

A Study of College EFL Students' Affective Reactions and Attitudes toward Two Types of Performance-based Oral Tests

Ying-Ying Chuang

Assistant Professor, Department of Applied Foreign Languages, Cheng Shiu University

Abstract

Performance assessment can be defined as the direct, systemic observation of an actual learner performance and the rating of that performance according to previously established performance criteria. Many scholars also argue that testing a person's language ability from the actual performance is presumed to be more trustworthy and reliable. Therefore, this study aims to explore Taiwanese college students' affective reactions and their attitudes toward using two different types of oral performance-based test in the English classroom. 102 non-English major students were administered two speaking tests at the midterm and the final: paired role-playing and individual interview. Both quantitative and qualitative data were collected from students' responses in the questionnaire. The results indicate that test anxiety occurs when college students engage in English oral performance tests: they feel more nervous before the test than during the test. In terms of test type, the majority of the students feel comfortable in paired role-playing test setting rather than individual interview; their oral performance however can be somewhat influenced by their partner. Students agree that the speaking tests which they took corresponded to what they learned in class, and the oral testing is necessary in English courses. In addition, many students believe that their performance more or less can be affected by test format/task type. The results also reveal that many students are willing to take oral tests instead of traditional written tests. In other words, more students get to realize how important of being able to use the language for meaningful communication orally. In this study, the author sheds light on using performance-based assessment as a necessary and practical way to enhance students' English speaking ability.

Key words: Performance-based test, speaking tasks, affective reactions, EFL speaking abilities

I. Introduction

The ability to speak in a foreign language is at the very heart of what it means to be able to use a foreign language. The development of oral proficiency has been one of the priorities in the language teaching profession in recent years (Hadley, 2000; Richards, 2008). Ideally, language testing should also cater to this direction in order to promote communicative language use as the major goal of the language learning. Harlow and Caminero's (1990) conducted a study regarding oral testing in 106 language programs of Spanish, French, and German at 67 universities in the United States. However, the results indicated that 43% of these language programs seldom conducted oral tests because of limited time, large classes, and heavy teaching loads. Harlow and Caminero (1990) concluded that oral skills were considered important in most language programs and institutions, even though they could be difficult to test.

In Taiwan, among the various types of classroom assessment, many English teachers believe that oral tests are the most difficult to create, administer, and score due to their large size of classes and too much teaching loadings (Liu, 2006; Teng, 2005). On the other hand, students are also afraid of taking oral tests due to their low confidence (Wang, 2003). For these reasons, some English teachers consider that the oral tests are not practical to conduct in the classroom, especially in the general English courses for non-English major students (Chuang, 2007). Besides, from the EFL learner's perspective, oral test is the most complex and difficult task among the four language skills since the test taker's preparations should include knowledge about the target language and the skills involved in using it (Bygate, 1987).

According to Krashen (1988), language learners' attitudes are also essential and should be considered in learning process. A number of research studies have examined the affective reactions of students to various types of language tests (Savignon, 1972; Madsen, 1982; Scott, 1986). Many studies have indicated that many factors influence learners' affective reactions to the oral tests, including test format, length, time constraints, testing environment, familiarity with test types, perceptions of test validity, and anxiety (Madsen, 1982; Scott & Madsen, 1983; Scott, 1986). However, little empirical research has been done to investigate the college EFL students' affective reactions while they are asked to be tested orally in the classroom (Chuang, 2007). In order to have a better understanding of college non-English major students' affective reactions and attitudes toward different oral test types, the author conducted this study in her English classes through employing two different types of performance-based oral test, individual interview and paired

role-playing, to assess students' speaking abilities and to see what students' attitudes and affective reactions would be.

II. Literature Review

Performance and Assessment

Carroll (1968) described that language performance as "the actual manifestation of linguistic competence ...in behavior" (p. 50). According to Bachman (1990), performance is observed as a result of the test-taker's language knowledge interacting with the characteristics of the test task and other non-linguistic characteristics of the test-takers, such as his or her strategic competence, knowledge of the topic, and personal attributes. In other words, performance assessment thus differs from traditional fixed-response tests (e.g. multiple-choice tests) in the degree to which it represents or simulates behavior in the real world. Therefore, in this study, a speaking test that involves learner's actual speaking performance, as opposed to completing selected response or limited production items, is considered as performance assessment, since the spoken product represents a performance sample of speaking.

In academic situations, teaching and testing are interrelated with each other. According to Heaton (1988), the goal of language testing is as a device to not only motivate and reinforce students' learning, but also to evaluate their performance and progress (p.5). For language teachers themselves, classroom tests can effectively help them to identify students' learning difficulties and weaknesses in order to make adjustment and improvement in their teaching (p.6). In other words, well-constructed speaking tests in classes can provide the opportunities for learners to demonstrate their ability through performing certain tasks in the target language, and the teachers receive feedback immediately through tests result as a valuable teaching device.

The speaking assessment has traditionally been a headache for language teachers. Even many well-established language tests do not include an oral component, such as the English subject test of the College Entrance Exams (CEE) in Taiwan. In addition, most of the major examinations in Taiwan include English language as one subject area to test, such as the Basic Competence Test for Junior High School Students (BCTest) and many educational standardized tests as well as the Taiwanese civil service examinations. They only use pencil-and-paper multiple-choice test form to evaluate test takers' language proficiency. However, the features of spoken language, such as pronunciation and speech fluency, cannot be assessed through pencil-and-paper tests.

