EVALUATING BASIC ENGLISH TEST ITEMS

FOR NON-ENGLISH STUDENTS FROM TEACHERS PERSPECTIVES

Prihantoro

Universitas Diponegoro

[prihantoro2001@yahoo.com](mailto:prihantoro2001@yahoo.com)

Abstract--This paper seeks to evaluate diverse Basic English tests items for Indonesian college students who are non-English major (all majors except English). The object of this research is a collection of English exam sheets (both mid and final terms) for non-English students in one state university in Central Java, Indonesia. Test items were categorized thematically. The most popular skills tested are reading comprehension and grammar. Writing is low. Listening is never tested. In order to understand the perspectives the English teachers, the teachers who composed the exam sheets were also interviewed. The result has shown that 95% of the items are indirect assessment (multiple choices, true false, matching, cloze quiz) are preferred as the correcting and scoring is much easier. Besides, this gives advantage for students as the answers are already provided. It fits language skills like reading and grammar. Direct assessment in the form of construction (e.g essay) is rare. The existing ones put very tight constraints (number of words, topic etc). Direct assessment is less preferred as correct/proper answers may vary and the scoring is time consuming. The interviews result show that classes they teach and their teaching loads drove this preference. However, changes might take place when the related program/department/university/ 1) strengthen the human resources (in terms of quantity and quality) or 2) delegates the responsibility to a committed language center.

Keywords: direct and indirect assessment, teacher perspectives, large classes

1. INTRODUCTION

This paper looks into test items in Basic English class, a mandatory class for all college students (except English Department students) in Indonesia. Creating test items for this course is considered challenging. First, in terms of entry criteria, this course can be taken without any requirement; where students have different degree of English language proficiency. Second, it is a mandatory class that students must take, even though it may not have a direct link to the core competence of the department; they have to take it whether they like it or not. Three, classes are usually large in size (more than 40 students per class).

With these facts, it is important to look at different aspects of this course, and suggests how to improve it. This paper discusses the test items designed by the teachers. Besides the evaluating the items, it is also important to understand how teachers reflect on the items; what motivate such items to surface; what has been considered advantage or disadvantages of the items.

This paper is organized as follow. Section 1 describes the background of this study. Section 2 briefly visits some basic concepts in language testing which are discussed in this study. Section 3 describes methodology. Findings are discussed in section 4, while section 5 proposes some possible solutions.

1. LANGUAGE TESTING

There are many definitions of language testing, especially when we test specific skills like reading, conversation, listening, grammar, translation, writing and etc. However, we are aware what testing is. It is an instrument to measure something. As quoted from Douglas (2009) language testing is an instrument to measure language ability.

There are at least two kinds of testing, formal and informal (Popham & Popham, 2005). Informal testing is also an important part of testing, but done informally; like when teachers ask questions to students to measure their understanding of what has been taught or just to discover their background knowledge. In fact, informal assessment is a part of learning process.

In this paper, however, we focus more on formal assessment (Fulcher, Practical Language Assessment, 2010). Formal assessment is important as it gives standards, in form of grades, which (should) become the reflection of the learning process. With the results of formal assessment, one who is not involved in the learning and teaching process may understand one’s competence.

1. Why, What and How

Before discussing the form of a formal testing, it is important to understand the reasons why the testing is given (Fulcher & Davidson, 2007). Consider diagnostic test (Bachman, 1990). This test is used to measure one’s competence and then decide whether the competence is up to standard or not. As an illustration, students registering for an International class must take written and oral test in English, as English will be the language of instruction. It is important for the students to pass the test as they are going to write reports, discuss and argue in the class by using. Another illustration is final term exam, for a business English class as an instance, where the reason for testing is to check whether learning objectives are achieved. These two illustrations are example to show the importance of reasons for conducting a test.

When the reasons for conducting a test have been established well, the next step is to identify what to be tested. (Hent, 1987) Is it the acquisition of certain grammatical structures (passive, causative, clauses, parts of speech)? Is it to measure vocabulary acquisition in certain domain (business, academic, sports)? Is it measure fluency in conversation (use of fillers, controlled ideas, flow, and responsiveness)? What competence to measure is usually present in learning outcomes (as a part of lesson plans).

After identification of what to be tested is cleared, it is important to set how the test should be organized (Fulcher & Davidson, Language Testing and Assessment, 2007). Does the test have to be held in oral? Should the oral test be divided into several segments (introduction, short talk, longer interactive talk)? Does it have to be written? What form of test items are appropriate (true-false, multiple choice, essay)? Answering the question of why, and what, is very important to determine the ‘how’; which is how we execute the test.

