Looking at an English Language Study with a Critical Eye

Kathryn Zidonis
Robert Dilenschneider

Over the years, there has been much criticism, in professional fields such as education and economics, both in and outside of Japan, with regards to how English language education in Japan has lagged behind its Asian neighbors. For example, a recent editorial in The Japan Times states, “As China, Vietnam and South Korea have moved ahead, Japan’s English education policies have languished. It may be a case of too little too late. Japan’s position in the future internationalized world will be determined by the nation’s English ability.” (The Japan Times 2009)

With regard to standardized exams, which often play a key role in ranking language ability among nations, in an article in The New York Times entitled, “Difficult Lesson: Learning English,” author, Miki Tanikawa states that although Japanese are among the world’s most avid TOEFL test-takers, exceeding second place South Korea by more than 130,000 and that, “Japan trails nearly all industrial countries and most of its Asian neighbors, including China.” The article cites Ikuo Koike, professor emeritus at Keio University and president of the Japan College English Teachers Association expressing his disappointment, “You cannot explain away Japan’s low scores.” (The New York Times 1996)

However, aside from the comparison of standardized test scores across the Asian continent, there may be some concern in the manner in which some independent studies have been conducted with regard to either administering and or assessing final results. Therefore, in an effort to recognize how some studies may be misleading and how they may be enhanced for educational utility, this critical review will look at the positive and negative features of a statistical study concerning the motivation levels of Chinese and Japanese English as a Foreign Language students. After a brief overview is explained, issues concerning the literature review, survey questions, scoring, and reporting of the study will be examined. Finally, a summary as well as recommendations to enhance the reliability, replicability, validity and generizability of this type of study will be discussed.
Overview

In his article, “Motivation as a Two-Sided Coin: Motivational Differences between College-level Chinese and Japanese Learners of EFL,” author Bill Teweles conducted a study to examine whether or not students’ motivation level and language test scores correlate with language proficiency. The study involved a total of 80 participants with two groups of 10 freshman students and two groups of 10 sophomore students learning English as a Foreign Language both at Hunan University in China and Okayama University in Japan.

The first questionnaire was called an “Attitudinal Questionnaire” designed to determine if students were “instrumentally” or “integratively” motivated. Of the ten questions presented in this questionnaire, eight were evenly divided and designed to measure instrumental and integrative types of assessments. They were presented in a five point Likert scale format worth a total of 40 points. The other two questions were not part of the scoring but were used to find additional information. For example, question number seven sought a “yes” or “no” response to inquire whether or not a student was taking English for college credit, while question number ten sought an open-ended response to allow students to express the reasons while they were studying English. The second, or follow-up, questionnaire was a “Motivational Intensity Scale” which consisted of six “yes” or “no” questions to measure how active or committed students were to learning the target language.

Together, both the Attitudinal Questionnaire and Motivational Intensity Scale were scored to determine which students had “HI Motivation” or “LO Motivation” and compared to the scores students received on multiple-choice, cloze and translation based tests in order to determine whether or not motivation correlates with language proficiency. The final results of the study reported that students from both universities showed high motivation and that Chinese students scored higher on language tests than their Japanese counterparts. However, motivation level and language test scores did not highly correlate with proficiency for students of either university.

Literature Review

In describing the manner in which English language learning is perceived in both China and Japan, the author provides solid examples from each culture. For example, the need for English ability is clearly expressed when he quotes a Chinese student as saying, “English is a useful communicational tool to study advanced Western technique” and a Japanese student as saying, “(English) has become a necessity in today’s modern society.” (p.18) In addition, the author clearly describes the importance of the need to learn English for testing purposes when he states, “As many as 90% in some urban centers of China and some 99% of Japanese youth study English in middle or high school for mostly test-related reasons (p.18).” Despite the importance of English in both countries,
however, the author also adeptly describes the extent English is used in both societies when he states the following:

Visitors to Japan often marvel at the variety of ways and means English plants itself onto the urban landscape. Its extensive appearance in the media and expression in fashion is a fact of life in most Japanese cities today. In contrast, English is used more sparingly in Mainland China; while pressures to use more English in advertising exist, it is rarely used in the decorative sense that it is in Japan. (p.18)

Using these and other examples, the author sets the stage for an interesting comparison to measure the motivation for learning English as a Foreign Language in each country.

