Is Communicative Language Teaching Being Tested Communicatively? An Analysis of English Tests in Oman

Maya Rashed Hamdan Al Mamari, Abdo Mohamed Al-Mekhlafi & Thuwayba Ahmed Al-Barwani

1 Ministry of Education, Sultanate of Oman
2 Sultan Qaboos University, Sultanate of Oman

Correspondence: Abdo Mohamed Al-Mekhlafi, Sultan Qaboos University, Sultanate of Oman. E-mail: Abdoali914@gmail.com

Received: August 9, 2018   Accepted: October 10, 2018   Online Published: October 12, 2018

Abstract
This paper investigates the degree of communicativeness of English final tests in Oman. It is a descriptive, content-based analysis of Five Final tests from 2011 to 2016. A framework and a checklist were developed for purposes of analysis. The framework was based on Bachman and Palmer's model of language ability (2010) and grade 10 English test specifications.

Research findings reveal: (i) A discrepancy between test specifications of English final tests and the actual content of the tests (ii) That communicative competence was not fully addressed in the English final tests (iii) That the input used in English final tests lacks authenticity (iv) The majority of the test components used suitable language in a limited context with limited instructions given for some components. There was also a lack of constructed responses (v) There was no integration between two skills or elements on the test, except for a random integration between vocabulary and reading.

Keywords: communicative testing, EFL, communicative competence, test authenticity, skill integration

1. Introduction
The idea of using language as a tool of communication originated from a theory developed by Hymes (1972) known as communicative competence, which emphasizes the importance of both language competencies and the ability to use them (Enache, 2005; Ireland, 2000; S. K. Kitao & K. Kitao, 1996; Razmjoo, 2011; Xin, 2007).

Communicative competence has since been elaborated and different models have been proposed. The two most influential models developed by Canale and Swain (1980) and Bachman (1990), extended communicative competence to include grammatical, social and psychological competences. Nevertheless, all models share the same basic idea of communicative competence, which is the ability to use appropriate language in real-life situations (Enache, 2005; S. K. Kitao & K. Kitao, 1996; Xin, 2007).

Subsequently these new views of language teaching have led to test development which reflects this view. Nunan (2009) indicates that the connection between teaching and testing is a fundamental principle of curriculum design. He also argued that communicative language teaching needs communicative language testing in order to achieve the intended communicative language outcomes. In terms of supporting the necessity of the match between teaching and testing, it is worth mentioning the relationship between comprehensible input and comprehensible output. According to Krashen’s comprehensible input hypothesis (1985), people acquire language if they receive enough comprehensible input beyond their current level of competence. At the same time, Swain’s comprehensible output hypothesis (1985) agrees with Krashen but adds that the development of learners’ communicative competence does not depend on comprehensible input alone. Instead, learners’ output also plays a significant role, as learners need to be pushed to produce the language. Additionally, Hatch and Wanger-Gough (1976) highlighted the interconnectedness between input and output, as they argued that if anybody wants to judge learners’ output, they should look back at the language input that the learners had to work on (cited in Liming, 1990). Accordingly, this interrelated relationship between input and output implies the necessity of having a match between teaching, which supplies the comprehensible input, and testing, which generates the comprehensible output. Therefore, it is expected that the assessment used in the communicative
teaching approach will be designed to evaluate the communicative skills of the language that the students have been taught.

Communicative language testing provides information about students’ ability to perform in the target language in context-specific tasks. It focuses on both learners’ knowledge of the language, how they use it and to what extent they can apply their knowledge in communicative situations (Bakhsh, 2016; Baseer & Alvi, 2014; Enache, 2005; Gopal, 2014; Harding, 2014; S. K. Kitao & K. Kitao, 1996; Miyata-Boddy & Langham, 2000; Razmjoo, 2011). When testing productive skills, more emphasis is given to appropriateness rather than the formation of correct grammatical sentences and when testing receptive skills, greater emphasis is placed on understanding the intention of the speaker or writer rather than getting specific details correct (Baseer & Alvi, 2014).

