Gender Differences in Test Anxiety in High-Stakes English Proficiency Tests

Josefina C. Santana
(jsantana@up.edu.mx)
Universidad Panamericana, Mexico

Cristina Eccius-Wellmann
(ceccius@up.edu.mx)
Universidad Panamericana, Mexico

Abstract

Test anxiety can be defined as a reaction of stress to evaluations. This article reports on a study of the gender differences in test anxiety between males and females taking an English language proficiency test – the *Test of English for International Communication* (TOEIC) at a Mexican university. Participants (N=231; 47% male and 53% female) were asked to complete Grandis's *Questionnaire on Test Anxiety in University Students* a few minutes before they took the TOEIC. Responses on the questionnaire were correlated to TOEIC scores by linear regression analysis. Results show that women's scores were significantly higher than the men's on the Test Anxiety questionnaire (F= 65.75 vs M = 53.56), significantly lower on the TOEIC (F= 731 vs M = 772), and that gender is an important factor in test anxiety.

1 Introduction

Standardized testing is becoming a norm in education. It is important, however, to understand exactly what tests can and cannot tell us. Equally important is the elimination of bias stemming from external factors, which may affect test scores and render them less useful. One important source of test bias may be *test anxiety* (TA; Zeidner, 1998), that is, "situation-specific anxiety experienced in evaluative situations" (Sommer & Arendasy, 2014, p. 116). Though test anxiety is a personal response to a testing situation, it can be exacerbated by the characteristics of the test itself. For example, high-stakes tests, or long tests with time constraints – such as language proficiency tests – bring together a number of characteristics which may set off TA (Aydin, 2009; Hurley & Padró, 2006). These characteristics may call into question the validity of the exam. Yet, high-stakes proficiency tests are increasingly being used throughout the world.

In countries where English is not the official language, it is becoming common to use English proficiency exams as a means to obtain access to employment, a job promotion, or a university degree. This is the case in Mexico. Most private universities have established English as a graduation requirement, and it is becoming the norm in state universities. Students are required to take a standardized international examination and they must obtain a minimum score in order to receive their final diploma.

The purpose of these tests is to evaluate if the test-taker has the language skills to carry out his or her professional activities in a globalized work environment. However, TA may have a negative

effect on test scores (Aydin, 2009; Gursoy & Arman, 2016; Hurley & Padró, 2006; Zeidner, 1998). Test-takers could potentially receive scores that do not reflect their true ability in the language, with impact on their professional lives. On the other hand, most studies into TA have found that TA is more prevalent among women than among men (Cizek & Burg, 2006; Grandis, 2009; Putwain & Daly, 2014; Rosario et al., 2008, among others). If TA affects language proficiency tests negatively, then women will have a disadvantage when taking tests.

This article reports on one aspect of a larger study of test anxiety with regard to high-stakes tests of English proficiency, specifically, the *Test of English for International Communication* (TOEIC). The study was carried out among undergraduates at a Mexican university who were taking a TOEIC test (N=231). The objective of the study is to analyze gender differences in TA among the test-takers as measured by a specific TA scale – the *Questionnaire of Test Anxiety among University Students* (CAFEU, for its Spanish acronym). The questionnaire was developed in Argentina and currently adopted to the study as described herein.

2 Research questions

The research questions were established as follows:
- What relation exists between TA as measured by the CAFEU and TOEIC scores?
- What is the level of significance between score means in males and females (total CAFEU and total TOEIC)?
- What aspects of TA as measured by the CAFEU differ between males and females?
- What differences are observed in the items of CAFEU with greater impact on test scores between males and females?

3 Theoretical framework

TA is a complex phenomenon with various causes and manifestations. This section will discuss the components and characteristics of TA. It will include the principal findings of previous studies on TA. Finally, it will include relevant studies on TA and language tests.

3.1 Test anxiety

TA can be understood as an emotional reaction to situations involving the evaluation of abilities. It is characterized by a recurring concern with a possible failure in a given task, and with the consequences of that failure (Furlan, Ferrero, & Gallart, 2014; Furlan, Pérez, Moyano, & Cassady, 2011). TA is not necessarily related to true ability; rather, it depends more on the individual's perception of his or her ability. Furthermore, TA is related to fear of evaluation, generalized dislike of tests, and lack of effective study skills (Hembree, 1988). Furlán, Ferrero and Gallart (2014) suggest that anxiety is not always negative. Moderate concern can lead to favorable anxiety, which motivates the subject to use a variety of strategies to overcome it. However, when anxiety is excessive, it may affect academic performance.

