

Sultan Qaboos University
Journal of Arts & Social Sciences



جامعة السلطان قابوس
مجلة الآداب والعلوم الاجتماعية

An Investigation into Assessment Preferences of Gulf Arab Students

Mashaël Al-Hamly

Associate Professor
Department of English Language
and Literature - College of Arts
Kuwait University
mashaël2@hotmail.com

Christine Coombe

Faculty Member
Dubai Men's College
Higher Colleges of Technology
ccoombe@hct.ac.ae

An Investigation into Assessment Preferences of Gulf Arab Students

Mashaal Al-Hamly and Christine Coombe

Abstract

The purposes of this study are to gain insight into Gulf Arab EFL students' preferences when being assessed in their university-level English courses and to investigate if gender and degree of test anxiety have an impact on these preferences. Data were obtained from responses on an adapted form of Birenbaum's (1994) Assessment Preferences Inventory. Results show that the students participating in this study favor assessments with multiple-choice (MCQ) questions as they are viewed as being easier to prepare for, easier to take, and thus will bring in relatively higher scores. Results also indicate that students are very much in favor of choosing how they will be assessed and data report a strong preference for multiple measures, and low stakes assessment. Implications for the classroom are also discussed.

Key Words: Assessment perceptions, Assessment preferences, English as Second Language

التحقق مما يفضله الطلبة الخليجيون في التقييم

مشاعل الحملي و كريستين كومب

المستخلص:

تهدف هذه الدراسة إلى الكشف عما يفضله طلبة الخليج العربي الذين يتعلمون الإنجليزية كلغة أجنبية عندما يجري تقييمهم في مقررات اللغة الإنجليزية على المستوى الجامعي، وإلى التحقق عما إذا كان الجنس ودرجة القلق حول الامتحان لهما تأثير على ميولهم. وقد تم جمع بيانات هذه الدراسة من ردود الطلبة على نسخة معدلة من استبانة بيرنباوم لتقييم ميول الطلبة (1994). وتشير النتائج إلى أن الطلبة المشاركين في هذه الدراسة يفضلون الأسئلة ذات الاختيار المتعدد وذلك لاعتقادهم أنها أسهل من حيث التحضير للاختبار وحل أسئلته والحصول على درجات أعلى نسبياً. وتشير النتائج أيضاً إلى أن كثيراً من الطلبة يفضلون اختيار الكيفية التي يتم تقييمهم بها وإلى وجود ميل قوي لديهم لاستخدام طرق مختلفة وأساليب منخفضة المخاطر في تقييمهم. هذا ويناقش البحث الآثار المترتبة لذلك على الفصول الدراسية.

الكلمات المفتاحية: انطباعات عن التقييم ، تفضيلات التقييم ، اللغة الانجليزية كلغة ثانية

Introduction

The field of assessment has been undergoing major reform since the appearance of alternative forms of assessment in the mid-1990s. Indeed, a shift has taken place from what some call “a culture of testing” to a “culture of assessment” (Kleinsasser, Horsch and Tastad, 1993). In more recent years, assessment specialists have been calling for “multiple measures assessment” as a more valid and reliable way to assess students’ language proficiency. Multiple-measures assessment is the realization that no one single type of assessment can give us all the information we need to accurately judge language ability (Coombe, Folsie and Hubley, 2007). These two reforms have now placed a strong emphasis on using a number of different assessment types while focusing on integrating assessment with instruction. Because of these recent changes in the assessment landscape, the perceived position of the student with regard to the assessment process has been changing from that of a passive, powerless, often oppressed, subject who is mystified by the process, to an active participant who shares responsibility in the process (Birenbaum, 1994:239).

According to Birenbaum (1994), the view that one assessment fits all is a naive and an unrealistic one. Assessment format has been the subject of numerous studies in the literature and students’ assessment preferences are considered by many to be a potentially valuable source of evidence of test validity (de Watering, Gijbels, Dochy and van der Rijt, 2008; Zeidner, 1987). Because students are the group of stakeholders most affected by assessment, it makes sense that their attitudes towards, preferences for and perceptions of assessment are critically important. Despite this importance, research examining students’ assessment preferences is sparse. This study will be an attempt to provide baseline knowledge about the assessment perceptions and preferences of tertiary-level Kuwaiti and Emirati students of English.

Review of Pertinent Literature

The learning processes and strategies of students in higher education have been widely investigated in the last two decades. The topic under study in this research, assessment preference and the effect of assessment format on performance has been widely investigated in the literature (Bennett, 1993; Birenbaum and Tatsuoka, 1987; Birenbaum, Tatsuoka and Gutvirtz, 1992; Shohamy, 1984; Traub, 1993). However, in light of the effect assessment has on students,

both as performers and as the objects (and often as Birenbaum points out ‘victims’ (1994)) of the decisions based on the assessment results, it is surprising to witness the paucity of research regarding students’ assessment attitudes and preferences. Most of the early studies that did investigate this issue did not relate it to students’ personal characteristics. The personal characteristics investigated in these studies included cognitive style (i.e., field dependence /independence) (Chapelle, 1988; Hansen 1984; Hansen and Stansfield, 1981; Lu and Suen 1993), test anxiety (Zoller and Ben-Chaim, 1988) and gender (Ben-Shakhar and Sinai, 1991; Grandy, 1987) and their main findings indicated that high field dependence and high test anxiety tended to debilitate performance on assessment of the constructed response type. With regard to gender, males were shown to outperform females on assessments of the MCQ variety.

According to the studies of Ben-Chaim and Zoller (1997), Birenbaum and Feldman (1998), Traub and MacRury (1990) and Zeidner (1987) students, especially the males (Beller and Gafni, 2000), generally prefer multiple choice formats, or simple and de-contextualized questions over essay type assessments or constructed-response types of questions.

Traub and MacRury (1990) reported that students have more positive attitudes towards multiple choice tests in comparison to free response tests because they think that these tests are easier to prepare for, easier to take, and thus will bring in relatively higher scores. In the study by Ben-Chaim and Zoller (1997), the examination format preferences of secondary school students were assessed by a questionnaire and structured interviews. Their findings suggest that students prefer written, unlimited time examinations and those in which the use of supporting material is permitted. The reasons for this being that time limits are found to be stressful and result in increased test anxiety on the part of the student.

