NCME 2012 Paper

Measurement of Opportunity to Learn and its Contribution to Achievement Gains for Students with Disabilities

Stephen N. Elliott & Alexander Kurz
Learning Sciences Institute
Arizona State University

Abstract

This study was designed to develop initial validity evidence for MyiLOGS, a tool to capture opportunity to learn (OTL) data, and to examine the relationship of this measure to students' end of year achievement. The applied methodology underlying this assessment tool is an extension of teacher logs via an online technology that provided teachers a self-report structure for logging key OTL indices immediately after their daily instruction. Based on the study's three-state sample of eight-grade teachers, we found teachers could use the tool with high integrity and covered approximately two-thirds of their state's academic content standards during an average of about 151 school days. The resulting OTL data indicated that students with disabilities received less instructional time and content coverage related to the state-specific standards compared to their classmates, while also experiencing more non-instructional time than their peers. These findings, among others, led to the following conclusions about MyiLOGS: (a) teachers can be trained to report reliably on various OTL indices that provide a valid account of classroom instruction as supported by third party observations; (b) the applied online technology offered teachers a feasible and time efficient method for collecting OTL data at the class and student level on a daily basis across the school year; (c) the resulting system shows promise for a large-scale collection of OTL data; and (d) future OTL research is needed to confirm OTL as a differentiated opportunity structure for students with disabilities and to establish further validity evidence for the collected indices. Limitations and future research are discussed to advance research on opportunity to learn particularly for students with disabilities who are being educator in general education classrooms.

Note

Significant portions of this conference paper are concurrently being used in a manuscript currently under review for publication. The content of this paper has not been previously presented at a national conference. If you wish to cite this paper, please consider contacting the senior author (steve_elliott@asu.edu) for an update on the more comprehensive manuscript and to acquire the most current citation.

Many educational stakeholders, policymakers, and researchers are interested in what children are learning in school and how teachers use class time. Every state in the country has content standards that they intend teachers to teach and students to learn, and they virtually all have 180 days of class time for teachers to deliver instruction. The research reported here focuses on the development of an instrument for quantifying the extent to which all students have the opportunity to learn (OTL) the intended curriculum as measured by instructional indicators of the enacted curriculum. This research was part of a larger investigation access to the general curriculum and the relationship of this access to students' performance on state wide assessments[1].

What is OTL?

For over 50 years, researchers have examined the inputs and processes necessary for producing important student outcomes (McDonnell, 1995). To this end, they have used a variety of indicators associated with the instructional dimensions of time, content, and quality of instruction (Kurz, 2011). Anderson (1986) acknowledged the use of varying conceptual definitions of OTL and suggested that "A single conceptualization of opportunity to learn coupled with the inclusion of the variable[s] in classroom instructional research . . . could have a profound effect on our understanding of life in classrooms" (p. 3686). Consistent with Anderson's suggestion, Stevens and Grymes (1993) proposed a "unified conceptual framework" of OTL to investigate "students' access to the core curriculum." that focused on four elements: content coverage, content exposure, content emphasis, and quality of instructional delivery. Although untested, their conceptualization of OTL has been adopted by a number of researchers (e.g., Abedi, Courtney, Leon & Azzam, 2006; Aguirre-Munoz et al., 2006; Wang, 1998; Herman & Abedi, 2004) and has resulted in OTL being characterized as a teacher effect related to the allocation of adequate instructional time for teaching a curriculum that requires various cognitive demands and instructional practices to produce student achievement.

We define 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 practices, and alternative grouping formats. The theoretical and empirical rationales for this definition are provided and used as the conceptual foundation for the constructs measured by *My instructional Learning Opportunity Guidance System* or MyiLOGS (Kurz, & Elliott, 2010), a new measure of OTL. The findings of this study have implications for educational equity, compliance with federal legislation, student achievement, and the validity of test score interpretations.

Why measure OTL?

Both legal and empirical rationales stress the importance of examining OTL for students with and without disabilities. Legally, federal legislations mandate students' access to the general education curriculum including its academic standards (Individuals with Disabilities Education Act, 1997; Karger, 2005). In addition, the mandated participation of all students in assessments that assess grade-level standards further necessitates instruction with the content of these standards to ensure the validity of certain test score inferences (No Child Left Behind Act, 2002; Wang, 1998).

An empirical rationale for the measurement of OTL for students with disabilities is derived from research on the limited use of allocated time for instruction (Vannest & Hagan-Burke, 2010), low exposure to standards-aligned content (Kurz, Elliott, Wehby, & Smithson, 2010), and inconsistent use of evidenced-based instructional practices (Burns & Ysseldyke, 2009), and poor instructional quality (Vaughn, Levy, Coleman, & Bos, 2002). Existing options for measuring aspects of OTL, such as the Surveys of the Enacted Curriculum (Porter & Smithson, 2001), are based on the concept of alignment, which addresses only one aspect of OTL at the class-wide level of instruction. Such measurement privileges content coverage and fails to account for OTL as a differentiated opportunity structure (Kurz, 2011), a vital concern for students with disabilities who are by law to receive instruction according to their individual needs.

