Assessing English Language Learners’ Opportunity to Learn Mathematics: Issues and Limitations

JAMAL ABEDI
University of California, Davis

JOAN HERMAN
University of California, Los Angeles

**Background/Context:** English language learner (ELL) students are lagging behind because of the extra challenges they face relative to their peers in acquiring academic English language proficiency, and the added burden of learning content in a language in which they are not proficient. The mandated inclusion of ELL students in the nation’s accountability system may not be productive if their academic needs are not defined and addressed.

**Purpose/Objective/Research Question/Focus of Study:** This study explores the relationship between students’ ELL status and their level of opportunity to learn (OTL) as a factor that may explain performance differences between ELL and non-ELL students.

**Setting:** The research was conducted on a sample of Grade 8 students in Southern California.

**Population/Participants/Subjects:** A total of 24 Grade 8 classes participated in the study. Within these classes, a total of 602 students served as subjects for this study.

**Research Design:** In this study, we examined factors that explain OTL for ELL students; therefore, the research design for this study is an ex post facto or causal-comparative design.

**Data Collection and Analysis:** Data were collected on student reading comprehension, student background, and motivation. Internal consistency approach was used for estimating reliability of the instruments. Comparison of performance of ELL and non-ELL was done using a hierarchical linear model approach.

**Findings/Results:** Results indicate that: (1) measures of classroom OTL are associated with student performance; (2) ELL students report a lower level of OTL as compared with non-ELLs, and such differential levels of OTL may indeed play a role in the lower performance
of ELLs; (3) a higher concentration of ELLs is associated with lower levels of OTL; and (4) English proficiency and self-reported ability to understand teachers’ instruction both appear to influence effective access to OTL.

Conclusions/Recommendations: The results of this study suggest that students’ ability to understand teacher instructions influences reported levels of OTL. Students who experience more difficulty report lower levels of OTL because they do not understand or thus recognize topics that have been addressed in class. The differential level of OTL may be a major factor contributing to the substantial performance gap between ELL and non-ELL students. However, there were limitations to this study. Among the most important is the meaning of OTL itself. What is a reasonable definition of OTL, and how far down into enacted instruction should it really go? Future studies need to carefully examine these factors and determine the best way to address the performance gap between ELL and non-ELL students.

The No Child Left Behind (NCLB) law’s requirement that all subgroups meet the same annual yearly progress (AYP) targets underscores both the mandate and the challenge of assuring that English language learners (ELLs) achieve the same high standards of performance that are expected of their native-English-speaking peers. The intent is laudable: States, districts, schools, and teachers must be accountable for the learning of their ELL students, and students themselves must be accountable as well. ELLs can no longer be invisible in the educational system; their learning needs must be met, and they too must make steady progress toward the goal of all students being judged proficient based on statewide testing by the year 2014.

OBSTACLES TO ENGLISH LANGUAGE LEARNER ACHIEVING STANDARDS

However, for English language learner students in particular, there are major obstacles in reaching proficiency by the target date of 2014, or even later (Abedi, 2004). Literature shows that ELL students lag behind their English-proficient peers in all content areas, particularly academic subjects that are high in English language demand. This performance gap could be explained by many different factors, including background experiences such as those indexed by parent education and poverty (Abedi & Gándara, 2006); the challenge of second language learning (Hakuta, Butler, & Witt, 2000; Moore & Redd, 2002); and a host of inequitable schooling conditions (Abedi, Herman, Courtney, Leon, & Kao, 2004; Gándara, Rumberger, Maxwell-Jolly, & Callahan, 2003). Furthermore, the language complexity of instructional materials and measurement tools presents obstacles to instruction and assessment for ELL students.
ELLs are more likely than other students to be taught by teachers without relevant teaching credentials and with little classroom experience (Gándara, Maxwell-Jolly, & Driscoll, 2005; Rumberger & Gándara, 2004). Within this context, ELL students must take on the challenges of learning English and U.S. culture, in addition to learning the academic content of subject matter curriculum. The process of developing new language skills is difficult and requires time and effort that otherwise could be spent on acquiring academic content knowledge. For example, a number of studies have shown that it takes between 5 and 7 years or even longer for most ELLs to gain sufficient mastery of academic English to join English-speaking peers in taking full advantage of instruction in English (Hakuta et al., 2000). During this time, learning cannot occur at the same rate as it would for a native speaker of English when instruction is offered only in English. Limited English proficiency may also make it difficult for ELLs to benefit fully from teachers’ instruction or to understand assessment questions sufficiently to show what they know.

Because of such limitations, NCLB assessment results thus far suggest that ELL subgroups are being left behind and that schools and districts that serve significant proportions of ELLs are less likely to meet their AYP goals and more likely to be subject to corrective action than schools serving fewer ELLs (EdSource, 2004). Research also shows that economically disadvantaged and culturally diverse subgroups including ELLs have had less access than other students to a challenging curriculum that would prepare them for success on today’s standards (Guiton & Oakes, 1995; Wang, 1998). ELLs also are grossly overrepresented among those dropping out (Silver, Saunders, & Zarate, 2008) and those failing to pass high school proficiency exams and thus are in danger of facing the dire consequences that accompany the absence of a high school diploma.

