TEST OF ACADEMIC ENGLISH PROFICIENCY (TAEP):
AN EFFORT AGAINST THE NEO-COLONIALISM IN EDUCATION

Masduki
masdki@yahoo.com

Cita Nuary Ishak
nuarycita@gmail.com
Language Center
Universitas Muhammadiyah Malang
Malang, Indonesia

Abstract
Neo-colonialism in the modern era no longer emerges physically. Rather, it appears as a mechanism control (Olssen and O’Neill, 2004) or an influential force. Such control comes in the form of certain products which are aimed for various political or economic purposes. In the realm of language education in Indonesia, especially in the context of English Language and Teaching, certain products such as language testing system have attained a position of dominance. Thus, the tests have been employed by many institutions or organizations in Indonesia to measure non-natives ability in English by using a standard determined by native speaker of English. In fact, our tracer study discovered that a large number of Indonesian workers communicate in English with non-native speakers of English. Therefore, an appropriate instrument to measure the English proficiency of non-native speakers is badly needed. The present research is aimed at addressing the immediate need. The research discusses the design, development, and the current use of Test of Academic English Proficiency (TAEP) which is initiated by Language Center of Universitas Muhammadiyah Malang. Since it was legally certified and acknowledged in 2016, TAEP has been used by various institutions and organizations for national and international purposes. Further development and collaboration will conclude this article.

Keywords: assessment, language testing system, TAEP

Introduction
English is not only spoken by people who come from places where the language has taken root, but it has also been used by many people from different countries. Along with globalization, English has come across different continents as it becomes the medium of the development of science and technology. The immediate consequence of the phenomena is that the language will inevitably be open to any kind of unpredictable linguistic changes, from which varieties of English emerge and known as ‘new Englishes’, (Crystal, 2003: 142).

New Englishes are shown to be present in countries where English is the second or foreign language. In Indonesia, new Englishes are also apparent among the speakers. Along with the growing number of its speakers, it is only natural that Indonesia contributes many variations to the newEnglishes. The fact that the use of English, though still foreign, keeps expanding throughout Indonesia proves that the language has a strong influence toward the country.

Historically, Indonesia was in the past controlled or ruled by a number of countries, with Portugal, Spain, and Netherland being the longest ones to invade the country. England colonization in Indonesia went on from 1811 to 1816, but their language is spoken until today. In contrast, Portuguese, Spanish, and Dutch left their traces only in borrowing words. Thus, it can be concluded that the spread and popularity of a language does not depend merely on the length of time
the language is used, but it relies on the power of its people. Crystal (2003: 9-10) argued that a language might indeed gain an international acceptance through military and political power of the people, but to expand and to maintain its international status, economical power plays the most significant role.

There is not any problem regarding how English was once used and eventually regained its triumph in Indonesia. It is only problematic when the practice continues into a neo-colonialism, or a modern colonization, which appears today as a “mechanism control or an influential force” (Olssen and O’Neill, 2004). Such control comes in disguise as certain products which are aimed for various political or economic purposes. Nwanosike and Onyije (2011) observed a similar account in Africa. They reported that education has been misused for gaining colonialism purposes and that this situation has led to a situation of dominancy and dependency. In the realm of language education in Indonesia, especially in the context of English Language and Teaching, certain products such as language testing system have attained a position of dominance. There are at least two consequences from this dominance: first, any other tests other than that made by the dominant language testing system is somehow regarded as not valid, nor relevant; then, second, the language testing system which is commonly perceived as the best test is used for almost all purposes.

As a rule of thumb, a test is not one-size-fits-all. Thus, there is no single best test. A test should be made precisely in line with the purpose of the test. However, most often various purposes of a language test are treated the same. Shohamy (1990: 79) argued that “most tests are developed under the assumption that all test takers have a similar background and learn the language in a similar context and for similar purposes.” With the dominance of a certain language testing system in Indonesia, it is common to find such reality where a single test would be overused for several purposes at the same time: job recruitment process, college admission, and graduation requirement.

