

〈研究ノート〉

〈Note〉

二つのプレイスメントテストの比較

TOEIC Bridge Test Sample : Results and Analysis — Comparing two placement tests —

ジェリー・ブッシュ^a

Jerre BUSH^a

ロバート・ジェイムズ^b

Robert JAMES^b

黒崎敬治^c

KUROSAKI Takaharu^c

^{a,b,c} 上武大学経営情報学部, 〒370-1393 群馬県多野郡新町270-1

^{a,b,c} Faculty of Management Information Sciences, Jobu University, Shimmachi, Gunma, 370-1393, Japan

受付 2005年10月5日

Received 5 October 2005

Abstract

The English Center of Jobu University administered a sample of the TOEIC Bridge test to first year students with a view to encouraging students to use the TOEIC study software that is available for student use, and to compare its performance as a placement instrument with the in-house placement test. A correlation study found a clear but not strong relationship between the two instruments. This paper also discusses the rationale for promoting business English in the curriculum at Jobu University.

Key words and phrases : placement test ; test validity ; proficiency examination ; correlation ; English for Specific Purposes (ESP)

TOEIC Bridge Test Sample : Results and Analysis

— Comparing two placement tests —

Jerre BUSH

Robert JAMES

KUROSAKI Takaharu

Introduction

In the Spring Semester of 2005 academic year, Jobu University Center for English Studies administered a shortened version of the TOEIC Bridge™ business English proficiency examination to freshman English students. The rationale for administering this test was twofold. Firstly, we hoped the use of this test would make students aware of the possibility of further study of business English at Jobu University. And secondly, we hoped that this test would provide some insight into the validity of the placement test presently used to stream freshman students, by acting as a second indicator of student proficiency levels.

1. The study of business English at Jobu University

Although our school is business oriented (with students majoring in Information Sciences and Business Studies) the English curriculum at present does not contain a specific business English orientation. It is therefore possible and even likely that English study at Jobu will be viewed as simply a general subject rather than one of relevance to the careers of the students. Use of the TOEIC Bridge test is one way in which we are trying to orient the students to English studies more relevant to their major study areas.

The school on occasions in the past has offered business English classes either as part of its elective component or within the required English curriculum. However, these courses were not popular or well-subscribed and a possible reason for this may be that the students could not see their relevance to their future career paths, or to their other subjects. Another reason is probably that the English proficiency level of the students is only at the elementary level. Studying specialized English becomes especially difficult and demotivating when the basic building-blocks of the language have not been sufficiently mastered. It is hoped that the possibility of gaining an external credential in business English might provide a source of

motivation for students who otherwise would see little benefit in further study of English at the tertiary level.

After the test was administered, students were given printed information about the TOEIC Bridge test, and invited to attend a series of introductory talks about the test and test preparation studies. The school has purchased software designed for TOEIC test preparation, which can be used by the students studying alone in their free time, without the need for an instructor. Sufficient interest in this approach by the students would provide a rationale for re-introducing business English classes into the general first or second year English programs.

2. Placement testing at Jobu University

At present, incoming freshmen are required to take an English placement test. This test is comprised of a grammar, a reading, an aural discrimination and a vocabulary section. The test was developed by the English teaching staff. Past uses of the test have found it to be a satisfactory instrument for placement purposes. The range of English abilities found within the English program does not appear great. This is hardly surprising given the standardized English teaching to which the majority of the incoming students have been exposed. Nevertheless, our teaching staff is generally satisfied that the test succeeds in identifying those students who are capable of working with more challenging material in the curriculum, and those who are weaker than average and benefit from classes that are paced more slowly and emphasize review and revision of basic language elements.

The four sub-components of the placement test mentioned above were chosen and designed to reflect the skill areas that the students will encounter in the English curriculum. One component that is obviously missing from the test, however, and which is of central importance within a general English curriculum, is the speaking skill. The large-scale nature of the placement testing procedure, and the time constraints under which it is administered (testing is limited to one hour and the streamed class lists are required on the same day), make it impractical to implement a speaking component at present. Given that caveat, however, the teaching staff regards the range and level of skills tested as being an argument in support of the content validity of the test.

The use of the TOEIC Bridge test for motivational purposes, as mentioned in the previous section, presents the opportunity to provide some form of criterion-related validity for the present placement test. According to Brown (1996), criterion-related validity is essentially one aspect of construct validity, in the sense that a relation is established between the test under consideration and some criterion measure, usually another test which is already regarded as having construct validity. For our purposes, a strong relationship in terms of students' scores between the in-house placement test and the Bridge test would provide a basis for confidence that our own test had a similar kind of validity as the criterion test.