However, at least two terms within the literature review need further explanation. For example, although the author’s intention is to measure the attitudes of learning English as a foreign language in terms of “integrative” and “instrumental” motivation, these terms are not explicitly defined. Rather than explain or describe what these terms mean in the beginning of the study, the author briefly mentions the work of other researchers in this field and their work with these types of motivations in second language research in the 1950s and 60s. It is not until much later in the article that the reader is presented with more information to give a better, yet still vague, understanding of how these two terms differ from one another. For example, Gass and Selinker define integrative motivation as motivation which, “comes from the desire to acculturate and become part of a language community,” and instrumental motivation as motivation that, “comes from the rewards gained from knowing another language.” (Gass & Selinker p.333) Similarly, Ellis states that Gardner and Lambert (1972), “distinguish ‘instrumental motivation’, which occurs when a learner has a functional goal (such as to get a job or pass an examination, and ‘integrative motivation’, which occurs when a learner wishes to identify with a culture of the L2 group.” (Ellis 1994 p.715) However, it is not until later in his article that the term “integrative” appears to be the type of motivation which occurs when a student wants to learn English to become part of something while the term “instrumental” seems to be the type of motivation which occurs when students want to learn English for a certain purpose. Nevertheless, if these terms were better defined at the beginning of the article, readers would be better able to understand the purpose of the study overall.

Questions

There are a few concerns with the types of questions as well as their placement within the questionnaires. For example, a least six questions within the Attitudinal Questionnaire could be interpreted as double-barreled because they use the word “or” in the middle. According to J.D.Brown, “Double-barreled questions are ones that really ask two or more
questions at the same time. Such multiple purposes may make a question difficult to answer and certainly will make accurate interpretation of the answers nearly impossible.” (J.D.Brown p.49) For instance, question six states, “English is necessary if one wishes to travel abroad or live in another country.” This is a double-barreled question because the phrases “travel abroad” and “live in another country” are two very different activities that can require two different levels of linguistic ability. That is, a person traveling abroad may only need to verbally communicate simple transactions for directions whereas a person living abroad may need not only verbal skills but also the ability to read and write in order to interpret phone bills, insurance forms and bank statements, for example. Likewise, with regard to the second questionnaire or Motivational Intensity Scale, the question, “Do you plan to continue learning or to use English after you graduate from college?” could be confusing to participants because the phrases “continue learning” and “to use English” are asking about two different actions. A student studying to be a hotel receptionist, for instance, may plan to use the English he or she has learned, but this does not necessarily mean that he or she plans to continue learning.

There are also two questions in the Attitudinal Questionnaire that are not only double-barreled but also contain loaded words. Once again, the use of loan words may be detrimental to a study because they are words, according to J.D.Brown, “that suggest an automatic positive or negative response because they are emotionally charged.” (J.D.Brown p.50) For example, question three states, “A truly educated person should be able to read or understand written or spoken English” and question nine states, “English is necessary in order for us to become truly ‘internationally minded’ or a ‘world citizen.’” In addition to being double-barreled, both questions contain the word “truly” which is a word that holds a positive meaning and may lead participants to a response which not genuinely their own. Therefore, these questions need to eliminate “truly” as a loaded word as well as express terms and phrases separately in order to acquire better responses.

Lastly, insofar as the ordering of the questions themselves, there does not seem to be a logical reason as to why a “yes” or “no” question was positioned in the middle of a Likert scale questionnaire as in the following:

6. English is necessary if one wishes to travel abroad or live in another country.
   Strongly agree Agree Not sure Disagree Strongly Disagree
7. I am taking English mainly to gain college course credit.
   YES / NO
8. English is important in order to understand Western thought.
   Strongly agree Agree Not sure Disagree Strongly Disagree

(p.32)
Because the content of question seven does not relate to either question six or question eight, and because question seven is not included in the scoring, perhaps it would be better to write this type of question at the beginning or end of the questionnaire in order to sequence the questions rationally and prevent any confusion on behalf of the respondents.