According to Chuang's (2007) study, college-level English teachers in Taiwan seldom conduct performance-based tests to assess students' speaking abilities. As a result, having good grade in English courses does not guarantee that the student's English oral proficiency has also achieved a certain level of competency—a good paper-and-pencil test-taker may not be a good language performer. In other words, indirect tests (such as paper-and-pencil and multiple-choice tests) certainly do not involve test-takers in speaking, requesting, responding, interacting, or in combining listening and speaking, or reading and writing.

Regarding language testing, Bachman (1990) stated that communicative testing must include all the components of communicative competence (pragmatic and interpersonal competence). Therefore, the focus in language pedagogy on communication in real-world contexts prompts the design of language assessment to be authenticity—to create more communicative assessment procedures which involve test takers' actual performance for the rater's judgment. In addition, Brown (2007) also pointed out that the principles of language assessment should be practicality (financial limitations, time constraints, ease of administration, and scoring and interpretation), reliability (consistent and dependable), validity (the test actually measures what it is intended to measure), authenticity (the target language task can be enacted in the 'real-world'), and washback (the effects of an assessment on teaching and learning). Without a doubt these principles are constructed in designing appropriate tests in order to ensure making accurate judgments about the competence of the learners.

Performance-based Assessment

Performance assessment can be defined as the direct, systemic observation of an actual learner performance and the rating of that performance according to previously established performance criteria. It has attracted considerable attention in the field of language testing in the past decade (Bachman, 2002; Bonk & Ockey, 2003; McNamara, 1996; Lumley, 2002). As compared to traditional fix-response tests, such as paper-and-pencil and multiple-choice tests, performance assessments are often regarded derived from performance-based tasks of speaking or writing comes closer to performance as it would be in a non-test form situation, and inferences made about a person's language ability from the actual performance are presumed to be more trustworthy and reliable (Bachman & Palmer, 1996).

O'Malley and Valdez-Pierce (1996, p. 5) stated that performance-based assessment includes

the following characteristics:

1. Students make a constructed response.
2. They engage in higher-order thinking, with open-ended tasks.
3. Tasks are meaningful, engaging, and authentic.
4. Tasks call for the integration of language skills.
5. Both process and product are assessed.
6. The depth of a student's mastery is emphasized over breadth.

Performance assessments however bring with its potential variability in tasks and rater judgments, as sources of measurement error. It means that the complexity of language performance assessments inevitably involves a range of factors which may influence the examinee's observed score.

Compared to the traditional language assessments, Alderson and Banerjee (2002) indicated that authentic tasks are more complex and lack of practicality—they are not only difficult to create but also hard to evaluate due to the reliability issues involved. That is, the complexity and multi-facetedness of language performance assessment inevitably introduce a range of factors that may influence the learner's observed scores. In an oral test, for instance, a number of sources of variance other than the speaking ability of the learner may contribute to the variance in test scores. Concluded from many research studies, as Weigle (2004) notes, two main sources of variance in an oral performance assessment are raters and tasks. In sum, testers have recognized that a variety of factors affect test scores of performance-based oral tests. As Purves (1992) notes that some of the related factors may interact with each other in complex ways because “different tasks present different problems, which are treated differently by students and judged differently by raters” (p. 112).

Speaking Tasks: Individual, Pair and Group Tasks

In language-learning contexts, tasks are usually defined in terms of language use. Nunan (1993) defines a communicative task which “involves learners in comprehending, manipulating, producing or interacting in the target language while their attention is principally focused on meaning rather than form” (p. 59). He also mentions the elements that the task designer has to design: input, goals, roles and settings. Bachman and Palmer (1996) have defined language use tasks in very similar terms. In sum, oral tasks can be seen as activities that involve speakers in using language for the

purpose of achieving a particular goal or objective in a particular speaking situation.

One of the key decisions in oral task design is what the learners will be asked to do with language. Regarding oral tasks, Brown and Yule (1983) made distinctions between four different types of informational talk: (1) description, (2) instruction, (3) storytelling, and (4) opinion-expressing/justification. Their main point was that different types of task has its own routines of organizing information for easy comprehension; with practice learners can improve their control of these routines and then enhance their language-use skills. Later, Bygate (1987) made more detailed distinctions between types of oral tasks:

1. Factually oriented talk: includes description, narration, instruction, and comparison.
2. Evaluative talk: includes explanation, justification, prediction, and decision.

Bygate (1987) indicated that learners' use of language is different in each of these categories. For instance, if one learner is good at explaining, it does not mean that he or she is also good at comparing things, telling a story or making a decision. In other words, as Luoma (2004) suggested, it is useful to test the types of talk separately, as they may give different information about the learners' skills.

The most typical way of arranging speaking tests is to evaluate learners one at a time, often in an interview format. Although individual testing is costly in terms of testers' time, it is flexible in that the questions can be adapted to each examinee's performance, and it gives the testers a plenty of control over what happens in the interaction. On the other hand, paired interviews or paired tasks also can be used. However, those settings have many advantages, there are also challenges. For instance, the examinees' talk is almost inevitably influenced by the other participant's personality, communication style and possibly also language level (Luoma, 2004). The concern is that all test takers may not get an equal opportunity to show their speaking skills at their best (Iwashita, 1999). Another problem of paired tasks that testers often feel uncertain about is the amount of responsibility that they give to the examinees if they are not trained in interview techniques.

Like pair work, group interaction tasks are also generally well received by learners (Fulcher, 1996), but perhaps due to administrative concerns about managing the sizes of groups and the mixture of ability levels in them, they are not often used in formal tests of speaking (Reves, 1991). However, in classroom assessment, group discussions, or individual presentations followed by group discussion, can be quite practical. They serve the purpose of practicing speaking and generating learner talk quite well.