Two questions arise when the TOEIC Bridge test is approached as a validating instrument. Firstly, is the Bridge test itself a valid test; and secondly, is it an appropriate instrument to use for validating the particular test under consideration?

Regarding the first question, the TOEIC Bridge test was developed by a well-respected educational testing institution as a test of general English proficiency¹, which implies a well-researched instrument. The shortened version of the test that was used in this study, however, was designed to be a sample test, to illustrate the form and content of the real test, rather than being an instrument developed to the same rigorous standards, and so its validity as a measure even of the English that it claims to measure, must remain an open question. Any conclusions that are reached concerning the criterion-related validity of the in-house test based on this sample test, therefore, must also be viewed as illustrative only and indicative of possible further research.

Is the TOEIC Bridge an appropriate instrument to validate our in-house test? Firstly, the purpose of both tests is to test general English. Although the TOIEC tests are predominantly associated with business English, the stated purpose of the test designers was to develop a test of general English competence². In terms of test design, there are some similarities between the tests: both are designed to take about one hour to complete. Both focus on the passive skills of listening and reading. On the other hand, the in-house test also contains vocabulary and grammar sections, while the Bridge test only claims to measure these competences indirectly, through the listening and reading sections (it offers sub-scores in language functions, listening, reading, vocabulary and grammar. The TOIEC test proper does contain a

grammar/structure section but this is not included in the Bridge version.)

Secondly, both tests are aimed at elementary level students. The Bridge test publishers recommend the test for those students whose TOEIC test score falls below about 450 to 470 points, this being the level of typical university graduates entering the workforce. In developing the in-house test, our texts and structure items were culled from elementary level syllabus items, while the vocabulary items were selected from the first three levels of the Cambridge English Lexicon (Hindmarsh, 1980). The above considerations led us to expect that the TOEIC Bridge test would be a reasonably valid criterion-measure of general English for our in-house test.

Finally, while it appears that the purpose of the two tests is quite different (placement versus proficiency measure), according to the publishers of the test, the Bridge test can also function as a placement test. The TOEIC website (English version³) claims that “The [TOEIC proper] test also finds many different applications in universities, being used as a placement test or as a part of entrance requirements.” Indeed, given that placement tests are essentially proficiency tests focused on specific ability bandwidths, the Bridge test (with its narrow range) may even function more effectively as a placement instrument than the TOEIC proper is claimed to do, for the appropriate kind of student. Whether the TOEIC Bridge test could function as a placement instrument in the context of the Jobu English program is an empirical question which is addressed in the correlation study reported below.

3. Business English and General English

Both the in-house placement test discussed here and the TOEIC Bridge tests are regarded as general English proficiency test. However, as has been mentioned, the TOEIC tests are generally associated with business English and are widely used in Japan as a language credential of particular usefulness for the business world. In this section, we will briefly review the relationship between business English and general English in order to further build a case for using the TOEIC Bridge as an introduction to business English for our students.

Within ELT (English language teaching) theory and practice, business English is regarded as being a subset of ESP (English for specific purposes), while ESP itself is regarded as a subset

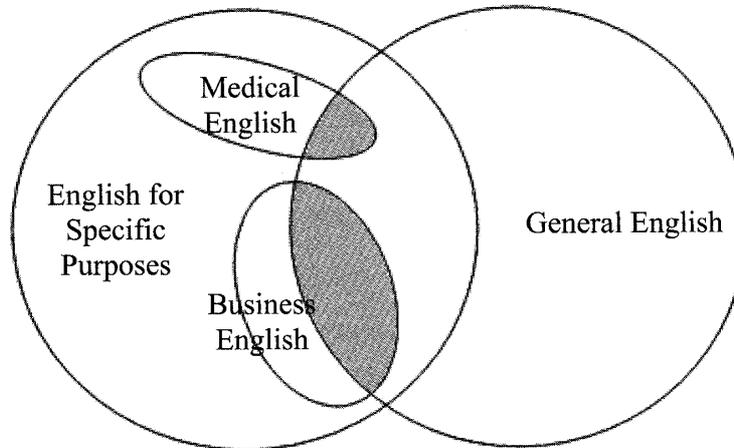
of general English. How these subsets are characterized, however, has changed as theorists have changed their views of the nature of language, discourse and communication. During the 1960s, as the study of ESP was emerging as an independent discipline, ESP was regarded as a set of specific registers, such as Electrical Engineering English, Biology English, etc., and theorists tried to identify the distinctive grammatical and lexical features of these registers. However, register analysis revealed that often there was in fact very little that was distinctive about registers in terms of grammar or lexis. So a second phase of development shifted attention beyond the sentence level to a focus on how sentences were combined in discourse. This newer approach assumed that specialized use of English would exhibit differing rhetorical patterns of text organization. Again, however, it was found that the typical discourse structures found in, say, academic texts, could also be found in general discourse. In later developments, ESP theorists moved away from rigorous analysis of linguistic features to attempts to identify and classify the situations in which the language would be used by the specialists concerned; then, additionally, to the types of skills and strategies learners would need to function in these situations; and finally towards more general learning needs of specialist learners.