Davies (1988) and S. K. Kito and K. Kito (1996) suggested talking about communicative testing as a continuum of which there is “more of” or “less of”. Davies (1988) indicated that communicative tests are “more integrative and less discrete point; more direct and less indirect; more criterion referenced and less norm-referenced” (p. 6). Morrow (1979) suggested seven features that any test should have in order to be called communicative. These include (a) interaction-based: the language user should take the receiver's requirements into account when delivering a message; (b) unpredictability: the test should contain unknown input; (c) context: language users should use the appropriate language for the context; (d) purpose: language users should choose the language that serves their objectives; (e) performance: the test should indicate whether learners can perform a set of authentic tasks; (f) authenticity: input should not be simplified for learners; and (g) behaviour-based: the test should measure what learners can achieve through language.

Similarly, Brown (2005) highlighted five requirements when designing any communicative test, including meaningful communication, authentic situations, unpredictable language input, creative language output, and integrated language skills. Furthermore, Bachman (1990) highlighted four features: (a) an information gap, in which learners are required to complete a task using different sources of input; (b) task dependency, which means that tasks are sequenced in a way that the information gathered from one task is used to answer subtasks; (c) integration of tasks; and (d) measuring different language abilities, including not only knowledge of grammar and vocabulary, but also knowledge of cohesion, functions and sociolinguistic aspects. Other researchers have stated that communicative tests should measure different language abilities, such as pragmatic, sociolinguistic and strategic competences, along with linguistic competence (Bachman & Palmer, 1996, 2010; Canale & Swain, 1980; Celce-Murcia et al., 1995; Hymes, 1972; Widdowson, 1978). For the purpose of this study, a modified Bachmann and Palmer’s framework (2010) was used in the analysis process. This modified framework consists of the following elements:

Knowledge of phonology, which is part of organizational, grammatical knowledge, and is concerned with pronunciation, word stress, and intonation patterns of lexical items in connected speech, and spelling conventions;

Conversational structure, which is a part of the knowledge of conversational organization, and concerns the rules used when partaking in a conversation, and which differs in different cultures, or languages;

Some aspects of pragmatic functional knowledge including:

Knowledge of manipulative functions, which includes instrumental functions employed in order to motivate other people to behave, or not behave, in a certain manner, such as suggestions, regulatory functions such as laws that are used to control other’s actions, and interpersonal functions that are used to establish interpersonal relationships, such as apologies;

Knowledge of imaginative functions, which enables students to use language to create an imaginary world, or to extend the world around them for humorous purposes, such as jokes and poetry;

Some aspects of pragmatic sociolinguistic knowledge, including:

Knowledge of dialects/varieties, which demonstrate the way in which different social or age groups use language, and how a setting can influence text and utterances. It also refers to the awareness of regional ways of using language;

Knowledge of natural or idiomatic expression, which is concerned with unnatural expressions that are usually grammatically correct, but are not the expressions that members of a specific community would use;

Knowledge of cultural references and figures of speech, in which cultural reference can be interpreted literally, but is intended to be understood for its extended meaning and figures of speech, such as metaphors and similes;
Some aspects of metacognitive strategies including:

- Goal setting, which involves deciding what one is going to do, such as selecting one or more tasks from a set of possible tasks;
- Planning, involves deciding how to use what one has, such as formulating one or more plans for implementing elements in a response to an assessment task.

1.1 Communicative Language Teaching in Oman

Two decades ago, EFL teaching in Oman adopted the communicative language teaching as its main approach. Accordingly, textbooks and classroom activities reflected this approach. In addition to the clear emphasis on communicative language teaching in the Omani curriculum, a series of classroom observations confirmed that teachers in Oman followed the communicative approach in their classroom practice. Richards and Rogers (1986) checklist was used for the purposes of these observations. Findings showed 94% of the communicative features were met during the lessons. All teachers used English as the medium of instruction, with no use of L1. They integrated all the skills, presented language in context (for example stories, role plays or interviews) and they created a positive atmosphere using competitions and incentives. They also organized students into groups, each of which had a leader to facilitate classroom situations and promote discussion and group and pair work. Teachers used dialogues and texts that focused on communicative functions and appropriate meaningful language use from the textbook or extra materials prepared by the teachers. They also tolerated students’ errors and encouraged them to check their own or one another’s work. The general focus of the lessons was on meaning, and most of the teachers tended to teach grammar implicitly by providing examples from the reading texts and stories to help students infer the rules.

