

Will You Explain the Question?

Linguistic modification for students

level the playing field without compromising

BY JAMAL ABEDI

[Preview]

- Recent research on the assessment and accommodation of English language learners (ELL) links students' content-based assessment outcomes with their language background and suggests that English language proficiency affects the instruction and assessment of ELL students.
- To control for the effect of language on content-based assessments for ELL students, the content vocabulary and linguistic structures of the instruments should be examined for unnecessary linguistic complexity.
- A series of studies show that assessment instruments can be linguistically modified to reduce complexity while the content task and content terminology are retained.

While such recent legislation as the No Child Left Behind Act (NCLB) mandates monitoring the progress of all students—including English language learners (ELL)—the validity and fairness of assessment tools for such students remains a major concern for educational researchers and practitioners. The rapid increase in the number of ELL students nationwide makes these concerns even more profound. Between 1990 and 1997, the number of U.S. residents not born in the United States increased by 30%, from 19.8 million to 25.8 million (Hakuta & Beatty, 2000). According to the National Clearinghouse for English Language Acquisition, more than 4.5 million Limited English Proficient students were enrolled in public schools in 2000–2001, representing nearly 10% of the nation's total public school enrollment for prekindergarten through grade 12 (Kindler, 2002).

Results from a series of UCLA/CRESST studies have clearly demonstrated the effect of language factors on assessments, particularly for ELL students. For example, the studies indicated:

- ELL students perform substantially lower than native English speakers in subject areas with higher levels of language demand

Jamal Abedi (jabedi@cse.ucla.edu) teaches research methodology at UCLA's Graduate School of Education and Information Studies and works on CRESST-related studies of the technical aspects of performance-based assessment.

- As the level of language demand decreases, the performance gap between ELL and non-ELL students' decreases
- The linguistic complexity of test items may threaten the validity and reliability of achievement tests, particularly for ELL students (see, for example, Abedi, 2002; Abedi & Leon, 1999; Abedi, Leon, & Mirocha, 2003; Abedi & Lord, 2001; Abedi, Lord, Hofstetter, & Baker, 2000).

The Research

In analysis of test data from four locations nationwide (referred to as sites 1 to 4), Abedi, Leon, and Mirocha (2003) found a large performance gap between ELL and non-ELL students in reading and writing. The performance gap was lower for science and lowest for math problem solving, content areas in which the test items were less linguistically challenging for ELL students. The performance gap virtually disappeared in math computation, where the language demands of the test items were minimal.



STOCK PHOTO IMAGE

The results of analysis from site 3, a large urban school district, illustrate the effect of language on students' test performance. Means and standard deviations of SAT 9 normal curve-equivalent (NCE) scores on reading, science, and math tests were computed, as were differences between ELL and non-ELL students on the proportion of correct item responses. Table 1 presents the means and standard deviations for students in grades 10 and 11 in reading, science, and math, organized by the students' ELL status.

By reducing the effect of language factors on content area test performance, the validity and reliability of assessments for ELL students can be improved and can result in fairer assessments for these students. To minimize the effect of language factors and consequently reduce the performance gap between ELL and non-ELL students, language modification of assessment tools should be considered.

Using items from the National Assessment of Educational Progress (NAEP), recent studies compared student scores on actual NAEP items with parallel modified items in which the content task and content terminology were retained but the language was simplified. One study of 1,031 grade 8 students in southern California found significant improvements in the scores of students in low- and average-level math classes using the linguistically modified version of the test (Abedi & Lord, 2001). Among the linguistic features that appeared to contribute to the differences were low-frequency vocabulary and passive-voice verb constructions (for a discussion of linguistic features and rationale for modifications, see Abedi, Lord, & Plummer, 1997).

Another study of 1,394 grade 8 students in schools with high enrollments of Spanish speakers showed that modification of the language of the items contributed to improved performance on 49% of the items; the students generally scored higher on shorter problem statements (Abedi, Lord, & Hofstetter, 1998). A third study tested 946 grade 8 students with various accommodations, including modified linguistic structures, provision of extra time, and provision of a glossary (Abedi, Lord, Hofstetter, & Baker, 2000). Among the different options, only the linguistic modification accommodation narrowed the score gap between ELL and non-ELL students. (Abedi, Lord, & Hofstetter, 1998; Abedi, Lord, Hofstetter, & Baker, 2000).