To understand why ELL students are lagging behind, it is important to look at the issues related to both the instruction and assessment of these students. The extra challenge these students face relative to their peers in acquiring academic English language proficiency, the added burden of learning content in a language in which they are not proficient, and the confounding of their language ability with their subject matter competency when they are assessed in English are all contributing factors to the performance gap between ELL and non-ELL students (Abedi, 2004; Abedi & Gándara, 2006). Accordingly, the current study explores relationship between students’ ELL status and their level of opportunity to learn in a small sample of urban classrooms in the first semester of a two-year Algebra I course.

Research clearly has shown that unnecessary linguistic complexity of assessment negatively impacts the reliability and validity of assessment for
ELL students (see, for example, Abedi, 2006; Abedi, Leon, & Mirocha, 2003; Solano-Flores & Trumbull, 2003). Minor changes in the wording of content-based assessments raise ELL student performance (Abedi, Lord, Hofstetter, & Baker, 2000; Cummins, Kintsch, Reusser, & Weimer, 1988; Duran, 1989; Riley, Greeno, & Heller, 1983), strongly underscoring the role of language proficiency in student performance. Quite obviously, language factors also influence the ability of ELL students to access and benefit from instruction. Lack of effective opportunity to learn—students’ access to and engagement in the academic content they need to perform well on tests and achieve standards—looms large as a possible barrier to the success of ELLs (Herman, Klein, & Abedi, 2000) yet has been little studied.

OPPORTUNITY TO LEARN AND ENGLISH LANGUAGE LEARNERS

Research on students’ opportunity to learn in general dates back to John Carroll’s coining of the term in the early 1960s, when it denoted whether students had sufficient time to learn (Carroll, 1963; Tate, 2001). Since then, escalating demands for accountability and higher standards of student performance have renewed interest in the concept. For example, hand in hand with the development of a first generation of standards-based accountability systems and their keen focus on student outcomes came concerns about students having access to the educational resources and programming they need to achieve established standards (Goals 2000: Educate America Act, 1994; National Council on Education Standards and Testing, 1992). Similarly, even earlier attempts to hold students accountable for passing minimum competency tests to receive a high school diploma established the firm legal principle that such accountability required curricular validity—that is, that there be evidence that schools are teaching what is tested (Debra P. v. Turlington, 1983; Phillips, 1994). The OTL concept thus has been expanded to encompass the quality of curriculum and instruction and their alignment with what is assessed (Brewer & Stacz, 1996; Burstein et al., 1995; Marzano, 2000; McDonnell, 1995; Porter, 1991; Smithson, Porter, & Blank, 1995; Stevens, 1996).

Today’s educational policies emphasizing standards-based reform, in fact, assume that by establishing and assessing standards and holding educators accountable for results, schools will alter students’ opportunities to learn (Herman, 2008)—that is, that schools will change the content and process of their curriculum and instruction to ensure that students develop the knowledge and skills they need for success. And, in reality, ample research shows that administrators and teachers in general
tend to respond to strong accountability demands by focusing their efforts on what is expected of students, by aligning curriculum and instruction with standards, and particularly with what is tested, and by identifying students who need extra help. The relationship between such focusing and student learning is more moot (Herman, 2008).

The situation for ELLs is more moot in that, as indicated, English language proficiency may inhibit ELLs’ access to effective learning opportunities, and little research has explicitly addressed the question of ELLs’ OTL and its role in ELLs’ performance. What research exists raises important equity issues.

Cultural and linguistic minority students have less exposure to content, and their instruction tends to cover less content relative to nonminority students (Masini, 2001). Moreover, research shows a dramatic under-representation in higher-level math courses, and over-representation in lower level mathematics courses among cultural and linguistic minority students, which affects their OTL (Gross, 1993; Jones, Davenport, Bryson, Bekhuis, & Zwick, 1986; Oakes, 1990). Gross noted that teachers of low-ability classes tend to emphasize drill and practice rather than higher thought processes, which are emphasized by teachers of high-ability courses. Gamoran, Porter, Smithson, and White (1997) found lower mathematics achievement among high school students in general-track classes as compared with those in college-preparatory classes, implying that the practice of tracking (or grouping by ability) denies students opportunities to learn. In 2005, Callahan’s research on tracking and ELLs’ achievement found that only 2% of ELL students who participated in the study took college-bound coursework.

Moreover, there is growing research suggesting that settings with higher concentrations of English learners are educationally wanting, in that they are associated with significantly less academic success and higher dropout rates. Although ELLs are disproportionately poor, and the relationship between socioeconomic status and student academic learning is well established (see, for example, Education Trust, 2006; EdSource, 2004), ELL status seems to exert a significant negative influence on performance beyond that of poverty and initial achievement status (Kim & Herman, 2008). Similarly, research shows a clear relationship between percentage of ELL student population and school dropout rates (Silver et al., 2008), raising the distinct possibility that ELL status may operate not only as an individual variable but also as an important social context variable in influencing student performance.

