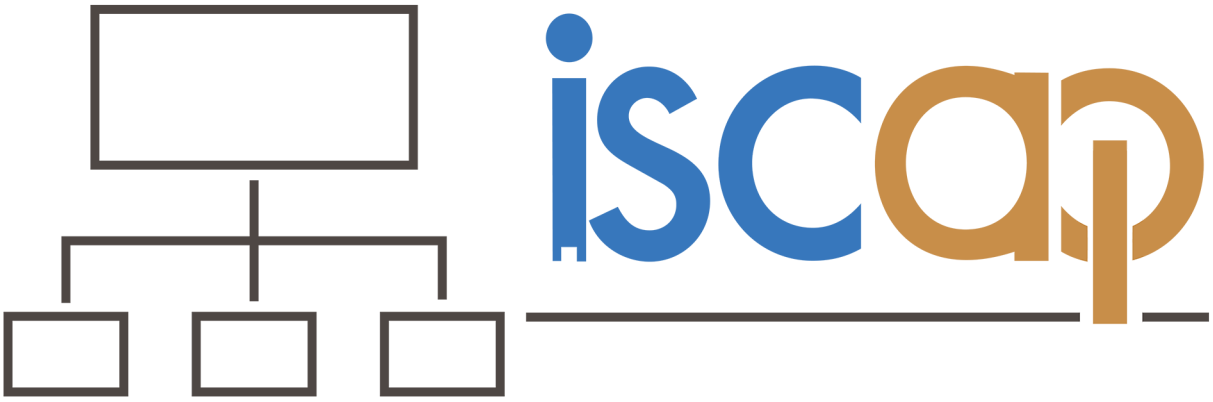
Extracted with Principal Axis Factoring, rotation with Promax with Kaiser Normalization.

Appendix C. Validity and Reliability Indicators

	Indicators			Factor Correlation Matrix (PAF), Square root of AVE on diagonal			
Factor	Cronbach's Alpha	CR (PCA*)	AVE (PCA*)	Satis	Stress	Cong	Risk
Satisfaction	.887	.907	.663	**.768**			
Stress	.902	.880	.746	-.564	**.866**		
Congruence	.740	N/A	.777	.491	-.110	**.760**	
Risk	.778	.681	.694	.079	-.011	.115	**.753**

** The EFA was performed with two methods of extraction in SPSS, principal component analysis (PCA) and principal axis factoring (PAF). Results are similar with both methods.*

INFORMATION SYSTEMS & COMPUTING ACADEMIC PROFESSIONALS



STATEMENT OF PEER REVIEW INTEGRITY

All papers published in the ***Journal of Information Systems Education*** have undergone rigorous peer review. This includes an initial editor screening and double-blind refereeing by three or more expert referees.

ISSN: 2574-3872 (Online) 1055-3096 (Print)