Regarding the difference between business English and other forms of ESP, Robinson, quoting Pickett (1989, p. 11) makes the important point that business English is ...

... a mediating language between the technicalities of particular businesses ... and the language of the general public. It is not purely for intra-group communication. This is not surprising since business and commerce are by definition an interface between the general public and the specialist producer ...

Thus, while business English needs to provide the tools for expressing the meanings necessary for conveying the roles, activities and concepts of the business world, the need also to provide an interface with the general public acts to prevent business English from straying too far from general English in terms of its structure, lexis and rhetorical functioning. So, typically, business English will be less specialized than, say, English for science and technology, medical English, or legal English.

These relationships can be summarized by the following diagram.



As the diagram illustrates, the wider “interface” between business English and general English ensures that business English will maintain a more general character than other specialized uses of the language, such as medical English. It may be worth remembering, however, that the interface between the public and specialized uses of language is not set in stone, and may change as society itself changes. This is particularly noticeable in English for science and technology, where the recent widespread dispersal of computers and telecommunications among the general public has resulted in a plethora of formerly specialist lexis becoming part of common usage. While a large portion of such innovations will prove ephemeral, some will no doubt find a more permanent place and thus contribute the growth and development of general English. Business English might be characterized as one form of specialized usage in which this particular process has advanced more than most specialist Englishes do.

A final point concerning business English: its more general nature is in fact borne out by considering the marketing of business English courses and the published material used in such courses. The tendency in language and business schools is towards offering “open door” business English courses aimed at providing skills of a general character that can be deployed in a wide range of business contexts. The publishing industry, no doubt responding to the needs of these customers, produces far more general business English texts than those for more specialized aspects of business. While texts can be found that address the needs of students in, say, the tourist industry, hotel and catering, secretarial and clerical work, and business negotiations, such texts are the exception rather than the rule.

Additionally, business English texts of the general variety are typically available in a wide range of levels, from elementary to advanced learners. This situation is certainly true of the TOIEC materials available, with texts typically claiming to prepare students for a variety of score levels on the test, such as TOEIC 400, 600 or 900 levels.

THE STUDY

As mentioned in the introduction, a comparison was made between the sample Bridge test and the in-house placement instrument, with a view of discovering the degree of correlation between the two tests. In addition, average scores on the two tests for classes involved in the study were compared to see if they produced similar results; again, this would provide some validation for the placement test.

1. Method

Freshman students enrolled in five classes of the required English classes (Oral English I & II), were administered the TOIEC Bridge sample test early in their first semester of study, (after approximately four weeks of study). This test consisted of a listening section and a reading section, each section being comprised of three questions or sub-sections. The listening section took up approximately five minutes, and the reading section ten minutes, for a combined total of 15 minutes.

The test was administered by classroom instructors in the context of a normal lesson. Answer sheets were collected after the test and scored by the present authors. Scores were inputted into the computer. Students' placement test scores were also inputted. A number of students had not taken the placement test for a variety of reasons, and these were excluded from the analysis.

Descriptive statistics were generated by the computer to determine whether the data met the statistical assumptions for a correlation analysis and also whether the two sets of scores seemed to be meaningfully similar.

For the correlation study, the Pearson product-moment correlation coefficient was used to

test the null hypothesis of no significant relationship between the data sets; and for the hypothesis of some degree of relationship, a positive direction was assumed. Since this was essentially an exploratory study, alpha level was set at $\alpha < .05$ based on the similarities of the tests in terms of purpose, level and design. Regression statistics and a regression line were also calculated to investigate the extent of the covariance between the two sets of data.

