Investigating Content and Face Validity of English Language Placement Test Designed by Colleges of Applied Sciences

Sharifa S. A. Al-Adawi[1] & Aaisha A. K. Al-Balushi[1]

[1] Rustaq College of Applied Sciences, Sultanate of Oman

Correspondence: Sharifa S. A. Al-Adawi, Rustaq College of Applied Sciences, Sultanate of Oman. E-mail: sharifa.rus@cas.edu.om

Abstract

An English placement test (PT) is an essential component of any foundation program. It helps place students into their suitable language proficiency level so that they do not spend time learning materials below or above their levels. It also helps teachers to prepare teaching materials to students of similar levels (Brown, 2004; Illinois, 2012). This paper aims to investigate the extent to which the English Language PT used at Colleges of Applied Sciences (CAS) is achieving this goal by exploring teachers' and students' perceptions of the current exam via questionnaires and interviews. Furthermore, it examines the format and content of the PT and students' PT score against their mid-term score. It was found that face validity of CAS PT ranged from low (teachers) to moderate (students). The majority of teachers and students emphasized the importance of including the listening and speaking components in the test. Moreover, a modified version of the reading section of the test needs to be incorporated into the test. Another suggested using a valid certified computerized test. For future research, it is recommended to design a new test, taking into consideration the findings of this research and pilot it to test its effectiveness. Furthermore, an analytical test of the current marking criteria is essential to check its clarity and consistency.

Keywords: placement test, language proficiency, test specification, writing rubric, construct & exam format

1. Introduction

Placement tests stream students into their suitable levels according to their performance in the test which is usually administered at the beginning of a university study or even before certain subjects or courses (Heaton, 1988; Brown, 2004). PTs vary in terms of the components incorporated from reading and writing sections to include speaking and listening and sometimes mathematics and computer skills (Heaton, 1988). At CAS, PT for foundation program is divided into three different PTs: English proficiency, mathematics and computer skills (CAS, 2011). It has been frequently noted by teachers at CAS that some students do not seem to fit into the level in which they have been placed. Their actual levels are either below or above the assigned levels which create unequal opportunities for these students in the classroom leaving them at disadvantage as teachers tend to pay more attention to the level of the majority of students. Boredom or frustration might be logical consequences in such cases (Brown, 2004). Despite the importance of PT in reducing the number of students facing challenge or likability of failing academic degrees because of poor language and study skills (Fulcher, 1997), it could be counterproductive. In other words, having invalid, unreliable test that misplaces students in wrong levels may create negative attitudes towards university study or even destroy the students' proficiency altogether no matter how poor it is. Taking this into consideration, and from the general observation of the effects of such tests, the idea for this study emerged. It came as a consequence resulting from having to experience some teachers' and students' frustrations regarding some students' levels in the classroom which creates challenges for both teachers and students. This research aims to 1) explore teachers' and students' perceptions of the current English PT at CAS by conducting a survey (questionnaires) and interviews. It also aims to 2) investigate English PT validity at CAS by comparing PT results and Midterm results of the same students and the number of failures in midterm due to deficiencies in PT. Finally, it aims to 3) suggest some changes for the sake of improving the current PT (Appendix 1) based on the results from 1 and 2 aims.

2. Literature Review

In the academic year 2008-2009, Colleges of Applied Sciences (CAS) started implementing the new National

English Programs. National English Foundation Program (NEFP) was one of these programs that began at CAS in Fall 2008. As it is with every plan for change, there was a challenge; the implementation of the new English Foundation Program encountered a number of challenges. One of the major challenges was finding a reliable criterion to classify entering students in the most accurate level that suits their present language skills and abilities. English Proficiency Placement Tests have been designed to help achieve this aim; however, since the implementation of the new NGFP placement tests, some students have repeatedly complained that these placement tests have misplaced them. Few students believed that their level is higher than their classmates in the same level group; others believed that they should be in a lower level as they have experienced some problems getting along with the rest of the group they have been classified to join. Some teachers, on the other hand, have also repeatedly criticized these tests on the grounds that they have encountered some problems of discrepancy of students in their groups and have questioned the readiness of some students for the level they are enrolled in.

About 80% of secondary school graduates who were enrolled in Higher Education start with Foundation Program (FP) before embarking on academic study (Al-Lamki, 1998). Foundation Program (FP) aims at providing students with English language proficiency along with other skills like numeracy skills required for university academic study (OAAA, 2009). In Colleges of Applied Sciences (CAS), FP results place students into four levels for English language courses which are levels A (equivalent to intermediate), B (equivalent to pre-intermediate), C (equivalent to elementary) and D (equivalent to beginner) (CAS, 2011). Thus, in order to place students accordingly in the FP, all enrolled students sit for a placement test in the first week of the academic year. The main purpose of the test is to measure students' proficiency level of English language and place them accordingly. Based on the placement tests scores, students are sorted into five groups as it is shown in Table 1.

Table 1. Results of English Placement Test (CAS, 2011)

Placement Test Score (out of 70)	Level the students will be placed in	Courses the students will study in FP
0-19	(D) Beginner (equivalent of IELTS 2.5 or below)	ENGL 3001
20-34	(C) Elementary (equivalent of IELTS 3.0)	ENGL 4001
35-49	(B) Pre-intermediate (IELTS 3.5)	ENGL 5001 & 5002
50-64	(A) Intermediate (IELTS 4.0)	ENGL 6001 & 6002
* 65-70	(eligiable for CH Test) Intermediate (IELTS 4.5)	Sit for English Challenge Test. Students who pass the Challenge Test, (and have also satisfied the requirements for Math and Computer Skills), will enter Year 1. Those who fail will enter level A (ENGL 6001 & 6002).

* Students who get (65-70) are eligible to do English Challenge Test (CH).