To cite an example, a test like TOEFL that is provided by Educational Testing Service (ETS) has been favored by institutions, universities, and companies in Indonesia to assess one’s English proficiency. The test has been used for any kind of purposes, which actually violates its own initial function, as stated by ETS: “…to measure the English proficiency of international students who wish to study in the United States. This is still the primary function of the test.” (ETS, 1990)

Thus, TOEFL test would be appropriate for those aiming to pursue further education in American universities. In Indonesia, the dependency on the test is growing larger and larger over times and this trend has caused a dependency on and the overuse of the test for various purposes, one of which in employee recruitment.

Our tracer study found that Indonesian non-native speakers of English in their real life communication do not always communicate with the native speakers of English in their home country. Rather, they largely communicate with non-native speakers of English, who come from China, India, Malaysia, Singapore, and Japan. The English used in their workplace, therefore, is new Englishes.

Based on the above mentioned reasons, it can be concluded that the emergence of new Englishes along with the growing dependency on certain test instrument triggers the need for a tool aimed at measuring English competence among non-native speakers of English. The present article aims at addressing the immediate need.
Methodology
The following sections discuss the method by which Test of Academic English Proficiency (TAEP) is designed. The approach employed in the present research is quantitative in nature, as it is intended to examine the quality of a test as an acceptable and appropriate instrument to measure language skills. The present research is a kind of Research and Development study.
There are four steps taken into consideration, but the main idea for developing the test is based on the purpose of the test itself, that is, to have a balance and relevant content in testing Indonesian non-native speakers of English proficiency in academic, business, and general settings. The four stages include (1) participant, (2) data collection, (3) expert judgment, and (4) try-out of the test.

Participant
The subjects of the research were classified into two groups. The first group consisted of six semester students (academic year 2013/2014) from Engineering, Accounting, Nursing, and Banking Department at the University of Muhammadiyah Malang, making up to 744 students in total. This group later participated in the pilot study. The second group consisted 1,030 graduates across all departments from different universities in Indonesia. This group later participated in the main trial for test administration.

Data Collection
The data were collected from the try-out of the test. The first try-out was conducted on August 2015. The quantitative data were then analyzed to see its item facility, discriminating power, validity, and reliability. The quantitative data were supported by the qualitative data obtained from the responses to the given questionnaire about the test they had just worked with.

Expert Judgment
According to Cohen (1993), the use of expert judgment is somewhat current trend in a language testing field. For developing TAEP, the experts were a professor in language testing and some colleagues who have completed their doctoral degree in their home country or abroad and whose major programs are in education and applied linguistics. The experts were asked to assist the test constructor through different stages ranging from giving comments and suggestions for the text selection, item writing, and item analysis. As far as the text selection and item moderation were concerned, they were given the first draft of the test developed in the present study and were invited to give comments and suggestions on it.

Try-out of the Test
The try-out of the test was done in two stages. The first stage is the pilot study or the preliminary try-out. It was then followed by the second stage, namely, the main trial.

The Pilot Study
The pilot study was carried out to the students of Engineering, Accounting, Nursing, and Banking Department at the University of Muhammadiyah Malang who were on the 6th semester at the time of the test. The procedure of the pilot study can be described as follows: First, the students were given test papers which had been moderated and revised. They were allowed to do the test without time limit and they might work on the test as long as they need. This resulted in different time of submitting their responses to the test administrator which was then utilized by the test administrator to interview some of them. The interview was conducted on the following topics: the appearance of the test paper, difficulty and confusion experienced by the test takers, if any, in understanding the prompts and questions, and reasons of the students for answering the incorrect responses. They were also invited to give comments for the improvement of the test.

