NCAASE NATIONAL CENTER ON ASSESSMENT AND ACCOUNTABILITY FOR SPECIAL EDUCATION

Advancing research on growth measures, models, and policies for improved practice

Opportunity to Learn: A Key Access and Validity Issue of Academic Assessments for Students with Disabilities*

By: Alexander Kurz & Stephen N. Elliott, Arizona State University

ABSTRACT

An underlying assumption in our current test-based accountability system is that all participating students have an opportunity to learn the tested academic content. For students with disabilities (SWDs), this assumption is stated clearly in federal legislation intended to ensure their access to the same academic standards that define the general curriculum of students without disabilities (SWODs). To date, few research studies have systemically examined this assumption. Operationalizing the concept of opportunity-to-learn (OTL) and assessing OTL via measures that can account for teachers' instructional provisions to the overall class and individual students have been a major obstacles to systematic inquiries into OTL. One NCAASE study was designed to (a) describe OTL for students with and without disabilities and (b) examine the relationship between OTL and students' end-of-year achievement and within-year growth. This research brief provides an overview of the construct of OTL and how it is being measured in the NCAASE study on OTL and student growth.

Key Features and Definition of OTL

OTL generally refers to schooling inputs and processes necessary for producing student achievement of intended outcomes (Kurz et al., in press). For nearly five decades, researchers have examined a number of OTL indices predictive of student achievement that can be grouped into three broad categories related to the *time, content,* and *quality* of classroom instruction (see Kurz, 2011). Key studies that support a multi-dimensional characterization of OTL are:

- Time: Scheerens and Bosker (1997) examined the effect of allocated time (i.e., time on instruction) on student achievement using 21 studies with a total of 56 replications across studies. The average Cohen's d effect size for time was .39.

- Content: Gamoran et al. (1997) examined the content overlap (i.e., alignment) between instruction and a test of student achievement in high school mathematics and noted that it accounted for 25% of the variance among teachers. More recently, Kurz et al. (2010) examined the relation between OTL and student achievement averages for classes taught by general and special education teachers. The correlation between OTL and (class averages of) student achievement was .64 ($p < .05$). A multi-level reanalysis of the Kurz et al. data via hierarchical linear modeling supported classroom type and classroom-level OTL as significant predictors of individual student achievement even after controlling for prior achievement at the student level.

- Quality: Walberg (1986) reviewed 91 studies that examined the effect of quality indicators on student achievement, such as frequency of praise statements, corrective feedback, classroom climate, and instructional groupings. He reported the highest mean effect sizes for reinforcement and corrective feedback with 1.17 and .97. Gersten et al. (2009) examined various instructional components that enhanced the mathematics proficiency of students with learning disabilities and identified two instructional components that provided practically and statistically important increases in effect size: teaching students the use of heuristics (i.e., general problem solving strategy) and direct instruction.

Based on a review and synthesis of these distinct OTL research dimensions, Kurz (2011)

provided a conceptual synthesis of OTL along three dimensions of the enacted curriculum—time, content, and quality—all of which co-occur during instruction. Based on this conceptual framework, Kurz et al (in press) defined OTL "*as the degree to which a teacher dedicates instructional time and content coverage to the intended curriculum objectives emphasizing high-order cognitive processes, evidence-based instructional practices, and alternative grouping formats*" (p. 10).

Measurement of OTL

Researchers have relied primarily on direct observation or annual surveys to assess OTL. The variability of classroom instruction, however, presents unique challenges for both these options (Rowan & Correnti, 2009). To make valid inferences about OTL for the entire school year, researchers must conduct a large number of observations to ensure *generalizability*. Large-scale assessment of OTL via observations thus is often cost prohibitive. Teacher surveys represent a cost-effective alternative. End-of-year surveys, however, assume accurate teacher recall, which has been questioned (Mayer, 1999; Rowan et al., 2004). Teacher logs that are completed for a particular number of school days are designed to address both challenges. Completed for multiple days across the school year, teacher logs can (a) reduce a teacher's response burden by focusing on a discreet set of behaviors, (b) increase the accuracy of *teacher recall* by focusing on a recent time period, and (c) increase generalizability through frequent administrations across the school year (Kurz, 2011).