Using placement test to place students into appropriate beginning courses has been widely encouraged in the literature over the past years. There are many researches that have emphasized the importance of placement test in helping students succeed in their studies later by placing them in the suitable courses that match their needs (Casazza & Silverman, 1996; Maxwell, 1997; McCabe, 2000). Therefore, many colleges across United States of America require a placement test for their students before entering the colleges (Gerlaugh, Thompson, Boylan & Davis, 2007). Similarly, in Oman, all public and private universities and colleges require students to take a placement test.

There is no doubt with respect to the importance of placement tests; however, there is a question with regard to the reliability and validity of those tests. Most placement tests offer numeric scores and each institution relates certain numeric score to a specific level. For example, Table 1 shows the classification of placement test scores and the

equivalent levels in CAS. Students are instantly placed in a level based on their scores in the placement test. This raises a concern regarding PT validity, does it test what it is supposed to test (Brown, 2004). In other words, does it measure students' actual level of English language proficiency? Literature on placement tests expresses similar concern. Arkoudis (2011) reported that; "The sector's blind faith in language testing inhibits the development of more robust ways of addressing English language outcomes for graduates" (p. 33).

According to Belfield and Crosta (2012), most placement tests are commonly used in a binary capacity. In other words, if a student gets a specific score, it means he should be placed in the level equivalent to that score. However, this practice is problematic and unjustified for three main reasons. To begin with, statistical errors regularly happen in marking those tests. That is to say, some errors might occur in counting which lead to misplacing students in wrong levels. Moreover, most of such tests consist of multiple-choice questions type. Tests with multiple-choice questions type increase the availability of luck element (Heaton, 1988; Brown, 2004). For instance, a student who has no idea about the answer for a question might randomly tick the correct answer and get a high score. As a result, this student might be placed in a high level based on the score despite his actual level of English language proficiency. Likewise, the PT used in CAS consists of 20 multiple-choice questions that weight 20 marks; 10 questions assess grammar and the other 10 assess reading, in addition to two writing tasks that weigh 50 marks. Thus, it can be argued that the availability of luck element in CAS placement test might be high. As a result, students might be misplaced in levels that do not represent students' actual level of English language proficiency. Moreover, the possibility of students cheating by copying answers from their friends in such tests increases (Heaton, 1988). Guessing and cheating may increase because of the format of the test and due to inefficiency of invigilators. As a consequence, students might get high scores and be misplaced in a level that does not match their actual level of English language proficiency.

Furthermore, excluding speaking and listening components from the exam despite the fact that students will be assessed in both skills in their midterm and final exams might have a certain influence in the scores acquired by the students in PT as different students may vary in terms of areas of strengths and weaknesses based on learning styles and strategies. For instance, Versant English placement test, a 50 minutes computer-based test that aims to evaluate students comprehension of the language on everyday topics, integrate all fours skills and its tasks vary from reading aloud to dictation (Versant guide, 2011). Another example is English Placement Test at the University of Illinois which focuses on students performing tasks such as responding to an oral interview, reading an article and listening to a mini lecture on the same topic in order to produce a well-written and organized essay (Illinois, 2012). Integration of the four skills is more likely to produce more accurate and valid results of students' proficiency level taking into account that language learning does not occur only through reading or writing. There are students who learn better through listening or speaking the language which corresponds to the theory of multiple intelligences and variety of learning styles (Rao, 2014). Consequently, Morante (1989), Scott-Clayton (2012) and others have strongly advised to not base the judgment of students' level on a placement test score especially if it does not integrate all language skills. They have recommended the practitioners to use multiple methods in addition to the obtained score to make decision regarding students' actual levels of English language proficiency.

This clearly raises questions on the validity of the English placement tests currently used at the Colleges of Applied Sciences. In other words, does it test what it is supposed to test (Brown, 2004)? Thus, the present paper primarily investigates this problem. Research papers such as Fulcher (1997), Research Notes (2005) Shin and Kin (2006) explored different principles and factors in relation to PT results in University of Surrey, University system of Gorgia and Columbia University, respectively. Wall, Clapham and Alderson (1994) conducted a study at the University of Lancaster that investigated the validity and reliability of PT offered at University of Surrey for all newly admitted students (Fulcher, 1997). They investigated the different types of validity which are face (perceptions of students), content, construct and concurrent validity. The results indicated a challenge in finding 'appropriate external criteria to conduct a concurrent study'. Fulcher (1997) followed their steps in investigating face validity and criteria for scoring which are both important principles of assessment. Fulcher's study results of face validity indicated that the majority of the 71 students who responded to the questionnaire thought that the test was fair enough, however, some expressed their concern regarding the ambiguity of some test items. Surprisingly, four students requested longer tests while few others wanted subject-specific content. Moreover, there was a general concern regarding the exam room environment which does not support reliability of results. A study conducted at the same University by Research Note (2005) stated that there is 'a relationship between placement test scores and grades in college-level courses to predict how students in a learning support program would have performed without remediation' (p. 1). To explain it further, students who were placed in learning support program, foundation program in our case, performed better and scored higher marks than those passing the PT with a cutoff score. Shin and Kin (2006), however, investigated some principles of assessment in relation to one particular

reading and writing PT. They investigated the reliability and validity of the design, evaluation and results analysis of the test. Empirically, they found positive results in terms of reliability and validity of the test which consist of two reading tasks (worth 12 marks and should be completed in 20 minutes) and a descriptive writing task (20 minutes). Generalization of the study results could not be accomplished due to the limitation of the study such as small number of participants and the likability of interference of practicality issues. All three mentioned studies are related to the scope of this paper in some way or the other.