Scouller (1998) investigated the relationships between students’ learning approaches, preferences, perceptions and performance outcomes in two assessment contexts: a MCQ exam requiring knowledge across the whole course and assessment essays requiring an in-depth study of a limited area of knowledge. The results of this study indicated that if students stated a preference for an essay type test then this is more likely to result in positive outcomes for their essays. Beller and Gafni (2000) overviewed several studies and found some consistent conclusions suggesting

that, if gender differences are found (and this was not always the case), female students generally preferred essay formats and male students showed a slight preference for MCQ formats (Gellman and Berkowitz, 1993). Another finding was that male students scored better than females on MCQ formats while females scored higher than men on open-ended questions. Birenbaum (1994) employed a questionnaire to determine students' assessment preferences. This questionnaire, the Assessment Preferences Inventory or API, was designed to measure three areas or what are referred to as dimensions of assessment. The first dimension was 'assessment-form related' and included aspects such as assessment type, item format/task type and pre-assessment preparation. The second was an 'examinee-related' dimension such as cognitive processes, students' role/responsibilities and cognitive aspects. The final area was a 'grading and reporting' dimension. One of the major findings in a study conducted by Birenbaum and Feldman (1998: 336) using the API was that students who employed a deep study approach tended to prefer essay-type questions while those with a surface study approach tended to prefer MCQs. Baeten et al. (2008) reported similar results and found that the degree to which a student suffered from test anxiety seemed to be a variable in their specific attitude towards assessment formats. In particular, they found that students with high degrees of test anxiety had more favourable attitudes towards MCQs while those with low test anxiety tended to prefer more open-ended formats.

To date, however, there has been very little research concerning the assessment preferences of tertiary-level EF/SL students, in general, and Gulf Arab or Middle Eastern students, in particular. One study by Casebeer and Iquraan (2011) investigated the beliefs of low-performing postsecondary students in Jordan concerning the assessment practices of their instructors using the SPAP (Student Perceptions of Assessment Practices Inventory), a self-report instrument with 30 Likert type items. Their results indicated that low performing students believe their instructors focus their assessment on three main areas: expectations, the communication of assessment practices and organizational assessment practices. Another major finding of this study was that instructors relied more on traditional, paper-pencil tests than on alternative assessments. No relationship with gender or subject area of the student with assessment practices was found to exist.

Overall, from the studies reviewed regarding students' assessment preferences, it seems that students prefer assessment formats which reduce stress and anxiety. It is also assumed that students will perform better on their preferred assessment formats. With these findings in mind there is a need for more research concerning the involvement of students in the assessment process.

Rationale for the study

It is important to examine tertiary level student beliefs concerning the assessment methods and practices used by their instructors. It is also necessary to understand these perceptions in order to assist students in their engagement in the educational process. Student beliefs concerning their instructor's assessment practices influence their learning strategies (Entwistle, 1991), which in turn impact the quality of their learning. There is now widespread recognition that students have a legitimate role in the assessment process and in their own learning (La Lopa, 2004; Leach, Neutze, and Zepke, 2001; Taras, 2002). Research findings have also revealed that students' perceptions about assessment significantly influence their approaches to learning and studying. Students hold strong views about different assessment and evaluation formats. Given this belief it is a major assumption that if students are provided with the type of assessment formats they prefer, they will be motivated to perform at their best.

Research Questions

This study posits the following research questions:

- Which type of assessment (traditional or alternative) do Gulf Arab students prefer?
- Which assessment formats do Gulf Arab students prefer?
- What types of assessments do high test anxious and low test anxious Gulf Arab students prefer?
- What teacher assessment practices do Gulf Arab students favor?

Methodology

Participants

A total of 116 students participated in the study, 88 females and 28 males. These students were final year students from Kuwait University in Kuwait and the Higher Colleges of Technology and Zayed University in the UAE. Students ranged in age from 17 to 25. The participants were in their final year of university study and enrolled in a variety of different degree programs.

Instruments

Student Profile Form

A form was developed for the purpose of collecting student demographic information. This form also functioned as a human subject consent form to participate in the study. This document was used to collect the following information for each student: name, date, and contact information, field of study, institution, university I.D. number, gender, and year of study. Students were also asked to read a 'permission statement' and sign it. The Student Profile Form can be found in Appendix A.

Adapted Assessment Preferences Inventory (AAPI)

An adapted version of the Assessment Preferences Inventory (AAPI) developed by Birenbaum in 1994 was used as the primary instrument for data collection in this study. The original API is a 67 Likert-scale questionnaire containing items referring to three content dimensions: assessment-form related dimensions, examinee related dimensions, and grading and recording. The items cover various alternatives of the traditional as well as the current alternative approaches to achievement assessment (Birenbaum, 1994: 246-7). For the purpose of this study, only two dimensions of the questionnaire were used and they measured assessment preferences and grading and reporting practices on a 5 point Likert-scale (from 1 = not applicable at all to 5 = to a great extent). Adaptations were made to the original API so that new assessment types were represented and hence the instrument reflected the current educational and assessment environment. The adaptations made were to add assessment types now in existence that were not prevalent when Birenbaum constructed his questionnaire. In particular, assessments delivered by technology (i.e., computer or iPad based assessments) and performance-based assessments were added. The adapted API (AAPI) posited 36 questions about students' preferences for different types of traditional, oral, written or alternative assessments as well as student preferences for grading and reporting practices. The adapted form of the API was translated into Arabic and students were free to choose either the English or Arabic version. A copy of the English version of the adapted AAPI can be found in Appendix B.

