Components of Engineering-English
Reading Ability?

J. D. Brown

In November 1979 a project was begun at UCLA to explore, develop and evaluate a methodology for creating ESP tests. Our initial interest focused on EST for engineers because of the large number of foreign students in the UCLA Engineering Department (one of the largest departments on campus). Based on Candlin et al (1978: 190-219), we further narrowed our inquiry to the EAP lecture-listening and reading comprehension skills involved in engineering English. The next step in this process was the development of three videotaped lecture listening tests (see Erickson 1983, Miller 1982 and Erickson and Molloy 1983) and three reading comprehension tests (see Kenyon 1982, Molloy 1983 and Erickson and Molloy 1983). Each of these tests was made up of 30-60 multiple-choice items drawn from "genuine" engineering lectures or readings. All of the items were necessarily developed in conjunction with engineering professors (see Brown 1981: 1-3 for a full description of this entire process).

Based on analysis of the three reading comprehension tests above, a "final" revised version was developed. Item analysis techniques were used to select the "best" items and the result was a sixty-item test made up of three passages with twenty items each. Brown (1983) indicated that this "final" version was reasonably reliable and valid for the norm-referenced testing of engineering English ability among Chinese students (for more detail see Materials below). As such, it will serve as one instrument for
further investigation of what some of the components of engineering reading comprehension might be.

In order to accomplish this, a closer look at the various item types employed in the test is in order. Initially we had very little idea of what to look for in engineering reading so we chose to use a variety of item types based on ESP theory, linguistic theory and input from three engineering professors. Ten different types of items were thus included. These were viewed as falling into one of two categories, linguistic or engineering, as follows:

A. Linguistic factors
   1. Cohesion (after Halliday and Hasan 1976)
      a. Reference items
      b. Substitution items
      c. Lexical cohesion items
      d. Conjunction items
   2. Non-technical vocabulary items

B. Engineering factors
   1. Fact items
   2. Inference items
   3. Lexis (after Cowan 1974: 389-399) and Inman 1978: 242-249
      a. Subtechnical vocabulary items
      b. Technical vocabulary items

The overall purpose of the present study was to investigate the above ten item types as possible components of engineering reading ability, their relationships to general English language proficiency and their relationships to each other. To those ends the following research questions were formulated:

1) Does the test significantly distinguish between Chinese engineering and TEFL students?
2) Which types of linguistic/engineering items best distinguish the ability to read engineering English?

3) To what extent is the ability to read engineering English accounted for by general English language proficiency?

Method

Subjects

All of the subjects in this study were university graduate students. They were 50 percent American native speakers of English and 50 percent Chinese non-natives. Each of these nationalities contained equal numbers of engineering students and students of teaching English as a second/foreign language (TESL/TEFL).

The American engineering students (n = 29) consisted of 23 males and six females. They ranged in age from 22 to 50 with a mean of about 28 and came from nine different specializations with the majority in chemical or mechanical engineering. The American TESL students (n = 29) were all students in the TESL section at UCLA. There were five males and 24 females who ranged in age from 23 to 52 with a mean of 30. The Chinese engineering students (n = 29) were sampled from all engineering students studying at the Guangzhou English language Center at Zhongshan University in the People's Republic of China (an EST program). There were 19 males and 10 females. They ranged in age from 24 to 42 with a mean just under 32. They represented eight different specializations with the majority in mechanical, electronic, computer or chemical engineering. The Chinese TEFL students (n = 29) were sampled from
all TEFL trainees in the Foreign Languages Department at Zhongshan University. There were 19 males and 10 females ranging in age from 24-42 with a mean of 26.

In terms of sampling, the American engineering group included all 29 volunteers, while 29 were randomly selected from the American TESL student volunteers. The Chinese engineering and TEFL students were selected in matched pairs, based on overall language proficiency and sex, from their respective groups.