Research concerning placement test at CAS level is limited and unpublished. In Spite of the number of meetings organized to improve assessment methods like placement tests, ideas remain at the meeting venue and seem restricted to that place due to the fact that CAS PT is an in-house designed exam. However, frequent observation of displaced students in higher or lower levels that do not suit their general abilities, urged the need for research in the area. And although it is widely-acknowledged that all tests might have a degree of error, there are some tests that are more successful than others in accurately placing students in their appropriate level. Therefore, this paper principally aims to investigate content and face validity of the English proficiency placement tests currently used at CAS (Appendix 1) by suggesting improvements regarding test construction. Basically, it attempts to answer the following two questions:

1).What are Rustaq-CAS teachers' and students' perceptions of the current English placement test? (Face validity)

2). Does English language placement test used in CAS, to identify students' proficiency level, accurately measure students' true language proficiency level? (Content Validity)

2. Methodology

There are a number of methods to assess how well these placement tests work at their intended task. The paper will investigate face validity of the test; teachers' and students' perceptions of the English placement test at the Colleges of Applied Sciences. Additionally, the paper achieves this by analyzing the test used and investigating how representative it is of what students' need or going to encounter in future study (Appendix 1). As far as the test is concerned, the paper aims to study the test's content, components, and format and assessment criteria. Regarding the test validity, the research will compare students' placing scores against their midterm scores.

2.1 Participants and Context

This study was conducted at Rustaq College of Applied Sciences for the foundation level in March 2014 to investigate the validity of CAS English placement test. Following Wall, Clapham and Alderson (1994), one hundred and sixty-one (70 level A, 51 level B and 40 level C) foundation students answered the questionnaire (Appendix 2) that aimed to have a clear picture of students' perception of the PT. Questionnaires were selected as they are efficient in terms of gathering a large quantity of data saving time and money (Sapsford, 2007). Moreover, 10 Rus-CAS teachers were interviewed (semi-structured interview, Appendix 3) to explore their opinion of the current placement test format and content and whether they had encountered a case where they believe PT was not valid enough to place students in the right proficiency language level and their reaction to such situations. Another detailed face to face interview (Appendix 3) was conducted (recorded) with a PT researcher to investigate the format and content of the PT. Interviews work well with a limited number of participants to gain thorough answers and elaborate on interesting ones, however, they could be time consuming while conducting and analyzing data (Sapsford, 2007). Furthermore, similar to Research Notes (2005) approach to testing Surrey PT validity, students' PT results were compared to their MT scores to test PT results validity. All participant students were selected randomly from Rustaq-CAS for easier access to the data.

2.2 Procedures

This study was conducted eight weeks into the semester which means there was an eight weeks period of instruction in language learning after the PT results were published and students placed in different levels accordingly. Questionnaires were created through the Google forms and piloted on about 10 students and then links were emailed to students and to teachers to send it to their students as shown in (Appendix 2). As soon as students send their answers, Google forms create and classify results in charts as presented in the results parts, therefore, it was a quick efficient program. Semi-structured interviews with teachers (Appendix 3), however, were recorded through note taking. The interview with a PT researcher was audio recorded for better access to data. Interview questions were piloted for clarity and insights on 2 teachers. Results from students' questionnaires are presented in charts while teachers' responses are described in details.

3. Results and Data Analysis

3.1 Students' and Teachers' Perception of CAS PT (Face Validity)

3.1.1 Students' Questionnaires

Results from the questionnaire regarding students' perception of current English proficiency PT at Rus- CAS were as follows:
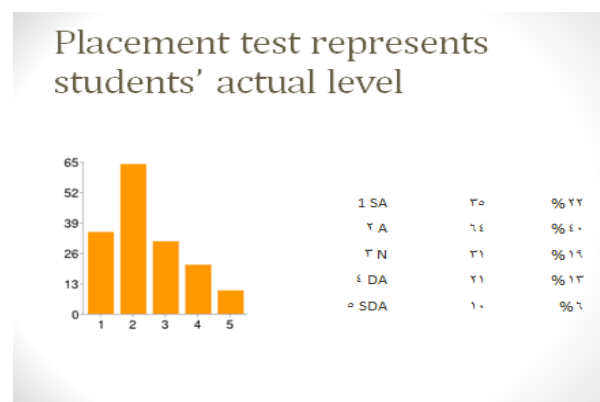
Graph 1. PT represents students' actual level

Graph 2. There is a need for L & S sections in PT

1). Statement one (Placement test represents our actual level) results, as shown in Graph 1, have revealed that 62% of participants agreed that the current PT placed them in the right level. However, (19%) were neutral and (19%) disagreed. T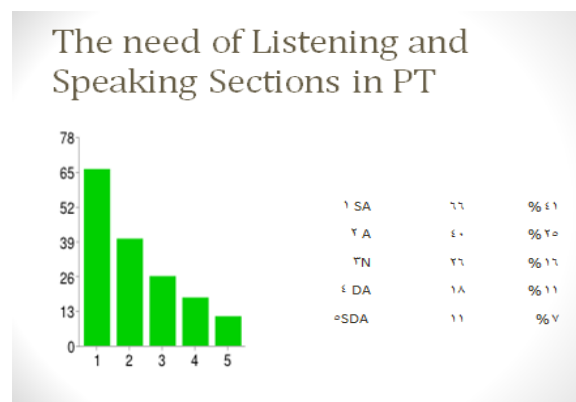his indicates that less than a quarter of the participants actually believe that they have been misplaced and that their actual level should be lower or higher than what their PT score indicated. This corresponds with Fulcher's (1997) results which indicated that the majority of the respondents agreed that the test was fair and looked right (face validity). Other factors might have affected the minority's responses such as students' physical or psychological factors. Therefore, issues of practicality needs to be taking into account while administering the exam such as light, room size, noise, exam start time and furniture to yield better insights on students' performance and test results.