Liebert and Morris (1967) identified two distinct factors of TA: worry and emotionality: "'Worry' (W) was conceptually identified as any cognitive expression of concern about one's own performance, while 'emotionality' (E) referred to autonomic reactions which tend to occur under examination stress" (p. 975), including physiological responses. Worry is a component conceptualized as cognitive TA (Cassady & Finch, 2014) that is concerned with those factors which can directly affect test performance, such as distractibility, and inadequate study skills, among others. However, the cognitive domain also considers inappropriate coping strategies, such as procrastination and avoidance.

Highly test-anxious persons tend to focus more on what is happening within themselves – fear of failure, for example – than on the test itself. They become easily distracted, or have difficulty remembering concepts and organizing their ideas. All this leads to poor performance (Sarason,

1980; Zeidner, 1998). Cassady and Johnson (2002) state that the anxious individual is incapable of avoiding irrelevant thinking during the examination. He or she is worried about the final result, compares himself or herself to others, and feels overwhelmed and inadequately prepared for the test. Sommer and Arendasy (2015) add that anxiety prevents the student from performing up to his or her true level. Distracting thoughts reduce cognitive resources available for the accomplishment of the task, making it more demanding. TA is much greater among students of average academic ability, compared to high ability students, but it is the low-ability students who report the greatest TA (Cizek & Burg, 2006).

Emotionality, on the other hand, includes physiological reactions (Cassady & Finch, 2014). Physiological anxiety may be manifested through increased heart rate and breathing, tense muscles, perspiration, dry mouth, upset stomach or even headaches, nausea and muscle contractures (Álvarez, Aguilar, & Lorenzo, 2012). Both Worry and Emotionality may explain the effect of TA on actual test performance, though Worry seems to play a stronger role (Hembree, 1988). This same author emphasizes the behavioral aspects of TA. His studies suggest that TA seems to be a behavioral construct (p. 74).

In sum, TA is a situation-specific response to an evaluation. Its manifestations can be cognitive, behavioral, and physiological. It is an individual response, but triggers can be related to culture, family background, grade level, general academic performance, or the test itself. Most studies have found that females report greater levels of TA than males. This point will be discussed in the following section.

3.2   Gender differences in test anxiety

As mentioned above, many studies have found that women tend to suffer from TA more than men do. Cassady and Johnson (2002), for example, state that the greatest differences can be found in the Emotionality component, with women reporting it more often than men do. A Pakistani study (Farooqi, Ghani, & Spielberger, 2012) found greater TA among female medical students, who were participants in the study. The authors explain this as a result of being raised in a patriarchal society.

Aydin (2017) found that TA tends to diminish from primary to secondary school, but that females show greater TA throughout their studies. Peleg, Deutsch and Dan (2016) found that parental expectations may be an important source of TA. On the other hand, Nuñez, Suarez and Bono (2016) mention two explanations for gender differences in TA. The first states that females feel more pres-sure to succeed academically, and thus, they have a greater fear of failing an exam. The second explanation is that men tend to express their feelings less; therefore, they tend not to report feelings of anxiety.

Finally, Zalta and Chambless (2012) explain gender differences in terms of "mastery" and "instrumentality". The first term refers ''the extent to which people see themselves as being in control of the forces that importantly affect their lives'' (Pearlin, Menaghan, Lieberman, & Mullan, 1981, p. 340). Instrumentality, on the other hand, "refers to a gender role orientation that is assertive, self-confident, and aggressive" (Zalta & Chambless, 2012, p. 488). Both concepts are crucial to the handling of stressful situations and men are socially taught to display them, whereas women are not. Most studies show that women suffer greater TA than men; this may be due to societal differences in the way men and women are educated.

3.3   Test anxiety and language tests

Some recent studies have looked into TA related to language achievement or proficiency tests. Salehi and Marefat (2014) researched *foreign language anxiety* (FLA), and TA among Iranian students at a language institute. Of the 200 participants, 193 were female, and their ages ranged between 14 and 47. The authors found a negative correlation between both FLA and test scores, and TA and test scores. That is to say, the participants suffered debilitating anxiety. Cheng et al. (2014) studied the effects of student motivation and TA on three different standardized high-stakes language tests.

They studied over one thousand Chinese-speaking participants and found that both motivation and TA have an effect on test performance. Female respondents showed greater TA, as did those participants who stated that the exam was important to them.