1. Direct and Indirect Assessment

Another type of assessment is direct and indirect assessment (Breland & Judith, 1979). This will be the core of the discussion in this paper. However, let us briefly review shortly the difference between the two types of assessments. Direct assessment is a type of open testing, like essay or interview. Responses required in this type of testing are open, which means it is not prepared by the teachers or test designers. Therefore, answers may vary, and there can be no exact answers. What teachers can do is measure whether the answers (to what extent) fit the variables of the tests from rubrics that have been prepared.

Indirect assessment is a type of testing where selected responses may be provided by the teachers or test designers. This includes true-false identification, matching sequence, or multiple choices. In this paper, we focus on indirect assessment, how teachers reflect on this in Basic English class, and the motivation to prefer this assessment as compared to indirect assessment.

3. METHODOLOGY

Data in this paper is obtained by using questionnaires where the respondents are from one state university in Central Java. The questionnaire will be discussed in the discussion section. I requested the list of English teachers from different programs/faculties and randomly sent requests for their participation in this research. I do not limit their status (permanent or part time). I also do not distinguish educational background (linguistics/literature/area studies/ Non-English departments). Those who have agreed to participate, received questionnaires. From 15 questionnaires, 13 are returned. One questionnaire is not returned because the respondent had to perform religious travelling, and another one was not returned for any reason. As the questionnaires were filled by handwriting, respondents are visited on the purpose of validation. Besides validating handwriting, another purpose of the visit is to request further clarification (when needed). The data in the questionnaires will be presented by both visual chart and prose description when necessary.

4. DISCUSSION

One of the main findings in this study is that the number of indirect assessment is weigh over direct assessment. The percentage of indirect assessment test items is 95% in total with 60% in multiple choices. Subjects tested by using this test items are mostly reading and grammar.

Figure 1. Distribution of Language Skills Tested

Figure 2. Proportion (%) of Test Items

Reading and grammars dominate language skills tested in the exam. According to the respondents, these two are selected as they are usually the objectives of the course, and they are skills realistically to be tested considering the size of the class. And what has been said about the preference of multiple choice?

Teachers have their practical considerations why this kind of selected response item is preferred (Roediger & Marsh, 2005). First, it reduces the anxiety of the students. Teachers believe that during direct assessment such as essay writing or oral interview, students feel anxious and as a result cannot perform well. Second reason is about the scoring. Scoring has always been in the mind of the teachers or test developers when they design tests. I survey how many credit hours teachers have to teach, and none of them teach below 10 credit hours (1 credit hour = 50-60 minutes). Most of the teachers teach 10-20 credit hours, but some of them teach more than 20 credit hours per week.

Figure 3. Teaching Loads per Week

I also ask them about their status. Three of them are part timers and the rests are permanent civil servants teaching in the University. Both are equally problematic. While the part timers also have other (teaching) job in different institutions, those who are in permanent positions are obliged to publish and also to do community service. Publication usually is the most challenging, as compared to teaching and community service. Often one has to spend months to publish in a national journal, as papers may have to go through revisions. With these teaching loads, indirect assessment is a logical choice. Scoring can be delegated to other people, as there is absolute answer within the choices.

Consider also the number of students in a Basic English class, which according to the survey mostly is within 50 to 100. Even the respondents mention that there are some classes where the students are beyond 100. Multiply this number with scoring time. If an answer sheet consisting the combination of direct and indirect assessment is scored within 15 minutes, finishing 100 sheets will take 1500 minutes or 25 hours.

Figure 4. Class Size

A large class is an issue (Bachman, 1990) here; why in a foreign language class, the number of students may reach more than 100. Some teachers talk to the authority (the department or faculty member), and reasons are complex, but to some extent, make sense. Some mention that relates to the financing. If they have to divide 100 students into three or four class, then the authority would have to think to hire more teachers. This also happens to other non-department class. As an illustration, take the example of civil engineering department where students in the earliest semester must take basic science classes like basic math, or physics. So, the issue is financing, which really makes sense for faculties or department with tight budgets. Another reason is room availability. In some departments, they only have rooms for large classes. It takes time to build rooms for language classes with 40 students, as they have to propose this first to the university. Even in some departments, scheduling is very tight as number of classes are very close to number of subjects. When a class is divided into two or three, they might be short of classroom.

When asked about the type of text items that they would use (assuming that classes are in the ideal condition from teaching loads number of students to facilities), teachers prefer the 70%-30% combination of indirect and direct assessments. Improvement in direct assessment is required to measure production skills like writing or speaking, while indirect assessment will fit best for understanding like reading or grammar.

Indirect assessment test items are also challenging in some aspects. First, it takes more time to prepare, as teachers or test developers would have to come with different alternatives as in the case of multiple choice. Second, the alternatives (or called distractors), are also not easy to find. They cannot be too different, but cannot also be too similar; achieving reliability in this kind item is not easy. Sometimes, test items need to be tested too before they are distributed to students. Instruction must be set very clearly in order not to confuse students. So, besides the scoring efficiency time benefit for teachers and selected responses benefit for the students, multiple choices item is challenging in terms of reliability.