Test Anxiety Inventory

The TAI is a frequently employed and thoroughly researched self-report psychometric scale developed to

measure individual differences in test anxiety in high school and college students (Spielberger, 1972, 1980). The one-page Likert-scale questionnaire asks subjects to record their degree of agreement/disagreement in twenty areas all related to the concept of test anxiety. Response choices are (1) almost never, (2) sometimes, (3) often, and (4) almost always. "Almost never" indicates low test anxiety and is scored "1"; "Almost always" indicates high test anxiety and is scored "4." The scoring weights are reversed on Item One only, since its intention is as a checks and balances item. All twenty items were used to determine the total score of the TAI. The minimum TAI total score (very low, if any, anxiety) is 20. The maximum TAI total score (very high anxiety) is 80. The objective of the TAI is to learn how frequently students experience anxiety symptoms before, during, and after tests.

The TAI was originally published in English and later translated into Arabic by the Arabic-speaking co-author. The TAI was translated in an effort to ensure total comprehension on the part of the subjects and to minimize the risk of miscomprehending any test items (Ollerand Perkins, 1978). The translation was validated by asking another highly educated native speaker of Arabic to translate the document back into English. Every attempt was made to provide an Arabic version that was as faithful a representation of the English as possible. The back-translated versions were compared with the original English to ensure that the translation was accurate. An Arabic/English linguist and the author checked the Arabic and the back-translated English versions of the instruments and compared them to the original English version. The same process was carried out with the Student Profile Form. A copy of the Test Anxiety Inventory can be found in Appendix C.

Data Collection Procedures

Researchers employed a quantitative approach to data collection consisting of responses to a questionnaire based on Birenbaum's (1994) Assessment Preferences Inventory (API). Data were collected in two separate administrations (at Kuwait University and at both the Higher Colleges of Technology-Dubai, Men's Campus and Zayed University in the UAE) over one session. First, students were asked to fill in the Student Profile Form, sign the consent form, and complete the TAI. The students were then asked to answer the AAPI. A protocol for the administration of all instruments was designed to insure that procedures would be standard throughout data collection.

Scoring

The TAI was manually calculated and the resulting score entered on the Student Profile Form. Because no published interpretations of the TAI were in existence, the researchers classified students scoring between 20-39, 40-60, and 61-80 as having either low, moderate, or high levels of test anxiety, respectively.

Data Analysis

Students' scores were statistically analyzed using Microsoft Excel and SPSS. Students' scores of the various items of the AAPI were calculated to figure out the mean score of each item. 'Not Applicable' choices were excluded from the calculations. Once the means were calculated, a comparison was made between items to identify students' preferences; the higher the mean (which on the AAPI is out of 5), the higher the preference of students for the item in question. Independent sample T-tests were conducted to investigate if observed differences in the student preferences on all variables were found to exist and whether or not they were statistically significant.

Results and Discussion

The research questions will be reported on individually based on the quantitative data obtained from the AAPI. Then these results will be compared to the pertinent literature to gain more insight into the most important concerns in student assessment preferences.

- Research Question 1: Which type of assessment (traditional or alternative) do Gulf Arab students prefer? Mean scores were calculated for all items on Part A of the AAPI that were classified as traditional and alternative assessments. Nine of the questions on the AAPI were classified as alternative (9, 10, 11, 12, 19, 20, 24, 25, 26) whereas the other 17 questions represented traditional forms of assessment (1-8, 13-18, 21-23). Results of the data collected to answer this question found that slightly higher mean scores (3.36) existed for all students for traditional forms of assessment over alternative forms of assessment (3.35). Table 1 shows students' individual preferences for types of traditional vs. alternative forms of assessment.

The pertinent literature on student preference of traditional forms of assessment over alternative assessments is inconclusive. The general overall findings indicate that students in various content areas prefer closed-response formats like MCQs (Ben-Chaim and Zoller, 1997; Birenbaum and Feldman, 1998; Traub and MacRury, 1990; Zeidner, 1987; Beller and Gafni, 2000; Scouller, 1998). However, it should be noted

that none of these studies investigated assessment preferences in an EF/SL environment.

In a meta-analysis on the use of self-, peer- and co-assessment in higher education, Dochy, Segers and Sluijsmans (1999) synthesized 63 studies and reported that there was "empirical evidence that the students perceived positive effects" from these three types of alternative assessment (p. 347). They went on to state

Table (1): Student preferences for traditional and alternative forms of assessment ranked from highest to lowest (N=116)

Q No.		Mean
14	Tests with multiple choice questions	4.07
26	Tests that students develop themselves	4.04
8	Take home exams where students are allowed to take the test home and use their class materials to answer the questions	3.97
23	Assessments where students reflect or give their opinions	3.90
16	Tests with 'True/False/No Information' questions	3.69
3	Written tests without a time limit, while permitting the use of supporting materials	3.65
15	Tests with matching questions	3.64
17	Tests with questions requiring students to write short answers of one or two sentences	3.53
24	Self-assessment where students give themselves a grade	3.50
1	Traditional tests, with a time limit	3.47
11	Projects where students work in groups (3+ students)	3.46
12	Projects where students work individually	3.38
21	Simple tasks having only one correct answer	3.37
10	Projects where students work in pairs	3.36
20	Assessments which require students to perform a real-life task	3.34
6	Oral tests, in the form of a group discussion where the instructor observes and assesses the contribution of each of the participants	3.27
13	Tests that students take on computer or iPad	3.25
7	Oral tests, in the form of a group discussion where the instructor observes and assesses the group as a whole	3.21
22	More complex tasks which have more than one possible answer	3.21
2	Traditional tests, without a time limit	3.13
9	Portfolios where students submit their work either finished or in progress	3.07
25	Peer assessment where students in the class grade other students	3.07
19	Assessments which require students to give oral presentations	2.97
5	Individual oral tests	2.85
18	Tests with questions requiring students to write longer answers like essays	2.66
4	Written tests, without a time limit with no supporting materials allowed	2.27

that “involving students in assessment is perceived as being valid, reliable, fair and as contributing to a growth in competence” (p. 347). Results of the current study indicate that one of the assessment types reviewed by Dochy, Segers and Sluijsmans (1999) that of co-assessment or what we term student-designed tests, ranks second with the present study’s overall student population.

- Research Question 2: Which assessment formats do Gulf Arab students prefer?