An Australian study by Rumsey, Thiessen, Buchan and Daly (2016) was carried out among persons migrating to Australia who must present an *International English Language Testing System* (IELTS) exam. The authors found that test-takers exhibited high levels of TA, brought about mostly by factors relating to the test, including its length, cost, and content. At the university level, Cakici (2016) studied gender differences in TA and FLA among 301 male and female students in an *English as a foreign language* (EFL) class in Turkey. She found that the females showed higher levels of TA and FLA than the males, but that these indices did not seem to have an effect on achievement.

Another Turkish study, by Gursoy and Arman (2016), was carried out among 138 high school students who were taking an EFL class. The study was both quantitative, using an anxiety scale to determine TA, and qualitative, gathering data through interviews. The participants showed a moderate level of TA, but mentioned in the interviews that several factors contributed to an increase in their anxiety. Among these factors were feeling unprepared for the exam, and having time limits.

Alikbari and Gheitasi (2017) studied levels of TA in 400 male and female Iranian high school students in an EFL class. They found above average levels of anxiety in over half the participants. Males and females showed similar levels of TA, but in females, it was more debilitating. As seen, there are few studies among universities students, and few studies analyzing TA in regard to language proficiency – rather than achievement – tests. However, these tests are becoming common in non-English speaking countries. University students are increasingly asked to take a proficiency test in order to enter the university, to take advanced courses, to graduate, or to have access to international exchange programs. As such, it is important to understand if proficiency tests lead to increased TA, and if TA has an effect on exam scores.

4 Study

4.1 Samples

The participants of the study were 231 undergraduates from a private university in western Mexico who took a TOEIC test to comply with a graduation requirement. 108 (46.75%) were male, 123 (53.25%) were female. They ranged in ages between 18 and 37 years old, and the mean age was 23.2 years. They were from three different fields of study: Administrative Sciences, which includes Finance, Marketing, Human Resources, Management, International Business, Hospitality, and Accounting (61.03% of the participants); Engineering, which includes Industrial Engineering, Civil Engineering, Mechatronics, Animation, Innovation and Design, and Systems Engineering (18.18%); and Humanities, which includes Law, Pedagogy, Psycho-pedagogy, Audiovisual Communications, Journalism, and Public Relations (20.77%). Percentages of participants from each field roughly represent the general population in this university.

To obtain their final diploma, the participants need to obtain a certain score on the TOEIC test (or its equivalent on other tests). The required TOEIC scores range between 750 and 795 depending on the field of studies. Participants took the TOEIC examination between May, 2015, and November, 2016. The participants entered the testing room, received and answered the TA questionnaire, and then took the TOEIC examination.

4.2 Instrument

There are a number of tests, scales, and questionnaires to measure different aspects of TA. This study used Grandis' (2009) *Questionnaire of Test Anxiety among University Students* (CAFEU, for its initials in Spanish; see Appendix 1). This is a 34-item questionnaire with a 5-point Likert scale designed to measure Worry (W; 13 items), Physiological Responses (PHR; 14 items) and Avoidance (A; seven items). The items are distributed as shown in Table 1.

Table 1. Items by category

Category	Items	Total
Worry	q7 When a teacher stands next to me, it makes me very nervous and I cannot keep working q8 I get nervous when I see the teacher arrive in the classroom with the exams q12 When I finish an exam, I feel sad, thinking that I did it incorrectly despite not having the results yet. q14 While I'm answering an exam, I think I'm doing it wrong. q15 I feel very bad if my classmates finish and start to turn in the exam before me. q16 While I am taking the exam, I think that the teacher is constantly looking at me. q17 I want to smoke during the exam. q21 Despite having studied, I think that I will not be able to pass. q27 If I sit at the front of the classroom, I get more nervous. q28 I get more nervous if the exam has a time limit to finish it and I do it worse or I do not finish. q32 When I finish the exam, I feel like I will not pass it. q33 I think that in the exam my nerves will make me forget everything I studied. q34 I take a long time to answer most of the questions and I'm almost always the last one to turn in the test.	13
Physiological response	q1 During exams, my hands perspire. q2 When I have been in an exam for a while, I feel discomfort in the stomach and need to defecate. q3 Even though I can see well, when I start reading the exam, my eyes blur and I don't understand what I'm reading. q5 When I finish an exam, I have a headache. q6 In exams, I feel that I lack air. I feel faint, very hot or a general malaise. q9 My hands feel stiff during exams q10 Before entering the test, I feel "something" in my stomach. q18 I really want to go to the bathroom during the exam q23 I cannot sleep the night before the test. q24 It makes me nervous if there are many classmates and / or teachers in the exam. q25 During the exam I felt dizzy and nauseous. q26 Before I turn in the exam, my mouth is dry; I cannot swallow. q29 When I'm doing an exam I feel like my heart beats faster. q30 When I enter the testing room, my legs shake.	14
Avoidance	q4 I am almost always late for an exam and when that happens I do not show up for it. q11 Just by reading or listening to the questions of the test, I go blank. q13 I think long before deciding whether to take an exam q19 I get sick and look for any excuse not to do an exam q20 For me it is a comfort or relief when, for whatever reason, an exam is postponed. q22 Before beginning the examination, I think that I know nothing and that I should suspend it. q31 I feel nervous in exams with few classmates.	7
Total		34

