ABSTRACT

As part of a study of the basic skills of secondary level education students, a study of prospective teachers' reading ability was conducted at Pennsylvania State University. Subjects were 47 students in secondary education. Two separate assessments of reading were used, the Nelson-Denny Reading Test and a criterion referenced test of reading based on an article in a professional journal. Results indicated that scores on the two tests correlated significantly only between the vocabulary section of each test and the total scores. Results also indicated that on the Nelson-Denny test the students' average score was close to the sixtieth percentile and on the criterion referenced test the students' average score was just over 82% correct. While female students performed better on the Nelson-Denny test, no differences by sex were found on the criterion-referenced test. When the results were examined according to students' subject area, prospective science teachers scored highest, followed closely by English majors. Math majors scored lowest on the reading tests. (MKM)
The Penn State Secondary Faculty was most concerned about prospective teachers' reading ability. Our concern is that we have no University-wide assessment of reading, hence no data on these students' reading levels. As a result, we included two separate assessments of reading: the Nelson-Denny Reading Test, a widely-used standardized test for adults, and a criterion-referenced test of reading, developed by Sandra L. Snyder (research assistant) and me.

The Nelson-Denny was chosen because it has demonstrated validity and reliability (.95 at the college level), and it is easy to administer and score. We used the short version, which allows 10 minutes for the vocabulary section and 20 minutes for the comprehension section. The Nelson-Denny, Form D, was administered in groups to students in the generic methods class, including students in English/communication, math, science, and social studies. From the Nelson-Denny, three scores are generated: Vocabulary, Comprehension, and Total. The Total score is figured from the raw scores according to the formula, T = V + 2C.

The Criterion-referenced Test (CRT) was developed as a test of professional reading skill. Whereas the Nelson-Denny used passages on general topics for its comprehension assessment, the criterion-referenced test is based on the reading of an article in a professional journal, "What's New in Ability Grouping?" (B. J. Wilson and D. W. Schmits) from the *Phi Delta Kappan*. The rationale for such a professional reading source is that a teacher should be assessed on reading material similar to that which s/he will be reading as part of his/her continuing professional development. The *Kappan* is recognized as a respected and well-written journal of general interest to teachers.
The criterion-referenced test establishes four objectives to be measured:

1. understanding of professional vocabulary used in the reading;
2. ability to answer literal level comprehension questions (known as low comprehension);
3. ability to answer inferential level comprehension questions (known as high comprehension);
4. ability to interpret information found in tables.

Thus, the test has four sections and we derive five scores from it:

1. Vocabulary -- 6 items
2. Low Comprehension -- 5 items
3. High Comprehension -- 9 items
4. Data Interpretation -- 3 items
5. Total -- 23 items

Comprehension Levels here are defined as in Barrett's Taxonomy of Reading Comprehension (Barrett, 1972). The four levels given by Barrett correspond roughly to Bloom's Taxonomy of Cognitive Levels (Bloom, 1956) for Levels 1, 2, and 3 (see Figure 2), while Level 4, appreciation, belongs in the Affective Domain. The CRT in reading tests only Levels 1 and 2 of Barrett's Taxonomy. Level 3, Evaluation, is tested by the writing sample, to be described in a later paper. For the writing sample, students were asked to respond in writing to one of two Level 3 questions related to the same reading selection.

This organization for assessing reading is based on the need for readers to move up the taxonomic levels as they read and consider a passage. Thus, our assessment procedure asks the prospective teacher to (1) read the article; (2) complete the CRT on reading; (3) complete the writing sample.
Figure 2

Barrett's Levels of Comprehension Related to the Taxonomy of Cognitive Levels.
The entire assessment can be completed by most students in a 75 minute period, although the test is not timed. Students are allowed to use as much time as necessary to complete the assessment. In addition, the CRT is given in our Instructional Support Center, an independent learning center, at which the student makes an appointment at his convenience to take the assessment.

These two assessments of reading clearly have several differences:
- criterion-referenced vs. norm-referenced
- not timed vs. timed
- individual scheduling vs. group testing
- professional material vs. general material

The questions we are concerned with here, then, include:

1. Are these tests providing us with the same information?
2. How do these prospective secondary teachers fare in reading, by content area and by sex?