When testing the oral proficiency in the language classroom, the teachers may simply

interview the learners and assess their responses, or use other techniques like role-playing, group discussion between learners, monologue, and picture-description and so on. But choosing an appropriate elicitation technique is only part of the problem; there are many other difficulties associated with design, administration and assessment. So serious are these difficulties, in fact, according to Ur (1996) that most language exams either do not include oral testing techniques or give them very low weighting in the final grade.

Teaching and Assessing Speaking Ability

In an EFL writing class, teachers provide forms of the writing texts or sentence formats for different purposes as good models for students. The teachers can easily check the student's writing pieces to see how correct they are by following certain grammar and structure rules. But what is the appropriate form of spoken language to teach? Unlike teaching writing, the teacher in a speaking class should not focus mainly on a student's correctness and completeness of oral production. According to Brown and Yule (1983), while listening to native speakers talking, the EFL learner should realize that "they do not produce ideal strings of complete, perfectly formed, sentences...they use language manipulatively to communicate with and make up what they say as they go along" (p.21).

What should an English speaking class prepare the learners to do? In general, the aim is to enable students using the target language as social functions, to express their ideas, to handle basic interactive skills, as well as to present their needs, such as making requests, showing directions, and expressing opinions. For those purposes, English teachers might design a construction syllabus in which students first develop simple conversation skills, and then build on those skills in order to further achieve more complex skills (Brown & Yule, 1983). In other words, in English class, EFL students can improve gradually, rather than simply learn another topic or set of things to say. Moreover, the teacher should employ a set of strategies to help students' performance progress, especially for those students who have difficulties in language learning—to find out their problems and give them helpful strategies to practice the language—such as building up their confidence in speaking without anxiety.

Brown and Nation (1997) suggested three elements that language learners must be exposed to in the EFL speaking classes: (1) form-focused instruction; (2) meaning-focused instruction; and (3) opportunities to improve fluency. Form-focused instruction emphasizes the details of grammar,

vocabulary, and pronunciation, and the use of the spoken language appropriately. Form-focused speaking is an effective way to enhance the learning of the beginning learners by helping them practice in repetition-drills for some simple, memorized phrases or sentences. The teacher must provide various interesting activities for student repetitive practice to motivate their learning. In addition, the teacher must provide opportunities and encouragement for learners through meaning-focused instruction to help them both produce and listen to meaningful oral communication. Developing fluency in speaking skills is an aim of many language learners; therefore, the teachers must provide lots of conversational practices and fluency activities for the learners to gain truly fluent use of the target language. Swain's (1985) study showed the importance of comprehensible output because it could provide opportunities for contextualization and meaningful use; however, it is unfortunately "generally missing in typical classroom settings" (p. 252).

Learner Affective Reactions to Oral Tests

Anxiety, one of the emotional reactions of students to tests, has been the objects of research studies. A variety of factors in addition to anxiety can affect a learner's performance on an exam. Galassi, Frierson, and Siegel (1984) conducted a study and found little connection between test anxiety and student performance on an achievement exam. They indicated that students concern about their performance, the testing environment, and time constraints. As Schwarzer, Ploeg, and Speilberg (1982) suggested that "test anxiety may play an indirect role in performance, depending on the amount of time available for competing an examination and the degree to which an individual has prepared for the task at hand" (p. 6).

The effect of time constraints on performance is clearly illustrated in Hill's (1983) study. He found that under time pressure, students identified as high test anxious made three times as many errors and took twice as long as low anxious students. However, when time limits were removed, the high anxious students performed as well as their low anxious peers and completed the test in approximately the same amount of time.

Some research studies in language testing have also compared student affective reactions to different test formats. Jones, Madsen, and Brown (1980) examined the affective reaction of students to seven different types of language tests. The results in a factor analysis of their data showed that one of these factors was identified as emotive reactions concerning how pleasant the experience of

taking the tests was, how easy the test were, how frustrating they seemed, how well the student felt they performed, and how well the students liked the tests.

Savignon (1972) reported that reactions to oral tests of communicative competence examined in her study were extremely positive, even though the tests were very difficult for students. In addition, Savignon concluded that although most students felt completely unprepared for the testing experience, they regarded the tests as actually testing those foreign language skills they were supposed to be learning.

Based on the studies discussed above, certain factors do exist to influence EFL students' affective reactions to the oral tests. Therefore, the aim of the present study tried to examine Taiwanese college students' affective reactions to the two types of oral tests, and their general attitudes toward the two types of performance-based assessments which were employed in their English classroom.

III. The Study

Participants

The data analyzed in this study was collected in 2007 during the spring semester. The participants of this study were college sophomore students of Cheng Shiu University who took the course of English Listening and Speaking as their sophomore general requirement course. Those participants were all non-English major students who were recognized as the EFL learners of low-intermediate level to intermediate level in their English language proficiency (they were placed in 'group A' from the college English placement test at the beginning of the semester).

Procedure

The selected students were administered oral tests at the midterm and the final exams in spring semester. In the midterm exam, students were interviewed individually in English as a teacher-student interaction—the teacher interviewed each student several questions, and then the student responded to the teacher. In the final exam, students were tested as paired role-playing. Both the midterm and the final exams, all students were graded on their fluency, comprehension, pronunciation, grammatical accuracy and vocabulary using a simplified scale based on the one used

in the Interagency Language Roundtable test (formerly Foreign Service Institute or FSI test). The scales, weighted equally, were then totaled for an overall score. The teacher introduced and explained this rating scale to the students at the beginning of the course; therefore, the students were assumed that they understood what components of English language abilities would be assessed during the oral tests. The content of the both two oral performance tests came from the textbook and teaching materials of this course which students were just taught at this semester.

