Hammill Institute on Disabilities

Remedial and Special Education
2015, Vol. 36(1) 58–64

DOI: 10.1177/0741932514551282
rase.sagepub.com
SAGE

Measuring Opportunity to Learn and Achievement Growth: Key Research Issues With Implications for the Effective Education of All Students

Stephen N. Elliott, PhD[1,2]

Abstract

The related constructs of opportunity to learn (OTL) and achievement growth are fundamental aspects of the current large-scale assessment and accountability system in operation in the United States. For purposes of this article, OTL is defined 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. Both OTL and achievement growth, although frequently talked about by education policy makers and teachers, are actually difficult to measure. Much of my research, in collaboration with other colleagues, over the past decade has concerned the measurement of these constructs in ways that facilitate teachers' understanding of them and to advance sound policies regarding the instruction and assessment of students with disabilities (SWDs). In this article, I examine both the measurement of OTL and achievement growth, share recent research on them, and outline an agenda for future research on these two important aspects of standards-based accountability systems for all students.

Keywords

instructional processes, opportunity to learn, achievement growth

Since completing work in 1997 on the National Academy of Sciences' Committee on Education Goals 2000 and Services to Student with Disabilities, I have been particularly interested in ways to improve the meaningful participation and performance of all students in statewide achievement tests and accountability systems. The culminating report of this Committee was the book *Educating One and All: Students With Disabilities and Standards-Based Reform* (McDonnell, McLaughlin, & Morison, 1997). In this book, the committee members documented the low participation rates and poor performance of students with disabilities (SWDs) in many state achievement tests and what little was known about the use of testing accommodations and testing modifications to facilitate these students' test performance. The use of accommodations and modifications in the 1990s, in fact, was widely viewed as facilitating meaningful access to tests at the expense of test score validity.

With the passage of federal legislation mandating access to the general education curriculum for SWDs (Individuals With Disabilities Education Act [IDEA], 1997, 2004) and the assessment of grade-level content standards further for all students (No Child Left Behind Act [NCLB], 2002), the world of large-scale assessment significantly changed in the early 2000s. SWDs were to be fully included in all assessments, testing accommodations became best practices, and test and item modifications were fundamental aspects of the test design and accessible testing practices. And yet in 2013, many SWDs and students who struggle to learn still exhibit poor performance in annual statewide achievement tests in comparison with expected standards of achievement! The expectation is that even fewer students will likely meet the proficiency standards in mathematics and English language arts on new tests based on the more rigorous Common Core State Standards (CCSS).

For the past decade, I have focused my research on the measurement of two related constructs: opportunity to learn (OTL) and achievement growth. Both these constructs, although frequently talked about especially by education policy makers, are actually difficult to measure. Most of my research has concerned the measurement of these constructs in ways that facilitate teachers' understanding of them and to advance better policies regarding the instruction and assessment of SWDs. In the remainder of this article, I

[1]Arizona State University, Tempe, USA
[2]Australian Catholic University, Brisbane, Australia

Corresponding Author:
Stephen Elliott, Arizona State University, 308K Payne Hall, Tempe, AZ 85287, USA.
Email: steve_elliott@asu.edu

examine both the measurement of OTL and achievement growth, share recent research on them, and outline an agenda for future research on these two important aspects of standards-based accountability systems for all students.

OTL

Results of our 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 are annually tested on. For SWDs, this assumption as previously noted is made explicitly in federal legislation (i.e., IDEA, 1997, 2004; NCLB, 2002) intended to ensure their access to the same academic standards that define the general curriculum for all students. Researchers by and large, however, have failed to systemically examine this assumption for SWDs, despite a persistent achievement gap between students with and without disabilities in state and national achievement tests. This failure is likely due to the challenges of measuring the concept of OTL. The lack of research is disconcerting, given the legislative and social imperatives for educating SWDs in general education settings to the greatest extent appropriate without clear evidence that teachers are able to differentiate OTL of the intended curriculum for them, when needed.

With the advent of a new and more rigorous (Herman & Linn, 2013) intended curriculum—that is, CCSS—the degree to which teachers can provide access to the curriculum and differentiate instruction for SWDs is a serious concern. This situation provokes a number of research questions that hinge on the conceptualization and measurement of OTL and concern (a) teachers' readiness and support to deliver high-quality instruction to SWDs in the general curriculum, (b) SWDs access to the complete general curriculum, and (c) the validity of test score inferences from large-scale achievement tests aligned with the CCSS.