2). Statement two (there is a need for listening and speaking components in the exam) results have showed, as indicated in Graph 2, that 66% of the participants believed that there is a need for both listening and speaking components in PT, while 16% were neutral and only 18% disagreed. This is a strong indication that many students think that having these two components might actually change their placed levels as they are tested in these components in the mid-term exam. Considering learning and multiple intelligences (Rao, 2014), it is advisable to have a variety of skills and components in the same assessment to cater for different learning preferences to avoid leaving any student at disadvantage because exam components are focusing on his weak areas and does not indicate his ability in language learning. Although the majority was happy with exam components and items, the 18% of students who disagree might be those students who are good at speaking or listening, the two skills that are missing in CAS PT, or they could be similar to Fulcher's (1997) respondents who wanted longer tests to have better judgments of their language skills.
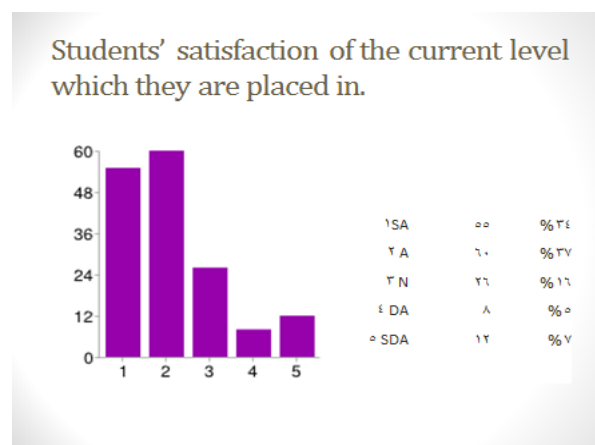
Graph 3. I am satisfied with my current level
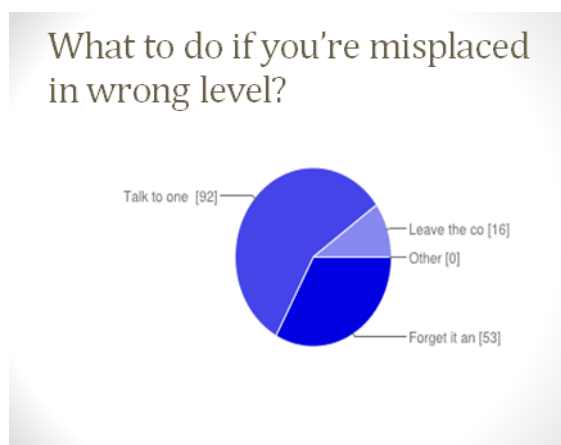
Graph 4. What to do if you are misplaced in a wrong level

3). Graph 3 shows that 71% of participants were satisfied with their placed levels, while 16% felt neutral and only 12% were dissatisfied. This indicates that after two months into the semester, 115 students felt satisfied and content with their current levels. The number of students who expressed their concern (in Q1) of being misplaced in a level that does not suit their language abilities dropped from 19% to 12% still dissatisfied with their level even though half of the semester period passed. This could be referred to their experience of being either bored placed in a lower-level class or challenged being placed in a higher-level class having to face embarrassment and feeling ignorant.

4). In answer to "What would you do if you are misplaced in the wrong level?" 92 students said that they will inform their teacher, administrator or coordinator, 16 students stated that they will leave college while 53 indicated that they will forget it and try to do their best. This shows that 53 students actually think that they cannot change the situation and that it does not matter, which is true to a large extent due to the centrality of the national programs. In other words, even coordinators and departments have to get permission from the Ministry of Higher Education in case they wanted to change a student level based on his or her performance.

3.1.2 Teachers' Perceptions of CAS-PT (Face Validity)

Six English language foundation teachers were interviewed to explore their perception of the current PT and suggestions to improve the test. Their experience in teaching English and assessment ranged from 1 year (an Omani teacher) to 25 years. All of the interviewees had encountered students misplaced in the wrong level in one of the A, B or C levels. Interviewees had encountered such cases and they were not surprised at all as one of them said "Colleges are told to accept all enrollees, so this is the first consequence of not eliminating undeserving students". Teachers have tried to work with those who are willing to improve to raise their levels. Four teachers suggested including speaking and listening components in the PT as they are important skills and some students are good speakers and listeners which again corresponds to learning and multiple intelligences theories (Rao, 2014). Exam designers could refer to already tested PTs for validity and reliability such as Versant (Versant Guide, 2011) and EPT (Illinois, 2012) which integrated the four skills effectively without extending exam time. Versant (2011), for example, tests students' general language abilities in all 4 language skills in 50 minutes only which is very practical. Moreover, EPT (Illinois, 2012) elicits a comprehensive range of students' writing and speaking skills by integrating listening and reading skills where students have to read articles, listen to a mini lecture and discuss in order to produce his or her written piece. Students are fully-engaged with the tasks during EPT. Making use of already published tests can provide a guide to design a valid and reliable in-house designed PT.

Another teacher, however, suggested modifying the reading component to make it a better judgement of students' reading ability rather than a measure of grammar or vocabulary. She and another teacher added that the writing questions, as with the rest of the exam, are easy and the criteria lax which give students a chance at scoring a higher level via "rehearsed and practiced essays that lack depth, originality, style and use of language features". Kim and Shin (2006) discussed Columbia's Reading and Writing PT has already been tested for reliability and validity. It consists of two reading tasks in which students have to read short texts and answer 12 questions in 20 minutes and a descriptive writing task graded according to a 5 point rating scale. Although CAS-PT included both components; reading and writing tasks, the reading component is very short consisting of a very short outdated text for which

students answer 5 wh-questions. The reading text is not long enough to test reading skills and the only sub-skill tested is scanning for details. Furthermore, the writing part, despite being a descriptive paragraph graded on 5 points rating scale like Columbia's PT, uses familiar topics to the students: Topics, which students have practiced or even memorized at school which explains the "rehearsed and practiced essays" argued by the interviewed teacher.

Furthermore, one interviewee advocated adopting or creating a computerized PT "which will increase the difficulty of questions as the students' progress answering correctly, thus placing them at appropriate levels. For example, low-level students would be placed quickly without needing to continue taking the entire test while advanced students would be challenged to answer more difficult questions to score an advanced level. However, such tests advocate objective types of questions which raise availability of the luck element by guessing (Heaton, 1988).