As noted, the current study explores such relationships among and between ELL status and students’ opportunity to learn in a small sample of California urban classrooms in the first semester of a two-year course
Our exploratory study aims to highlight factors that may help to explain the performance gap between ELLs and non-ELLs, the understanding of which can inform future policy and practice. Our investigation addresses the following research questions: (1) Is there a relationship between measures of classroom OTL and student performance in middle school mathematics? (2) Do ELLs and non-ELLs receive the same levels of OTL? (3) What factors may account for differences in OTL and performance for ELLs and non-ELLs?

METHOD

We describe the research context, sample, instruments, and composition of variables in the following sections.

STUDY CONTEXT

Sample

The study draws on a sample of 9 algebra teachers from three urban middle schools in a large district in Southern California. To explore how the concentration of ELLs might relate to OTL and performance, the sample purposively represents classrooms in which ELLs constituted the clear majority (more than 75% of the class), in which proportion of ELLs and non-ELL was more evenly mixed, and those in which non-ELLs were the clear majority (ELLs less than 25% of the class). A total of 24 classes volunteered to participate in the study, and, within these classes, a total of 602 Grade 8 students served as subjects.

The Two-Year Algebra Course

As noted, our study context is the first semester of a two-year Algebra I course spanning Grades 8 and 9. The Grade 8 portion offers fertile ground for OTL research because the expanded two-year course is intended to provide an equitable learning opportunity for all students to be successful in algebra and is designed for all eighth graders—ELL and non-ELL—who are not ready for the pace of a one-year course. The two-year algebra course also is the math course in which the majority of Grade 8 ELL students are enrolled, so it is not a “language minority ghetto.” Most of the district’s eighth-grade population is enrolled in the two-year algebra course, and about half of the enrollees are non-ELL students, albeit often with a Spanish home language background. Throughout the district, all teachers use the same text for the course and
are held to the same district pacing plan, a plan that establishes a common schedule for addressing specific content goals relative to state standards.

**Demographic Information**

Of the total student sample, 54% were female, slightly less than half of the sampled students were ELLs (44%), and the rest (56%) were non-ELLs. Information gathered from the school’s student data revealed that four levels of English language development (ELD) were represented in the ELL sample, with the majority of the ELL students (77.9%) classified as ELD5, the level just prior to reclassification as English proficient. A large majority of the sample qualified for free lunch, and about 80% were of Hispanic descent.

**Achievement Information**

Table 1 shows the distribution of students’ performance on standardized tests of reading and mathematics for the spring prior to the study. Based on normal curve equivalence scores, the data show that almost all students in the sample scored below the national norm group on standardized tests of reading and mathematics (mean test score of less than 50.00). As might be expected, ELLs scored considerably lower than English proficient students—English only (EO), initially fluent English proficient (IFEP), and reclassified fluent English proficient (RFEP). Lower levels of English language proficiency (lower ELD levels) were associated with lower performance.

<table>
<thead>
<tr>
<th>ELL Designation</th>
<th>SAT9 Reading</th>
<th></th>
<th>SAT9 Math</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>M</td>
<td>SD</td>
<td>N</td>
</tr>
<tr>
<td>English Only (EO)</td>
<td>89</td>
<td>38.83</td>
<td>23.86</td>
<td>86</td>
</tr>
<tr>
<td>Initially Fluent English Proficient</td>
<td>40</td>
<td>46.25</td>
<td>24.61</td>
<td>40</td>
</tr>
<tr>
<td>Reclassified Fluent English Proficient</td>
<td>193</td>
<td>33.38</td>
<td>17.94</td>
<td>191</td>
</tr>
<tr>
<td>English Language Development,</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level 5 (ELD5)</td>
<td>194</td>
<td>15.37</td>
<td>12.08</td>
<td>193</td>
</tr>
<tr>
<td>ELD Level 4</td>
<td>12</td>
<td>13.92</td>
<td>12.86</td>
<td>12</td>
</tr>
<tr>
<td>ELD Level 3</td>
<td>11</td>
<td>5.36</td>
<td>11.64</td>
<td>11</td>
</tr>
<tr>
<td>ELD2 Level 2</td>
<td>25</td>
<td>2.60</td>
<td>15.92</td>
<td>25</td>
</tr>
<tr>
<td>Total</td>
<td>564</td>
<td>26.63</td>
<td>33.98</td>
<td>558</td>
</tr>
</tbody>
</table>

*Note: ELD1 students (n=2) were very likely exempt from taking the SAT9.*
A majority of the students (72%) reported Spanish as being at least one of their spoken home languages before they started going to school, according to the Student Background Questionnaire that was administered as part of our study.

Participating Teachers

The participating teachers’ experience ranged from 2 to 11-plus years in middle school with a continuum of training, credential, and educational backgrounds. Three of the 9 participating teachers had earned a teaching credential beyond temporary status, and 3 others had education beyond a bachelor’s degree.

