Abstract*

Opportunity-to-Learn: The Key Access and Validity Issue for All Academic Assessments

Stephen N. Elliott
Learning Sciences Institute
Arizona State University
Steve_elliott@asu.edu

Background

Current large-scale assessments and accountability systems are predicated on the assumption that all participating students have the opportunity to learn what they are expected to know and tested on. For students with disabilities (SWDs), this assumption is made explicit in federal legislation intended to ensure their access to the same academic standards that define the general curriculum of students without disabilities (SWODs). Despite a persistent achievement gap between students with and without disabilities on state and national achievement tests, researchers have failed to systemically examine this assumption for SWDs. This failure is partly due to the conceptual and methodological challenges of 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. The lack of research is particularly disconcerting given the strong legislative and social imperatives for educating SWDs in general education settings to the greatest extent appropriate without clear evidence about the degree to which teachers are able to differentiate OTL for a diverse group of learners.

Key Features and Definition of OTL

OTL generally refers to schooling inputs and processes necessary for producing student achievement of intended outcomes. 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* (e.g., Borg, 1980; Brophy & Good, 1984; Porter, 2002). 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. Walberg reported the highest mean effect sizes for reinforcement and corrective feedback with 1.17 and .97 respectively. 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 in the context of an accountability-based curriculum framework. Accordingly, OTL can be operationalized along three key dimensions of the enacted curriculum—time, content, and quality—all of which co-occur during instruction. That is, teachers distribute OTL of what

we want students to know and be able to do by allocating instructional time and content coverage to intended objectives using a variety of pedagogical approaches. Based on this conceptual framework, Kurz and Elliott (2009) 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.*

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 options (Rowan & Correnti, 2009). To ensure *generalizability* of classroom observations, researchers must sample a large number of lessons to make valid inferences about OTL for the entire school year. The high costs associated with this approach have fostered the adoption of teacher surveys, typically conducted at the end of the school year (Porter, 2002). End-of-year surveys, however, assume accurate *teacher recall*, which has been questioned, in particular as the number of OTL indicators increases (Mayer, 1999; Rowan et al., 2004). Teacher logs represent an alternative approach that is intended to (a) reduce a teacher's response burden by focusing on a discreet set of behaviors, (b) increase 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 traditional application of surveys or logs requires teachers to report on OTL indices at the class level, which is predicated 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 according to their specific abilities and needs. Second, SWDs often receive their subject-specific instruction from multiple sources. Capturing OTL in an inclusion classroom, for example, may miss additional pullout sessions by a special education teacher. These "additive instructional scenarios" are common, even outside of special education (Croninger & Valli, 2009), and must be considered to accurately measure OTL. 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 *My Instructional Learning Opportunities Guidance System (MyiLOGS)*, an online measure of OTL. Promising 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 2012). Specifically, feedback from special and general education teachers, who integrated MyiLOGS into their regular instructional practices for up to eight months, indicated that it was easy to use, time efficient, and provided a job-embedded, personalized PD experience. Teachers noted that the regular instructional planning and reporting required by MyiLOGS offered valuable opportunities for reflection on their teaching. Intended end users in three states (AZ, PA, & SC) further acknowledged the formative benefits of the MyiLOGS instructional feedback reports and also responded favorably to a draft of the MyiLOGS Instructional Growth Plan for setting personalized instructional improvement goals. After this initial research, Kurz and Elliott (2012) refined MyiLOGS feedback report and scoring methods. For more psychometric details on MyiLOGS, download the MyiLOGS Guidebook at www.myiLOGS.com or read Kurz, et al. (2012).

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 permitted comparisons between classwide OTL scores *across* settings and between classwide and student-specific OTL scores *within* settings. This initial study was specifically designed to

address the following research questions: (1) To what extent do general and special teachers provide their 8th-grade mathematics and reading classes with an opportunity to learn the intended curriculum? (2) 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, Elliott, Lemons, Kettler, Zigmond, and Kloo (2013). 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 classwide 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 standards of the general curriculum. This finding was apparent when OTL was conceptualized in terms of instructional time and in terms of content coverage. Despite the fact that 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. The current findings raise the practical question about the extent to which students with and without disabilities receive a basic and often assumed instructional provision: adequate time on and coverage of the content standards they are expected to know. In the absence of normative data, however, it is difficult to put the current findings into context. Future large-scale assessment of OTL will be necessary to provide norm-referenced interpretations of such findings. In the meantime, teachers, schools, and districts interested in improving standardized test scores need to increase their instructional time prior to testing. Second, the comparisons of OTL indices at the class and student

levels highlight that classwide OTL indices are not sufficient for describing OTL of SWDs nested within that class. Our 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. Moreover, the current findings provide some evidence for the so-called "OTL gap" (Abedi, Leon, & Kao, 2008), which has been suggested to exist for certain student subgroups. That is, certain students may receive less OTL than others as a function of belonging to a certain subgroup (e.g., SWDs, English Language Learners). More large-scale research is needed to determine the extent to which these gaps are systemic and the reasons that these gaps are occurring. Currently, we do not know 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).

An important implication for practice lies in the remediation of potential OTL gaps through the development of teacher-level interventions. A concurrent online teacher self-report log, such as MyiLOGS, can be used to provide teachers with ongoing feedback about aspects of their classroom instruction. Given the established effects of self-recording and self-monitoring on behavior change (Elliott & Gresham, 2008), the recording and review of one's personal OTL data have the potential to induce change—especially if considered in the context of instructional coaching. The evaluation of various teacher interventions affecting malleable factors of instruction (e.g., time on standards, content coverage) seems to be an important area for future research. In addition, a tool like MyiLOGS provides a unique opportunity for multiple teachers to collaborate on instructional provisions for certain classes or students. Future research focused on the use of OTL data in conjunction with student achievement data appears to be particularly salient. Such an approach would allow teachers to use data on instructional inputs, processes, and outcomes to inform instruction and promote students' academic growth.

Another implication for practice concerns the validity of test score interpretations used to determine student achievement as a consequence of instruction. Given some evidence that OTL is a differentiated opportunity structure, student achievement data are confounded by varying "dosages" of OTL related to intended and ultimately assessed curricula. That is, a student's poor test performance can be due to, or in spite of, having had the opportunity to learn the intended and hence assessed curriculum. 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. Indicators of OTL such as the ones operationalized in this study could be used to ascertain more directly and validly the instructional provisions of teachers.

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. More research is needed prior to drawing casual conclusions about the OTL provisions for individual SWODs or the extent to which student-specific OTL indices for students with and without disabilities are different or similar to classwide OTL indices. If the findings of our initial research are found to be generalizable, concerns about both access and test score validity of students with disabilities will be confirmed. Taking steps now to improve the OTL for all students seems prudent and is achievable.

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.

Borg, W. R. (1980). Time and school learning. In C. Denham & A. Lieberman (Eds.), *Time to learn* (pp. 33-72). Washington, DC: National Institute of Education.

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., & Yel, N. (2012). *Assessing students' opportunity to learn the intended curriculum: Initial validity evidence for an online teacher log*. Unpublished manuscript.

Kurz, A., Elliott, S.N., Lemons, C.J., Kettler, R.J., Zigmond, N., & Kloo, A. (2013). *Opportunity to learn: A differentiated opportunity structure for students with disabilities in general education classrooms*. Unpublished manuscript.

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.

***Portions of this abstract are taken from a series of manuscripts authored by Kurz, Elliott, and colleagues based on the MAAPS Project findings. This abstract is intended to highlight the main presentation points on OTL and to stimulate discussion about conference participants. Citations of any content in this abstract should be of the original sources.**