3.2 PT Content Validity

3.2.1 Analysis of Students' PT Scores against MT Scores

Following Research Notes (2005) steps in validating a learning support program for students who needed remediation at the University System of Georgia, this study has adapted their approach to validate not the CAS foundation program but CAS PT. Therefore, Students' midterm results were compared to students' PT results taking into account there was a two months period of instruction between the two tests. The results showed a corresponding positive correlation to that resulted from Research Note (2005).
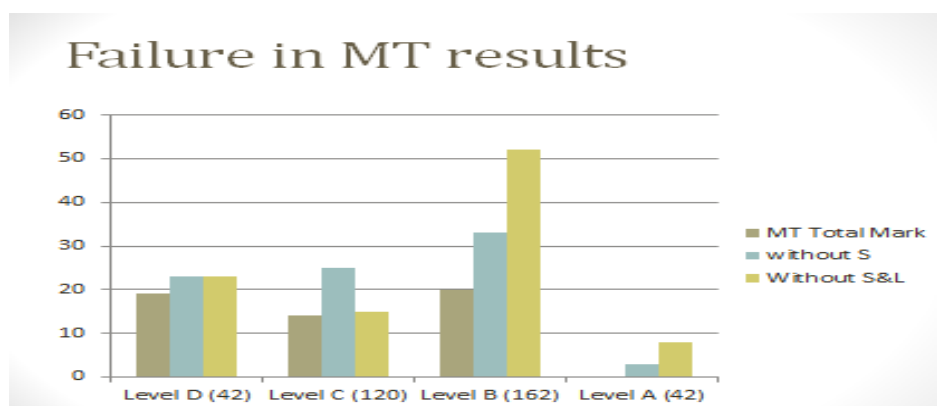
Graph 5. Number of failures in MT results

This graph shows the failure percentage in the mid-term exam in each level. It can be clearly observed that the failure percentage increased when the speaking and listening (S & L) results were excluded from the total mark. In level D, for example, 19 students failed the MT getting less than 20% out of 40%. That number, however, increases to 22 when both speaking and listening marks are excluded. Level B showed similar results with 20 students failing the MT and the number shot up to 52 students failing when both S & L scores were excluded bearing in mind that the total percent of the MT in that case goes down to 26% as S & L worth 8% each. Therefore, it could be directly interpreted that speaking and listening components need to be part of the placement test exam as they make a difference with regard to students' pass rate. Students seem to perform better in speaking and listening due to different factors such as constant practice in the classroom, lenient and vague criteria of marking speaking, the number of times listening audio is played (four times) and learning preferences. However, such results are likely to indicate low reliability due to the 8 weeks of teaching interval between the two exams. In other words, being instructed in language learning for such a long period of time is likely to have positive effects on the students' acquisition of the language. Furthermore, the two tests are different in their purpose, format and content which made this comparison results unreliable and required validation by conducting the following interview with a PT researcher.

3.2.2 An Interview with a PT Researcher

Dr. Y completed his Ph.D in Assessment and as he has a wide solid background of assessment and due to the fact that he has analyzed students' PT papers, this face-to-face interview was conducted with him and he showed a depth of knowledge of the PT format and content. To begin with, Dr. Y indicated that CAS PT lacks essential

components such as listening and speaking which are valuable to students majoring in IT and IBA as they are going to need them constantly in the workplace. Furthermore, he claimed that the other constructs, reading and writing, are under-represented. The reading part does not test reading comprehension as much as it tests grammar and vocabulary which yielded similar responses from the foundation teachers interviewed. Also, the writing part asks superficial questions and promotes rehearsed and practiced answers. However, he said that the writing grading rubric is very analytical and detailed which is good as holistic marks say nothing about written production. However, he failed to notice that some words are vague and need to be well-defined such as 'some' or 'few' mistakes as some, few or many are vague quantities and they do not refer to a specific number.

In answering the question whether this exam reflects students' actual performance, his answer was no. He said "I have had three level B students in my level C class the previous semester. Their level and performance was better than their peers." He had communicated the situation to the foundation year coordinator, but most of the time it is hard to do anything about it. He claimed that there are some reasons behind the mismatch between students' performance and PT score. The first reason is lack of speaking and writing constructs in the exam which may have a strong effect on the results as in placing them in a higher or lower level. Secondly, the percentage of the multiple choice questions in the exam is 90% which is impractical as it does not represent real-life situation as was advocated by Brown (2004) and other writers as well. To answer questions in real life, you are not really given a set of options to choose from and in order to match teaching to assessment, other types of questions should be thought of which are authentic or at least simulate real-life situations. Thirdly, PT is not testing the teaching objectives or learning outcomes. Many students complain that what they are assessed in does not resemble what they have been taught which is a claim by some teachers as well (content validity).

Dr. Y was interested in examining students' PT papers because he wanted to investigate students' side of the story as the first part, examining content and format, is a pre-stage to investigate students' answers and reactions to their PT score. He argued that a badly-designed PT has a strong impact on students' reaction to learning, classes and future especially because exams are usually viewed negatively (Heaton, 1988; Brown, 2004). Dr. Y stated that CAS PT might be a valid exam but it cannot be reliable. And in order to solve this problem of misplacing students in wrong levels, Dr. Y suggested incorporating speaking and listening constructs in the exam. Moreover, he suggested investing available resources on training Omani teachers on writing and grading PTs. Omani Teachers will in turn train new teachers of multi-cultural and educational backgrounds as each of them defines constructs and grades them differently and as these teachers are here for a short period, "invest on the citizens", he said. He has also suggested adopting *ABTIS* test designed by the British Council to assess English skills and he claimed that it is a valid exam of international standards; however, as we prefer designing our own exams, currently this is not an option.