Most Preferred Types of Assessment

When ranking the highest to lowest scores across assessment types, the following six preferences were found to exist as shown in Table 1. The highest overall preference was for assessments that employ multiple-choice questions (MCQs) with a mean of 4.07. The second highest preference was for tests that students develop themselves (4.04). The third highest rated assessment type (3.97) for all students was for take-home exams where students are allowed to complete the test at home and use their class materials to help answer the questions. Assessments where students reflect or give their opinions was ranked fourth overall with a mean of 3.90. The fifth and sixth highest rated types of assessment were two forms of objectively-scored questions types: matching (3.69) and true/false questions/no information (3.64) respectively. It is interesting to note that overall five out of six of the most preferred test types were forms of traditional assessment. The only form of alternative assessment that students ranked in their top 6 was student-designed tests.

Least Preferred Types of Assessment

As shown in Table 1, should be noted that when studying what forms of assessment (traditional or alternative) students least preferred, it was found that the students did not prefer one form over the other. The

three least preferred types of assessment were forms of traditional assessment followed by three types which were forms of alternative assessment.

The least preferred type of assessment was found to be written tests administered without a time limit which did not permit students to use supporting materials (2.27). This type of assessment was followed by tests that have questions requiring students to write longer answers such as essays (2.66). The third least preferred test is individual oral tests with a mean of 2.85.

Assessments which require students to give oral presentations scored fourth (2.97). Peer assessments where students in the class grade each other’s tests and the use of portfolios where students submit their work either in its finished format or in progress were the fifth and sixth least preferred types of assessment on the student’s preference list with a mean of 3.07 for both types.

Major Differences between Males and Females

In order to find out if there is a significant difference between male and female students in terms of assessment type preference, the researchers analyzed students’ preferences for the six most and least preferred types of assessments. Tables 2 and 3 show students’ preferences for the highest and lowest ranked types of assessment in relation to gender.

As indicated in Table 2, male students most prefer assessments where they can reflect on or give their opinion about a topic (4.07), followed by MCQ tests (3.93) and student developed tests (3.85). These most preferred types of assessment by male students were followed by ‘take-home exams (3.70), exams which do not have a time limit and do not allow for use of supporting materials (3.63), and true/false tests (3.3). Females ranked students developed tests (4.11), MCQs (4.10), and take-home exams (4.06) as their three most preferred types of assessment. Assessments where they can reflect on or give opinion (3.85), true/

Table (2:) Students preferences for the highest ranked types of assessment

Rank	Assessment Types	All	Males	Females	t-value	df	p
1	MCQs	4.07	3.93	4.10	0.73	113	.47
2	Student developed tests	4.04	3.85	4.11	1.05	113	.29
3	Take-home exams	3.97	3.70	4.06	1.25	111	.22
4	Assessments where students reflect/give their opinion	3.90	4.07	3.85	-.89	113	.37
5	True/false questions No information	3.69	3.30	3.76	1.57	110	.19
6	No time-limit test with no supporting materials	3.65	3.63	3.64	0.49	112	.96

Table (3): Students choices for least preferred types of assessment

Rank	Assessment Types	All	Males	Females	t-value	df	p
1	Written tests, no time limit no supporting materials	2.27	2.36	2.24	0.49	112	.96
2	Long-answer tests like essays	2.66	1.93	2.87	3.52	112	.001
3	Individual oral tests	2.85	3.52	2.64	3.01	112	.003
4	Oral presentations	2.97	2.89	2.95	.212	114	.832
5	Peer assessment	3.07	3.37	2.95	-1.34	112	.182
6	Portfolios	3.07	3.69	2.87	-2.79	109	.006

Table (4): Preferences of high and low anxious students for the highest ranked types of assessment

Rank	Assessment Types	High anxious	Low anxious	t-value	df	p
1	MCQs	4.30	4.05	-.56	28	.58
2	Student developed tests	4.44	3.60	-1.71	27	.099
3	Take-home exams	4.80	3.89	-2.34	26	.027
4	Assessments where students reflect/give their opinion	3.80	4.21	1.06	27	.30
5	True/false questions No information	4.20	3.45	-1.50	28	.14
6	No time-limit test with no supporting materials	4.10	3.79	-.67	27	.51

false tests (3.76), and exams which do not have a time limit and do not allow for use of supporting materials (3.64) were ranked next on the female students list of most preferred assessment types.

As shown in Table 3, some significant differences were found on three types of assessment between male and female students' rankings. Female students' lack of preference for individual oral tests and portfolios where students submit their work either finished or in progress was significantly higher than male students with a p-value of .003 and .006 respectively. On the other hand, male students showed a significantly lower preference for tests with questions requiring students to write longer answers such as essay than their female counter parts with a p-value of .001.

Both genders differed in their ranking of the other least preferred test types. The least preferred assessment type for female students was written tests without time limits and without the use of supporting materials, followed by individual oral tests, portfolios, written tests with questions requiring students to write longer answers such as essay, oral presentations, and peer assessments where students in class grade each other. Male students least preferred written tests with questions requiring students to write longer answers such as essays, followed by written tests without time limits and do not allow the use of supporting materials, oral presentations, peer assessments, individual oral presentations, and portfolios. As regards the pertinent literature on gender differences in assessment preferences, the results of this

study deviate slightly from previous findings. According to the studies of Ben-Chaim and Zoller (1997), Birenbaum and Feldman (1998), Traub and MacRury (1990) and Zeidner (1987) students, especially the males (Beller and Gafni, 2000), generally prefer multiple choice formats, or simple and de-contextualized questions over essay type assessments or constructed-response types of questions.

Results of the current study indicate that Gulf Arab males stated a preference for assessments where they can reflect and give their opinion as their top choice. This was followed closely by MCQ tests. Student designed tests ranked third with the current study's male student population. Females in the study recorded a preference for student designed tests followed closely by MCQ tests.

- Research Question 3: What types of assessments do high test anxious and low test anxious Gulf Arab students prefer?

For the purposes of this study, students who scored between 61 and 80 on the Test Anxiety Inventory were categorized as high test anxious while those who scored 20 to 39 were classified as low test anxious. In total, 20 students in this study were found to suffer from low amounts of test anxiety whereas 26 were classified as high text anxious.