Other TA scales exist, but this instrument was selected, because it was created specifically for university students in a Latin American context, and thus seemed a good fit. The questionnaire was written in Latin American Spanish and thus did not require translation. To create her questionnaire, Grandis (2009) evaluated 1,368 students of both sexes from the 42 different degree programs at the Universidad Nacional de Rio Cuarto in Argentina. She found that TA was greater among students

of Human Sciences, especially in regard to Worry; but that Engineering students responded higher to items related to Avoidance.

4.3 Hypotheses

4.3.1 Simple linear regression analysis between CAFEU and TOEIC

A series of hypotheses will determine if the dependent variable (TOEIC) is significantly related to the independent variable. In the following linear regressions, the dependent variable is the mean points obtained on the *TOEIC* test when it correlates to the mean points on the anxiety questionnaire (CAFEU) – the independent variable – from the total number of participants as analyzed by gender. (Sub-indices M and F stand for male and female students.)

$TOEIC = \beta_0 + \beta_1\ CAFEU$
$TOEIC_M = \beta_0 + \beta_1\ CAFEU_M$
$TOEIC_F = \beta_0 + \beta_1\ CAFEU_F$

For each of the linear regressions, the null hypothesis of no relationship is proposed. This means that the slope of the regression line is zero. The alternative hypothesis states that the slope of the regression line is not equal to zero. A linear regression t-test was carried out, with a significance of 0.05 (Hernández-Sampieri, Fernández-Collado, & Baptista-Lucio, 2014).

H_0: $\beta_1 = 0$
H_a: $\beta_1 \neq 0$ with a significance of 0.05.

4.3.2 Comparison by gender of TOEIC means and CAFEU means

A second analysis, by selection variable – male and female gender – statistically compared the mean points on CAFEU and TOEIC. For the means CAFEU/TOEIC by gender, the hypothesis of equal means is proposed; whereas the alternative hypothesis is that they will not be equal.

H_0: $CAFEU_M = CAFEU_F$ H_0: $TOEIC_M = TOEIC_F$
H_a: $CAFEU_M \neq CAFEU_F$ H_a: $TOEIC_M \neq TOEIC_F$

A student's t-test for equal means was carried out, considering if the variances are equal or not (Levene Test), with a significance of 0.05.

4.3.3 Test anxiety factors by gender

The TA questionnaire of CAFEU is divided into three factors: Worry (W), Avoidance (A) and Physiological response (PHR). The means by factor by gender were calculated and compared, using the Student's t-test with the null hypothesis for equal means and for the alternative hypothesis of non-equality (in this case, the Levene Test of equality of variances was also taken into account), with a significance level of 0.05.

H_0: $W_M = W_F$ H_a: $W_M \neq W_F$
H_0: $A_M = A_F$ H_a: $A_M \neq A_F$
H_0: $PHR_M = PHR_F$ H_a: $PHR_M \neq PHR_F$

5 Results

As a first step, it was desirable to know if the CAFEU was reliable. To this end, a Cronbach's alpha was determined. A Cronbach's alpha greater than 0.8 shows that the test instrument is reliable. All statistical analyses were run on SPSS (version 23). Cronbach's alpha is 0.936 for 34 items, which indicates a good reliability in the application of the CAFEU. A linear regression analysis was carried out to find the relation of the TOEIC to the CAFEU. Table 2 shows the relationship between the dependent and independent variables.