According to the teachers, these all are classical problems, and they, and also the faculty or department members are actually aware of this. But why these classic problems do not end? One answer is because in many department, (or all departments) Basic English is not the core competence. This is a very logical concerns. Basic English falls to the type of *Mata Kuliah Umum* (MKU) which is general subject in literal translation. Some other subjects that fall to MKU are subjects like sports, religion, or Pancasila (five pillars of Indonesian statehood). They do not directly contribute to core competence of the departments (physics, chemistry, engineering, psychology), but students cannot graduate without passing these courses. Why not removing these courses from the curriculum when they do not have major contribution to core competence? Why not replacing them with more practical course such as entrepreneurship and internship?

It is not that easy regarding the impacts it may probably cause. Permanent teachers might have problems with the removal if MKU courses are the only courses they rely on. In the university where the questionnaires are distributed, permanent teachers are recruited from English department, so when such courses are removed they can still teach in English department. But consider this, not all university have English department. When such courses are removed, the permanent teachers will have difficulty in meeting the minimum threshold of teaching loads. Where should they teach? Then why not making this subject elective? This is actually a good idea, but quite challenging for implementation. Who can guarantee that students want to take this? Basic English is usually only offered without further progress (intermediate, advanced English), so that they will have to go outside to acquire English. There are so many English language institute out of the universities provide more interesting offers for English language development. If this subject is elective, why not starting from the scratch outside of the university? Unless the offer (teaching methods, facility, teaching aids, curriculum) is better, students will definitely choose to learn outside.

5. CONCLUSION

I here offer two solutions, but I totally understand that in reality, they are really difficult to implement and can only be successful with the support of students, teachers and including budgets and policies from higher authorities. In a university with English department, it might not be a good idea to request them to teach Basic English. It is better to recruit their own teachers with specification on TESOL and sub-specialization in English for Specific Purposes. This has actually been done in some universities; let’s say an English teacher who works under the department of agriculture. But the problem is, the teacher is also requested to teach in other department. It is better that an English teacher in college focus on only one department with 12 credit hours max. In this way, the direction of research, publication, teaching and test material developments will focus on the nature of the department. If one stay consistent under one department, say department of geology, s/he will have time to develop teaching and test materials specific to geology. There will be enough time to research geological texts to feed those materials. One can also develop a teaching method and write papers, and publish on teaching English under geological domain. The department should assist the teacher by providing assistance such as teaching aids, access to geological lab, involving in the curriculum design and etc.

Another solution is to request the help of the language center. Each university usually has a language center. Ideally, teachers working in the language center should be recruited by the university, not department. However, university may delegate the teachers to different department, but with careful considerations. Teachers are recruited on the basis of TESOL and English for Specific Purpose, as it has been commented previously. Therefore, when recruiting teachers from literature or area studies (common sub-specifications in English Departments), they must be provided with further training. In this way, the oversee of the teachers is under the authority of the language center. Therefore, the language center has to be excellent. It is not enough for the language center to have resources only for conversations or international tests like TOEFL or IELTS. They must have resources from many different disciplines to assist teachers design teaching and test materials. In my personal opinion, permanent position should only attach to language center managements, while for the teachers it is better to hire English teachers on yearly contract basis. There is a dilemma for state universities. It is quite difficult these days for state universities to propose a new position to the government; therefore teachers in language centers are often recruited from inside circle in the university with remuneration besides they basic salary. It is not good for their career development. Too many teachings will affect publication, and not all publications are related to teaching. Some have expertise on literatures, some others on linguistics, some others on area studies, other might be on translation and etc. But recruiting teachers might be related to budgeting, where some universities have different ways in caring for their language center.

# Reference

1. Bachman, L.-F. (1990). *Fundamental considerations in language testing.* Oxford: Oxford University Press.
2. Breland, H.-M., & Judith, L.-G. (1979). A comparison of direct and indirect assessments of writing skill. *Journal of Educational Measurement 16.2*, 119-128.
3. Douglas, D. (2009). *Understanding Language Testing.* Routledge: London and New York.
4. Fulcher, G. (2010). *Practical Language Assessment.* London: Hodder Education.
5. Fulcher, G., & Davidson, F. (2007). *Language Testing and Assessment.* London and New York: Routledge.
6. Hent, G. (1987). *A Guide to Language Learning: Development, Evaluation and Research.* Wadsworth: Forein Language Teaching Press.
7. Popham, W.-J., & Popham, J.-W. (2005). *Classroom assessment: What teachers need to know.* New York: Pearson/Allyn and Bacon.
8. Roediger, H. L., & Marsh, E.-J. (2005). The positive and negative consequences of multiple-choice testing. *Journal of Experimental Psychology: Learning, Memory, and Cognition 31(5),* , 1155.