In general, high anxious students scored higher means of preference than low anxious students on the highest ranked types of assessment. A significant difference of .027 and .099 was found between the two

Table (5): Preferences of high and low anxious students for the least ranked types of assessment

Rank	Assessment Types	High anxious	Low anxious	t-value	df	p
1	Written tests, no time limit nosupporting materials	2.6	2.3	-.56	28	.579
2	Long-answer tests like essays	2.70	2.32	-.79	27	.438
3	Individual oral tests	2.40	3.30	1.76	28	.089
4	Oral presentations	2.40	3.45	1.93	28	.064
5	Peer assessment	3.11	2.89	-3.99	26	.693
6	Portfolios	4.0	2.58	-3.04	26	.005

Table (6): Students preferences for assessment practices

No of Qs of Part B	Assessment Practices	All	Males	Females	t-value	df	p
1	Detailed description of assessment.	4.70	4.38	4.80	2.23	100	.028
2	What and how to prepare for the test.	4.57	4.50	4.57	0.34	115	.734
3	Examples of test questions	4.53	4.25	4.61	1.65	115	.102
4	Practice questions	4.41	4.21	4.45	1.04	114	.303
5	Assess class participation and discussions	4.09	3.68	4.22	-1.07	114	.287
6	Assess homework and exercises completed	3.78	4.0	3.73	2.27	115	.025

anxiety groups in favour of the take-home exams and student developed tests respectively by the high anxious students.

Not surprisingly, high test anxious students also preferred MCQ tests (4.30), true/false (4.20) as well as written tests without a time limit and which permit the use of supporting materials (4.10). Low anxious students, however, preferred more assessments where students reflect or give their opinion (4.20).

As for the least ranked types of assessment, Table 5 shows that low anxious students prefer significantly less than high anxious students' portfolios (a p-value of .005). Not surprisingly, high anxious students were found to prefer less than low anxious students' individual oral tests (2.40) and assessments which require students to give oral presentations (2.40). Low anxious students stated a lesser preference than the high anxious students for written tests which do not have a time limit and which do not allow the use of supporting materials (2.30), tests with questions requiring students to write longer answers like essays (2.32), and peer assessment where students in the class grade other students (2.89).

In the pertinent literature on the effects of test anxiety on assessment preference, Baeten et al. (2008) found that the degree to which a student suffered from test anxiety seemed to be a variable in their specific attitude towards assessment formats. In particular, they found that students with high degrees of test anxiety had more favorable attitudes towards MCQs while those with low test anxiety tended to prefer

more open-ended formats. Our results indicate that high test anxious students prefer what they consider to be low stress take-home tests where they can work from home and use class materials to answer the questions. Student-designed tests and MCQ tests ranked very highly as well with those students who exhibited high degrees of test anxiety.

- Research Question 4: What teacher assessment practices do Gulf Arab students favor?

Parts B and C of the AAPI investigated teacher test preparation/grading practices and general assessment practices respectively. In general, on Part B, both male and female students responded quite positively to the listed test preparation and grading practices. As indicated by the high mean score of 4.7, students, in general, are very much in favor of transparency citing practices like handing out a detailed description of the course and its assessment from the very beginning of the academic year as valuable ones. A significant difference of .028 and .025 was found between male and female students where female students expressed stronger preference for being handed out at the beginning of the course, a detailed description of the way they will be assessed; and for teachers to assess, as part of the course grade, the homework and exercises the students have completed respectively. Other teachers' practices mentioned on Part B of the AAPI were more preferred by females than male students. Female students preferred teachers to clarify what will be on the test and how to prepare for it

Table (7): Preferences of high and low anxious students for assessment practices

No of Qs of Part B	Assessment Practices	High anxious	Low anxious	t-value	df	p
1	Detailed description of assessment.	5.0	4.63	1.08	26	.289
2	What and how to prepare for the test.	5.0	4.55	-1.41	28	.169
3	Examples of test questions	4.90	4.35	-1.49	28	.148
4	Practice questions	4.70	3.74	-1.94	27	.063
5	Assess class participation and discussions	4.0	3.75	-.56	28	.583
6	Assess homework and exercises completed	4.30	3.85	-.88	28	.386

Table (8): Students preferences for reporting and grading

No of Qs/of Part C	Reporting and grading	All	Males	Females	t-value	df	p
1	Questions presented in order of difficulty (from easy to difficult)	3.37	3.58	3.32	-.77	101	.442
2	Questions presented in order of study topics	4.08	4.15	4.01	-.59	111	.550
3	One big test during the semester	2.80	2.96	2.80	-.53	108	.559
4	Several smaller quizzes throughout the semester	3.89	3.84	3.94	.38	112	.705
5	Achievements assessed by different types of tasks	4.06	3.70	4.17	2.19	113	.031
6	Choice of preferred type of assessment.	4.14	4.18	4.13	-.19	110	.842

Table (9): Preferences of high and low anxious students for reporting and grading

No of Qs of Part C	Reporting and grading	High anxious	Low anxious	t-value	df	p
1	Questions presented in order of difficulty (from easy to difficult)	3.22	3.25	.047	27	.963
2	Questions presented in order of study topics	4.67	3.94	-2.15	25	.042
3	One big test during the semester	2.90	3.39	.92	26	.366
4	Several smaller quizzes throughout the semester	4.80	3.47	-.320	27	.003
5	Achievements assessed by different types of tasks	4.70	3.89	-2.29	27	.030
6	Choice of preferred type of assessment.	4.50	3.95	-1.35	27	1.87

(4.57), to give out prior to the test examples of the kind of questions to be asked on the test (4.61), to hand out a list of practice questions, from which the teacher will choose the actual test questions (4.45); and to assess, as part of the course grade, students' participation in class discussions (4.22).

No significant differences were found between high and low anxious students in relation to teachers' assessments practices. Students with high degrees of test anxiety were more in favor of transparent testing processes than their low test anxious counterparts. The high anxious students had stronger preferences for receiving detailed descriptions of what they will be assessed on (5.00), what and how to prepare for it (5.00), getting examples of test questions (4.90) and practice questions (4.70) as well as on being assessed for homework and exercises completed (4.30) and class participation and discussions (4.0). This is consistent with the literature that higher degrees of test transparency seem to lower the anxiety levels of

students.