To further assess the similarity of the tests, a comparison was made of the class average scores on the two tests for the five classes used in the study.

All calculations were performed using Microsoft Excel

2. Results

a. Descriptive statistics.

Histograms of the students' scores on the two tests were generated and compared. As can be seen from Figure 2, the scores for both tests form a relatively normal distribution curve, suggesting that both tests had an appropriate fit to the level of the students, and also that the subsequent correlation study would be statistically appropriate.

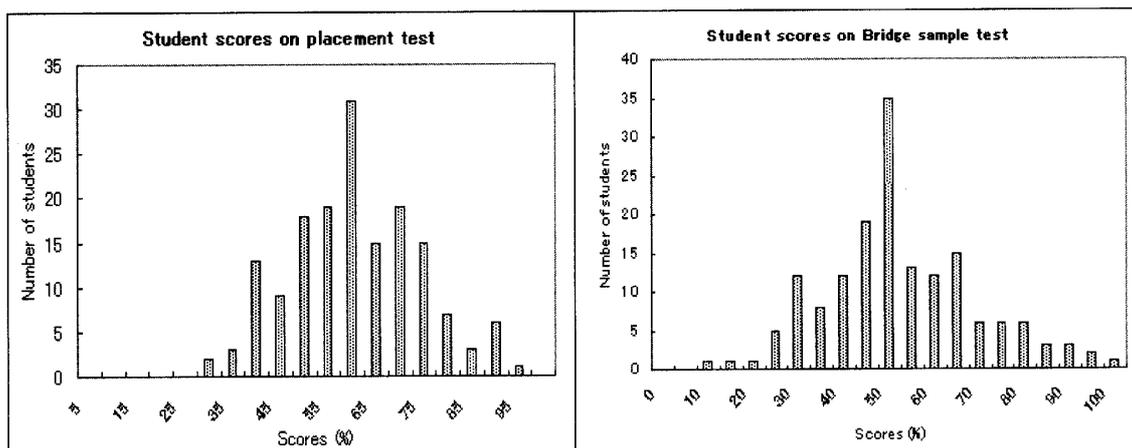


Figure 2: Student scores on the placement test (left) and the Bridge sample test (right).

Table 1 shows the statistical characteristics of the above distributions. As the table shows, in both of the tests, measures of central tendency (mean, median and mode) are almost identical. There is also little kurtosis and skew in the distributions. The standard deviation on the Bridge score is somewhat higher than the placement test, as is the range. Overall, however,

our assessment of the data sets was that they provided an adequate basis for the comparisons that we were interested in.

Table 1. Descriptive statistics for the placement and Bridge tests.

<i>Statistic</i>	<i>Placement scores (%)</i>	<i>Bridge scores (%)</i>
Mean	58.5	50.4
Median	58	50
Mode	60	50
Standard Deviation	14.2	17.1
Kurtosis	-.1	-.2
Skewness	0	.3
Range	79	86
Count	161	161

b. Correlation study.

The Pearson product-moment correlation coefficient for the two sets of scores was computed, and a scatter plot generated. Figure 3 show the scatter plot with the regression line added. The regression line indicated a fairly linear relationship between the two groups of test scores; however, there is a fair amount variation evident, as indicated by the standard error of estimate of 14.14 shown in table 2 below.

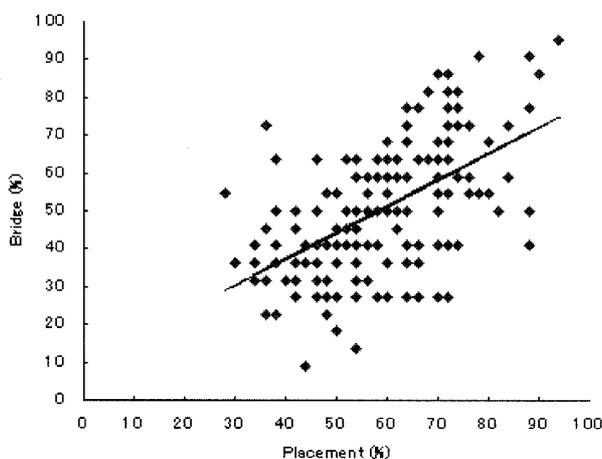


Figure 3 Scatterplot and regression line for the two tests

The correlation study of the two tests produced the results shown in Table 2. As this table shows, a correlation coefficient (r_{obs}) of .56 was obtained for the correlation between the two data sets. The r_{crit} for this study was .1638, allowing us to reject the null hypothesis of no relationship between the two distributions. The r^2 for the correlation was .32, indicating that 32 % of the variance in the Bridge test scores was explainable by placement test scores, with 68 % of the variance remaining unexplained.