Definition and Key Features of OTL

OTL generally refers to schooling inputs and instructional processes necessary for producing student achievement of intended outcomes. For decades, researchers have examined OTL indices of *time, content*, and *quality of classroom instruction* and in many cases have found these to be predictive of student achievement (e.g., Borg, 1980; Brophy & Good, 1986; Porter, 2002). Based on the synthesis of this research, 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.*

Researchers until recently have relied primarily on direct observation or annual surveys (e.g., Survey of the Enacted Curriculum) to assess OTL. The variability of classroom instruction, however, presents unique challenges for both options (Rowan & Correnti, 2009). 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., Kurz, Elliott, Wehby, & Smithson, 2010; Rowan, Camburn, & Correnti, 2004), 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.

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 Instructional Learning Opportunities Guidance System (MyiLOGS; Kurz, Elliott, & Shrago, 2009; www.myilogs.com) is an example of a daily online teacher log that was originally designed to document access to the general curriculum for SWDs. Since its original conceptualization, it has been used in nearly 2,000 general and special education classrooms for mathematics and English language arts instruction of all students.

MyiLOGS is built on a unified model of OTL (Kurz, 2011). This model features indices for instructional time on standards, content coverage, and instructional quality (see Figure 1). These three indices result in five MyiLOGS OTL Summary Scores: Percentage of Instructional Time Used (IT), Percentage of Content Covered (CC), Cognitive Processes Emphasized (CP), Instructional Practices Emphasized (IP), and Grouping Formats Emphasized (GF).

What Do We Know From OTL Research

To date, there have been several studies completed where OTL was measured using MyiLOGS and involved general education classrooms with SWDs (Elliott, Roach, & Kurz, 2014; Kurz, Elliott, Lemons, Zigmond, & Kloo, 2014; Kurz, Elliott, Kettler, & Yel, 2014; Kurz et al., 2010). From these studies, we have learned the following:

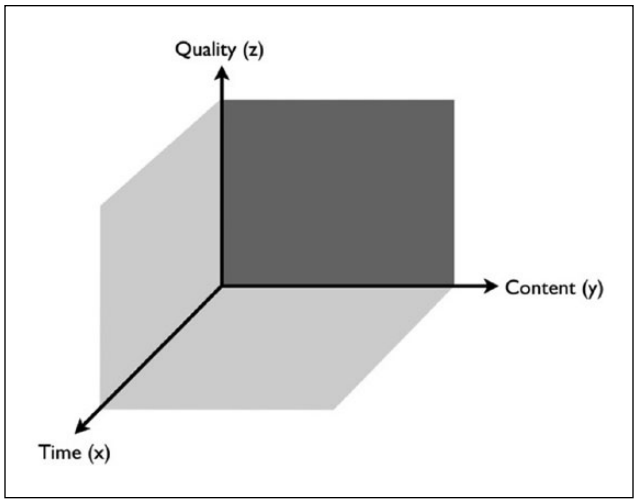

Figure 1. Conceptual model of opportunity to learn (OTL).

1. Teachers report covering approximately 67% of the academic standards based on an average of 151 school days of logged instruction prior to their state achievement test. An examination of classwide OTL indices by class type further indicated that general and special educators were able to address only 74% and 59% of all standards and there was a greater emphasis on higher-order thinking skills in general education classrooms than in special education classrooms.
2. Teachers' report 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 with the overall class.
3. Teachers who logged instructional time, content covered, cognitive process expectations, instructional practices, and grouping format for 7 months and measured math achievement four times within the year found that these five OTL indices accounted for 32% of the variance for SWDs end-of-year achievement scores and 24% in achievement growth scores; both statistically significant. The vast majority of the variance in achievement was accounted for by time on standards and content standards covered.
4. Alignment is a poor proxy for OTL, especially for special needs students.

Implications for Future Research

The findings of these recent OTL studies have several research implications as one thinks about the implementation of a new intended curriculum in most states and the instruction and use of new large-scale assessment of the achievement of SWDs.