The use of teacher surveys or more frequently administered logs in the context of special education presents three additional challenges. First, the typical application of surveys or logs is based on the assumption that no instructional differentiation is taking place at the individual student level. This assumption has been challenged (e.g., Rowan et al., 2004; Kurz et al., 2010), especially for SWDs who should receive individualized instruction. Second, SWDs often receive their subject-specific instruction from multiple sources. Capturing OTL for a SWD in a general education classroom only, for example, may miss additional pullout sessions by a special education teacher. Finally, the reliability of self-report is frequently questioned especially when there are significant consequences associated with the results of these reports. As a result, third-party independent observations are likely to be necessary to establish the accuracy of OTL reports.

Based on the 3-dimensional conceptualization of OTL and the associated measurement challenges above, Kurz, Elliott, and Shargo (2009) developed the *Instructional Learning Opportunities Guidance System* (MyiLOGS), an online teacher log. Available evidence from an initial usability and validity study supported the measure's technical qualities and high acceptance by intended users in authentic delivery settings (Kurz, Elliott, Kettler, & Yel, in press). For more psychometric details on MyiLOGS, download the MyiLOGS Guidebook at http://www.myilogs.com.

Initial Research on OTL with MyiLOGS

To assess OTL for students with disabilities and their respective classes, personnel on the MAAPS Project (Elliott, Kettler, & Zigmond, 2009-2011), a USDE funded project, trained general and special education teachers to report on five OTL indices for their overall class and two SWDs nested within that class. This initial study was designed to address the following research questions: (a) To what extent do general and special teachers provide their 8th-grade mathematics and reading classes with an opportunity to learn the intended curriculum? (b) To what extent do the respective classwide OTL scores differ from the student-specific scores of SWDs nested within these classes? The methodology and detailed results of this study are provided in Kurz et al. (in press). The major findings from this study were:

- At the class level, general and special education teachers reported spending about two-thirds of their allocated class time on teaching the academic standards of the general curriculum, another fourth on custom objectives, and about one twentieth on non-instructional activities/tasks. In addition, teachers reported covering approximately two-thirds of the academic standards based on an average of 151 school days. Moreover, teachers generally emphasized Understand/Apply expectations as well as Independent Practice during their instruction. An examination of classwide OTL indices by class type further indicated a greater emphasis on higher-order thinking skills in general education classrooms than in special education classrooms. Moreover, classwide differences between general and special educators related to Time on Standards and Content Coverage indicated effect sizes above .50. Given that students in both types of classes were held to the

same general curriculum standards irrespective of educational setting, it is problematic that teachers in special education classes provided less instructional time on, and coverage of, the academic standards.

• Teachers' reported OTL provision differed for the overall class and SWDs nested within that class. Comparisons in the context of class type indicated that differences between classwide and student-specific OTL scores were most pronounced in general education classrooms. Based on general education classrooms, which represented a full inclusion model, SWDs experienced less time on standards, more non-instructional time, and less content coverage compared to the overall class. Given that the majority of general education classes were comprised of SWODs, the general education teachers' classwide OTL indices were most likely a closer reflection of OTL for SWODs. Each setting further represented the sole source of subject-specific instruction for target students, which raises OTL concerns for these SWDs and their participation in the same large-scale assessments as their peers. Moreover, these findings do not support the commonly held assumption in OTL research that class-wide OTL indices are sufficient for describing OTL of all students nested within that class. At least for this limited sample of SWDs nested in general education classrooms, OTL appears to be a differentiated opportunity structure when comparing classwide OTL data to that of individual SWDs.

• Findings indicating SWDs receiving instruction in general education classrooms with SWODs are reported by their teachers to actually receive fewer opportunities to learn state standards is contradictory evidence that teacher self-report measures are strongly influenced by social desirability effects. It is clearly not socially desired nor is it consistent with federal policy that SWDs receive less access to the general curriculum than their peers without disabilities. Thus, we take this as evidence that teacher reactivity to completing MyiLOGS is minimal.