This communicative teaching approach in the English curriculum in Oman is supposed to be associated with communicative testing to achieve the intended objectives. With regard to the testing system in Oman, the Ministry of Education uses continuous assessment together with two final tests at the end of each of the two semesters in the academic year. In 2004/2005 the Tests and Examinations Administration Department (TEAD) introduced a new form of continuous assessment, which formed part of the examination process. To this end, teachers are encouraged to use different tools such as presentations, written work, projects, portfolios, independent reading, questioning in the classroom and day-to-day observations. However, it was not clear whether the tests were designed communicatively.

Despite the communicative teaching methodologies adopted in the Government schools in Oman, students continue to graduate with limited language ability and therefore are unable to succeed in their academic pursuits or compete for jobs in the private sector. In 2011/2012, after analysis of test results, the Ministry of Education indicated that English was one of the subjects in which students from all over the Sultanate achieved the lowest percentage score (Al-Shukri, 2012). These poor communication skills and low performance in English tests despite the use of the communicative teaching approaches suggests the need to pay more attention to English language tests in order to assess their degree of alignment with the communicative teaching that is taking place in the classrooms.

1.2 Research Purpose

The main purpose of this paper, therefore, is to examine the degree of communicativeness of Omani English tests by addressing the following research question: To what extent do the Omani English tests exhibit the features of communicative testing?

2. Method

2.1 Research Design

This study is a descriptive research using a qualitative content analysis research design. Hsieh and Shannon (2005) defined this type of design as being “a research method for the subjective interpretation of the content of text data through the systematic classification process of coding and identifying themes or patterns” (p. 1278). Similarly, Patton (2002 cited in Hsieh & Shannon, 2005) defined this form of research as “any qualitative data reduction and sense-making effort that takes a volume of qualitative material and attempts to identify core consistencies and meanings” (p. 453).

2.2 Instruments

A checklist and a framework were developed to analyze the language tests.
2.3 Population and Sample
The sample of the study consisted of the second semester grade ten final tests from the previous five years - 2011/2012 to 2015/2016.

2.4 Instrument Validation
The framework and checklist were validated by a jury of 11 reviewers in order to assess their clarity, accuracy, and relevance. Minor changes on the framework and the checklist were suggested by the reviewers including modifying the format of the framework, using consistent language and altering the order of some of the items. Categories of the checklist were also altered, based on the reviewers’ comments.

Additionally, the reviewers recommended certain modifications to the authenticity checklist, including providing a brief description of the checklist, and its aim; altering the formation of some phrases; simplifying the analysis key and deleting certain items that were either repeated, or irrelevant.

2.5 Procedures
Zhang and Wildemuth (2005) provided eight steps for conducting qualitative content analysis research, that were adhered to in this study, and these include preparing the data, defining the unit of analysis, developing categories and coding schemes, testing the coding scheme on a sample, coding all the tests, assessing the coding consistency, drawing conclusions from the coded data, and reporting the findings.

2.6 Preparing the Data
Five final grade 10 English tests from the second semester were chosen for analysis. Grade 10 was selected since the teaching approach adhered to at this grade is communicative teaching, which should be associated with communicative testing procedures. Additionally, the final tests were chosen for analysis because they are the only formal tests prepared by the Ministry of Education for all grade 10 students in Oman. The final grade 10 English tests consist of twelve different questions, including three on listening, two on vocabulary, two on grammar, three on reading, and two on writing. Students are expected to answer all questions in two and a half hours. Table 1 provides further details on the specifications of the Omani grade 10 English tests.