Materials

Two measures were used in this study: the engineering reading test and a cloze test. The engineering reading test was based on three passages of 600-800 words on the following topics: 1) The mechanics of deformable bodies, 2) refractories and 3) thermodynamic analysis of heat pumps. These were topics felt by the engineering professors to be common to most engineering students. Brown (1983) indicated that the test was reliable (K-R20) at approximately .84 for the American students and .85 for the Chinese students in this study. The validity of this measure was demonstrated through analysis of variance techniques. The test was shown to separate students who did have knowledge of engineering English from those who did not. Nevertheless, the question of validity will be reviewed here.

The cloze test used here was on the relatively neutral topic of "Man and His Progress." It contained 399 words with an every
7th word deletion pattern for a total of 50 blanks. It was scored here using the acceptable-word method, i.e., any word acceptable to native speakers in a given blank was counted as correct. This method applied to the cloze test \((n = 55)\) was found to be reliable \((K-R20)\) at .95 (Brown 1980: 311-317). It was also viewed in that study as a valid measure of overall English language proficiency because it was found to have a high degree of relationship \(r_{xy} = .90\) with the much longer UCLA English as a Second Language Placement Examination. Hence, the cloze test is used here as a reliable and valid test of general English language proficiency.

**Procedures**

Both tests were administered to the Chinese subjects under controlled classroom conditions. The engineers and TEFL trainees were tested separately, but under very similar conditions. They were allowed two hours and thirty minutes to finish the engineering reading test and 30 minutes to complete the cloze test. All students finished both tests well within these time limits.

The American subjects were a bit problematic in that there was no mechanism for convincing them that they must, or even should, take the engineering reading test. Hence, it was necessary to rely on volunteers. These groups were also allowed two hours and thirty minutes to complete the test. The cloze test was not administered to this group.

All tests were scored by computer analysis of dichotomously coded (i.e., right or wrong) answers for each item. No correction for guessing was applied because speed was not felt to be a factor in score variation; the time limit was more than adequate for all groups.
Analysis

The analyses will be discussed here as they apply to each research question in turn. Please note that the alpha level for all statistical tests was set at .01.

1) Does the test significantly distinguish between Chinese engineering and TEFL students? The question being asked here was basically whether or not the difference in mean performance between the two groups occurred by chance alone. Thus, rejection of the null hypothesis (of no significant difference at \( p < .01 \)) would indicate 99% probability that this difference was caused by other than chance factors. To accomplish this, a straightforward two-tailed \( t \) test (\( df = 1, 56 \)) was performed. Though the sample sizes were equal, variances were not assumed to be equal. Hence, they were estimated separately rather than pooled.

2) Which types of linguistic/engineering items best distinguish the ability to read engineering English? Perhaps mercifully, no inferential statistics were applied here. Instead, the purpose of this analysis was strictly descriptive. The ten item types were examined as average percentages of correct responses within each type. These percentages were calculated for various combinations of nationality and major, as well as for different composites of item types.

3) To what extent is the ability to read engineering English accounted for by general English language proficiency? The essential question at issue here was whether any observed correlation differed significantly (\( p < .01 \)) from zero with 99% probability. To do this, Pearson product-moment correlation coefficients (\( r_{xy} \)) were calculated. Shared variance (\( r_{xy}^2 \)), which is obtained by squaring the
correlation coefficient, can be interpreted as the percentage that two sets of scores overlap, or the percentage that the results on one test account for the results on another test. In this manner, the extent to which general English language proficiency accounted for the results on the engineering reading test was determined. A correction for attenuation was used to eliminate the effect of unreliable variance in the tests themselves on the analysis of the relationship between the two variables, general English language ability and engineering reading ability.

Results

The results will be presented here as a technical report. They will be discussed in more practical terms in the Conclusion below. For the sake of clarity, they will also be organized under research question headings.

1) Does the test significantly distinguish between Chinese engineering and TEFL students?

As a prelude to other statistical tests, the descriptive statistics should first be examined. The mean ($\bar{x}$), range, standard deviation(s), standard error of measurement (s.e.m.) and number of subjects (n) for each group is presented in Table 1. Note that the mean scores for both groups of native speakers were higher than those for non-natives and that average performance for engineers was higher than for non-engineers. This phenomenon will be discussed in more detail below.