INSTRUMENTS

The study drew on data from a number of sources, including specially developed study instruments and existing measures: (1) teacher and student surveys of OTL, (2) an assessment of achievement in initial Algebra I content, (3) the fluency subscale from the Language Assessment Scale (LAS), and (4) a student background questionnaire that provided data on prior preparation in mathematics and ability to understand directions. The OTL surveys were used to provide measures of the level of students’ opportunity to learn the specific topics specified for the first semester of the district’s two-year Algebra 1 curriculum. Similarly, the achievement instrument measured student content knowledge in curriculum topics that were expected to be addressed during the time period of the study. The English language assessment instrument helped define students’ level of English knowledge/proficiency and was used to help explain variability in the ELL sample. Finally, the student background questionnaire provided data on students’ background that helped further explain performance differences between ELL and non-ELL students. We also collected students’ grades in their first semester of algebra for use as an additional variable that could add to our overall understanding of student performance. As described, we also collected archival data on students’ prior standardized test performance in reading and math, which served as a covariate for analyses examining the relationship between OTL and student learning.

OTL Measures

Devising an efficient measure of OTL was a major undertaking in our study. Four OTL variables have been prevalent in research: content cov-
verage, content exposure, content emphasis, and quality of instructional delivery (Herman et al., 2000; Porter, 2002; Stevens, Wiltz, & Bailey, 1998). Content coverage refers to the actual coverage of core curriculum topics specific to a particular grade level or subject area. Content exposure refers to the amount of time teachers allocate to covering the content. Content emphasis refers to the emphasis given to certain topics that are part of the core curriculum. The quality of instructional delivery or strategies refers to how coherently and effectively teachers engage students so that they can understand and acquire what is being taught. As with the variety of OTL variables, a variety of tools have been used to address OTL, including teacher surveys or ratings, content analyses of instructional materials, collection and analysis of instructional artifacts, and teacher interviews (see, for example, Colker, Toyama, Trevisan, & Haertel, 2003; Gamoran et al., 1997; Herman & Klein, 1996, 1997; Herman et al.; Muskin, 1990; Porter, 2002; Wang, 1998). Because of limited study resources, we confined the study to the most common OTL variable in the literature, content coverage, specifically addressing the alignment between course content and test content, and to the most efficient forms of data collection, teacher and student surveys.

Development of surveys. Survey development started with the standards for the two-year algebra course and the identification of the 28 topics that were supposed to be covered during the first semester of Grade 8, the time period of this study. The teacher survey simply listed the 28 topics, using terminology used in district documents, and asked teachers to indicate whether each class had addressed each topic. We listed the same 28 topic areas in the student survey and asked students to indicate which of the content areas their class had covered. The language used in the student survey incorporated a linguistic modification procedure to simplify English language demands, principally vocabulary, so that the questions were readable for students at lower levels of English proficiency, yet were consistent in meaning with the original (see for example, Abedi, 2006). Responses from both teachers and students provided cross-validation data to examine the validity of the measure. We computed the percent of agreement between teachers’ and students’ responses by counting the number of topic matches (if both the teacher and a student indicated that the topic was covered) divided by the total number of content areas (28) multiplied by 100. The percent of teacher-student agreement for all students was 64.9%. For ELLs, the percent of agreement was 54.1%, and for non-ELLs, the percent of agreement was 73.4%.

To analyze the effect of OTL on achievement, we focused on 11 of the 28 content/skill areas that were represented in the algebra test items that we used for measuring students’ algebra knowledge in this study (see the
Achievement Measures section for more detail on the test). We computed percent of agreement between teachers’ and students’ responses in these 11 content/skill areas. The percent of agreement for all students in these 11 areas was 69.5%; for ELL students, it was 62.8%, and for non-ELL students, the percent of agreement was 75.5%.

**Classroom-level OTL.** Because we believe that OTL, in the context of these algebra classes, is largely controlled by teachers for their classes as a whole (and mediated through a single textbook), we thought it was most appropriate to consider OTL as a class-level variable. For each class, an OTL index was computed by summarizing student survey responses for each topic and then summarizing student responses across topics. That is, student scores for each topic were coded as either 1 or 0, depending on whether the student indicated that topic as taught thus far in the course; a class would receive a score of 0.50 for a particular topic if 50% of the students in that class marked that they had studied the topic in Grade 8. A total class-level OTL measure then was computed as the sum of the scores on the 11 topic areas addressed by the study outcome measure. Thus, the total class-level OTL measure for a class would be 11 if all the students agreed that all 11 content/skill areas had been taught. The remaining topics were addressed in a similarly composed OTL measure of non-tested topics.

**Student OTL reports versus teacher OTL reports.** The decision to use student OTL reports rather than teacher reports had several motivations. The student measure was based on a composite score of all students in the class (over 25 students in each class), but the teacher measure would only be based on a single response: that from the teacher; therefore, the student measure was considered more reliable. Moreover, the students’ OTL measures showed a stronger relationship to student performance than did the teacher measure (see the following section).

**Evidence of validity.** Because it was based on 11 content areas that are conceptually related through the given curriculum, aligned with the district pacing plan, and highly correlated, we expected our OTL measure to be unidimensional. A principal components analysis using the 11 content areas confirmed our expectation, a single underlying factor.