Part C of the AAPI examined various assessment preferences for allocating and reporting on grades (Table 8). Students in this study reported a marked preference for multiple measures assessment. Generally speaking, students preferred to be assessed through a variety of different assessments/tasks (4.06) rather than on one or two. A significant difference was found for this preference between male and female students of .031 where female students had stronger preference (4.17) than their male counterparts (3.70) for multiple measures assessment. Female students also showed a slightly stronger preference (3.94) than male students (3.84) for having several smaller quizzes throughout the semester than having one big exam. Male students, on the other hand, had stronger preferences than female students on the other four items on Part C of the AAPI.

The concept of continuous, multiple-measures assessment figured prominently on Part C of the AAPI

for high vs. low anxious students (Table 9). Based on responses to two questions (4, 5), high anxious students in this study showed a significant preference for continuous assessments throughout the semester and the use of multiple measures assessment (p. value of .003 and .030 respectively). The third significant difference (p value of .042) found related to how high anxious students preferred the content of their assessments to be presented. High anxious students stated a preference that the questions on their assessments be presented in the order that they were studied.

The final question on Part C of the AAPI measured the students' desire to be able to choose their own assessment methods. Overall, students reported a strong preference for being able to choose their preferred type of assessment (4.14). Males had slightly stronger feelings than females about having a say in their own assessment choice (males 4.18/females 4.13) (Table 8) and high anxious students stated a stronger preference (4.50) than those students with low test anxiety (3.95) about assessment choice (Table 9).

Implications for the Classroom

Much attention has been paid in recent years on assessment for learning as a superior method of educational assessment (Black and William, 1998). This approach is robustly committed to using assessment explicitly to improve student learning (as opposed to simply measuring it) by actively involving students in the processes of assessment. As such, much emphasis has been placed on the active participation of students in their own assessment. Yet, research in this area clearly shows that students' voices are not being heard, especially in regards to assessment preference. From the findings of this study and similar studies conducted previously, there are a number of implications that emerge. First, we as EF/SL educators need to listen to students' voices about how they want to be assessed. We can gain much information about them by familiarizing ourselves with the types of assessments they favor and the types of assessment they disfavor. Exploring their perceptions about why they like and dislike certain types of assessment can provide us with valuable information and could be viewed as positive 'wash back'. By exploring these views we can become acquainted with possible misunderstandings on the part of our students where certain types of assessment are concerned.

The results of this study also indicate that students favor having a voice in selecting their own types of

assessment. Specific likes reported by students in this study include having several smaller assessments rather than one larger one. This focus on having a multiple measures assessment scheme is in keeping with the current recommendations of the larger testing community which advocates a number of low-stakes traditional and alternative forms of assessment.

Areas for Further Research

Before definitive conclusions can be drawn on whether incorporating student assessment preference into our assessment repertoire is a good idea, there are a number of areas for future research. As far back as 1994, Birenbaum called for further research that extends the scope of personal characteristics and their impact on assessment preference. Some of the characteristics that were mentioned included: thinking styles, tolerance of ambiguity, causal attributions, procrastination, etc. (p. 81). These variables and more need to be investigated to examine their link to assessment preference.

Future research on Gulf Arab students should examine their learning approaches (deep vs. surface learning) to determine whether these have an impact on student assessment choices. The assumption made by many of the previous researchers that students will perform better on their preferred assessment formats also needs to receive empirical attention. These studies should also incorporate qualitative elements to better understand the reasoning involved behind assessment preference.

A surprising finding of the current study is the low preference mean scores for individual oral tests from both males and females. Strangely Gulf Arab students perform very well in these types of tests, often receiving strong IELTS bands in this skill area. Research into why students rate this type of assessment so low would shed light on these seemingly contradictory findings. Before the results of this study can be generalized, further research is needed to answer these questions.

Conclusion

Results of the current study support earlier findings carried out in different educational contexts and with different subject matter areas. It has been argued that if we provide students with the assessment type they prefer, the perceived validity of the assessment will improve, thus motivating students to perform at their best. By hearing our students' voices on what they like and don't like vis-a-vis assessment preferences, we, as

educators, will be better able to improve our assessment practices and achieve a higher quality of learning for our students.