To determine the effect of accommodations on grade 8 students in math, Abedi, Lord, Hofstetter, and Baker (2000) applied four different types of accommodation (i.e., linguistically modified English version of the test, standard NAEP items with glossary only, extra time

[Results from Site 3]

Grades 10 and 11 Normal Curve Equivalent (NCE) Means and Standard Deviations

	READING		SCIENCE		MATH	
	M	SD	M	SD	M	SD
GRADE 10						
SWD only	16.4	12.7	25.5	13.3	22.5	11.7
ELL only	24.0	16.4	32.9	15.3	36.8	16.0
ELL & SWD	16.3	11.2	24.8	9.3	23.6	9.8
Non-ELL & non-SWD	38.0	16.0	42.6	17.2	39.6	16.9
ALL STUDENTS	36.0	16.3	41.3	17.5	38.5	17.0
GRADE 11						
SWD only	14.9	13.2	21.5	12.3	24.3	13.2
ELL only	22.5	16.1	28.4	14.4	45.5	18.2
ELL & SWD	15.5	12.7	26.1	20.1	25.1	13.0
Non-ELL & non-SWD	38.4	18.3	39.6	18.8	45.2	21.1
ALL STUDENTS	36.2	19.0	38.2	18.9	44.0	21.2

Note: ELL = English language learners SWD = students with disabilities
M = Mean SD = Standard Deviation

The data in table 1 show a trend of higher NCE scores for non-ELL students as compared with ELL students. The trend of the performance gap between ELL and non-ELL students decreases as the level of language demand of the test items decreases. Among these content areas, reading has the highest level of language demand since language is central to the construct being measured. In the science and math tests, understanding of the science and math content-not the language-is the aim of assessment. ELL students in grade 10 had a mean reading score of 24.0 (SD = 16.4) as compared with a mean reading score of 38.0 (SD = 16.0) for non-ELL students, a difference of 14 score points. The difference between ELL and non-ELL NCE mean scores for science was 9.7, substantially less than the difference of 14 score-points for reading. For math, the difference between ELL and non-ELL was 2.8. As the data show, the performance gap between ELL and non-ELL students diminishes as the level of language demand of test items decreases. For grade 11 students, the ELL/non-ELL performance difference was 15.9 in reading, 11.2 in science, and for math, the performance difference between ELL and non-ELL was 0.0.

[Linguistic Modifications]

Certain language features can unnecessarily complicate test items. These examples show original and modified test items

UNFAMILIAR WORDS

■ Idioms, words from unfamiliar contexts, long words

Original: Max is in charge of the raffle.

Modified: Julia is in charge of the bake sale.

Original: Circle the clumps of eggs in the illustration.

Modified: Find the groups of eggs. Draw circles around the eggs.

COMPLEX SENTENCES

■ Subordinating clauses

Original: Because the box was a cube with six equal sides, Jenny calculated the area by....

Modified: The box is a cube with six equal sides. Jenny calculates the area of the cube by....

■ Idioms, conditional clauses, passive voice

Original: If two batteries in the sample were found to be dead,....

Modified: He found two broken pencils in the sample.

■ Logical connectors, conditional clauses, and adverbial clauses

Original: If rulers cost \$1.23 each, including tax, and Fred has \$9.00, how many can he buy?

Modified: Rulers cost \$1.23 each, including tax. Fred has \$9.00. How many rulers can Fred buy?

only, and glossary plus extra time). Students were also tested using standard NAEP items with no accommodation. Some of the accommodations were effective and increased students' performance in both ELL and English-proficient groups. Glossary plus extra time was the most effective form of accommodation in this study. This accommodation helped to increase performance of ELL students by 1.62 score points, or 13%; it also increased performance of English-proficient students by 2.81 points, or 16% (out of a possible 35 points). Considering that the non-ELL students in this study, who happened to be members of the low-performing student population, also benefited from the linguistic modification of test items it is possible that clarifying the language of instruction and assessment may help not only

ELL students but also low-performing students in general. On the other hand, these findings may raise issues concerning validity of accommodations that help both ELL and non-ELL students (see Abedi, Lord, Hofstetter, & Baker, 2000).