Insert Table 1 about here
In direct answer to the first research question, the $t$ test performed indicated that the difference of 9.59 points between the mean scores for Chinese engineers and TEFL students was, in indeed, significant at $p < .01$, two-tailed ($df = 1, 56; t_{obs} = 4.89, t_{crit.} = 2.70$).

2) Which types of linguistic/engineering items best distinguish the ability to read engineering English?

The performances of American and Chinese engineers and TESL/TEFL students were calculated in average percentages for the different item types. Those for linguistic and engineering items are presented in Table 2. Those for all ten individual item types are reported in Table 3. Notice that the pattern found in mean performance (i.e., in Table 1, all natives outperformed non-natives and all engineers out-performed non-engineers) is also found here regardless of which item type or combination of item types is examined. This phenomenon will also be considered below.

3) To what extent is the ability to read engineering English accounted for by general English language proficiency?

The correlation coefficients ($r_{xy}$) calculated between scores on the engineering reading test and those on the general English proficiency measure, as well as their squared values ($r_{xy}^2$) are presented in Table 4. None of these correlations appears to be due to chance alone ($p < .01$). However, a relationship between scores can be statistically significant without being particularly meaningful. This too will be discussed below.
Discussion

In direct answer to the first research question, it appears that the test does significantly distinguish between Chinese engineers and TEFL graduate students. This adds support to the arguments in Brown (1983) that the engineering reading test is valid, i.e., it separates those who are more likely to know engineering English from those who are much less likely to be familiar with it. However, returning to Table 1, notice that this statement only holds true within nationalities. Across nationalities, it turns out that the American non-engineers (TESL) scored higher on average than the Chinese engineers on the reading test. How is this possible if the specific nature of engineering English is indeed specific?

To explore this problem, the individual item types were examined. Notice that in Table 1, the mean scores descend from high to low from left to right. In Table 2, the same pattern emerges for the broad categories of item types: linguistic and engineering factors. Table 3 further reveals that the same pattern maintains for each and every item type. Does this indicate that it may not matter what kind of question is asked as long as it is based on a genuine engineering passage?

Perhaps, but a closer examination of Table 3 seems to indicate that some item types (see*) are more efficient than others. For example, let us rather abitrarily examine those items that
separated the highest group score from the lowest by 50 percentage points or more. Note that this will include "substitution", "inference" and "technical vocabulary" items. Perhaps these are somehow better item types for such a test while items like "reference", "conjunction" and "non-technical" vocabulary appear to be least efficient.

On the whole, Table 2 seems to show that the engineering items are more efficient than the linguistic ones. There are only 34 percentage points between high and low group scores for the linguistic items, while the same figure for the engineering items is 49 points. Using only engineering items might also be more justified from a theoretical standpoint as more "authentic" engineering tasks after Widdowson's (1978: 80) distinction between genuine and authentic.

Collectively, the items have yet another interesting characteristic. To some degree at least, they appear to be tapping general English language proficiency. Remember that the squared correlation coefficients reported in Table 4 indicated that 62 percent of the variation among Chinese engineers was accounted for by general English language proficiency. At the same time, this figure for Chinese TEFL students was 42 percent and for both groups combined it was only 30 percent. These results may seem confusing at first glance. How is it possible that general language proficiency plays a more important role in engineering reading when the groups are separated than when they are combined? One possible explanation is that, when the two groups are combined, error variance (e.g., in the form of differences among individuals'
majors) may enter the picture. When the error variance increases, the correlation squared (or percentage of shared variance) drops. Nevertheless, it was the Chinese engineering students who were of interest here because it was for such students that the test was originally designed.