Improved measurement of OTL. First, 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. There is a substantial need for better measurement of OTL and the development of assessment tools that are user-friendly, and yield reliable and valid information about class and individual students' OTL. Some combination of self-report online instructional log and a periodic direct third-party observation tool seems viable. Trust, but verify is called for and is possible with the measurement of OTL. Some of this work is underway, but much more is needed. Better measurement of OTL will open up four important additional areas of research:

Policy and systems-level data research. The research reviewed suggested that SWDs did not receive equal, let alone equitable, OTL compared with their classmates along three key dimensions—time, content covered, and cognitive process level—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 students without disabilities 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.

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. These findings raise practical questions about the extent to which students with and without disabilities receive adequate time on and coverage of the content standards they are expected to know. In the absence of normative data, 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.

Teacher job-embedded professional development (PD) and intervention research. An important implication for research on teachers' practices lies in the remediation of potential OTL gaps through the development and implementation of PD and teacher-level interventions. 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. 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.

Research on instruction and student achievement. Initial research based on an integrated model of OTL indicated that teaching that maximizes instructional time, content coverage, and instructional quality matters in the achievement of students, especially SWDs who typically learn at a slower pace and require some different instructional methods. More research is needed that measures daily instructional practices and their relationship to the learning of students with special needs. Such work has the potential to advance understanding of instructional and learning progressions, and also guide the development of instructionally embedded, interim assessment methods that can provide more frequent and powerful feedback to both teachers and students.

Test score validity and accountability. Finally, there are substantial research implications concerning the validity of test score interpretations when students have not had a "reasonable" opportunity to learn tested content. In fact, in the forthcoming revision of the *Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education [AERA, APA, & NCME], 2014) OTL takes on a foundational role in test fairness and test score validity for all test takers. These new standards stress that access is the opportunity for a student to learn the content of the intended and assessed curricula. In the current educational framework, this means students have meaningful opportunities to acquire the knowledge and skills featured in the content standards of their state, and ultimately assessed on the state's end-of-year achievement test. Teachers are encouraged to teach to the standards, not the test, and create engaging instruction for all students to increase the opportunity for learning to occur.

Given some evidence exists that OTL is a differentiated opportunity structure, student achievement data are confounded by varying "dosages" of OTL related to the 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 needed.

In summary, if the findings of the OTL research discussed are found to be generalizable, concerns about both instructional access and test score validity of SWDs will be confirmed. With the continued emphasis on educational assessment and accountability with new tests based on the CCSS, it is critical that such assessments are valid and appropriate, and in line with the intended and enacted curricula. Taking steps now to improve OTL research on instructional practices for all students receiving instruction within the CCSS and assessed by tests aligned with the CCSS seems prudent and is achievable.

Achievement Growth

Logically connected to OTL is the achievement of SWDs. For several decades now, the achievement of SWDs has been a concern of both educators and researchers (Carlberg & Kavale, 1980; Schulte, Osborne, & Erchul, 1998). Many states report that more than 70% of SWDs perform below proficiency on annual statewide reading and mathematics tests. For example, in a three-state study of schools that failed to make adequate yearly progress (AYP) targets, Eckes and Swando (2009) found that the most frequent reason for schools' AYP failure was the performance of the SWDs subgroup.

The centerpiece of current federal accountability legislation, the NCLB (2001), requires reporting of school-level outcomes as well as the disaggregation of achievement test scores for subgroups who have historically performed poorly relative to other students. Individual states, however, determine many of the specifics of the accountability and reporting mechanisms used to comply with this legislation. As a result, states vary considerably in their assessment instruments, testing and reporting mechanisms, definition of grade-level proficiency, and means for schools to demonstrate progress toward universal proficiency (Linn, 2008). We know relatively little about the implications of states' assessment choices in terms of the reliability of building-level scores and the validity of the inferences and decisions about schools made on the basis of these scores (Linn & Haug, 2002; Zvoch & Stevens, 2005). Unfortunately, we know even less about the reliability and validity of scores for targeted subgroups (Kiplinger, 2008).