Table 2. T-test to measure the relation of the dependent variable TOEIC to independent variable CAFEU

Model		Non-Standardized coefficients		Standardized coefficients	t	Sig.
		B	Standard error	Beta		
1	(Constant)	847.638	30.208		28.060	.000
	CAFEU	-1.621	.479	-.218	-3.384	.001

TA has a negative impact on the TOEIC score, as demonstrated by the negative coefficient. The p-value indicates that the relation between TOEIC and CAFEU is significant. The linear regression line is TOEIC = 847.638 – 1.621CAFEU. That is, for every 10 points more on the CAFEU, the TOEIC score is reduced by 16 points.

Of special interest in this study are the differences between gender male (M) and female (F). Results for variable CAFEU and variable TOEIC indicated differences for CAFEU and TOEIC scores by gender. In order to know if these differences are significant, a Student's t-test for equal means was carried out, considering homoscedasticity (Test of Levene). The CAFEU scores were significantly higher for F (m = 65.75 with s = 21.665) than for M (m = 53.56 with s = 13.803) with a t = - 5.16 and p = 0.000. The TOEIC scores were significantly higher for M (m = 772.31, with s = 142.277) than for F (m = 731.02, with s = 142.846) with a t = 2.197 and p = 0.029.

There is a difference in scores in both TOEIC and CAFEU by gender. In order to know if these differences are significant, a Student's t-test for equal means was carried out (see Table 3), considering homoscedasticity (Test of Levene).

Table 3. Student's t-test for equal means

		Test of Levene of equality of variances		T-test for equal means		
		F	Sig.	t	gl	Sig. (bi-lateral)
CAFEU	Equal variances are not as-sumed	16.912	.000	-5.161	209.761	.000
TOEIC	Equal variances are assumed	.026	.873	2.197	229	.029

There are significant differences in mean points, both in the variable CAFEU and in the variable TOEIC, by gender. If TOEIC is related to CAFEU, are they so by gender? A test of linear regression showed, that $TOEIC_M$ and $CAFEU_M$ for male gender are not related; slope (0.014) was not significantly different from zero. That is, TA does not significantly affect the score obtained on the TOEIC for males.

For the female gender, with a constant of 858.49, a negative slope of -1.939, and a p=0.001, $TOEIC_F$ and $CAFEU_F$ are related. The regression line is: $TOEIC_F$=858.490 -1.939 $CAFEU_F$; that is, the $CAFEU_F$ score negatively affects $TOEIC_F$ results. For example, for every 10 points more on the CAFEU, the TOEIC score is reduced by 19 points. CAFEU means are significantly different for M and for F, indicating that TOEIC scores may depend on TA. Thus, it was considered important to analyze by factor Worry (W), Avoidance (A), and Physiological Response (PHR), and by gender.

The means of W, A and PHR show differences (mW_F = 28.24 & mW_M = 22.05; mA_F = 10.63 & mA_M = 9.56; & $mPHR_F$ = 26.91 & $mPHR_M$ = 21.98). With a Student's t-test to compare means, considering the Test of Levene for equality of variances, it was determined whether those differences were significant. There are significant differences by gender in all the factors of the CAFEU questionnaire for TA ($W_F > W_M$ with a p = 0.000; $A_F > A_M$ with a p = 0.019; and $PHR_F > PHR_M$ with a p = 0.000). Because there are significant differences in all the factors by gender, it becomes

important to know which of the items show, by factor (W, A, and PHR), relation to the TOEIC. To this effect, a multiple linear regression analysis will discard those items, by factor, which have no significant effect on the dependent variable TOEIC.

Tables 4, 5, and 6 correspond to the feminine gender (F). Table 7 shows TOEIC score relation to items measuring Physiological Response.

Table 4. Multiple linear regression analysis of PHR items and TOEIC

Model		Non-standardized coefficients		Standardized errors	t	Sig.
		B	Standard error	Beta		
1	(Constant)	778.695	22.544		34.541	.000
	q2 (PHR)	-28.748	11.272	-.226	-2.550	.012

Dependent variable: TOEIC
Selection variable: F

For the female gender, question 2, related to factor PHR correlates to TOEIC thus:
TOEIC = 778.695 – 28.748 q2
q2: When I have been in an exam for a while, I feel discomfort in the stomach and need to defecate.
Table 5 shows TOEIC score relation to items measuring Avoidance.

Table 5. Multiple linear regression analysis of A items and TOEIC

Model		Non-standardized coefficients		Standardized coefficients	t	Sig.
		B	Standard error	Beta		
1	(Constant)	797.590	24.203		32.954	.000
	q22 (A)	-43.556	13.592	-.280	-3.205	.002

Dependent variable: TOEIC
Selection variable: F

For the female gender, q22, related to factor A correlates thus:
TOEIC = 797.590 – 43.556 q22
q22: Before beginning the examination, I think that I know nothing and that I should suspend it.
Finally, Table 6 shows TOEIC score relation to items measuring Worry.