Table 1. Test specifications of the final Grade 10 English Test

<table>
<thead>
<tr>
<th>Question</th>
<th>Type</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Listening 1</td>
<td>Multiple Choice (Dialogue)</td>
<td>6 marks</td>
</tr>
<tr>
<td>Listening 2</td>
<td>Wh-Qs (Informative Text)</td>
<td>5 marks</td>
</tr>
<tr>
<td>Listening 3</td>
<td>Matching (Short Texts w/Words)</td>
<td>4 marks</td>
</tr>
<tr>
<td>Vocabulary 1</td>
<td>Gap-fill (Words provided) (Text)</td>
<td>2.5 marks</td>
</tr>
<tr>
<td>Vocabulary 2</td>
<td>Word Completion (Sentences)</td>
<td>2.5 marks</td>
</tr>
<tr>
<td>Grammar 1</td>
<td>Multiple Choice (Sentences)</td>
<td>2.5 marks</td>
</tr>
<tr>
<td>Grammar 2</td>
<td>Gap-fill (No words provided) (Paragraph)</td>
<td>2.5 marks</td>
</tr>
<tr>
<td>Reading 1</td>
<td>Matching (Texts w/Texts)</td>
<td>5 marks</td>
</tr>
<tr>
<td>Reading 2</td>
<td>Multiple Choice (Evaluative Text)</td>
<td>6 marks</td>
</tr>
<tr>
<td>Reading 3</td>
<td>Wh-Qs (Short Answers) AND Multiple Choice (Narrative Text)</td>
<td>9 marks</td>
</tr>
<tr>
<td>Writing 1</td>
<td><strong>EITHER</strong>: Information Points (Paragraph)</td>
<td>5 marks</td>
</tr>
<tr>
<td></td>
<td><strong>OR</strong>: Picture Cue (Describe and Comment)</td>
<td></td>
</tr>
<tr>
<td>Writing 2</td>
<td>Task Instructions (Letter/email)</td>
<td>10 Marks</td>
</tr>
</tbody>
</table>

MOE (2010).

2.7 Defining the Unit of Analysis
Since the final grade 10 English tests are divided into different questions, the test questions were selected as the units of analysis. For the purpose of this analysis the term ‘component’ was employed instead of ‘question’.
2.8 Developing Categories and Coding Schemes

Zhang and Wildemuth (2005) indicated that categories of analysis can be derived from three sources, including the data, previous related studies, and theories. Thus, the categories of analysis used in the self-analysis checklist investigating the communicative competence in the tests were the types of knowledge discussed in Bachman and Palmer’s (2010) model of language ability. The coding scheme used was primarily deductive as it was based on a specific model and related studies.

2.9 Testing the Coding Scheme on a Sample

Inter-coder reliability was used to determine the reliability of the test analysis checklist. Three independent coders were trained and familiarized with the analysis of the checklists. These coders included one regional supervisor involved in writing tests, one evaluation specialist from the Specialized Teachers’ Training Centre in Oman, and an English teacher who taught grade 10. The coders were encouraged to engage in discussions of the checklists until they achieved a complete understanding of the criteria of analysis. The coders were then trained to use the checklists to categorize a sample of final grade 10 tests from the second session of 2015/2016. Following this, the coders separately analyzed one of the five tests to be used in this study, which was that of the first session of 2015/2016. Coefficients of agreement were then calculated using the following formula:

\[
\text{Coefficient of agreement} = \frac{\text{Number of agreement times}}{\text{Number of agreement times} + \text{Number of differences}}
\]

Consequently, three coefficients were found for the checklist. The overall coefficient of agreement was found to be .93.

2.10 Assessing the Coding Consistency

In order to determine intra-coder reliability, the same tests were analyzed and recorded within a three week interval. The correlation between the analyses and the recordings were computed using the same procedures used to obtain inter-coder reliability. The degree of consistency was found to be .91.

3. Findings and Discussions

3.1 Findings

When the communicative features were used to analyze the tests, results revealed that all communicative features are not exhibited by the tests analyzed for this study. Table 2 shows the frequencies and percentages of occurrence of the three main features in the Omani tests.