The relationship between tested and non-tested OTL topics provided further evidence of the validity of the measure. Recall that we asked students to rate their OTL for all the topics of algebra, derived from those ratings a measure of OTL for the 11 topics that were included on the test, and summarized the remaining topics in an index of non-tested OTL. If our OTL measure is capturing curriculum content, we would expect higher student performance on tested OTL than on non-tested OTL; that is, when instruction and assessments are aligned, performance should be
higher than when they are not. Table 2 tests this hypothesis by showing the correlations between OTL ratings and performance for both tested and nontested OTL ratings. These correlations are reported separately by student OTL ratings and those by teachers.

As the data in Table 2 show, test-aligned OTL questions are more strongly correlated with math scores than nonaligned OTL. For example, the correlation between test-aligned OTL (as rated by students) and the 19-item math test was .72, as compared with correlations of .197 for the 17 OTL content areas not assessed by the math test. As indicated earlier, student-reported OTL has a stronger relationship to students’ performance than the teacher-reported OTL, and this was one of the main reasons that we used student-reported OTL. Based on the data in Table 2, the correlation between teacher-reported OTL and math scores was .533, as compared with a correlation of .720 with the student-reported OTL. The results also showed that algebra grades are not a good criterion in examining the effects of OTL on students’ learning. For example, the correlation between test-aligned OTL and algebra grades was .037, which suggests that less than half a percent point of the variance of algebra grades was explained by the knowledge from OTL. Reasons for this low correlation may point to the subjectivity involved in assigning grades.

**Achievement Measures**

The achievement test used in this study to provide the outcome score was a 20-item algebra test specially compiled for this study. In the interest of available resources—both time and budget—we used available test items that matched course goals from the National Assessment of Educational Progress and Trends in International Mathematics and Science Study to construct our measure. The 20-item algebra test covers 11 of the 28 target content areas. One of the items was deleted because of technical problems.

As an estimate of the reliability of the test, the internal consistency coefficient (alpha) was computed. For the entire test (all 19 items), the

| Table 2. Correlations of Class-Level OTL With Class-Level Math Achievement (n = 24) |
|---------------------------------|------------------|-----------------|
|                                 | Math Score       | Algebra Grade   |
| Student OTL on test content     | .720**           | .037            |
| Student OTL other content, not tested | .197            | -.112           |
| Teacher OTL                     | .533**           | .321            |

** p < 0.01, two-tailed.
alpha was .604 ($n = 602$). To examine the impact of linguistic complexity on the reliability of the algebra test, we grouped items as linguistically complex (9 items) or less linguistically complex (10 items). Those items determined linguistically complex had complex sentence structures, used complicated verb tenses and grammatical features, and were dense in sophisticated or unconventional vocabulary. For the complex items, the alpha was .227, and for the noncomplex items, the alpha was .562, with a difference of .335. The results of internal consistency analyses indicate that the entire test suffers from low reliability, which may be the result of several factors. One of the main reasons for low reliability for this test was the problem of restriction of range; students included in the sample for this study were considered low-performing students. In addition, because the test measures a number of areas within algebra, it may well not be unidimensional. The relatively small number of test items may be another contributing factor to the overall low reliability of the test. The level of the test items’ linguistic complexity may also have had a profound impact on the reliability of the test, consistent with our earlier studies suggesting that language factors may be a source of measurement error and may reduce the reliability of the tests (Abedi, 2006).

**English Language Proficiency Measures**

The study used the Language Assessment Scale’s (LAS) fluency subscale score to measure student language proficiency (ELP). Although fluency is only one component of the English language proficiency test, it was decided that this measure would be included in the absence of a more comprehensive measure of ELP. The reliability coefficient (internal consistency) for the 10-item LAS frequency subscale was .697 ($n=602$), again lower than optimal.

**Student Background Questionnaire**

Students were asked to respond to questions about their language background and prior instruction after completing the OTL survey. Students were also asked about their country of birth and time in the United States, and they were asked to self-assess their comprehension of their teachers’ instructional activities (e.g., directions, instruction, tests, and math tests).

Information from this questionnaire was used to create an individual measure of prior academic preparation, which was a latent composite score derived from three groups of student background variables. These variables, formed based on how they relate to prior opportunities to
learn, were: (1) prior OTL in specific math content, (2) years in school in the United States, and (3) access to learning resources. These variables were selected based on the recommendations from prior studies (Abedi & Gándara, 2006; Abedi et al., 2004; Abedi et al., 2003; Abedi & Lord, 2001; Solano-Flores & Trumbull, 2003).

In addition, student responses to a question about their ability to understand their teacher’s directions was used as a proxy for understanding of instruction. That is, understanding directions seems to be a prerequisite for easy accessibility to any instructional activity.

RESULTS

Reporting of study results will be structured around the three research questions stated earlier. In response to the first research question, the relationship between opportunity to learn and student performance was examined. The second research question explored whether English language learners and non–English language learners experienced differential levels of OTL. The third question investigated study variables that could explain possible differential levels of OTL across the ELL and non-ELL groups. Exploratory analyses in response to the final questions addressed the relationship among language proficiency, ability to understand instruction, OTL, and performance.

Q1. IS THERE A RELATIONSHIP BETWEEN MEASURES OF CLASSROOM OTL AND STUDENT PERFORMANCE?