Main Trial
The main trial, or the try-out, was conducted on April 2016, with the total number of participants reaching up to 1,030 students. They were the second group participants. The test-takers in the try-out were highly comparable to the prospective test-takers to whom this test is constructed. That is, non-native speakers of English. The procedure of the try-out was started with an announcement by the testers that the participants were given a Test of Academic English Proficiency (TAEP). Then, the test was distributed and the try-out participants were given 60 minutes to do the grammar section and the reading comprehension section. However, they were given extra time when they could not finish the test in the given time. After all the responses and the test papers were submitted. The participant were given questionnaire asking about the test they had just worked with.

Finding and Discussion
The data of the present study were analyzed to see the item facility (IF) or facility value (FV) showing how easy or difficult the particular item is proved in the test (Anderson et al., 1995; Heaton, 1988). The formula used to compute the IF of the objective type test is taken from Heaton (1988). He asserts that the IF should range from .20 to .70.

Further, the analysis of the item discrimination is also done. This is intended to see the extent to which the item discriminates the test takers, separating those who are more able from those who are less able. In order to analyze the item discrimination index found in the test developed in the present study, Ebel’s (1997) classification is used. He classifies the discriminating index into four groups: very good item (.40 and up), reasonably good item (.30-.39), marginal item (.20-.29), and poor item (below .19).

Then, the empirical validity is obtained by comparing the results of the try-out of the test with students’ score on reading comprehension provided by the lecturer of the subject and the existing score of English proficiency resulted from the already standardized test. This method was done following Heaton (1988) who claims that one of the ways to establish empirical validity of a test is by comparing the results of the test with criterion measure, one of which is the teacher’s scores. The computation of the correlation of the two sets of scores was done using Microsoft Excel Program based on the Pearson Product Moment formula. In addition to that, the effort to see whether or not the test has high reliability is done by employing the formula of Kuder Richardson (KR21).

The results of the computation of IF or FV range from .21 to .40. Based on the suggested criteria, the items of the test are acceptable in terms of the IF values. No item falls below .20 and above .70. Then, based on the application of Ebel’s classification to the data collected from the try-out, it was found that there are four poor items that should be revised. The other items were considered good since the “good” test takers tend to do well on the items and the “poor” ones tend to be failed in the same items. Thus, the items can distinguish the “good” from the “poor” test takers.

The computation of the correlation of the two sets of scores using Microsoft Excel Program based on Pearson Product Moment formula results in correlation coefficient .98. This shows that the validity of the test is very high. Thus, the test can really measure what it is supposed to be measured. Finally, the internal consistency of the items (the reliability) is measured using the formula of KR21. The usual reliability suggested by Lado (in Hughes, 1989) is .90. and the result of the computation is .93. This index confirms that TAEP has a high reliability.
Furthermore, “in order for a particular language test to be useful for its intended purposes, test performance must correspond in demonstrable ways to language use in non-test situations,” (Bachman and Palmer, 1996: 9). Since the sole purpose of TAEP is to assess Indonesians’ proficiency in academic, business, and general English, the content of TAEP must be relevant with the settings and contexts of the purposes. Thus, the test items are made varied, yet balanced, from academic, business, to general topics.

**Conclusions**

August 17th of this year marked the 72nd of Indonesia independency. However, the country has not yet been absolutely free from the new kinds of neo-colonialism or modern colonization. In some aspects, including education, we still heavily depend on products made by other countries. With the spirit to be free from any kind of dependency, it is time for us to begin investing in our own people and our products. To achieve this aims, TAEP was made as an effort against control mechanism in language testing system in Indonesia.

As TAEP has passed all language test criteria, it has been legally certified and gained its registered intellectual copy rights. At national level, apart from schools and universities, TAEP had been administered by Language Center of University of Muhammadiyah Malang and the Assessment Division of PPM Manajemen, Jakarta, to test the English skills of applicants of Bank of Indonesia and the Financial Service Authority of Indonesia (OJK). At international stage, TAEP has collaborated with programs from Erasmus Mundus and been used as one of the candidate selection processes. Through the availability of TAEP, we hope to strive for mental revolution as proposed by the President of Indonesia and to be independent from all kinds of modern colonization.

**References**