The NCLB expectation is that all student subgroups including SWDs will meet the same achievement targets. Although NCLB considers SWDs as one undifferentiated group, specific exceptionality categories represent very different kinds of learners whose average performance differs significantly (Morgan, Farkas, & Wu, 2009; Wei, Lenz, & Blackorby, 2012). So while NCLB focuses on the

achievement gap for SWDs as a whole, I believe that it is important to study and understand differences in the achievement gap associated with specific exceptionality subgroups. Another implication of federal policy is that student subgroups that perform significantly lower in achievement on entrance into the accountability system (Grade 3) must attain greater rates of growth to meet expectations for proficiency and close the achievement gap. Regrettably, we know little about trends in growth for the SWDs as a whole and even less for students in specific exceptionality subgroups (Wei, Blackorby, & Schiller, 2011; Wei et al., 2012). Some research, however, is starting to emerge.

What Do We Know From Achievement Growth Research

Much more information about the achievement growth of SWDs is needed and how various approaches to modeling school effects fare when applied to outcomes for SWDs. Several recent growth studies are instructive when considering growth analysis of all students, but relatively few studies provide an in-depth examination of the achievement growth of students with various types of disabilities. I briefly review these studies before focusing on their implications for future research.

A foundational, although cross-sectional, study by Bloom, Hill, Black, and Lipsey (2008) provides a useful context for thinking about students' growth in achievement. Specifically, these researchers used norm group data from seven nationally standardized achievement tests to characterize typical growth in reading, math, science, and social studies as grade-level effect sizes (ES). In a pattern that was consistent across tests and content areas, students showed large annual gains (e.g., ES > 1.0) in the early grades (K-2) that decreased gradually in the later grades. They also examined performance gaps between several policy-relevant target groups (e.g., Black vs. White, male vs. female students) at each grade and content area, but did not examine gaps for SWDs. They found marked differences in achievement gap ES across different combinations of grade level and target groups. Although Bloom et al. did not examine growth or variations in the achievement gaps across grades for SWDs, their findings point to the importance of understanding grade-by-grade growth trajectories when setting growth benchmarks or evaluating the impact of interventions in closing the achievement gap between students with and without disabilities.

Surprisingly, investigators rarely have examined achievement growth in multiple disability categories (e.g., Wei, Lenz, & Blackorby, 2013; Wei et al., 2011), but when they have, considerable heterogeneity appears in intercept and/or slope of growth trajectories for different exceptionalities. For example, Wei et al. (2013), using the nationally representative Special Education Elementary Longitudinal Study (SEELS) survey, estimated mathematics achievement growth trajectories for students in 11 specific disability categories from age 7 to 17 using quadratic growth models. They found that, compared with national SEELS norms, students in all disability categories had lower math achievement levels at each age. Wei et al. (2013), however, were not able to provide direct comparisons of exceptional children's growth to students without disabilities because these students were not included in the SEELS survey. They also reported that mathematics growth rates slowed during high school and were similar across disability categories. In examining growth trajectories for specific disability categories, Wei et al. (2013) found that students with speech or visual impairments had the highest average math achievement and students with multiple disabilities or intellectual disability had the lowest achievement. At the midpoint of the age range, the ranking of disability subgroups from highest to lowest average performance was speech impairments, visual impairments, emotional disturbances, learning disabilities, other health impairments, orthopedic impairments, hearing impairments, traumatic brain injury, autism, intellectual disability, and multiple disabilities.

Addressing the limitations of the Wei et al. study, Stevens, Schulte, Elliott, Nese, and Tindal (in press) at the National Center on Assessment and Accountability for Special Education conducted a study of the mathematics achievement growth of 92,045 North Carolina students with and without disabilities in Grades 3 to 7. In particular, Stevens et al. examined the growth of SWDs in seven exceptionality categories. These categories, according to NC state terminology, were as follows: emotional disturbance, hearing impairment, educable mentally handicapped, other heath impairment, speech-language impairment, specific learning disability, and autism. Students in all subgroups showed significant growth that decelerated over grades as well as significant variability in achievement by student subgroup both at the initial assessment in Grade 3 and in rates of growth over time. Race/ethnicity, gender, parental education, free lunch status, and English language proficiency were also significant predictors of achievement. ES estimates showed substantial year-to-year growth that decreased over grades. Sizable achievement gaps that were relatively stable over grades were observed between students without disabilities and students in specific exceptionality categories.