Table 2. Correlation and regression output for the placement and Bridge tests

<i>Statistic</i>	<i>Results</i>
r_{obs}	.56
r^2	.32
Standard Error	14.1
Slope	.7
Intercept	9.5
Observations	161
r_{crit}	.1638
Standard error of estimate	14.14

c. Class-based comparison of test scores

Table 3 shows the class averages for the five classes in the study, for both tests

Table 3. Class-based averages and correlation between the scores

<i>Class</i>	<i>Placement scores</i>	<i>Bridge scores</i>	<i>Correlation</i>
Class A	69.2	53.5	.3
Class B	56.8	48.3	.27
Class C	44.1	32	-.06
Class D	75.1	67.1	.28
Class E	47	49.7	.31

The correlation between the two tests at the level of class averages is significantly weaker

than was found for the entire population of the study. These class-based correlations, with the exception of Class C, are about .3 as compares with the .56 achieved for the entire 161 subjects in the study. This suggests that streaming the students by placement score results in ability bands somewhat different to those that would be obtained by using the Bridge test for the same purpose. Also of note is the anomalous result that Class E obtained a higher score of the Bridge test than on the Placement test. The correlation, however, was the highest of the five classes, while Class C achieved a negative correlation.

3. DISCUSSION

The distribution of placement and Bridge test scores illustrated in Figure 2 and the descriptive statistics in Table 1 show both the test to perform adequately in terms of distributed scores and suggest a measure of similarity between them. However, the correlation coefficient obtained of .56 is too weak to support the claim that the two tests are measuring exactly the same traits.

What could explain this weaker than expected correlation? Firstly, the tests, while both generic tests of competence may, by the nature of their primary function, give weight to measuring differing competencies. The placement test aims at measuring competencies in terms of skill areas addressed, while the Bridge test, with only two forms of items, necessarily has to emphasize those areas that are predictive of other skill areas not tested directly.

Another possibility, perhaps suggested by the weak correlations found in the class-based comparisons, is that students at such low proficiency levels do not have an integrated proficiency base to draw upon, but rather have pockets of knowledge which allows them to perform well in some areas, while remaining deficient in others.

Despite the fact that no strong correlation between the tests was found, given the shortness of the sample test (15 minutes and 22 items only), and the specialized nature of the in-house placement instrument, the correlation found is considered by us to be reasonable and offers some support for the notion that our own test is a valid test of general English.

These results suggest further exploration of the relation between our in-house test and the

TOEIC Bridge test could provide useful information for the English program. In particular, a stronger relationship with the full Bridge test would enable us to generate estimates of potential scores for our students on this external examination from our own placement test. Such information could also have a beneficial effect on student motivation.

Another possibility would be the development of an alternative in-house instrument as a more general proficiency test, which could be used to measure students progress during their English studies at Jobu. Such an instrument, free of the security and time constraints that are necessarily imposed on a placement test, could be developed with a view of providing predictive information about students' performance on either one or both of the TOEIC tests.

Notes

1. Both TOEIC test were developed by Educational Testing Service
2. The TOEIC website claims the TOEIC Bridge test is "is aimed at measuring the proficiency levels of beginning and intermediate learners. " (See "What is the TOEIC Bridge test?" at the TOEIC website)
3. http://www.toeic.or.jp/toeic_en/

References

- Hatch, E. & Lazaraton, A. (1991). *The research manual*. Boston: Heinle & Heinle
- Hindmarsh, R. (1980). *Cambridge English lexicon*. Cambridge: Cambridge University Press.
- Brown, J.D. (1988). *Understanding research in second language learning*. Cambridge: Cambridge University Press.
- Brown, J.D. (1996). *Testing in language programs*. New Jersey: Prentice Hall Regent.
- Hutchinson, T. & Waters, A. (1987). *English for specific purposes*. Cambridge: Cambridge University Press.
- Pickett, D. (1989). The sleeping giant: investigations in business English. *Language International*, 1 (1), 5-11.
- Robinson, P. (1991). *ESP today: A practitioner's guide*. Hemel Hempstead: Prentice Hall International.