4. Conclusion and Future Research

This paper has tackled the effectiveness of the placement test used by CAS reflecting on teachers' and students' perceptions, comparing PT scores against MT scores and seeking expert analysis of the exam content and format. A notable number of students failed the MT when the speaking and listening components were excluded from the total mark of the students. Incorporating speaking and listening constructs have been suggested by different parties: teachers, students and a researcher to raise exam content validity. In general, face and content validity of CAS English PT range between low to moderate level. A PT examiner recommended training Omani teachers on exam writing, exam grading and writing test specification as it is a good investment for future. Another teacher suggested using computerized tests as they are better means of assessing students' proficiency level. In addition, they save time and energy in terms of whether there is a need for students to complete the exam or not based on their performance and in terms of grading PT. However, there are few limitations to this study. First of all, students may score low in the PT not because of the unreliability or invalidity of the exam only but due to practicality issues, physical and psychological factors during the exam such as stress, heat and sickness, which this paper has not investigated. Secondly, comparing MT scores against PT scores might not be the best way of assessing PT effectiveness because of the differences between the two exams in terms of purpose, content and format. Not to ignore the 7-8 weeks period between the two tests in which students have 20 hours of English language instruction every week and this is doomed to affect students' performance positively. Therefore, for future research, asking a group of students to sit for the same PT twice in different days to see whether it will produce similar results or there will be a significant difference is an essential step. It is also recommended to design a new test taking into consideration the results of this research and pilot it to test its effectiveness. Furthermore, an analytical test of the current marking criteria (Appendix 4) is essential to check its clarity and consistency. Moreover, environmental and health factors should be considered and investigated as they are essential factors in the field.

Acknowledgement

References

Al-Lamki, S. M. (1998). Barriers to Omanization in the private sector: The perceptions of Omani graduates. The *International Journal of Human Resources Management, 9*(2), 378-400. http://dx.doi.org/10.1080/095851998341143

Arkoudis, S. (2011, 12 October). *English language development faces some testing challenges*. The Australian Higher Education, University of Melbourne, p. 33. Retrieved on April 20, 2014 from http://www.theaustralian.com.au/higher-education/english-language-development-faces-some-testing-challenges/story-e6frgcjx-1226164268886

Belfield, C., & Crosta, P. M. (2012). *Predicting success in college: The importance of placement tests and high school transcripts* (CCRC Working Paper No. 42). Retrieved on April 20, 2014 from http://ccrc.tc.columbia.edu/Publication. asp?UID=1030

Brown, H. D. (2004). *Language assessment principles and classroom practice.* White Plains, NY: Longman.

CAS. (2011). *Foundation Program 2011-2012*. Muscat: Ministry of Higher Education.

Casazza, M., & Silverman, S. (1996). *Learning assistance and developmental education*. San Francisco, CA: Jossey-Bass.

Flucher, G. (1997). An English Language Placement Test: Issues in Reliability and Validity. *Language Testing, 14*(2), 113-138. http://dx.doi.org/10.1177/026553229701400201

Gerlaugh, K., Thompson, L., Boylan, H., & Davis, H. (2007). National study of developmental education II: Baseline data for community colleges. *Research in Developmental Education, 20*(4), 1-4.

Heaton, J. B. (1988). *Writing English language tests.* London: Longman.

Illinois. (2012). *English Placement Test* (EPT). Retrieved on March 2, 2014, from www.publications.uiuc.edu

Kim, H., & Shin, H. (2006). A reading and Writing Placement Test: Design, Evaluation and Analysis. *Teachers college, Columbia University Working Papers in TESOL & Applied Linguistics, 6*(2), 1-36.

Maxwell, M. (1997). *Improving student learning skills.* Clearwater, FL: H & H Publishing.

McCabe, R. (2000). *No one to waste: A report to public decision-makers and community college leaders.* Washington, DC: Community College Press.

Morante, E. (1989). Selecting tests and placing students. *Journal of Developmental Education, 13*(2), 2-4, 6.

OAAA. (2009). *Oman academic standards for general foundation programs*. Muscat: OAAA.

Rao, R. (2014). Integration of Bloom's Taxonomy and Multiple intelligences in English classroom. *THE ELT PRACTITIONER, I*(1), 1-3. Retrieved on September 10, 2015, from https://sites.google.com/site/theeltpractitioner/archive/vol-1-no-1/3

Sapsford, R. (2007). *Survey research*. London, UK: SAGE publication.

Scott-Clayton, J. (2012). *Do high-stakes placement exams predict college success?* (CCRC Working Paper No. 41). Retrieved on July 1st, 2015, from http://ccrc.tc.columbia.edu/media/k2/attachments/high-stakes-predict-success.pdf

Pearson. (2011). *Versant English Placement Test: Technical paper*. Retrieved on March 2, 2014, from http://www.versanttest.com/technology/VersantEnglishTestValidation.pdf

Appendix

Appendix 1

Ministry of Higher Education
Colleges of Applied Sciences Placement Test for Foundation English

Candidate's Booklet
Time: 1 hour 30 minutes

Name	
ID e.g. Resident Card	

There are two parts to the test: Part One, Language and Reading and Part Two, Writing. Answer all questions in both parts. Write your answers in the spaces provided in this booklet.

ينقسم هذا الامتحان إلى جزئين: الجزء الأول عن اللغة و القراءة بينما الجزء الثاني عن الكتابة
أجب عن جميع الأسئلة واكتب جميع إجاباتك في الأماكن الموضحة في هذا الكتيب

1.   (A)   (B)

2.   (A)   (B)

3.   (A)   (B)

4.   (A)   (B)

5.   (A)   (B)

Questions 6 – 10

>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>> - - (6) - - <<<<<<<
<<<<<<< <<<<<<<<< <<<<<<<<<<<<<<<<<<<<<<<<<<<<<
<<<<<<<<<<<<<<<<<<<<<<<<<<<< - - (7) <<<<<<<<<<<<<<<

Part One: Language Knowledge and Reading

For **Questions 1-15**, choose the words that best fit each space in the texts below.[1] Put a circle around each of the correct answers.

للأسئلة من1إلى 15: اختر الكلمة الأنسب لكل فراغ في القطعة التالية.
ضع دائرة حول الإجابة الصحيحة.