The results of the Stevens et al. study confirmed their perspective that there is a substantial variation in growth trajectories by student subgroup that may be masked by the aggregation of subgroups into a single special education category. Although many similarities were found in growth trajectories for students in different exceptionality groups, average achievement gaps in comparison with students without disabilities ranged from approximately one quarter standard deviation for students with speech-language impairments to about two standard deviations for students with educable mental handicaps.

Implications of Research for Future Investigations

The research on achievement growth for subgroups of SWDs, to many educators and policy makers' surprise, really is just starting to emerge. The early findings show a significant, but decelerating achievement growth for students over grades. The results reported by Stevens et al. fill in a number of gaps in the literature by analyzing an operational state test's database, providing well-defined comparison groups, and expressing differences in growth over grades and differences between groups with ES benchmarks. The findings also extend a very small number of studies of achievement growth for multiple student exceptionality groups. Additional research, however, is needed that expands on these results with the same and additional exceptionality groups. It is also important that some portion of research on student achievement growth trajectories is conducted with state assessment data used for NCLB accountability reporting, so that educational program planning, decision making, and policy formation can be informed and enhanced. It is therefore important, given the high stakes surrounding accountability testing, that we more fully understand the growth of all students on state tests as well as the growth of specific subgroups of students including exceptional children.

Conclusion

The laudable goal of closing the gap between SWDs and students without disabilities is an important policy target and rich area for researchers interested in special populations. The policy, however, does not fully embrace empirical evidence about student instruction or achievement growth. For those students who are "behind" to catch up with other students, it is necessary that they are provided more instructional time, make progress at faster, more accelerated growth rates than their higher performing and often more advantaged peers.

Expectations for high-quality and aligned instruction and academic growth cannot rest solely on policy intent or a desire to set lofty goals, but must be contextualized by empirical findings that describe the instruction and growth that occurs for students as a function of their exceptionality and background. I believe, through the continued pursuit of better measures of OTL and the use of ES interpretations of student growth, the results reported by Stevens et al. provide rich additional context for characterizing and interpreting student growth for SWDs and students without disabilities. As Stevens and his colleagues noted, it is not only important to express how much growth we would like to see from a pedagogical or policy perspective but also how much growth typically occurs, how that growth is conditioned on instructional opportunities, student exceptionality, characteristics, and background, and what are reasonable expectations for future student achievement.

As you have read, progress has been made in the inclusive assessment research arena, but there remain many questions to answer and studies to be conducted if we are to realize our potential to teach and validly assess the achievement of all students, especially those with disabilities. To reinforce this call for more research, it should be noted that in June 2014, as this article was being finalized, the U.S. Department of Education's Institute for Education Sciences called for the creation of a National Research and Development Center on Standards in Schools. This national center is to focus on the role that opportunity to learn new content standards has on the achievement growth of all students, especially those SWDs and English language learners. This new center has the potential to address many of the questions I have posed in this article; however, contributions from many other researchers will also be needed. I look forward to continuing to work in this arena and hope some of the readers of this article will too in the near future.

Author's Note

Some of the content of this article was orally presented by the author at the Institute of Education Sciences' Technical Work Group on *Researching College- and Career-Ready Standards to Improve Student Outcomes*, Washington, D.C., August 20, 2013. The opinions expressed are those of the author and do not necessarily represent views of the Institute or the U.S. Department of Education.

Declaration of Conflicting Interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: MyiLOGS, a measure mentioned in this article, is copyrighted by Arizona State University (ASU). ASU contracts with IPM Innovations, a limited liability company of which the author is a part owner, to conduct MyiLOGS training and distribution.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The manuscript is a thought piece, but the bases for many of the thoughts about measuring OTL and achievement growth are the result of research being conducted for the National Center on Assessment and Accountability for Special Education, funded by the Institute of Education Sciences, U.S. Department of Education, through Grant R32C110004 awarded to the University of Oregon.

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: National Center for Research on Evaluation, Standards, and Student Testing, University of California.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: Author.

Bloom, H. S., Hill, C. J., Black, A. R., & Lipsey, M. W. (2008). Performance trajectories and performance gaps as achievement effect-size benchmarks for educational interventions. *Journal of Research on Educational Effectiveness*, *1*, 289–328.

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, NY: Macmillan.