<table>
<thead>
<tr>
<th>Communicative features</th>
<th>Frequency</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Communicative competence</td>
<td>130</td>
<td>21.7</td>
</tr>
<tr>
<td>Authenticity</td>
<td>65</td>
<td>23.8</td>
</tr>
<tr>
<td>Integration of skills</td>
<td>3</td>
<td>10</td>
</tr>
</tbody>
</table>

A closer look at these features reveals that there is a relatively higher use of the authentic features compared to the other two features. The table also shows that integration of skills is the least exhibited feature in the analyzed tests.

Further analysis of the tests revealed a discrepancy between the grade 10 test specifications, and the actual content of the tests. An investigation of the types of knowledge present in the actual tests based on an adapted framework revealed that the most emphasized types of knowledge in the tests were organizational knowledge, especially grammatical knowledge and ideational knowledge, specifically the students’ ability to understand the language used to inform. In addition, other types of knowledge were covered by the tests, but were not mentioned in the framework, such as spelling conventions, lexical chains related to content schemata, manipulative, instrumental functions, and appraising.

Additionally, using clear and legible hand-writing and correct punctuation conventions, together with other functions such as predicting, comparing, and evaluating, and some aspects of strategic knowledge, including evaluating, editing, and improving drafts of texts, were mentioned in the framework, but were not covered by the tests. Finally, communicative competence was not addressed in its totality within the tests, including aspects of
pronunciation, word stress and intonation patterns of lexical items, conversational structure, regulatory and interpersonal functions, knowledge of imaginative functions, knowledge of dialects/varieties, knowledge of natural or idiomatic expression, knowledge of cultural references and figures of speech, goal setting, and planning.

Moreover, it was found that the input used in the final grade 10 English tests was mainly inauthentic, as the sources of the input were not cited, and the language used was direct and simple. Different topics related to students’ lives were discussed, including topics relating to their academic lives, and their everyday lives, together with universal topics, such as medicine, films, and the environment. It was also noted that there was a limited number of topics that discussed other countries and the Omani context. The language employed was mostly suitable for the students’ level. Some components used a limited context, while other components did not provide students with sufficient instructions to answer the items. Most of the components required selective and limited responses, except for the writing component that required constructed, extended responses.

Finally, there was no integration between two skills or elements, except for the random integration between vocabulary and reading in a small number of items. The study also found that there was no integration between more than two skills, or elements.

3.2 Discussion

Many aspects of Bachman and Palmer’s (2010) model of language ability were absent. This supports the findings of previous studies, including that undertaken by Nguyen and Le (2013), who analyzed the content of 10 tests for Grade Six students at five different schools in Vietnam; Kharrant (2013), who analyzed the content of 1418 and 1419 tests administered to ESP students at King Khalid University in Saudi Arabia; Razmjoo (2011), who analyzed two final tests for Grade Three students at public and private high schools in Iran; Bernardo (2011), who investigated the communicativeness of the 22 English language tests of 22 different instructors from 22 colleges.

All of these researchers found that the tests they analyzed did not address communicative competence holistically although Bachman and Palmer (1996; 2010) stated that a consideration of language ability in its totality must be present when developing, or using, any language assessment. Additionally, Maraheel (2004) stated that pragmatics competence “enable(s) language learners to identify their problems and their contribution of each component of language competency to the totality of communicative competence” (ibid., p. 40).

With regard to authenticity, many researchers have stressed the importance of using authentic input (Brosnan, Brown & Hood, 1984; Caroll, 1980; Flucher, 2000; Al Kubaidi, 2009; McKay, 2006; Tomlinson, 2003; Wang, 2004). This current study found that most of the materials used in the final English tests were not authentic which concurs with Bernardo’s (2011) findings that the reading texts in the 22 English language tests that he analyzed from 22 colleges and universities in the Philippines were also not authentic.