In the final exam, the teacher divided the 102 students into 51 pairs according to their student ID numbers before the oral test. It means that students knew who would be their partner just right before the test. Students were assigned as the speaker A and the speaker B in each group, and each student was given a cue card to do a controlled role-playing as the test task. Through cued dialogues of the cue card, students in pairs perform as the following:

Speaker A	Speaker B
(You meet speaker B outside the bank down the street.)	(You meet speaker A outside the bank down the street.)
A: Greet speaker B.	A: _____.
B: _____.	B: Greet speaker A.
A: Ask speaker B where he is going and what he is going to do.	A: _____.
B: _____.	B: Say you are going to open a new bank account.
A: Suggest some ideas which are good for speaker A regarding the bank account.	A: _____.
B: _____.	B: Agree with speaker A's suggestion and show appreciation.
A: Suggest somewhere to go together for a lunch later.	A: _____.
B: _____.	B: Reject speaker A's suggestion of the food/place for a lunch. State a reason why you don't like the food/place. Make a different suggestion.
A: Accept speaker B's suggestion.	A: _____.
B: _____.	B: Express pleasure.

To analyze the data of this study, all the participants were asked to fill out an affective-reactions questionnaire (please see Appendix) which was based on the one developed by Jones, Madsen, and Brown (1980) to evaluate students' affective variables and attitudes toward second language tests. Their several studies have insisted that speaking in the target language is the most anxiety-provoking aspect of learning the language for learners (Madsen, Brown, & Jones, 1991). In order to examine if Taiwanese college non-English major students hold the same attitudes to confirm their studies, the author decided to employ their questionnaire as the instrument for the study. However, the author only selected the items which were with relation to speaking from Jones, Madsen, and Brown's questionnaire, and the author also modified those items in order to fit in with the purposes of the present study. The author also categorized those revised items into four sections for further data analysis process: students' test anxiety, the suitability and equity of the oral tests, test formats, and students' satisfaction with the two tests and their won performance as well.

The questionnaire was translated into Chinese language by the author, and then administrated in class after the final exam. Students were asked to mark all 16 items on a five-point Likert-style scale, while item 11 to 13 also included the space for students to write down their explanations regarding their ratings (in Chinese). Both English and Chinese versions of questionnaire were checked by two English teachers and two senior college students in advance to see if any items were confusing or unclear in order to make sure that the instrument of the study could be as valid and effective as possible.

Data Analysis

In this study, means and standard deviations were computed for all 16 items on the affective-reactions questionnaire. Frequency and percentile were also used as descriptive statistics to compare Taiwanese college non-English major students' attitudes toward the two types of oral performance-based tests as the quantitative data. Finally, students' written explanations and opinions on the particular questionnaire items as the qualitative data for the ratings would also be discussed later.

IV. Results and Discussion

The purpose of the present study was to investigate Taiwanese college non-English major students' affective reactions and attitudes toward two different performance-based oral test formats: individual interview and paired role-playing. 102 college sophomore students of non-English major participated in this study. They were tested by individual English oral interview in their mid-term exam and paired role-playing as their final exam. In addition, they were required to fill out a questionnaire regarding the oral tests at the end of the semester. Among them, there were 47 male (46.1%) and 55 female students (53.9%).

Quantitative Results

Item 1 to 4 of the questionnaire in this study is chiefly concerned with students' anxiety which takes place in language testing. The results are shown in Table 1. Item 1 regards whether or not students feel nervous before the oral test, and the result shows that more than half (51.9%) of the total students chose 'agree' while only 14.7% chose 'disagree.' Item 2 regards whether students feel nervous during the test, and 44.2% said yes while 23.5% said no. Therefore, in this case, it seems like that these students feel more nervous before the test than during the test according to their experiences.

Table 1

Frequency and Percentage of Students' Responses toward Test Anxiety (N = 102)

Item	SD	GD	N	GA	SA	M	SD
1. I felt nervous before the test.	4 3.9%	11 10.8%	24 23.5%	43 42.2%	20 9.7%	3.62	1.04
2. I felt nervous during the test.	5 4.9%	19 18.6%	33 32.4%	38 37.3%	7 6.9%	3.23	.99
3. I feel comfortable when I take oral test with another student.	6 5.9%	14 13.7%	41 40.2%	33 32.4%	8 7.8%	3.23	.98
4. I felt my performance was somewhat influenced by my partner.	2 2.0%	9 8.8%	25 24.5%	46 45.1%	20 19.6%	3.72	.95

Note. SD = Strongly Disagree; GD = Generally Disagree; N = Neutral; GA = Generally Agree; SA = Strongly Agree; M= Mean; SD = Standard Deviation

In addition, the result of item 3 shows that 40.2% of students agreed that they feel more comfortable when they are not tested individually in oral test, while only 20% disagreed. Item 4 regards if the students thought the partner was one of the factors which influenced their performance in the paired role-playing oral test. Surprisingly, almost two-third of the students (64.7%) agreed that their oral performance was affected by their partners in the language test situation.

Item 5, 6, 7, 8, and 9 of the affective questionnaire are concerned with students' feelings about the suitability and equity of the oral tests in English class. The results are shown in Table 2. Item 5 would like to see if students knew what they were supposed to do in their mid-term and final oral exams (what and how they would be tested in general). The result indicates that 41.1% of the students showed agreement while one-fourth (24.5%) still expressed that they did not realize what or how they would be tested in the test. In item 6, nearly two-third of the students (63.7%) agreed that these two oral tests (interview and role-playing) were related to what they had been taught in their English class, while only 5.9% disagreed. The result of item 7 shows very large proportion (80.4%) of the students who believe that oral performance should be tested in the English language course. In other words, those students agree that speaking is an important language skill, and it is necessary for the learners to actually speak out and then their oral production can be evaluated.