It can be hypothesized that the more opportunity to learn that students have, the higher their performance should be. To the extent that OTL predicts performance, it helps to explain learning outcomes and thus potential gaps between ELLs and non-ELLs. To examine this hypothesis, the relationship between student-reported OTL and student performance was examined. Given the nested nature of the data, a two-level hierarchical linear model (HLM) was used, with student data as level 1 variables and classroom data as level 2 variables. The level 1 variables included prior achievement in the form of SAT9 math scores, prior academic preparation, and math grade, and the level 2 variables included classroom-level OTL.

Table 3 summarizes the result of analyses for Model 1. The results show that the classroom-level OTL measure was significantly and positively related to the outcome variable ($t = 4.90, p = .000$). Additionally, student-level math grades ($t = 4.23, p = .000$) and SAT9 math scores ($t = 5.94, p = .000$) affected the outcome variables. However, after accounting for the
classroom-level OTL measure, the prior academic preparation factor showed no significant effect.

Q2. DO ELLS AND NON-ELLS RECEIVE THE SAME LEVELS OF OTL?

To shed light on this question, we computed the mean classroom OTL index by student ELL status. Table 4 shows the descriptive statistics across the ELL classification categories. As indicated earlier, the classroom OTL score could range from 0 to 11, with 0 representing students’ indication of no opportunity to learn the content/skill areas that were represented in the algebra test questions, and 11 suggesting the opportunity to learn all 11 areas on the test. Therefore, the higher the class-level OTL measure, the more topics the students indicated they had opportunity to learn. As data in Table 4 show, the mean OTL across all classrooms was 8.43 (SD = 2.13, N = 24 classrooms, 602 students within these classrooms), out of a perfect score of 11, suggesting that the students reported the opportunity to learn a majority of the topics included on the test.

Table 4. Descriptive Statistics for OTL Measure by ELL status

<table>
<thead>
<tr>
<th>ELL Status</th>
<th>M</th>
<th>N classes</th>
<th>N students</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>OTL all cases</td>
<td>8.43</td>
<td>24</td>
<td>602</td>
<td>2.13</td>
</tr>
<tr>
<td>OTL ELL students</td>
<td>7.29</td>
<td>24</td>
<td>263</td>
<td>2.49</td>
</tr>
<tr>
<td>OTL non-ELL students</td>
<td>9.31</td>
<td>24</td>
<td>339</td>
<td>1.23</td>
</tr>
</tbody>
</table>

However, the mean OTL status varies by student ELL status. The data in Table 4 suggest that non-ELL students reported a higher level of opportunity to learn than the ELL students. The mean OTL for non-ELL students was 9.31 (N = 339, SD = 1.23) as compared with the mean OTL of 7.29 (N = 263, SD = 2.49) for ELL students. The difference between the
OTL results across the ELL categories was highly significant ($t = 13.02$, $df = 600$, $p < .001$), suggesting that the non-ELL students reported more opportunity to learn than the ELL students. Moreover, the standard deviation for the ELL group ($SD = 2.49$) was twice the standard deviation for the non-ELL group ($SD = 1.23$), suggesting that the ELL students were less consistent across classes than the non-ELL group about whether they had OTL.

Correlations among class-level variables provide similar evidence of the relationship between ELL status and learning opportunities. As the data in Table 5 show, there were strong negative relationships between the percentage of ELL students in a class and both student-reported OTL (-.680) and teacher-reported OTL (-.642). There were likewise strong positive relationships between levels of language proficiency (as indicated by LAS scores) and OTL measures (.801 for student-reported OTL and .622 for teacher-reported OTL, respectively).

Table 5. Class-Level Correlations for OTL and Language Status Variables ($n = 24$)

<table>
<thead>
<tr>
<th></th>
<th>Student OTL</th>
<th>Teacher OTL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent LEP</td>
<td>-.68</td>
<td>-.64</td>
</tr>
<tr>
<td>Mean LAS score</td>
<td>.80</td>
<td>.62</td>
</tr>
</tbody>
</table>

Q3. WHAT FACTORS MAY ACCOUNT FOR DIFFERENCES IN OTL AND PERFORMANCE FOR ELLS AND NON-ELLS?

We hypothesized that language proficiency itself might be contributing to OTL, in that students who are not fully proficient in English might have difficulty understanding and fully benefiting from textual materials and teachers’ instruction. Such a relationship could play out in OTL in at least two ways: Because of language difficulties, teachers in classes with higher proportions of ELLs might proceed through the curriculum at a slower pace, resulting in less OTL relative to the full set of topics addressed by the test, or ELLs may not perceive OTL because effectively, they have not been able to fully understand or benefit from curriculum and instruction even though they have been exposed to it.

*Student Understanding of Instruction*

Results from the background questionnaire further underscored the relationship between students’ English language ability and OTL. Table 6 shows descriptive data demonstrating the relationship between students’ self-reporting of their ability to understand directions (in English)
and reported levels of OTL. As noted earlier, we used this question as a proxy for understanding of instruction, in that failure to understand directions clearly impedes students’ participation in instructional activities. Survey responses indicated that 58% of the students said that they could understand the teacher’s directions in English “very well,” 35% of students said “well,” and the remainder said “not well” or “not at all.”