How is it possible that 62 percent of differences among these engineering students in apparent engineering reading ability can be accounted for by general language proficiency? And what accounts for the other 38 percent of score variance? Something specific to engineering English? Engineering content knowledge? Student motivation? Or, just random testing error? These and many other questions arose in the process of doing this study; some are included here in the hope that they will aid in the ongoing search for an understanding of what ESP is and how best to teach and test it:

1) Would the same results be obtained if this study were replicated?

2) How accurately do the item types reflect the criteria upon which a criterion-referenced interpretation should be based?

3) What psychological constructs other than general English language proficiency contribute to the ability to read engineering English?

4) How can the other ESP skills most effectively be tested?

5) How do these skills relate to each other within EST as well as to the same skills in the more general English for academic purposes?
References


Components of Engineering English Reading Ability?

J.D. Brown

Tables
Table 1: Descriptive Statistics

<table>
<thead>
<tr>
<th>Statistic</th>
<th>American Engineers</th>
<th>TESL</th>
<th>Chinese Engineers</th>
<th>TEFL</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{X}$</td>
<td>50.52</td>
<td>44.79</td>
<td>36.97</td>
<td>27.38</td>
</tr>
<tr>
<td>range</td>
<td>38-56</td>
<td>22-53</td>
<td>17-54</td>
<td>17-43</td>
</tr>
<tr>
<td>s</td>
<td>3.91</td>
<td>6.94</td>
<td>8.22</td>
<td>6.63</td>
</tr>
<tr>
<td>s.e.m.</td>
<td>2.18</td>
<td>1.69</td>
<td>3.39</td>
<td>3.57</td>
</tr>
<tr>
<td>n</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
</tr>
</tbody>
</table>
Table 2: Linguistic and Engineering Items

<table>
<thead>
<tr>
<th>Item Type</th>
<th>Americans</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Engineers</td>
<td>TESL</td>
<td></td>
<td>Engineers</td>
<td>TEFL</td>
</tr>
<tr>
<td>Linguistic</td>
<td>91%</td>
<td>77%</td>
<td></td>
<td>69%</td>
<td>57%</td>
</tr>
<tr>
<td>Engineering</td>
<td>85</td>
<td>74</td>
<td></td>
<td>65</td>
<td>36</td>
</tr>
</tbody>
</table>
Table 3: Individual Item Types

<table>
<thead>
<tr>
<th>Item Type</th>
<th>Americans</th>
<th>Chinese</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Engineers</td>
<td>TESL</td>
</tr>
<tr>
<td><strong>Linguistic Factors</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference</td>
<td>82%</td>
<td>67%</td>
</tr>
<tr>
<td>Substitution*</td>
<td>100</td>
<td>79</td>
</tr>
<tr>
<td>Lexical Cohesion</td>
<td>94</td>
<td>64</td>
</tr>
<tr>
<td>Conjunction</td>
<td>85</td>
<td>80</td>
</tr>
<tr>
<td>Non-technical Vocabulary</td>
<td>97</td>
<td>95</td>
</tr>
<tr>
<td><strong>Engineering Factors</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fact</td>
<td>89</td>
<td>81</td>
</tr>
<tr>
<td>Inference*</td>
<td>90</td>
<td>69</td>
</tr>
<tr>
<td>Subtechnical Vocabulary</td>
<td>71</td>
<td>65</td>
</tr>
<tr>
<td>Technical* Vocabulary</td>
<td>80</td>
<td>52</td>
</tr>
<tr>
<td>Rhetorical Functions</td>
<td>92</td>
<td>91</td>
</tr>
</tbody>
</table>

* Difference of more than 50 percentage points between high and low mean scores.
Table 4: Correlation Coefficients (Corrected for Attenuation)

<table>
<thead>
<tr>
<th>Statistic</th>
<th>All Chinese</th>
<th>Chinese Engineers</th>
<th>Chinese TEFL</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r_{xy}$</td>
<td>.55*</td>
<td>.79*</td>
<td>.65*</td>
</tr>
<tr>
<td>$r_{xy}^2$</td>
<td>.30</td>
<td>.62</td>
<td>.42</td>
</tr>
</tbody>
</table>

*p < .01*