Research Implications and Conclusions

The findings of this initial study have two critical implications. First, the students – both with and without disabilities - were reported by their own teachers to receive limited opportunities to learn the intended academic content standards. Although teachers were expected to address all academic content standards, general and special educators were able to address only 74% and 59% of all standards, respectively. Second, the comparisons of OTL indices at the class and student levels highlight that class-wide OTL indices are not sufficient for describing OTL of SWDs nested within that class. These initial findings suggest that SWDs did not receive equal, let alone equitable, OTL compared to their overall class along three key dimensions of the enacted curriculum. These concerns were particularly applicable to SWDs nested in general education classrooms. Differences in Non-Instructional Time were most notable. Additional research is necessary to determine why SWDs experience more Non-Instructional Time and the extent to which SWODs experience similar differences, both in terms of direction and magnitude. These findings do not explain why these SWDs received less instructional time and content coverage of the academic standards and why they experienced more non-instructional time than their overall class (e.g., disciplinary reasons, pullout services).

Another implication concerns the validity of test score interpretations that link student achievement to classroom instruction. Given some evidence that OTL is a differentiated opportunity structure, student achievement data are confounded by varying "dosages" of OTL. That is, a student's poor test performance can be due to, or in spite of, having had the opportunity to learn the tested content. If test score inferences go beyond what students know and are able to do and include interpretations that seek to attribute student achievement to adequate or effective instruction, then additional evidence to support the validity of those interpretations is recommended.

Lastly, the results of this initial research indicated that teachers can report on OTL at both the class and student levels and that, as characterized by such reporting, students with and without disabilities do not appear to have the opportunity to learn all the content that they are likely to be tested on in a statewide achievement test. When comparing classwide OTL indices to student-specific indices for SWDs in general education classes, the question arises as to how SWDs can be expected to achieve at the same levels as their peers without disabilities when they do not receive a comparable opportunity to learn content which they are expected to know on large-scale achievement measures.

Preliminary NCAASE Results

Following up on the Kurz et al. (in press), Elliott, Kurz, Tindal, and Yel (2013) began analyzing preliminary data for the NCAASE study on OTL and student growth. Year 1 data for this study involved general education teachers (N = 69) and students (N = 261) from Arizona and Oregon schools grades 4th-8th for an entire academic year. Of the students, 136 were identified as having an IEP; the remaining 125 were not known to have a disability. MyiLOGS was used daily for 7 months to log instructional time, content covered, and cognitive process expectations, instructional practices, and grouping format. easyCBM (Tindal, 2009), a curriculum-based measure consisting of three, 16-item scales (numbers and operations, algebra, measurement) was administered online 4 times (September, December, February, and May). Multiple regression analyses were used to examine the relationship of the five OTL scores, grade cluster, and disability status on (a) the end-of-year achievement and (b) achievement growth within year.

The regression analysis indicated a model with all five OTL indices accounted for 26% of the variance in end-of-year mathematics achievement. When using a change score as the growth criterion, the complete OTL model accounted for 22% of the variance. Both of these r2 were statistically significant. Subsequent regression analyses using the five OTL indices to predict both end-of-year status and change scores in mathematics for students with disabilities accounted for 32% of the variance in achievement status scores and 24% in achievement growth scores; both statistically significant. By comparison, the variance accounted for with students without disabilities was 11% and 17%, and statistically non-significant. The vast majority of the variance in achievement was accounted for by time on standards and content standards covered.

Acknowledgements

Publication Information:
*This research brief draws from the Kurz (2011) chapter and the Kurz et al. (in press) and Kurz, Elliott, Kettler, and Yel (in press) articles.

Following is the correct citation for this document.

Kurz, A., & Elliott, S. N. (2013). *Opportunity to learn: A key access and validity issue of academic assessments for students with disabilities.* Retreived from the National Center on Assessment and Accountability for Special Education (NCAASE) website: http://ncaase.com/publications/in-briefs

Funding Sources:
This research was funded in part by a USDE grant awarded to Elliott, S.N., Kettler, R.J., & Zigmond, N. from the U.S. Department of Education - Modified Alternate Assessment Participation Screening (MAAPS) Consortium (2009-2011; S368A090006), and in part by a Cooperative Service Agreement from the Institute of Education Sciences (IES) establishing the National Center on Assessment and Accountability for Special Education – NCAASE (PR/Award Number R324C110004). The findings, perspectives, and conclusions from this work do not necessarily represent the views or opinions of the U.S. Department of Education.