One can hypothesize that the higher the level of understanding of instruction in English, the more proficient a student is in English, and the more proficient a student is in English, the more he or she benefits from OTL. The results of our analyses presented in Table 6 support this hypothesis. Students who indicated that they understand directions in English “very well” had a substantially higher class-level OTL measure mean (M = 8.74, SD = 1.74) as compared with those who believed they did not understand directions in English—that is, those who said “not well at all” (M = 6.50, SD = 1.58). The trend of decreasing OTL scores by the decreasing level of self-reported English proficiency is quite evident from the data in Table 6.

### CONCENTRATION OF ELLS WITHIN CLASSES

The relationship between ELL status, ability to understand instruction in English, and OTL, as discussed in relation to Tables 5 and 6, is a complex one and suggests that concentration of ELLs within classrooms may influence access and learning. Whereas Model 1 in Table 3 shows the relationship between student prior preparation, class-level OTL, and student performance, the results in Table 7 add classroom-level ELL, the proportion of ELL students in a class, as an additional level 2 variable. We included this variable to test whether the proportion of ELL students in the classroom might affect instruction and thus be confounded with the classroom-level OTL measure.
Table 7 presents a summary of HLM analyses of this second model. As data in Table 7 show, both the classroom-level OTL measure ($t = 2.07$, $p = .050$) and proportion of ELLs within the class ($t = 2.95$, $p = .008$) were significant predictors of student math test scores, suggesting that the classroom-level OTL and proportion of ELLs in a classroom were both associated with student performance in math. Similar to the data presented for Model 1, students’ grades ($t = 4.55$, $p = 0.000$) and SAT9 scores ($t = 5.72$, $p = 0.000$) were predictors of students’ performance on the 19-item math test. Once again, after entering OTL as a classroom-level variable, the student-level prior preparation factor was not a strong predictor of the outcome variable ($t = 0.90$, $p = 0.376$).

**DISCUSSION**

Our study was motivated by concerns for the performance of English language learners, particularly in light of current assessment mandates, and questions about how opportunity to learn might help both to explain and to support policy and practice aimed at improving performance of ELL students. Although clearly exploratory, study results add fuel to our initial concerns and underscore some of the complexity of measuring and assuring OTL for ELL students.

**MEASUREMENT OF OTL**

The results showed that a relatively simple composite of student survey responses showed reasonable OTL measurement qualities. In response to questions about the specific topics that had been addressed in a course, students’ ratings showed reasonable consistency across students in a class, with 70% agreement for the full OTL measure. Student ratings within a
class also showed reasonable agreement with teacher ratings; on average, the agreement level was 65%. The measure had desirable characteristics of unidimensionality, although admittedly, it covered only a limited aspect of OTL: whether specific topics had been addressed during a course. Further, the relationship between student performance and survey results for OTL ratings for tested content (as compared with content that was not tested) offered additional evidence of validity. That is, in comparing the relationship between the full OTL measure and student performance and that for the OTL index reflecting only tested content, we found strong and statistically significant correlations for the latter, but a small and statistically insignificant relationship for the former. The relationship between OTL indices based on teacher report and those based on student reports with student performance further supported the decision to base the index on students. That is, not only was the student-based measure more reliable, being based on 25 or more student responses, but it also showed a much stronger relationship to student performance, suggesting that it was a better measure of what students effectively had had an opportunity to learn.

**EVIDENCE OF DIFFERENTIAL OTL**

The strong relationship that the study found between classroom-level OTL and student performance is both a substantive finding of the study and additional evidence of the validity of the measure. The results of HLM analyses suggest that OTL has important effects on student performance, even after controlling for students’ prior ability and prior academic preparation. In light of the strong relationship found between OTL and student performance, the observed differences in class-level OTL reported by ELL students as compared with non-ELL students is reason to give pause. Descriptive results indicate that on average, ELL students reported OTL on approximately 7 of 11 assessed topics, whereas non-ELL students from the same set of classes reported OTL on approximately 9, a difference of more than 20% and a sizeable effect size of more than 1. Similarly, the study found high and significant correlations between both teacher- and student-reported OTL indices and classroom-level language status variables. For example, we found a correlation of .80 between mean classroom scores on LAS, a measure of language proficiency and student-reported OTL, and a .62 correlation with teacher-reported OTL, suggesting that students with lower language proficiency have less access to learning opportunities.

When we examined the relationship between the concentration of ELLs in classrooms and both student-reported OTL and student
learning, we similarly found significant relationships: Although correlational analyses showed a strong negative relationship between the percentage of ELLs in a class and students’ OTL in class, the HLM analyses indicated a strong negative relationship between such percentages and student learning. Importantly, the HLM analyses indicate that classroom OTL is an important variable in addition to class-level language status. Although the two variables may be confounded, they do suggest that classrooms with high concentrations of ELLs perform better when they have higher levels of OTL.

The collective results from the study, in short, suggest that current debates about bias in testing for ELL students ought to give at least as much attention to bias in OTL. Our data suggest that differential OTL may indeed play a role in the depressed performance of ELLs.