Questions 1 – 5

>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>> - - (1) - - <<<<<<< <<<<<<<<<
<<<<<<<<< <<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<< - - (2) - -
<<<<<<<<<<<<<<<<<<<<<<<<<<<< <<<<<<<<<<<<<<<<<<<<<<<<<<
<<<<<<<<<<<<<<<<<<<<< - - (3) - - <<<<<<<<< <<<<<<< <<<<<<<
<<<<<<<<< <<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<< <<<<<<<<<<<
<< - - (4) - - <<<<<<<<<<<<<<<<<<<<<< <<<<<<<<<<<<<<<<<<<<<<<<<
<<<< <<<<<<<<<<<<<<<<<< <<<<<<< - - (5) - - <<<<<<<<<<<<<
<<<<<<< <<<<<<<<< <<<<< <<<<< <<<<<<<<<<<<<<< <<<<<< <<

<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<
<<<<<<<<<<<<<<< <<<<<<<<<<<<<<<<<<<<<<<< - - (8) - -
<<<<<<<<<<<<< <<<<<<<<<<<<< <<<<<<<
<<<<<<<<<<<<<<<<<<<<<<<<<<< <<<<<<<<<<<<<<<<<<<<<<<<
 - - (9) - - <<<<<<<<<< <<<<<<<<<<<<
<<<<<<<<<<<<<<<<<<<<<<<<<
<<<<<<<<<<<<<<<<<<<<<<<<<< - - (10) - -
<<<<<<<<<<<<<<<<<<< <<<<<<<<< <<<<< <<<<<
<<<<<<<<<<<<<<<<<<<< <<<<<<<<<<<<<
<<<<<<<<<<<<<<.

6.   (A)   (B)   (C)   (D)

7.   (A)   (B)   (C)   (D)

8.   (A)   (B)   (C)   (D)

9.   (A)   (B)   (C)   (D)

10.   (A)   (B)   (C)   (D)

13.   (A)   (B)   (C)

Questions 11 – 15 (Vocabulary)

14.   (A)   (B)   (C)

15.   (A)   (B)   (C)

```
>>>>>>>>>>>>>>>>>>>>>>>>>>>> - - (11) - -  <<<<<<< <<<<<<<< <<<<<<<<<
<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<< - - (12) - -
<<<<<<<<<<<<<<<<<<< <<<<<<<<<<<<<<<<<<<<<<<<<
<<<<<<<<<<<<<<<<<<< - - (13) - - <<<<<<<<<<<<<
<<<<<<<<<<<<<<<<<<<< <<<<<<<<<<<<<<<< <<<<<<<<
<<<<<<< <<<<<<<<<<<<<<<  - - (14) - - <<<<<<<<<<<<<<<<
<<<<<< <<<<<<<<<<<<<<<<<<<<<<<<<<< <<<<<<<<<<<<<<
<<<<<<<<<<<  - - (15) - - <<<<<<<<<<<<<<<<<<< <<<<<<<<<
<<<<< <<<<< <<<<<< <<<<<<<<<<<<<< <<<<<<<<<<<<<<>
```

Questions 1-15 from or adapted from: University of Cambridge (2004) ESOL Examinations (photocopiable) Oxford University Press.

Questions 16 – 20

For **Questions 16-20**, circle the best answer.
للأسئلة من 16 إلى 20: ضع دائرة حول الإجابة الصحيحة

```
<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<
<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<
<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<
<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<
<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<
<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<
<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<
<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<
<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<
<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<
```

11.   (A)   (B)   (C)

12.   (A)   (B)   (C)

```
<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<
<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<
<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<
<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<
<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<
<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<
<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<
<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<
<<<<<<<<<<<<<<<<<.
```

(A)   (B)   (C)   (D)

Part Two: Writing

Answer **Questions 21 and 22**. Write your answers below. You may ask the teacher for extra paper. If you want to plan your answer use rough paper.

أجب عن السؤالين 21 و22 واكتب إجابتك في الفراغ التالي
يمكنك أن تطلب المزيد من الورق إن لم يكفيك الفراغ
إذا أردت ان تخطط لإجابتك إستخدم أوراق أخرى

16   Why ....? Because
     (A)   (B)   (C)   (D)

17   Which ....?
     (A)   (B)   (C)   (D)

Question 21

18   How ....?
     (A)   (B)   (C)   (D)

Describe your ....
Please tell us about .... Try to write about 100 words.

19   What ....?
     (A)   (B)   (C)   (D)

صِف
اكتب بعض المعلومات.
اكتب 100 كلمة

20   Who ....?

_____   _____
_____   _____
_____   _____
_____   _____
_____   _____
_____   _____
_____   _____
_____   _____
_____   _____
_____   _____
_____   _____
_____   _____
_____   _____
_____   _____
_____   _____
_____   _____
_____   _____
_____   _____

Question 22

| Describe ....
... .

Tell us .... Try to write about
100 words.

صِف ....
أكتب 100 كلمة |

_____   _____
_____   _____
_____   _____
_____   _____
_____   _____
_____   _____
_____   _____
                                             _____

Appendix 2
Questionnaire

(The Effectiveness of Foundation English Language PT at CAS)

This survey is part of a research project examining the effectiveness of foundation English language placement test at the colleges of Applied Sciences. We do appreciate your time and effort in answering this survey. Your responses are confidential and information provided will be used for research purposes only.

* Required

Foundation Level: *

Level A

Level B

Level C

This is a required question

Section One: Choose one answer: Do you agree that: *

1. Placement test represents our actual level.

      1   2   3   4   5

      1   2   3   4   5

Strongly Agree		Strongly Disagree

*This is a required question

Section Two: Choose One answer: *

If I am misplaced in a wrong level, I should:

Forget it and stay in the level.

Talk to one of the faculty (teacher).

Leave the college and go to another college.

Other:

This is a required question

Never submit passwords through Google Forms.
Powered by
     This form was created inside of COLLEGES OF APPLIED SCIENCES.

Strongly Agree		Strongly Disagree

*This is a required question

2. Placement test places students in the correct level.

      1   2   3   4   5

Strongly Agree		Strongly Disagree

*This is a required question

3. Placement test assesses all our skills.

      1   2   3   4   5

Strongly Agree		Strongly Disagree

*This is a required question

4. I think there is a need for listening and speaking sections in the placement test.

      1   2   3   4   5

Strongly Agree		Strongly Disagree

*This is a required question

5. Now, I am happy that I have been placed in this level because it improves my skills.

100%: You made it.