References

Abedi, J., Leon, S., & Kao, J. C. (2008). *Examining differential item functioning in reading assessments for students with disabilities* (CRESST Report No. 744). Los Angeles, CA: University of California, National Center for Research on Evaluation, Standards, and Student Testing.

Brophy, J., & Good, T. L. (1986). Teacher behavior and student achievement. In M. C. Wittrock (Ed.), *Handbook of Research on Teaching* (3rd ed., pp. 328-375). New York:

Croninger, R. G., & Valli, L. (2009). "Where is the action?" Challenges to studying the teaching of reading in elementary classrooms. *Educational Researcher, 38*(2), 100-108.

Elliott, S. N., & Gresham, F. M. (2008) Social skills improvement system. San Antonio, TX: Pearson.

Elliott, S.N., Kettler, R.J., & Zigmond, N. (2009-2011). *Modified Alternate Assessment Participation Screening (MAAPS) Consortium.* USDEducation, (S368A090006).

Gamoran, A., Porter, A. C., Smithson, J., & White, P. A. (1997). Upgrading high school mathematics instruction: Improving learning opportunities for low-achieving, low-income youth. *Educational Evaluation and Policy Analysis, 19*(4), 325-338.

Gersten, R., Chard, D. J., Jayanthi, M., Baker, S. K., Morphy, P., & Flojo, J. (2009). Mathematics instruction for students with learning disabilities: A meta-analysis of instructional components. *Review of Educational Research, 79*(3), 1202-1242.

Kurz, A., & Elliott, S.N. (2012). *MyiLOGS: My Instructional Learning Opportunities Guidance System Version 2.0*. Tempe, AZ: Arizona State University.

Kurz, A., & Elliott, S. N., Kettler, R. J., & Nedim, Y. (in press). Assessing students' opportunity to learn the intended curriculum using an online teacher log: Initial validity evidence. *Educational Assessment.*

Kurz, A., Elliott, S. N., Lemons, C. J., Zigmond, N., Kloo, A. & Kettler, R. J. (in press). Assessing opportunity-to- learn for students with and without disabilities. *Assessment for Effective Intervention.*

Kurz, A. (2011). Access to what should be taught and will be tested: Students' opportunity to learn the intended curriculum. In S. N. Elliott, R. J Kettler, P. A. Beddow, & A. Kurz (Eds.), *Handbook of accessible achievement tests for all students: Bridging the gaps between research, practice, and policy* (pp. 99-129). New York: Springer.

Kurz, A., Elliott, S. N., & Shrago, J. S. (2009). *MyiLOGS: My Instructional Learning Opportunities Guidance System.* Nashville, TN: Vanderbilt University.

Kurz, A., Elliott, S. N., Wehby, J. H., & Smithson, J. L. (2010). Alignment of the intended, planned, and enacted curriculum in general and special education and its relation to student achievement. *Journal of Special Education, 44*(3), 1-20.

Mayer, D. P. (1999). Measuring instructional practice: Can policymakers trust survey data? *Educational Evaluation and Policy Analysis, 21*(1), 29–45.

Porter, A. C. (2002). Measuring the content of instruction: Uses in research and practice. *Educational Researcher, 31*(7), 3-14.

Rowan, B., & Correnti, R. (2009). Studying reading instruction with teacher logs: Lessons from the Study of Instructional Improvement. *Educational Researcher, 38*(2), 120-131.

Rowan, B., Camburn, E., & Correnti, R. (2004). Using teacher logs to measure the enacted curriculum: A study of literacy teaching in third-grade classrooms. *Elementary School Journal, 105*(1), 75-101.

Scheerens, J., & Bosker, R. (1997). *The foundations of educational effectiveness.* New York: Pergamon.

Walberg, H. J. (1986). Syntheses of research on teaching. In M. C. Wittrock (Ed.), *Handbook of Research on Teaching* (3rd ed., pp. 214-229). New York: Macmillian Publishing Company.