What Can We Say With Confidence?

1. Linguistically modified assessment tools help reduce the performance gap between English learners and other students. The modification includes reducing the use of low-frequency vocabulary and language structures that are incidental to the content knowledge being assessed.
2. Because the research suggests that the linguistic modification of test items does not alter the construct being



STOCK PHOTO IMAGE

measured, we suggest all students should be given content-area assessments that use clear language and provide sufficient time for them to show what they know and can do.

3. We recommend that the development of future large-scale content-area assessments should use clear language that is free of unnecessary complexity. This strategy should and can be a part of good instructional planning and assessment practice.
4. The language demands of academic materials and assessment tools should be identified and provided to teachers so they can ensure that students have the language resources to demonstrate their content-area knowledge and skills.

Ethical and Valid

Language and performance are confounded. Solving math word problems and responding to science and social sciences test items presents a double challenge for students whose language proficiency is limited, and the added cognitive load can affect individual performance negatively. To perform well on content-based assessments with high language loads, students—both native and non-native speakers—have to know less-frequently used vocabulary and sentence structure, in addition to the content being assessed. As things stand, to meet NCLB's high standards for adequate yearly progress for

ELL students, content-area teachers will have to broaden their students' language proficiency in addition to eliciting content mastery.

The results of the CRESST-sponsored studies summarized here suggest that language modification of test items can be used as an effective and valid accommodation for English language learners. It is effective because it reduces the performance-gap between ELL and non-ELL students. It is valid because it does not alter the construct being measured; that is, it does not affect the performance of high-performing non-ELL students. Further, the linguistically modified test version is a feasible form of accommodation because it does not cause any additional burden on the instruction or assessment of the students being tested. PL

References

- Abedi, J. (2002). Standardized achievement tests and English language learners: Psychometrics issues. *Educational Assessment*, 8(3), 231–257.
- Abedi, J., & Leon, S. (1999). *Impact of students' language background on content-based performance: analysis of extant data*. Los Angeles: University of California, National Center for Research on Evaluation, Standards, & Student Testing.
- Abedi, J., Leon, S., & Mirocha, J. (2003). *Impact of student language background on content-based performance: analysis of extant data* (CSE Tech. Rep. No. 603). Los Angeles: University of California, National Center for Research on Evaluation, Standards, & Student Testing.
- Abedi, J., & Lord, C. (2001). The language factor in mathematics tests. *Applied Measurement in Education*, 14(3), 219–234.
- Abedi, J., Lord, C., & Hofstetter, C. (1998). *Impact of selected background variables on students' NAEP math performance* (CSE Tech. Rep. No. 478). Los Angeles: University of California, Center for the Study of Evaluation/National Center for Research on Evaluation, Standards, & Student Testing.
- Abedi, J., Lord, C., Hofstetter, C., & Baker, E. (2000). Impact of accommodation strategies on English language learners' test performance. *Educational Measurement: Issues and Practice*, 19(3), 16–26.
- Abedi, J., Lord, C., & Plummer, J. (1997). *Language background as a variable in NAEP mathematics performance* (CSE Tech. Rep. No. 429). Los Angeles: University of California, Center for the Study of Evaluation/National Center for Research on Evaluation, Standards, & Student Testing.
- Hakuta, K., & Beatty, A. (Eds.). (2000). *Testing English-language learners in U.S. schools*. Washington, DC: National Academy Press.
- Kindler, A. L. (2002). *Survey of the states' limited English proficient students & available educational programs and services, 2000–2001 Summary Report*. Washington, DC: National Clearinghouse for English Language Acquisition & Language Instruction Educational Programs.