Table 6. Multiple linear regression analysis of W items and TOEIC

Model		Non-standardized coefficients		Standardized coefficients	t	Sig.
		B	Standard error	Beta		
1	(Constant)	850.568	32.178		26.434	.000
	q15 (W)	-25.908	10.574	-.214	-2.450	.016
	q34 (W)	-25.453	9.525	-.233	-2.672	.009

Dependent variable: TOEIC
Selection variable: F

For the female gender, q15 and q34 related to factor W correlate thus:
TOEIC = 850.568 – 25.908 q15 – 25.453 q34
q15: I feel bad if my classmates finish and start turning in the exam before I do.

q34: I take a long time to answer most of the questions and I almost always turn in the exam last.

For the female gender (F), Tables 4, 5, and 6 show that the TOEIC is related, by factor, to the following questions on the CAFEU: q2, q22, q15 and q34. Both items related to Worry have a greater negative affect on TOEIC score than the question related to Avoidance, or the item related to Physiological Response. Table 7 for the male gender (M) corresponds to the factor physiological response. PHR was the only factor which showed significant relation between TOEIC and two items.

Table 7. Multiple linear regression analysis of items related to PHR and TOEIC score

Model		Non-standardized coefficients		Standardized coefficients	t	Sig.
		B	Standard error	Beta		
1	(Constant)	769.254	37.644		20.435	.000
	q3 (PHR)	-44.525	16.323	-.258	-2.728	.007
	q30 (PHR)	55.125	25.127	.207	2.194	.030

Dependent variable: TOEIC
Selection variable: M

For the male gender, factor PHR: TOEIC = 769.254 – 44.525 q3 + 55.125 q30. The questions state: q3: Even though I can see well, when I start reading the exam, my eyes blur and I don't understand what I'm reading. q30: When I enter the testing room, my legs shake.

For q30, the regression coefficient is negative, whereas for q3, the regression coefficient is positive. For the masculine gender, the TOEIC score is not significantly related to any of the items for either A or W, but it is related, though not significantly, to Physiological Response.

6 Conclusion

The CAFEU has shown to be reliable, since its Cronbach alpha value is 0.936. The data show significant differences by gender in each of the aspects analyzed: mean points in CAFEU, mean scores in TOEIC, and mean points in the different factors measured by CAFEU: Worry, Avoidance, and Physiological Response. It was shown that TOEIC significantly relates to CAFEU for the female gender with a negative slope (in the linear regression analysis, which shows that greater TA leads to lower TOEIC score); this is not so for the masculine gender. For the female gender, by CAFEU factor, TOEIC scores relate to one or two questions for each factor. In particular, two items for Worry relate to TOEIC scores. Worry, as mentioned by Cassady and Finch (2014), is conceptualized as cognitive TA, and affects test performance. Avoidance is also related to the cognitive domain; thus, it also affects test scores negatively.

On the other hand, Physiological Response belongs to the domain of Emotionality, which has a lesser effect on TA (Hembree, 1988). For the male gender, scores on TOEIC relate to only two questions of the factor PHR. This may explain why their test scores are not affected as strongly as the women's scores are. The results beg the question: if so many differences are found in TOEIC and CAFEU (and their factors) by gender, should men and women be evaluated by the same measure of TOEIC points? The findings show that women are at a disadvantage when taking a TOEIC test. Their higher level of TA has a negative impact on their test scores, rendering these biased. It is important to consider this when using test results as a graduation, college entry or job entry requirement.

At the university where this study was carried out, alternatives to exams are being offered to students who need to prove language proficiency. One example is a presentation of a case study, where students present in English on a topic related to their field of studies. This is useful because, in most cases, the students are proficient in the language; it is the test itself which leads to anxiety and failure. On the other hand, exam preparation courses should consider the effect of TA on results

and should offer suggestions and practice techniques to counteract anxiety. Preparation exams should help students who fear they are inadequately prepared to face the TOEIC. Knowing what they have to face goes a long way toward diffusing anxiety. TA questionnaires can be administered to students at the beginning of the course and special attention can be given to the elements that cause greater anxiety. Time should be taken to address specific issues of TA. For example, if students mention that stomach discomfort is frequent, the value of a light but nutritious meal before the test needs to be emphasized. For students who worry about being the last to finish, the tutor can mention that it is not important to turn in the exam quickly; rather, it is better to administer one's time carefully and to full advantage.