Table 2

Frequency and Percentage of Students' Responses Regarding the Suitability and Equity of the Oral Test (N = 102)

Item	SD	GD	N	GA	SA	M	SD
5. I understood what I was supposed to do during the oral test.	1 1.0%	24 23.5%	35 34.3%	34 33.3%	8 7.8%	3.23	.94
6. I thought the test corresponded to what I learned in class.	1 1.0%	5 4.9%	31 30.4%	47 46.1%	18 17.6%	3.74	.84
7. I believe oral tests are necessary in English courses.	1 1.0%	2 2%	17 16.7%	45 44.1%	37 36.3%	4.13	.83
8. I believe these two oral tests were accurate evaluations of my ability in English speaking.	4 3.9%	35 34.3%	36 35.3%	25 24.5%	2 2.0%	2.86	.90
9. I believe I had adequate opportunities in demonstrating my ability to speak English.	4 3.9%	37 36.3%	24 23.5%	29 28.4%	8 7.8%	3.00	1.06

Note. SD = Strongly Disagree; GD = Generally Disagree; N = Neutral; GA = Generally Agree; SA = Strongly Agree; M = Mean; SD = Standard Deviation

Unfortunately, in item 8 and 9, only one-fourth of the students (26.5%) thought these two oral tests could assess their true English language speaking ability accurately, and only 36.2% agreed that they could demonstrate their English ability from these tests. On the other hand, 38.2% in item 8 and 40.2% in item 9 showed their disagreement that the two oral tests which were managed in this study could let them distinctly and authentically demonstrate their speaking ability in English. It means that many students believe their English language oral proficiency level cannot be judged only through these two types of the tests—individual interview and paired role-playing.

The major concern of item 10 to 13 is the test format. Item 10 wonders whether students think the test format and type will affect their scores. As the table 3 shown, 61.8% of the students agree that their scores would be different if they were tested by individual instead of paired-format. In other words, those students believe that the scores which the teacher marked according to their performance during the test reflect their English language oral proficiency level; their performance however was somewhat affected by the test format and test type—individual or with another student. Only 9.8% of the students think their scores would not be different either they were tested in

individual or pair situations. The result of item 11 shows an interesting phenomenon: the majority of the students (41.2%) said that they felt more comfortable when they took the oral test with another classmate than with the teacher, while 33.3% showed that they would rather interact with the teacher in oral test than with their classmates.

Table 3

Frequency and Percentage of Students' Responses regarding Test Format (N = 102)

Item	SD	GD	N	GA	SA	M	SD
10. If I took the oral performance test individually, I would get a different score.	1 1.0%	9 8.8%	29 28.4%	47 46.1%	16 15.7%	3.67	.88
11. I felt more comfortable when I take an oral test with another student than with the teacher.	3 2.9%	31 30.4%	26 25.5%	35 34.3%	7 6.9%	3.11	1.01
12. I would rather take a written test than an oral test.	11 10.8%	27 26.5%	34 33.3%	19 18.6%	11 10.8%	2.96	1.13
13. I thought the oral performance test was too difficult.	2 2.0%	26 25.5%	35 34.3%	36 35.3%	3 2.9%	2.90	.93

Note. SD = Strongly Disagree; GD = Generally Disagree; N = Neutral; GA = Generally Agree; SA = Strongly Agree; M= Mean; SD = Standard Deviation

To compare oral-type with written-type tests in English class (item 12), more than one-third of the students (37.3%) said they would prefer the oral test to the written test. Another one-third remained neutral while 29.4% still chose taking written test as their preference. It seems like that if students can choose the test type to evaluate their learning outcomes, many of them would like to take an oral test as a new way to assess their oral performance rather than to be tested through the traditional paper-pencil test in language class. The result from the item 13 regards test difficulty; however, the majority of the students (38.2%) still expressed that oral test in English was too difficult for them, while only 27.5% disagreed. Language teachers may wonder why those students have such reactions occurred; therefore, the reasons and opinions which students stated in their questionnaires will be summarized at the end of this session.

Finally, the last three questions in the questionnaire are concerned with students' satisfactions with the two oral exams they experienced shown in Table 4. The result of item 14 indicates that

while one-third of the students enjoyed taking part in the oral test of the role-playing style and nearly one-fourth of the students did not seem to like this test type, there were still a large part of students (41.2%) who showed their neutral opinion.

Table 4

Frequency and Percentage of Students' Responses about Satisfactions with Oral Test (N=102)

Item	SD	GD	N	GA	SA	M	SD
14. In general, I liked to take part in role-playing test.	6 5.9%	20 19.7%	42 41.2%	30 29.4%	4 3.9%	3.05	.91
15. In general, I think I did well and felt satisfied with my performance in tests.	7 6.9%	35 34.3%	38 37.3%	18 17.6%	4 3.9%	2.77	.95
16. To compare with myself, I believe I will definitely perform better in my next t speaking test.	1 1.0%	1 1.0%	34 33.3%	44 43.1%	22 21.6%	3.83	.81

Note. SD = Strongly Disagree; GD = Generally Disagree; N = Neutral; GA = Generally Agree; SA = Strongly Agree; M= Mean; SD = Standard Deviation

However, item 15 examined if students were satisfied with their oral performance in these two spoken-type tests, the majority of the students (41.2%) did not think that they did well in the tests, and they were not satisfied with their oral performance in general. Only 21.5% of the students showed positive reaction that they felt satisfied with their achievement. The last question wonders if students have self-confidence that they will perform better than they had done in the oral tests. Nearly two-third (64.7%) showed their positive view that they think they will definitely do better in their future oral tests while only 2% of the students did not think so. It means that having the experience of oral testing in this semester, most of the students are sure that they will improve to have better performance in their future oral test opportunities.