**Scoring**

The scoring procedure expressed in this study is difficult to understand. As described previously, the first Attitudinal Questionnaire consisted of 10 questions containing eight Likert scaled questions worth a total of 40 points. As a result, the study reported that Hunan university students scored 33.45 instrumental motivation points (83.63%) and 32.85 integrative motivation points (82.13%) while the Okayama university students scored 29.95 (74.87%) and 31.65 (79.13%) for the same instrumental and integrative tests, respectfully. However, there is no mention of how these results were calculated. Another problem which exists is the issue of how both questionnaires coordinate with each other. The only information about any type of scoring procedure is as follows:

The numerical difference between “yes” and “no” responses on the second questionnaire was also intended to help determine assignment to a “HIGH” or “LOW” motivation level, thirty-two points (70%) or higher on both questionnaires signifying “HI MOtivation.” Similarly, a respondent earning twenty-eight points or fewer on the initial questionnaire and not scoring four or more points on the second part would be considered “LO MOtivation.” (p.16)

There is no indication of the point values assigned to the second Motivational Intensity Scale which contains six “yes” or “no” questions. As a result, it is not clear whether or not the scores of the Attitudinal Questionnaire and the Motivational Intensity Scale were averaged together or separately. In order to score a student as having “HI Motivation,” are 32 points needed from both the Attitudinal Questionnaire and the Motivational Intensity Scale? Or, alternatively, do the majority of points result from the Attitudinal Questionnaire while the Motivation Intensity Scale, with six questions, constitutes as one additional point for a “yes” answer?

In an effort to better understand the scoring procedure, we referred to the Microsoft Excel like graphs in the middle of the study that reported the responses of Chinese and Japanese students. We determined that if the Attitudinal Questionnaire scored a student below the 32 points required for “HI MOtivation,” then the Motivational Intensity Scale could increase the score one point for each “yes” response. For example, a score of 30 points from the Attitudinal Questionnaire could be improve to an overall score of 33 points if the same respondent answered “yes” to at least three of the six questions on the
Motivational Intensity Scale (30 points on Attitudinal Questionnaire +3 “yes” responses on the Motivational Intensity Scale=33 total points). This logic seems to follow for 75 of the 80 participants, or the entire freshman sample, as well as 35 of 40 sophomore students at both Hunan and Okayama universities. However, there seems to be a discrepancy for at least five sophomore students. For instance, an Okayama student who received a score of 31 on the Attitudinal Questionnaire and indicated four “yes” responses on the Motivational Intensity Scale was judged as being a student with “HI MOtivation” (31+4=35). On the other hand, another student who received the same score of 31 on the Attitudinal Questionnaire and indicated four “no” responses on the Motivational Intensity Scale was also judged as being a student with “HI MOtivation” (31–4=27). According to the pattern followed by the other 75 out of 80 students, this student should have been judged as a student with “LO MOtivation” because his or her score of 27 was below the required 32 for “HI Motivation.” Therefore, in all likelihood, it is possible that the researcher who conducted this study inaccurately interpreted and or reported his findings.

**Reporting**

The findings of the attitudinal questionnaire were displayed in eight separate tables for both universities: two tables for the freshmen and two tables for the sophomore students. Each table listed four separate columns which displayed student numbers, instrumental scores, integrative scores and total score. For the Motivational Intensity Scale, the number of students who responded “yes” or “no” from both freshmen and sophomores at each university were listed. As a result, findings from both questionnaire and scale used in this study were able to be interpreted.

However, there is no mention of the type of analysis used to correlate the results of the Attitudinal Questionnaire or the Motivational Intensity Scale with the results of the language tests. Other than a footnote stating, “StatView 512 was used to calculate all correlation coefficients,” there is no mention of t-tests, ANOVAs, or any other type of analysis used in this study. (p.27) Also, rather than have correlation results reported in an A.P.A. table, findings from this study were displayed in a graph format similar to what may be produced from a Microsoft Excel program. This may be acceptable for some journals, however, whether within the report of the study or within the graphs, no p values were shown to indicate the level at which findings were significant.