One of the main concerns preventing educators and test designers from using authentic materials is their difficulty in terms of language level. This was indicated by Thomas (2014), who summarized other researchers’ views that some of the challenges in using authentic materials are that their language may be too difficult, the vocabulary too specialized, and the grammar structures too complex. Therefore, such texts would not be suitable for the level of the students, especially for beginners, and would therefore raise their level of anxiety. However, Widdowson (1978) argued that the process of simplifying vocabulary and syntax could actually complicate the message of a text.

Another factor that may hinder the use of authentic materials is the time required for preparation and sourcing of appropriate texts. Thomas (2014) suggested organizing the useable materials into portfolios for easy access; she also stressed the importance of practice in order to facilitate the faster sourcing and preparation of the materials. Very few studies have investigated the authenticity of textbook topics (Siegel, 2014), while no studies have investigated the authenticity of the topics that must be included in the tests, with the exception of Nguyen and Le (2013), who analyzed the content of 10 tests from five schools in Vietnam, and found that topics such as family, school, and community were included. McKay (2004) highlighted the importance of including topics relating to local culture, as this enables students to explain their culture to other people, and to communicate their ideas. Some test designers tend not to include topics relating to other cultures, as such topics may seem difficult and uninteresting, and students may be confused by the information in the text, while the teachers themselves may not possess enough background information concerning the topic (ibid.). McKay (ibid.) discussed the importance of including information about international target culture, since it will be necessary for non-native speakers of English to interact with native, and non-native, speakers of English in the context of cultural encounters relating to trade, tourism, and social contact. She also assumed that the key to this is the appropriate selection of the
cultural information included in the tests.

Shimada (1997) indicated that a sentence in isolation is often meaningless from the communicative perspective. Similarly, Thrasher (2000 cited in Bernardo, 2011, p. 38) rejected the use of isolated sentences as such items do not simulate the way in which students will use the language in real life, they may result in negative washback, as they give students the incorrect impression that language consists of independent parts that can be learned in isolation; and they are unnatural due to reduced context.

No particular study investigating the effectiveness of clear instructions was located, however Bachman and Palmer (1996; 2010) indicated the importance of using explicit and comprehensive instructions, as test designers need to base their interpretations on students’ responses, not on their understanding of the instructions. The instructions in the final English tests were brief and clear, since most of the required responses were either selected or limited. However, the instructions in the writing test, which required constructed responses, should have been more detailed and explicit. Similar to the results of this current study, Bernardo (2011) also found that the tests he analyzed did not provide clear instructions in terms of how the answers would be assessed.

Razmjoo (2011) and Bernardo (2011) found that the tests in their studies were quantitative, rather than qualitative. This current study also found that almost all of the questions in the tests analyzed required selective, short answer responses, with the exception of the writing questions that required constructed, extended responses. The use of multiple choice and short answer tests is not recommended by many researchers, since they are not communicative (Bernardo, 2011; Gopal, 2014; Ireland, 2000; Razmjoo, 2011; Wang, 2004). Al Kubaidi (2009) argued that such test formats are not derived from authentic interaction, nor are they a representation of real-life situations. Additionally, they involve a high degree of guesswork and ambiguity. In contrast, using constructed responses will simulate students’ creativity and engagement. Researchers attributed the use of traditional test formats, to their objectivity in scoring, and their coverage of a wide range of construct, while constructed responses require subjective scoring, and involve undefined construct (Al Kubaidi, 2009; Ireland, 2000).

With regard to skill integration, many researchers have highlighted the importance of skills integration in communicative testing (Bachman, 1991; Brown, 2005; Hinkel, 2010; Mendelsohn, 1989; Widdowson, 1978). This current study found that the final grade 10 English tests assessed skills separately. This finding concurred with the findings of Nguyen and Le (2013), Razmjoo (2011), and Bernardo (2011) that the tests they analyzed constituted discrete points tests, with no skills integration. When skills are assessed separately, they result in lengthy tests that may cause students to lose concentration and may also be a source of agitation and stress.

References


Centre.


**Copyrights**

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/).