SOURCES OF BIAS

The findings are suggestive of the reasons underlying such bias, or at least the differential levels of OTL. Our results indicate, not very surprisingly, that students’ ability to understand instruction, as exemplified by self-reported difficulty in understanding teacher directions in English, influences reported levels of OTL. Students who experience more difficulty report lower levels of OTL. Do ELL students who do not fully understand their teachers’ instruction in English report lower levels of OTL because they do not understand or thus recognize topics that have been addressed in class? Might the difference be due to their teachers slowing down the instructional pace to accommodate ELLs language needs, or due to their teachers not having sufficiently high expectations or not being sufficiently expert to engage ELLs in effective teaching and learning activities? Our study results raise such questions, including a sample that included uncredentialed teachers, but cannot provide specific answers.

LIMITATIONS AND IMPLICATIONS

Our study had some limitations; however, we think these generalize to other studies as well. One issue is the quality of the outcome measures used to gauge student learning. For this study, we were careful to construct a measure that only included items measuring content that was supposed to have been covered in the curriculum. Although expediency limited how well we could specify the content and select appropriate items, the technical quality of the resulting instrument was disappointing. Apparently, in seeking alignment through publicly available
measures, we sacrificed reliability. We suspect that in most studies, the compromise is in the opposite direction—that is, the measures are reliable, but their alignment with curriculum or specific OTL content is suspect (see, for example, Webb, 2007). Either way, the curriculum/outcome alignment problem is nontrivial. OTL research requires high-quality measures that are closely aligned with expected OTL. This in turn requires careful specification of content, equally careful selection or development of items, and validation of test scales and subscales. Further, although our study used an overall measure of learning, ideally one would want to be able to derive subscales to look at the relations between OTL in specific content areas and learning in those areas. Our study attempted analyses linked to individual items, but issues of stability and individual item quality hindered our success.

A second problematic issue is the meaning of OTL itself. What is a reasonable definition of OTL, and how far down into enacted instruction should it really go? Should we define OTL as exposure—exposure to the “right” content, at the “right” level of cognitive complexity, and using the “right” process? Should the concept credit intent—teachers trying to do right by all their students, regardless of language learning, and/or thinking that they are? Or to what extent should the concept really incorporate the quality of teaching and learning activity and students’ engagement in effective instruction? The issues underlying such questions may help to explain the differences we found between teacher and student reports of OTL. Recall that teachers’ estimates of OTL were higher than those of students, and yet, the student indices showed a stronger relationship to student performance than the teacher indices. In responding to our survey, teachers noted topics that they considered they had addressed in some way during the course; students responded with topics about which they remembered learning: That is, if the topic was not sufficiently noteworthy to remember or they didn’t really learn or understand it, it did not count in their view.

Our findings with regard to the relationship between language proficiency and OTL suggest the limits of only looking at exposure. Exposure clearly does not ensure effective access to curriculum and appropriate opportunities to learn. For most students, particularly English learners, exposure may be necessary, but not sufficient, for learning. To get at quality OTL for quality learning, measures must be aligned with the content and cognitive demand of what students are expected to learning, but also with research-based principles of effective ELL pedagogical practices. For example, research suggests a variety of strategies, including use of verbal and procedural scaffolding, comprehensible input, explicit language instruction, use of collaborative group work, negotiation of
meaning, and emphasis on oral skills and written and reading skills (Aguirre-Muñoz et al., 2006; August & Hakuta, 1997; Williams, 2001).

We view this small study as an interesting beginning. With this study and in future studies, we hope to encourage the development of fuller measures of OTL. In addition, we wish to explore the “within” as well as the “between” classroom differences in OTL for ELL and non-ELL students and to better grasp the sources of inequity in OTL for ELL students.

Acknowledgments

The authors had an equal level of contribution in preparing this article; therefore, the names are presented alphabetically. The work reported here was supported in part under a grant from the U.S. Department of Education, Institute for Education Sciences (IES), U.S. Department of Education, Award No. R305B960002. The findings and opinions expressed here do not reflect the position or policies of the Institute for Education Sciences or the U.S. Department of Education. The author acknowledges the contributions of several people. Rita Pope contributed substantially with comments and assistance in structuring and revising the paper. Seth Leon provided valuable assistance with the data analyses. Correspondence concerning this article should be addressed to Jamal Abedi.

References


JAMAL ABEDI is a professor at the School of Education of the University of California, Davis, and a research partner at the National Center for Research on Evaluation, Standards, and Student Testing (CRESST). Abedi’s research interests include studies in the area of psychometrics and test and scale developments. His recent works include studies on the validity of assessments, accommodations, and classification for English language learners (ELLs). Abedi is the recipient of the 2003 national Professional Service Award in recognition of his “outstanding contribution relating research to practice” by the American Educational Research Association, and the 2008 Lifetime Achievement Award by the California Educational Research Association.

JOAN HERMAN is director of the National Center for Research on Evaluation, Standards, and Student Testing (CRESST) at UCLA. Her research has explored the effects of accountability and assessment on schools and teachers and the design of assessment systems to support school planning and instructional improvement. Her recent work focuses on the quality and consequences of teachers’ formative assessment practices. Herman currently serves as editor of Educational Assessment, member at large for AERA, and a member of the Committee to Revise the Standards for Educational and Psychological Measurement. She also chairs the Para Los Niños school board.