Finally, it must be noted that there is no TA scale specifically created for language proficiency exams. These however, have characteristics, such as length, time constraints, strict rules, and associated high stakes, which can trigger TA. As proficiency exams become increasingly common, an avenue for further research should be the creation and validation of such a scale.

Acknowledgement

References

Aliakbari, M., & Gheitasi, M. (2017). Investigating students' test anxiety and attitude toward foreign language learning in secondary school in Ilam. *Journal of Pan-Pacific Association of Applied Linguistics*, *20*(2), 51–69.

Álvarez, J., Aguilar, J., & Lorenzo, J. (2012). La ansiedad ante los exámenes en estudiantes universitarios: Relaciones con variables personales y académicas. *Electronic Journal of Research in Educational Psychology*, *10*(1), 333–354.

Aydin, S. (2009). Test anxiety among foreign language learners: A review of literature. *Journal of Language and Linguistic Studies 5*(1), 127–137.

Aydin, U. (2017). Test anxiety: Do gender and school-level matter? *European Journal of Educational Research*, *6*(2), 187–197. doi: 10.12973/eu-jer.6.2.187

Cassady, J. C., & Finch, W. H. (2014). Confirming the factor structure of the Cognitive Test Anxiety Scale: Comparing the utility of three solutions. *Educational Assessment*, *19*(3), 229–242. doi.org/10.1080/10627197.2014.934604

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

Cheng, L., Klinger, D., Fox, J., Doe, C., Jin, Y., & Wu, J. (2014). Motivation and test anxiety in test performance across three testing contexts: The CAEL, CET, and GEPT. *TESOL Quarterly, 48*(2), 300–330. https://doi.org/10.1002/tesq.105

Cizek, G. J., & Burg, S. S. (2006). *Addressing test anxiety in a high-stakes environment: Strategies for classrooms and schools*. Thousand Oaks, CA: Corwin Press.

Farooqi, Y. N., Ghani, R., & Spielberger, C. D. (2012). Gender differences in test anxiety and academic performance of medical students. *International Journal of Psychology and Behavioral Sciences*, *2*(2), 38–43. doi: 10.5923/j.ijpbs.20120202.06

Furlan, L., Pérez, E., Moyano, M., & Cassady, J. (2011). Propiedades psicométricas y estandarización de la Escala de Ansiedad Cognitiva frente a los Exámenes a la población universitaria argentina. *Evaluar*, *10*, 22–31.

Furlan, L., Ferrero, M., & Gallart, G. (2014). Ansiedad ante los exámenes, procrastinación y síntomas mentales en Estudiantes Universitarios. *Revista argentina de ciencias del comportamiento, 6*(3), 31–39.

Grandis, A. M. (2009). *Evaluación de la ansiedad frente a los exámenes universitarios* (Unpublished doctoral dissertation). Faculty of Medical Sciences, Universidad Nacional de Córdoba, Argentina.

Gursoy, E., & Arman, T. (2016). Analyzing foreign language test anxiety among high school students in an EFL context (Note 1). *Journal of Education and Learning, 5*(4), 190–200. doi: 10.5539/jel.v5n4p190.

Hembree, R. (1988). Correlates, causes, effects, and treatment of test anxiety. *Review of Educational Research*, *58*(1), 47–77.

Hernández-Sampieri, R., Fernández-Collado, C., & Baptista-Lucio, M. P. (2014). *Metodología de la Investigación, 6ª Edición*. Mexico City: Mc-Graw-Hill.

Hurley, M., & Padró, F. F. (2006). Test Anxiety and high stakes testing. *International Journal of Learning*, *13*(1), 163–170.

Liebert, R. M., & Morris, L. W. (1967). Cognitive and emotional components of test anxiety: A distinction and some initial data. *Psychological reports*, *20*(3), 975–978.

Núñez-Peña, M. I., Suárez-Pellicioni, M., & Bono, R. (2016). Gender differences in test anxiety and their impact on higher education students' academic achievement. *Procedia-Social and Behavioral Sciences*, *228*, 154–160. doi.org/10.1016/j.sbspro.2016.07.023

Pearlin, L. I., Menaghan, E. G., Lieberman, M. A., & Mullan, J. T. (1981). The stress process. *Journal of Health and Social behavior*, *22*, 337–356.