Appendix 3

The effectiveness of English Language PT at CAS
Sharifa Al Adawi & Aaisha Al Balushi

Interview questions

This survey is part of a research project examining the Effectiveness of English language placement test at the colleges of Applied Sciences. We do appreciate your time and effort in answering this survey. All responses and answers are confidential and information provided will be used for research purposes only.

Name:

Years of experience: ∨ in teaching:

      ∨ in assessment (Writing exams):

Foundation Level you are teaching now:

1) Have you looked at English placement test which is run in the college?

2) What do you think about it?

3) Does it cover all language skills?

4) Do you think it represents the students' actual levels of language proficiency? Why?

5) Have you ever encountered a student that you thought he/she might be misplaced in that level?

6) If yes, how did you reach that conclusion?

7) What was your reaction?

8) What are the drawbacks of the current placement test?

9) Do you think that the placement test should be modified?

10) Do you have any suggestions for changing the placement test? What are the suggestions?

Thank you for your time

Appendix 4

Part Two: Writing: Rating Scale and Procedure for Assessing Writing

1. Assess for **Task Achievement** first.
2. If the script meets the **2** criteria, continue to mark the script. If not enter 1, 2 or 0 on the mark sheet.
3.
4. **TASK ACHIEVEMENT**: This is a 'gateway' criterion i.e. a criterion that establishes whether or not the script should be marked in full. The criterion attempts to answer the question: has the student attempted to construct a text, in more or less their own words, that attempts to address the question?
5. **ORGANISATION**: This criterion covers the organisation of text above the level of the sentence, so at paragraph and text level. It assesses structure i.e. division into functional sections (introduction – body – conclusion), functionality i.e. the extent to which the sections perform their jobs and linearity i.e. the extent to which the ideas flow in – to Western readers - a logical manner.
6. **GRAMMAR**: This criterion covers range, complexity and accuracy of grammar use. 'Range' refers to whether or not the text displays the structures actually necessary to carry out the task in. For example a task that required a writer to describe a sequence of events in a linear way (first A , then B, then C) would require use of the Past Simple and perhaps the Past Continuous. A more complex task would be to describe a sequence of events in a non-linear way (i.e first B,

then back to A and then forward again to C) and this would require use of the Past Perfect, as well as the other two tenses. 'Complexity' is related to range. In this rating scale the common – but contentious and misleading - distinction between 'simple' and 'complex' grammar is used. Simple structures would include one clause sentences, or *co-ordinated* two clause sentences (e.g. He lives in Muscat but he works in Nizwa), noun phrases with no more than three components (e.g. determiner-adjective-noun), single verb phrases, adverbs rather than adverbial phrases, Past Simple for narrative, Present Simple for description. Complex structures would include subordinated clause sentences, multi-item noun and verb phrases, multi-item adverbial phrases, conditional structures, Past Perfect in narrative, Present Perfect on the narrative/descriptive boundary, modal verb phrases.
7. **VOCABULARY:** This criterion also covers range in relation to the task but also considers three common consequences of a lack of range: the use of circumlocution, the simplification of ideas through the use of general words, misuse of words. It also considers register, by which is meant a choice of words appropriate to the type of text and the context of use. In an exam we expect standard academic English and should penalise colloquially.
8. **MECHANICS:** This covers hand-writing, spelling, punctuation and capitalization.

Mark	Organization	Grammar	Vocabulary	Mechanics	Task achievement		
5	Shows a topic sentence, supporting sentences with logical 'flow', coherence throughout, appropriate use of cohesion	Shows the range required for the task i.e. has the grammar required to perform the task, largely accurately.	Shows the range of vocabulary required for the task, largely accurate collocation, appropriate register	Largely accurate spelling, punctuation, capitalization throughout. Neat, clear hand-writing	A paragraph, at/ above minimum length, addressing the rubric, relevant throughout, reasonably comprehensive in coverage		
4	Shows a topic sentence, supporting sentences with generally logical 'flow' (may meander a little), generally coherent and cohesive	Shows the range required for the task but with errors in complex structures.	Shows the range of vocabulary required for the task, despite lapses in collocation or register	Generally accurate spelling, punctuation, capitalization thoughout. Legible hand-writing	A paragraph, at/ above minimum length, addressing the rubric but occasionally irrelevant or lacking coverage of some issues		
3	Shows a topic sentence, supporting sentences that lack logical 'flow', with lapses in coherence or cohesion.	Frequency of errors shows insufficient range or accuracy for the task. Meaning will be fairly clear but will not be conveyed consistently through grammar.	Shows insufficient range of vocabulary for the task. General meaning may be clear but lexis will be inadequate to express precise meaning. There may be lapses in collocation or register.	Frequent errors in spelling, punctuation or capitalization	**OR** Difficult hand-writing (i.e. legible but with difficulty)	A paragraph, at/ above minimum length, addressing the rubric, occasionally irrelevant and lacking coverage of some issues	
2	Shows no clear topic sentence, supporting sentences that lack logical 'flow' throughout, giving rise to coherence problems and limited cohesion	Frequency of errors in complex and basic structures shows insufficient range for the task and this will affect understanding: the text may be	Shows insufficient range of vocabulary for the task and this will affect even general understanding. There are likely to be errors in collocation and register	Frequent errors in spelling, punctuation and capitalization	**OR** Very difficult hand-writing (largely illegible)	Short (50% or less)	**OR** on-topic but not addressing the rubric (so frequently irrelevant)
1	The text displays no discernible topic sentence, no logical 'flow' throughout, lacks coherence and cohesion	Frequency of errors shows wholly insufficient range for the task and understanding will be severely affected. The reader may understand little of the text.	Shows wholly insufficient range for the task and understanding will be severely affected. The reader may understand little of the text.	Frequent errors in spelling, punctuation and capitalization. Very difficult hand-writing (largely illegible)	Short (50% or less) On-topic but not addressing the rubric (so frequently irrelevant) or off-topic.		
0	No assessable sample						