Furthermore, concerning the motivation and language correlation reports, there appears to be at least two discrepancies between what was written and what was displayed within the graphs. For example, with regard to the correlation of grammatical scores, the following was reported:
Meanwhile, the level of correlation between scores on a standard grammatical proficiency test were considerably higher, ranging from a low of .614 for sophomores on the last Multiple-Choice test to a high of .874 on the first translation. (p.27)

According to graphs, the first mistake is that the score of .614 was a result of a cloze test and not a result the “Multiple-Choice” test. The second mistake is that the score of .874 is incorrect and is actually a score of .825.

Finally, although the author states that the number of participants in this study is, “too small to provide an accurate indicator of how a great a factor attitudes and motivation are in the EFL context,” he nevertheless concludes, “general verbal ability shows a consistently higher correlation with performance on a variety of performance tests … than a Motivational assessment.” (p.28) However, there are two problems with this assessment. First, the article lacks a clear explanation as to how motivation level and proficiency were correlated. Second, the only proficiency scores of language tests presented in this study were the ones revealed in the Microsoft Excel type graphs which, as mentioned previously, have been either calculated or reported incorrectly.

**Summary and Recommendations**

Author Bill Teweles presents an interesting study to examine the motivation levels in learning English as a Foreign Language at a Chinese and a Japanese university. He clearly explains how students from both countries understand the necessity and importance of learning English as well as demonstrates the degree in which both nations differ in their use of the language. However, there are some shortcomings in this study as well. First, with regard to the questionnaires, terminologies such as “instrumental” and “integrative motivation” are not clearly explained. Also, the wording of several questions is confusing because they are double-barreled and/or carry loaded words. In addition, with regard to manner in which motivational scores correlate with language test scores, there are no copies of the language tests given, no information as to the type of analysis used, and no p values shown to verify at what level the final results proved significant. This lack of data, along with scoring and reporting inaccuracies, seriously challenge the credibility of this study.

Nonetheless, this type of study could produce interesting information of how the motivation to learn English as a Foreign Language differs between two economically prosperous Asian nations. Therefore, in order to improve the reliability and replicability of this study, questions should be refined in order to avoid confusion and gather specific inquiries. In addition, it is essential to elaborate on the scoring and analyses procedures as well as provide copies of the language tests so that the correlation of the final results can be understood and, if necessary, reproduced. With regard to the external validity and
generalizability, or the degree to which the results can be applied to the outside world, the number of participants should be increased in order to diminish the possibility that the final results occur through chance. In addition, although this study focused on only one university in China and one in Japan, the overall findings may have a greater impact and appeal to educational institutions if several universities from various areas of each country were included.

For example, in a drive to enhance students’ English communication skills, MEXT, Japan’s Ministry of Education, Culture, Sports, Science and Technology, laid out a new teaching guideline to teach English primarily in English, as well as increase the number of English vocabulary to be taught at each academic stage. In addition, efforts are being implemented to match the international learning environment. As reported by The Japan Times, “The level of the vocabulary would match that taught in China and South Korea.” (The Japan Times, 2009) Therefore, because of the frequency that South Korea comes into play with Japan and China when it comes to English education, it would make for an even more academically useful study to expand the research to include Korea, as well as other nations in Asia.

Learning about English education from other Asian nations may be the biggest asset for Japan. Marcus Noland, a Senior Fellow at the Institute for International Economics, notes, “South Koreans outscore Japanese in TOEFL scores,” and, “have more exposure to the outside world” as more of the population undergoes training overseas. In contrast, Japan exercises less international practices. (Research Institute of Economy, Trade and Industry 2002). In recognition of this, MEXT stated that they would be, “Concentrating mainly on Asian countries, reports on the state of English education in other countries concerning teaching methods, teaching materials, techniques for evaluation, and approaches to teacher training.” (MEXT 2003) However, in collecting any data, language educators should be mindful to examine studies with a critical eye.
Bibliography