Carlberg, C., & Kavale, K. (1980). The efficacy of special versus regular class placement for exceptional children: A meta-analysis. *The Journal of Special Education*, *14*, 295–309. doi:10.1177/002246698001400304

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

Eckes, S. E., & Swando, J. (2009). Special education subgroups under NCLB: Issues to consider. *Teachers College Record*, *111*, 2479–2504.

Elliott, S. N., Kurz, A. K., Tindal, G., & Yel, N. (2014, April). *The relationship of opportunity to learn and interim assessment results to end-of-year mathematics achievement*. Paper presented at the annual convention of the National Council on Measurement in Education, Philadelphia, PA.

Elliott, S.N., Roach, R.T., & Kurz, A. (2014). Evaluating and advancing the effective teaching of special educators with a dynamic instructional practices portfolio. *Assessment for Effective Intervention*, *39*, 83-98. doi:10.1177/1534508413511491

Herman, J., & Linn, R. (2013, January). *On the road to assessing deeper learning: The status of smarter balance and PARCC Assessment Consortia* (CRESST Report 823). Los Angles: University of California at Los Angeles.

Individuals With Disabilities Education Act Amendments of 1997, 20 U.S.C. §1400 *et seq*.

Individuals With Disabilities Education Act Amendments of 2004, 20 U.S.C. §1400 *et seq*.

Kiplinger, V. (2008). Reliability of large-scale assessments and accountability systems. In K. E. Ryan & L. A. Shepard (Eds.), *The future of test-based educational accountability* (pp. 93–114). New York, NY: Routledge.

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, NY: Springer.

Kurz, A., & Elliott, S.N. (2009). *My instructional Learning Opportunity Guidance System (MyiLOGS)*. Nashville, TN: Vanderbilt University.

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

Kurz, A., Elliott, S. N., Kettler, R. J., & Yel, N. (2014). Assessing students' opportunity to learn the intended curriculum: Initial validity evidence for an online teacher log. *Educational Assessment*, *19*, 159–184. doi:10.1080/10627197.2014.934606

Kurz, A., Elliott, S. N., Lemons, C. J., Zigmond, N., & Kloo, A. (2014). Opportunity to learn: A differentiated opportunity structure for students with disabilities in general education classrooms. *Assessment for Effective Intervention*. doi:10.1177/1534508414522685

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. *The Journal of Special Education*, *44*(3), 1–20.

Linn, R. L. (2008). Methodological issues in achieving school accountability. *Journal of Curriculum Studies*, *40*, 699–711.

Linn, R. L., & Haug, C. (2002). Stability of school building accountability scores and gains. *Educational Evaluation and Policy Analysis*, *24*, 29–36.

McDonnell, L. M., McLaughlin, M. J., & Morison, P. (Eds.). (1997). *Educating one & all: Students with disabilities and standards-based reform*. Washington, DC: National Academy Press.

Morgan, P. L., Farkas, G., & Wu, Q. (2009). Five-year growth trajectories of kindergarten children with learning difficulties in mathematics. *Journal of Learning Disabilities*, *42*, 306–321.

No Child Left Behind Act of 2001, Pub. L. No. 107–110, § 115, Stat. 1425 (2002).

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

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*, 75–101.

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

Schulte, A. C., Osborne, S. S., & Erchul, W. P. (1998). Effective special education: A United States dilemma. *School Psychology Review*, *27*, 66–76.

Stevens, J. J., Schulte, A. C., Elliott, S. N., Nese, J. F., & Tindal, G. (in press). Mathematics achievement growth of students with and without disabilities on a statewide achievement test. *Journal of School Psychology*.

Wei, X., Blackorby, J., & Schiller, E. (2011). Growth in reading achievement in a national sample of students with disabilities ages 7 to 17. *Exceptional Children*, *78*, 89–106.

Wei, X., Lenz, K. B., & Blackorby, J. (2012). Math growth trajectories of students with disabilities: Disability category, gender, racial, and socioeconomic status differences from ages 7 to 17. *Remedial and Special Education*, *34*, 154–165.

Zvoch, K., & Stevens, J. J. (2005). Sample exclusion and student attrition effects in the longitudinal study of middle school mathematics performance. *Educational Assessment*, *10*, 105–123.