References

- Baeten, M., Struyven, K., and Dochy, F. 2008. Student's assessment preferences and approaches to learning in new learning environments: A replica study. Paper presented at the annual conference of the American Educational Research Association, March 2008, New York City.
- Beller, M. and Gafni, N. 2000. Can item format (multiple choice vs. open-ended) account for gender differences in mathematics achievement? *Sex Roles: A Journal of Research*, 42, 1-21.
- Ben-Chaim, D. and Zoller, U. 1997. Examination-type preferences of secondary school students and their teachers in the science disciplines. *Instructional Science*, 25 (5), 347-367.
- Ben-Shakhar, G. and Sinai, Y. 1991. Gender differences in multiple-choice tests: The role of differential guessing tendencies, *Journal of Educational Measurement*, 28, 23-35.
- Bennett, R.E. (1993). On the meaning of constructed response, in Bennett, R.E. and Ward W.C. (Eds.), *Construction Versus Choice in Cognitive Measurement*, Hillsdale, NJ: Erlbaum.
- Birenbaum, M. 1994. Toward adaptive assessment: The student's angle. *Studies in Educational Evaluation*, Vol. 20, pp. 239-255.
- Birenbaum, M. 1997. Assessment preferences and their relationship to learning strategies and orientations, *Higher Education*, 33, 71-84.
- Birenbaum, M. and Feldman, R. A. 1998. Relationships between learning patterns and attitudes towards two assessment formations, *Educational Research*, 40 (1), 90-97.
- Birenbaum, M. and Gutvirtz, Y. 1995, Feb.. On the relationships between assessment preferences, cognitive style, motivation and learning strategies. Paper presented at the 11th conference of the Israeli Research Association, Jerusalem: The Hebrew University.
- Birenbaum, M. and Tatsuoka, K. 1987. Open ended versus multiple choice response formats—it does make a difference! *Applied Psychological Measurement*, 11, 385-395.
- Birenbaum, M., Tatsuoka, K. and Gutvirtz, Y. 1992. Effects of response format on diagnostic assessment of scholastic achievement. *Applied Psychological Measurement*, 16, 353-363.
- Black, P. and Wiliam, D. 1998. Assessment and classroom learning. *Assessment in education: Principles, policy and practice*, 5 (1), 7-74.
- Casebeer, C. and M. Alquraan. 2011. Low performing Jordanian postsecondary students: What are their beliefs concerning the classroom assessment practices of their instructors? *International Journal of Academic Research in Business and Social Sciences*, Vol. 1, Special Issue (pp. 24 – 41).
- Chapelle, C. 1988. Field independence: A source of language test variance, *Language Testing*, 5, 62-82.
- Coombe, C., Folse, K. and Hubley, N. 2007. *A practical guide to assessing English language learners*. Ann Arbor, MI: The University of Michigan Press.
- Dochy, F., Segers, M. and Sluijsmans, D. 1999. The use of self-, peer and co-assessment in higher education: A review. *Studies in Higher Education*, 24 (3), 331-350.
- Entwistle, N.J. 1991. Approaches to learning and perceptions of the learning environment. Introduction to the special issue, *Higher Education*, 22, 201-204.
- Gellman, E. and Berkowitz, M. 1993. Test-item type: What students prefer and why. *College Student Journal*, 27 (1), 17-26.
- Grandy, J. 1987. Characteristics of examinees who leave questions unanswered on the GRE general test under rights-only scoring. ETS Research Report 87-38, Princeton, NJ: Educational Testing Services.
- Hansen, J. 1984. Field dependence-independence and language testing: Evidence from six Pacific island cultures, *TESOL Quarterly*, 18, 311-324.
- Hansen, J. and Stansfield, C. 1981. The relationship between field dependent-independent cognitive styles and foreign language achievement, *Language Learning*, 31, 349-367.
- Kleinsasser, A., Horsch, E., and Tastad, S. 1993, April). Walking the talk: Moving from a testing culture to an assessment culture. Paper presented at the annual meeting of the American Educational Research Association. Atlanta, GA.
- La Lopa, J. M. 2004. Developing a student-based evaluation tool for authentic assessment. *New Directions in teaching and learning*, 100, 31-36. doi: 10.1002/tl.168,
- Leach, L., Neutze, G. and Zepke, N. 2001. Assessment and empowerment: some critical questions. *Assessment and Evaluation in Higher Education*, 26 (4), 293-305. doi: 10.1080/02602930120063457
- Lu, C. and Suen, H. K. 1993, April). The interaction effect of individual characteristics and assessment format on the result of performance-based assessment. Paper presented at the Annual Meeting of the American Educational Research Association, Atlanta, GA.

- Oller, J.W. and Perkins, K. 1978. Language proficiency as a source of variance in self-reported affective variables. In J. Oller and K. Perkins, editors, *Language in education: Testing the tests*. Rowley, MA: Newbury House.
- Scouller, K. 1998. The influence of assessment method on students' learning approaches: Multiple choice question examination versus assignment essay, *Higher Education*, 4, 453-472.
- Shohamy, E. 1984. Does the testing method make a difference? The case of reading comprehension, *Language Testing*, 1, 147-170.
- Spielberger, C.D. 1972. Conceptual and methodological issues in anxiety research. In Spielberger, C.D. editor, *Anxiety: Current trends in theory and research*, New York: Academic Press.
- Spielberger, C. D. 1980. *Test Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.
- Struyven, K., Duchy, F. and Janssens, S. (2005). Students' perceptions about evaluation and assessment in higher education: A Review. *Assessment and Evaluation in Higher Education*. 30 (4), 325-341.
- Taras, M. 2002. Using assessment for learning and learning from assessment. *Assessment and Evaluation in Higher Education*, 27 (6), 501-510. doi: 10.1080/0260293022000020273
- Traub, R.E. 1993. On the equivalence of the traits assessed by multiple choice and constructed-response tests, in Bennett, R.E. and Ward, W.C. (Eds.), *Construction Versus Choice in Cognitive Measurement*, Hillsdale, NJ: Erlbaum.
- Traub, R.E. and MacRury, K. 1990. Multiple-choice vs. free response in the testing of scholastic achievement, in Ingenkamp, K. and Jager, R.S. (eds.) *Test und tends 8: Jahrbuch der pedagogischen Diagnostik*, Weinheim und Base, BeltzVerlag (pp. 128-159).
- de Watering, Gijbels, D., Duchy, F., and van der Rijt, J. 2008. Students' assessment preferences, perceptions of assessment and their relationships to study results. *Higher Education*, 56, 645-658.
- Van de Watering, G. and van der Rijt, J. (2006). Teachers' and students' perceptions of assessments: A review and a study into the ability and accuracy of estimating the difficult levels of assessment items. *Educational Research Review*, 1 (2), 133-147.
- Zeidner, M. 1987. Essay versus multiple choice type classroom exams: The students' perspective. *Journal of Educational Research*, 80 (6), 352-358.
- Zollner, U. and Ben-Chaim, D. 1988. Interaction between examination-type anxiety state and academic achievement in collect science: An action-oriented re-
search, *Journal of Research in Science Teaching*, 26, 65-77.

Appendix A: Student Profile Form

- Date:.....
- Name:
- Contact Phone Number:
- Course #:
- Student University I.D. Number:
- A. Gender: (circle) 1 = Male 2 = Female
- B. Uni. Year: (circle) 1 = First 2 = Second 3 = Third 4= Fourth

I give my permission to Higher Colleges of Technology Dubai to use my responses to questionnaires, test questions, and interview questions for research purposes. I understand that my name will not be revealed.

Signature of candidate.....

For Tester Use Only:

- C. TAI Scale: -----
- 1 = 20-262 = 27-333 = 34-39
- 4 = 40-465 = 47-536 = 54-60
- 7 = 61-668 = 67-739 = 74-80

Appendix B: Adapted Assessment Preferences Inventory (AAPI)

Part A: To what extent would you want your achievements in English class to be assessed with the following methods?