Peleg, O., Deutch, C., & Dan, O. (2016). Test anxiety among female college students and its relation to perceived parental academic expectations and differentiation of self. *Learning and Individual Differences*, *49*, 428–436. doi.org/10.1016/j.lindif.2016.06.010

Putwain, D., & Daly, A. L. (2014). Test anxiety prevalence and gender differences in a sample of English secondary school students. *Educational Studies, 40*(5), 554–570.

Rosario, P., Núñez-Pérez, J. C., Salgado, A., González-Pineda, J. A., Valle, A., Joly, C. & Bernardo, A. (2008). Ansiedad ante los exámenes: relación con variables personales y familiares. *Psicothema*, *20*(4), 563–570.

Salehi, M., & Marefat, F. (2014). The effects of foreign language anxiety and test anxiety on foreign language test performance. *Theory and Practice in Language Studies, 4*(5), 931–940. doi: 10.4304/tpls.4.5.931-940

Sarason, I. G. (1980). *Test anxiety: Theory, research, and applications*. Mahwah, NJ: Lawrence Erlbaum.

Sieber, J. E. (1980). Defining test anxiety: Problems and approaches. In I. G. Sarason (Ed.), *Test anxiety: Theory, research, and applications*, (pp. 15–40). Mahwah, NJ: Lawrence Erlbaum.

Sommer, M., & Arendasy, M. E. (2014). Comparing different explanations of the effect of test anxiety on respondents' test scores. *Intelligence, 42*, 115–127. doi.org/10.1016/j.intell.2013.11.003

Sommer, M. & Arendasy, M. (2015). Further evidence for the deficit account of the test anxiety–test performance relationship from a high-stakes admission testing setting. *Intelligence*, *53*, 72–80. doi.org/10.1016/j.intell.2015.08.007

Zalta, A. K., & Chambless, D. L. (2012). Understanding gender differences in anxiety: The mediating effects of instrumentality and mastery. *Psychology of Women Quarterly*, *36*(4), 488–499. doi.org/10.1177/0361684312450004

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

Appendix 1

CAFEU **Questionnaire** (translated from original Spanish. Participants received Spanish version).

Date

Student ID number

Instructions:

In each of the items or situations described, indicate with a score of 1 to 5 how often you notice the proposed symptoms. The following criteria must be used for this:

1 - Never
2 - Only on one occasion
3 - On more than one occasion
4 - Many times
5 - Always

1. During exams, my hands perspire.
2. When I have been in an exam for a while, I feel discomfort in the stomach and need to defecate.
3. Even though I can see well, when I start reading the exam, my eyes blur and I don't understand what I'm reading.
4. I am almost always late for an exam and when that happens I do not show up to take it.
5. When I finish an exam, I have a headache.
6. In exams, I feel that I lack air. I feel faint, very hot or a general malaise.
7. When a teacher stands next to me, it makes me very nervous and I cannot keep working.

8. I get nervous when I see the teacher arrive in the classroom with the exams

9. My hands feel stiff during exams

10. Before entering the test, I feel "something" in my stomach

11. Just by reading or listening to the questions of the test, I go blank.

12. When I finish an exam, I feel sad, thinking that I did it incorrectly despite not having the results yet.

13. I think long before deciding whether to take an exam

14. While I'm answering an exam, I think I'm doing it wrong

15. I feel very bad if my classmates finish and start to turn in the exam before me

16. While I am taking the exam, I think that the teacher is constantly looking at me

17. I want to smoke during the exam

18. I really want to go to the bathroom during the exam

19. I get sick and look for any excuse not to do an exam

20. For me it is a comfort or relief when, for whatever reason, an exam is postponed

21. Despite having studied, I think that I will not be able to pass

22. Before beginning the examination, I think that I know nothing and that I should suspend it.

23. I cannot sleep the night before the test.

24. It makes me nervous if there are many classmates and / or teachers in the exam.

25. During the exam, I felt dizzy and nauseous.

26. Before I turn in the exam, my mouth is dry; I cannot swallow.

27. If I sit at the front of the classroom, I get more nervous.

28. I get more nervous if the exam has a time limit to finish it and I do it worse or I do not finish.

29. When I'm doing an exam I feel like my heart beats faster.

30. When I enter the testing room, my legs shake.

31. I feel nervous in exams with few classmates.

32. When I finish the exam, I feel like I will not pass it.

33. I think that in the exam my nerves will make me forget everything I studied.

34. I take a long time to answer most of the questions and I'm almost always the last one to turn in the test.