The correlations between the Nelson-Denny and the CRT are our best evidence on question 1. We would predict high correlations between the two tests and their comparable sections. However, the only significant correlations were between the vocabulary sections of each test and the total scores. This suggests that vocabulary is tested similarly in both tests, perhaps also contributing to the significant total score correlations. The comprehension scores do not show high correlations with one another or the other sections of the tests. Nor does Data Interpretation, table reading, correlate highly. It seems reasonable to conclude that there are some elements of reading measured in both tests, but there is a large area of assessment which is not the same. The tests are not measuring the same things entirely.
The difference in measurement is especially important to the faculty who believe that comprehension is a major goal of professional reading. Our analysis of the comprehension section of the Nelson-Denny is that while it contains both Level 1 and Level 2 questions, it does not contain as heavy an emphasis on inferential (Level 2) comprehension as does the CRT.

The second major question asks how well our prospective secondary teachers do in reading. Looking at the whole group of 47 students, the means on the two tests look like this:

<table>
<thead>
<tr>
<th></th>
<th>Nelson-Denny (mean percentiles)</th>
<th>Criterion-Referenced Test Mean Possible</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vocabulary</td>
<td>57.2</td>
<td>Vocabulary 5.09 6</td>
</tr>
<tr>
<td>Comprehension</td>
<td>58.5</td>
<td>Low Comprehension 4.07 5</td>
</tr>
<tr>
<td>Total</td>
<td>59.0</td>
<td>High Comprehension 7.65 9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Data Interpretation 2.11 3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total 18.91 23</td>
</tr>
</tbody>
</table>

Both sets of results are encouraging. On the norm-referenced Nelson-Denny, our students' average score is close to the 60th percentile. On the CRT, our students' average score is just over 82 percent right. These figures are encouraging, given the bad press teachers have been given. Further research on practicing teachers and including additional variables will be needed before more definitive answers can be found.

The sex question is part of our evaluation here because it has become traditional to expect females to do better than males on reading tests. Again, our results show differences between the tests. True to the predictions, the norm-referenced Nelson-Denny shows significant differences between the sexes in vocabulary and in the total score. The comprehension difference approaches significance:
Nelson-Denny Reading Test
(mean percentiles)

<table>
<thead>
<tr>
<th></th>
<th>Vocabulary</th>
<th>Comprehension</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>49.5</td>
<td>52.1</td>
<td>51.2</td>
</tr>
<tr>
<td>Female</td>
<td>64.6</td>
<td>64.6</td>
<td>66.5</td>
</tr>
</tbody>
</table>

On the CRT, on the other hand, no differences by sex were found on any of the five variables. None even approached significance. If the standard norm-referenced test is biased against males, then the CRT does not reflect that bias.

When we examine these reading results by subject area, similar mixed results are found. We compared four groups: English/communication, math, science, and social studies. On the Nelson-Denny, none of the three variables showed significant differences by subject area. On the CRT, only two of the five variables showed significant differences. It is striking that vocabulary, which correlates well between tests, does not identify differences. Instead, the High Comprehension (Level 2) section and the Total score show differences at high levels of significance. Science teachers scored highest, followed closely by English teachers. Math teachers scored lowest, with social studies teachers slightly above them. The significance in the differences is between the scores of science and math teachers. This suggests that these are meaningful differences. We will be watching to see if this pattern is found in subsequent testing cycles. The low standing of the math and social studies teachers, compared to the English and science teachers, may be related to the small N's. However, it may also suggest the way they have been trained to respond by their disciplines. In the CRT, the ability to operate at a higher level of comprehension is weighted heavily in the total score (9 of 23).
We were somewhat surprised, in retrospect, that English teachers didn't score exceptionally well on both tests. Somehow we assumed that all the reading in English classes would prepare them better to handle reading tasks. While these English teachers read well enough, they are not at the top on the CRT.

We conclude, then, that the reading levels of prospective teachers in the Penn State secondary program is reasonably high. We can feel reasonably sure that these teachers can read professional materials successfully. While the English teachers' scores are higher than some of the other subject groups, the differences are smaller than expected. Although further research needs to be conducted with larger N's and including additional variables, these preliminary results support our belief that the reading skills of our prospective teachers can withstand public scrutiny.

In addition, the assessment procedures demonstrate the value of multiple tests of the same skill area. We gain valuable information about our students' reading skills from both tests. We can provide information to our skeptical public which they can understand, and we can provide precise diagnosis of reading difficulties to prescribe remedial teaching for our weaker students.

References