Qualitative Results

The following section includes some abstracts which were collected and translated into English from the students' responses in the questionnaire item 11, 12, and 13 (most of them were written in Chinese).

Item 11: *I felt more comfortable when I take an oral exam with another student than with the teacher. Please explain why.*

“I agree, especially with my classmate I know him or her well. We can discuss what we are going to say during the test, or simulate the situation before the test. It is helpful to reduce the test anxiety a lot.”

“Testing with my classmate is more informal...with TEACHER, who dare says that he or she is not afraid of?”

“I like to interact with my teacher in oral test...she helps me out or tries to elicit me to speak more during the test. I felt at ease when I was interviewed by my teacher.”

“I like to take the oral test with my classmate rather than my teacher. I know my classmate better, and our conversation was going very smooth in the test.”

“I was too nervous to speak out when I looked at my teacher during the test.”

Item 12: *I would rather take a written test than an oral test. Please explain why.*

“Taking written test seems more secure for me...memorizing the vocabulary, grammatical rules, sentence structure...not too difficult...but I know I will totally go nuts when I have to be tested orally.”

“Passing written test doesn't mean you can speak perfect English. I prefer having a fluent English speaking ability, so I think oral test is necessary in English class.”

“Paper-pencil test is boring...I hope English class can provide me more opportunity to speak out in English, or have real conversation with foreigners.”

“I feel more nervous if I have to take oral test in English...but I think developing speaking skills is important in language learning.”

“I wish I can speak beautiful English...oral test is more practical than written test.”

Item 13: *I thought the oral performance test was too difficult. Please explain why.*

“I did prepare a lot before the test, but I totally forgot what I have to say during the test. Oral test is too difficult for me mentally...too nervous.”

“I knew the spelling, I know the grammar, I knew what teacher saying and asking...but I didn't know why teacher seemed not understand what I was answering...Is that my problem?”

“I don't know if my pronunciation is correct or not, and I worry about putting all words together meaningfully; sometimes I cannot find the right words to use, so I am afraid that the sentences I

speak are so broke...”

“I only have limited vocabulary—it’s hard to express my idea orally in a limited test time. I even hardly use complete sentence...too bad.”

“I hate oral test at all. I just couldn’t find the right words to use...I always broke down again and again during the oral test. I didn’t look straight to my teacher or my partner...why I became ‘big tongue’ during the test?”

“Oral test is much easier than written test for me. I don’t have to worry my spelling and grammar errors (the teacher seems focus on my idea and not care too much on my grammar use), and I just say what I like to say. More oral practice makes me feel confident in speaking English.”

V. Conclusion and Pedagogical Implication

Oral skills cannot be ignored in the EFL classroom since people often judge a person’s language knowledge/competence and proficiency level based on how well his or her speaking performance in that target language. In other words, a learner’s end product of language learning is to be capable to speak the target language fluently (Chuang, 2007). Thus, the present study was conducted to look at Taiwanese college students’ affective reactions and attitudes toward using oral performance-based tests to evaluate their learning outcome.

The results of this study indicate that test anxiety occurs when college students engage in the oral performance-based tests: they feel more nervous before the test than during the test. Therefore, the result did confirm Madsen, Brown, & Jones’ (1991) studies that speaking in the target language is the most anxiety provoking aspect of learning the language for learners. In addition, this study reveals that the majority of the students feel more comfortable when they take oral test with a partner; their performance however can be somewhat influenced by their partner. In these two oral test experiences, students agreed that the tests corresponded to what they learned in class, and the oral performance tests are necessary in English courses. Nevertheless, many students thought their English abilities cannot be measured only through individual interview and paired role-playing tasks. Also, many students believe that their performance more or less can be affected by test format/task type, and most of them said they prefer interact with another classmate rather than with the teacher in oral test setting.

In addition, the results also reveal that college students are willing to take the oral test rather than the traditional written test in the classroom. From this viewpoint, it seems that college students

get to realize how important that language learners should be able to use the language for meaningful communication orally. However, many students still expressed that English oral test was difficult for them. Finally, the results show that although most of the students did not satisfied with their performances in these two test experiences, they still enjoyed taking part in oral tests and positively believed that they are able to perform better in their future speaking opportunities.

Some pedagogical implication extended from this study will be discussed here. Oral tests should be conducted in the language classroom to evaluate language learners' learning outcome and improvement through their actual performance. If English teachers like to employ oral tests to assess students' oral abilities, they should let students become familiar with those test types in order to avoid test anxiety. It means that different types of test should be practiced as speaking tasks and activities to elicit students' speaking in class. What are the reasons for getting students to speak in the English classroom? Firstly, speaking tasks which teachers employ in the classroom can provide *rehearsal* opportunities for students to practice real-life conversation in the safe environment. Secondly, speaking activities in which students try to use the target language they have learned provide feedback for both teachers and students. Thirdly, the more students have chance to *activate* the carious elements of the target language they have stored in their brains, the more automotive their use of these elements become. This means that students will be able to "use words and phrases fluently without very much conscious though" (Harmer, 2007, p. 123). Therefore, classroom activities that develop learners' ability to express themselves through speech would therefore seem an important component of a language course.