- N/A = Not Applicable
- 5 = to a great extent
- 4 = to a certain extent
- 3 = unsure
- 2 = to a small extent
- 1 = not at all

Circle the response that best describes your opinion.

Appendix B: Adapted Assessment Preferences Inventory (AAPI)

Part A: To what extent would you want your achievements in English class to be assessed with the following methods?

N/A = Not Applicable

5 = to a great extent

4 = to a certain extent

3 = unsure

2 = to a small extent

1 = not at all

Circle the response that best describes your opinion

1.	N/A	5	4	3	2	1	Traditional tests, with a time limit
2.	N/A	5	4	3	2	1	Traditional tests, without a time limit
3.	N/A	5	4	3	2	1	Written tests without a time limit, while permitting the use of supporting materials
4.	N/A	5	4	3	2	1	Written tests, without a time limit with no supporting materials allowed
5.	N/A	5	4	3	2	1	Individual oral tests
6.	N/A	5	4	3	2	1	Oral tests, in the form of a group discussion where the instructor observes and assesses the contribution of each of the participants
7.	N/A	5	4	3	2	1	Oral tests, in the form of a group discussion where the instructor observes and assesses the group as a whole
8.	N/A	5	4	3	2	1	Take home exams where students are allowed to take the test home and use their class materials to answer the questions
9.	N/A	5	4	3	2	1	Portfolios where students' submit their work either finished or in progress
10.	N/A	5	4	3	2	1	Projects where students work in pairs
11.	N/A	5	4	3	2	1	Projects where students work in groups (3+ students)
12.	N/A	5	4	3	2	1	Projects where students work individually
13.	N/A	5	4	3	2	1	Tests that students take on computer or iPad
14.	N/A	5	4	3	2	1	Tests with multiple choice questions
15.	N/A	5	4	3	2	1	Tests with matching questions
16.	N/A	5	4	3	2	1	Tests with 'True/False/No Information' questions
17.	N/A	5	4	3	2	1	Tests with questions requiring students to write short answers of one or two sentences
18.	N/A	5	4	3	2	1	Tests with questions requiring students to write longer answers like essays
19.	N/A	5	4	3	2	1	Assessments which require students to give oral presentations
20.	N/A	5	4	3	2	1	Assessments which require students to perform a real-life task
21.	N/A	5	4	3	2	1	Simple tasks having only one correct answer
22.	N/A	5	4	3	2	1	More complex tasks which have more than one possible answer
23.	N/A	5	4	3	2	1	Assessments where students reflect or give their opinions
24.	N/A	5	4	3	2	1	Self-assessment where students give themselves a grade
25.	N/A	5	4	3		1	Peer assessment where students in the class grade other students
26.	N/A	5	4	3	2	1	Tests that students develop themselves

Other assessment types (please specify): _____

Part B: To what extent would you like your teacher to?

N/A = Not Applicable

5 = to a great extent

4 = to a certain extent

3 = unsure

2 = to a small extent

1 = not at all

1.	N/A	5	4	3	2	1	Hand out at the beginning of the course, a detailed description of the way you will be assessed.
2.	N/A	5	4	3	2	1	Clarify what will be on the test and how to prepare for it
3.	N/A	5	4	3	2	1	Give out prior to the test examples of the kind of questions to be asked on the test
4.	N/A	5	4	3	2	1	Hand out a list of practice questions, from which the teacher will choose the actual test questions
5.	N/A	5	4	3	2	1	Assess, as part of your course grade, your participation in the class discussions
6.	N/A	5	4	3	2	1	Assess, as part of your course grade, the homework and exercises you've completed

Part C: To what extent would you want?

N/A = Not Applicable

5 = to a great extent

4 = to a certain extent

3 = unsure

2 = to a small extent

1 = not at all

1.	N/A	5	4	3	2	1	The test questions to be presented in order of difficulty (from easy to difficult)
2.	N/A	5	4	3	2	1	The exam questions to be presented in order of the study topics
3.	N/A	5	4	3	2	1	There to be one big test during the semester
4.	N/A	5	4	3	2	1	There to be several smaller quizzes throughout the semester
5.	N/A	5	4	3	2	1	Your achievements to be assessed by a variety of tasks of different types.
6.	N/A	5	4	3	2	1	To be able to choose your preferred type of assessment.

Appendix C: Test Anxiety Inventory (TAI)

DIRECTIONS: A number of statements which people have used to describe themselves are given below. Read each statement and then circle the number to the right of the statement to indicate how you generally feel. There are no right or wrong answers. Do not spend too much time on any one statement but give the answer which seems to describe how you generally feel.

	ALMOST NEVER	SOMETIMES	OFTEN	ALMOST ALWAYS
1. I feel confident and relaxed while taking tests.....	1	2	3	4
2. While taking examinations I have an uneasy, upset feeling	1	2	3	4
3. Thinking about my grade in a course interferes with my work on tests	1	2	3	4
4. I freeze up on important exams	1	2	3	4
5. During exams I find myself thinking about whether I'll ever get through school	1	2	3	4
6. The harder I work at taking a test, the more confused I get	1	2	3	4
7. Thoughts of doing poorly interfere with my concentration on tests	1	2	3	4
8. I feel very jittery when taking an important test	1	2	3	4
9. Even when I'm well prepared for a test, I feel very nervous about it	1	2	3	4
10. I start feeling very uneasy just before getting a test paper back	1	2	3	4
11. During tests I feel very tense	1	2	3	4
12. I wish examinations did not bother me so much	1	2	3	4
13. During important tests I am so tense that my stomach gets upset	1	2	3	4
14. I seem to defeat myself while working on important tests	1	2	3	4
15. I feel very panicky when I take an important test	1	2	3	4
16. I worry a great deal before taking an important examination	1	2	3	4
17. During tests I find myself thinking about the consequences of failing	1	2	3	4
18. I feel my heart beating very fast during important tests	1	2	3	4
19. After an exam is over I try to stop worrying about it, but I just can't	1	2	3	4
20. During examinations I get so nervous that I forget facts I really know	1	2	3	4