Regarding language assessments in the classroom, the author suggests that direct test items are more appropriate than indirect test items. In other words, in direct test items, the teacher can ask students to *use* language to do something, instead of just testing their knowledge of how the language itself works. In addition, when designing direct test items for students in language class, the tasks should be those which students are familiar with (they have been practicing in their lessons.) Thus, they can demonstrate properly how well they can use the language. Finally, teachers need to bear in mind that oral testing should be reserved for summative evaluation purposes in the language classroom (Hadley, 2000). In terms of students' anxiety, the author suggests that the more students practice orally in class with their classmates or their teacher, the less anxiety and stress they have, especially when they are familiar with the test type/task in advance, their anxiety can be minimized.

In this study, the author sheds light on using oral performance-based assessment as a

necessary and practical way to enhance students' English learning, especially speaking abilities. Even though students get nervous feelings when they have to be tested orally, they give positive views on oral tests and believe that speaking skills are important. Teachers should keep in mind that task/test design is essential because it makes the construct assessed in the test more tangible and because, for the learners, it provides the context for the talk and guidelines about what they must *do* with the language. Accordingly, to choose or design the performance-based tasks in language tests, teachers must consider their purpose of the test: to make sure the tasks can really elicit oral performances from the students, as well as can provide *positive washback* (the benefit that tests offer to learning) to the learners.

Reference

- Alderson, J. C., & Banerjee, J. (2002). Language testing and assessment (Part 2). *Language Teaching*, 35, 79-113.
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Bachman, L. F. (2002). Some reflections on task-based language performance assessment. *Language Testing*, 19, 453-476
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- Brown, D. (2007). *Teaching by principles: An interactive approach to language pedagogy*. New York: Person Education, Inc.
- Brown, R. S., & Nation, P. (1997). *Teaching speaking: Suggestions for the classroom*. Retrieved January 31, 2008, from <http://www.jalt-publications.org/tlt/files/97/jan/speaking.html>
- Brown, G., & Yule, G. (1983). *Teaching the spoken language*. Cambridge: Cambridge University Press.
- Bonk, W. J. & Ockey, G. J. (2003). A Many-facet Rasch analysis of the second language group oral discussion task. *Language Testing*, 20(1), 89-110.
- Bygate, M. (1987). *Speaking*. Oxford: Oxford University Press.
- Carroll, J. B. (1968). The psychology of language testing. In A. Davies, (ed). *Language Testing Symposium: A Psycholinguistic Perspective*. London: Oxford University Press.

- Chuang, Y. Y. (2007). A study of college English teachers' preferences and perceptions of using the types of tasks for assessing Taiwanese EFL students' foreign language oral proficiency. *Journal of Applied Foreign Languages*, 8, 40-77.
- Fulcher, G. (1996). Does thick description lead to smart tests? A data-based approach to rating scale construction. *Language Testing*, 13(2), 208 - 238.
- Galassi, J. P., Frierson, H. T., & Siegel, R. G. (1984). Cognitions, test anxiety, and test performance: A closer look. *Journal of Consulting and Clinical Psychology*, 52, 319-320.
- Hadley, A. O. (2000). *Teaching language in context*. Boston: Heinle & Heinle Publishers.
- Harlow, L., & Caminero, R. (1990). Oral testing of beginning language students at large universities: Is it worth the trouble? *Foreign Language Annals*, 23(4): 489-501.
- Harmer, J. (2007). *How to teach English*. England: Pearson Education Limited.
- Heaton, J. B. (1988). *Writing English language tests*. New York: Longman.
- Hill, K. T. (1983). Interfering effects of test anxiety on test performance: A growing educational problem and solutions to it. *Illinois School Research and Development*, 20(1), 8-19.
- Iwashita, N. (1999). Negative feedback and positive evidence in task-based interaction. *SSLA*, 25, 1-36.
- Jones, R. L., Madsen, H. S., & Brown, B. L. (1980). *Evaluation of affective variables in second language testing*. Paper presented at the 14th Annual TESOL Convention, San Francisco, CA.
- Krashen, S. D. (1988). *Second language acquisition and second language learning*. Prentice-Hall International.
- Liu, Y. L. & Wen, S. M. (2007, April). *Rating reliability on the assessment of speaking performance*. Paper presented in the International Conference on English Education, Taipei.
- Lumley, T. (2002). Assessment criteria in a large-scale writing test: What do they really mean to raters? *Language Testing*, 19, 246-276.
- Luoma, S. (2004). *Assessing speaking*. Cambridge: Cambridge University Press.
- Madsen, H. S. (1982). Determining the debilitating impact of test anxiety. *Language Learning*, 32, 133-143.
- Madsen, H. S., Brown, B. L., & Jones, R. L. (1991). Evaluating student attitudes toward second language tests. In E. K. Horwitz & D. J. Young (Eds.), *Language anxiety: From theory and research to classroom implications* (pp. 65-86). Englewood Cliffs, NJ: Prentice Hall.

- McNamara, T. F. (1996). *Measuring second language performance*. London: Longman.
- Nunan, D. (1993). Task-based syllabus design: Selecting, grading, and sequencing tasks. In G. Crookes and S. Gass (Eds.), *Tasks in a Pedagogical Context: Integrating theory and practice* (pp. 55-68). Clevedon: Multilingual Matters.
- O'Malley, J.M. & Valdez Pierce, L. (1996). *Authentic Assessment for English Language Learners*. Reading, MA: Addison-Wesley.
- Purves, A.C. (1992). Reflections on research and assessment in written composition. *Research in the Teaching of English*, 26(1), 108-122.
- Reves, T. (1991). From testing research to educational policy: A comprehensive test of oral proficiency. In J. C. Alderson and B. North (Eds.), *Language Testing in the 1990s* (pp. 178-188). London: Modern English Publications and the British Council.
- Richards, J. C. (2008). *Teaching listening and speaking: from theory to practice*. Cambridge: Cambridge University Press.
- Scott, M. L. (1986). Student affective reactions to oral language tests. *Language Testing*, 3(1), 99-118.
- Scott, M. L. & Madsen, H. S. (1983). The influence of retesting on test Affect. In J. W. Oller (Ed.), *Issues in Language Testing Research*, Rowley: Newbury House Publishers, Inc.
- Savignon, S. J. (1972). *Communicative competence: An experiment in foreign language teaching*. Philadelphia, Pennsylvania: The Center for Curriculum Development Inc.
- Schwarzer, R., Ploeg, H. M., & Spielberg, C. D. (1982). *Test anxiety: An overview of theory and research*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. M. Gass, & C. G. Madden (Eds.), *Input in second language acquisition* (pp. 235-253). Rowley: Newbury House Publishers, Inc.
- Teng, K. H. (2005). *Perceptions of Taiwanese students to English learning as functions of self-efficacy, motivation, learning activities and self-directed learning*. Unpublished doctoral dissertation, the University of Idaho.
- Ur, P. (1996). *A course in language teaching: Practice and theory*. Cambridge: Cambridge University Press.
- Wang, Y. C. (2003). *Communication-orientation in Freshman English curriculum: A New response*. In the Papers from the 12th Conference on English Teaching and Learning in the Public of China (pp. 588-598). Taipei: Crane Publishing Co., Ltd.

Weigle, S.C. (2004). Integrating reading and writing in a competency test for non-native speakers of English. *Assessing Writing*, 9(1), 27-55.

Submission Date : 2008/05/12

Revision Date : 2009/03/26

Acceptance Date : 2009/09/11

Appendix:
Questionnaire

This questionnaire is part of a research study being carried out by an English instructor from Cheng Shiu University. The results will be used to examine using oral performance tests in English courses. For this reason it is very important that your responses be as frank as possible. Your answers are confidential and will not interfere in any way with your evaluation.

Instructions:

Please mark with an “X” the description on the agreement/disagreement scale which best expresses your opinion. Note that in some questions you are asked to explain your response.

1 = Strongly Disagree

2 = Generally Disagree

3 = Neutral

4 = Generally Agree

5 = Strongly Agree

You are male female

You consider your English speaking ability is excellent fair acceptable poor

1	I felt nervous before the test.	1 2 3 4 5
2	I felt nervous during the test.	1 2 3 4 5
3	I feel comfortable when I take an oral test with another student.	1 2 3 4 5
4	I felt my performance was somewhat influenced by my partner.	1 2 3 4 5
5	I understood what I was supposed to do during the oral test.	1 2 3 4 5
6	I thought the test corresponded to what I learned in class.	1 2 3 4 5
7	I believe that oral tests are necessary in English courses.	1 2 3 4 5
8	I believe these two oral tests were accurate evaluations of my ability in English speaking.	1 2 3 4 5
9	I believe I had adequate opportunities in demonstrating my ability to speak English.	1 2 3 4 5
10	If I took the oral performance test individually, I would get a different score.	1 2 3 4 5
11	I felt more comfortable when I take an oral exam with another student than with the teacher. Please explain why: _____	1 2 3 4 5
12	I would rather take a written test than an oral test. Please explain why: _____	1 2 3 4 5
13	I thought the oral performance test was too difficult. Please explain why: _____	1 2 3 4 5
14	In general, I liked to take part in role-playing oral test.	1 2 3 4 5
15	In general, I think I did well and felt satisfied with my oral performance in these two spoken-type tests.	1 2 3 4 5
16	To compare with myself, I believe that I will definitely perform better in my next speaking test or opportunity.	1 2 3 4 5

Thanks for your cooperation.

大學生對於兩種英語口說實作評量測驗的看法及態度之研究

莊嫻纓

正修科技大學應用外語系助理教授

摘 要

實作評量是一種新的評量方式：它取代傳統紙筆測驗題型式的考試，而直接以測量學生實際的表現為評量標準。實作評量可用在教學活動中來達成總體性評量的目的，比傳統評量形式來的生動、客觀、有效。在外語教學中，許多學者也認為，直接從學習者的語言表現來實際評量其語言能力與程度，是比較可可靠且有意義的。因此，本研究的主要目的是探究臺灣的大學學生，對於在英語課堂中使用口說實作評量測驗的情感反應與看法。本研究的對象為大二學生，作者於其通識英語課程期中考時，實施雙人角色扮演口語對話，及期末考實施個人與老師一對一面談對話。研究中之質量數據，皆取至於學生的問卷結果。本研究結果顯示，在情感反應上，學生認為比起雙人角色扮演口語對話，個人與老師一對一面談對話是比較容易緊張的，並且認為他們的口語表現會受到其對話同伴的影響。學生也贊同口語考試內容與課程是相關的，而學生的口語表現或多或少會受到測驗題型與形式的影響。研究結果也證明，大學學生認為口說實作評量在英語課堂中是必要的：傳統紙筆測驗是無法準確評量學生說英語的能力與技巧的。並且他們也瞭解：能用英語交談來傳遞有意義的信息，是語言學習的最終目的。作者希望藉由此研究，來說明及證明運用口說實作評量在大學生英語課堂上的重要性與實用性，最後也對大學英語課堂中的教學，提出實施英語口說評量上的建議。

關鍵字：語言實作評量、口說測驗題型；情感反應；英語口說能力