Data on students' opportunity to learn the standards of the general curriculum is also critical to the validity of test score inferences from state test results used for accountability. The inclusion of all students with disabilities in test-based accountability is challenging, but still must result in reliable test scores that permit valid inferences about student achievement and the extent to which teachers and schools can be held accountable for this achievement. Not surprisingly, current accountability provisions in many states are making inferences about the instructional effectiveness of teachers and schools on the basis of test scores. Such test score interpretations, however, go beyond inferences about what students know and are able to do. These interpretations generally seek to attribute high student achievement to adequate instruction and low student achievement to inadequate instruction. These test score interpretations are subject to additional validity evidence including evidence of students' opportunity to learn the material that is being tested (AERA, APA, & NCME, 1999; D'Agostino et al., 2007; Kurz & Elliott, 2011). Thus, the importance of measuring OTL is not only critical to the validity of test score inferences but also to the overall premise underlying standards-based reform.

How Can OTL be Measured?

Key dimensions of OTL are time, content, and quality (Kurz, 2010). Thus, one necessary multifaceted construct measured by an OTL instrument must be time. To provide students with the opportunity to learn the intended curriculum, teachers must invest instructional time dedicated to the respective knowledge and skills implicated in the intended curriculum. As such, previously used indicators of time such as "allocated time" are not suitable for operationalizing this OTL construct. Rather data are needed on time spent on teaching the academic standards of the general curriculum and, if applicable, any intended skills prescribed by a student's IEP. Prior research on time and learning further provides empirical support for examining student engagement and portion of work/task completion in conjunction with instructional time.

Content taught is also a multifaceted construct that must be integrated into the measurement of OTL. To provide students with the opportunity to learn the intended curriculum, teachers must cover the content implicated in the intended curriculum. Of interest is a teacher's content coverage of the academic standards of the general curriculum and, if applicable, any intended skills prescribed by a student's IEP. OTL indicators related to only "tested content" don't provide enough information. The desirable target of classroom instruction should be the broader intended curriculum, which subsumes the content of the assessed curriculum.

Instructional quality is the final multifaceted construct to be measured as part of a unified definition of OTL. To provide students with the opportunity to learn the intended curriculum, teachers can employ a range of instructional practices that have received empirical support. These include the use of guided feedback (e.g., Brophy & Good, 1986), reinforcement (e.g., Walberg, 1986), direct instruction (e.g., Gersten et al., 2009), student "think alouds" (e.g., Vaughn et al., 2000), and visual representations (e.g., Gersten et al., 2009). In addition, researchers have identified grouping formats other than whole class (e.g., Elbaum et al., 2000) and cognitive processing demands (e.g., Porter, 2002), as important qualitative aspects of instruction. With respect to cognitive processing demands, several classification categories ranging from lower-order to higher-order cognitive processes have been suggested, most notably in Bloom's taxonomy of education objectives (Bloom, 1976). Three quality indicators have been identified: cognitive processing expectations, evidence-based instructional practices, and grouping formats. Little evidence exists to suggest one of these indicators is better than the other.

The multidimensional MyiLOGS model of OTL further represents each instructional dimension as a continuum that originates at zero. On the basis of theory and research, we established five OTL indicators and provided suggestions for operationally defined indices. These indicators and indices are defined in Table 1.

Table 1

Instructional Dimensions, Indicators, Definitions, and Operational Indices of OTL

Dimension	Indicator	Definition	Index
Time	Instructional Time	Instructional time dedicated to teaching the general curriculum standards and, if applicable, any intended IEP objectives.	***IT***: Average amount of instructional minutes spent on intended curriculum objectives per day.
Content	Content Coverage	Content coverage of the general curriculum standards and, if applicable, any intended IEP objectives.	***CC***: Percentage of addressed intended curriculum objectives.
Quality	Cognitive Processes	Emphasis of cognitive process expectations along a range of lower-order to higher-order thinking skills.	***CP***: Sum of differentially weighted percentages of instructional time dedicated to each cognitive process expectation.
	Instructional Practices	Emphasis of instructional practices along a range of generic to empirically supported practices.	***IP***: Sum of differentially weighted percentages of instructional time dedicated to each instructional practice.
	Grouping Formats	Emphasis of grouping formats along a range from individual to whole class instruction.	***GF***: Sum of differentially weighted percentages of instructional time dedicated to each grouping format.
Note. Emphasis can be operationalized as the amount of instructional minutes.			

What Should be Considered When Assessing OTL in Classrooms?

Data on the five OTL indices specified in Table 1 can be collected through a variety of methods including teacher self-report and direct observation. For purposes of our conception of OTL, the instructional dimensions of OTL related to time, content, and quality can be operationalized and subsequently documented using (a) observers who conduct classroom observations or code videotaped lessons, or (b) teachers who self-report on their classroom instruction via annual surveys or daily logs. Third-party observations and teacher surveys are the most frequently used methods, each with a unique set of advantages and challenges (Rowan & Correnti, 2009).

Third-party or independent observations are often considered the "gold standard" for classroom research, but the high costs associated with this method limit its large-scale application for measuring OTL. Moreover, the complexity and variability of classroom instruction across a school year (Jackson, 1990; Rogosa, Floden, & Willett, 1984) raise generalizability and representativeness questions. For example, how many observations are necessary to generalize to a teacher's entire enacted curriculum? Annual surveys, on the other hand, are relatively inexpensive but rely exclusively on teacher memory for the accurate recall of the enacted curriculum. To address these assessment and measurement challenges, Rowan, Camburn, and Correnti (2004) suggested a third alternative, namely the use of periodically completed teacher logs. Teacher logs are 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 by using a representative sample of school days across a school year.

As part of their Reform Up Close study, Porter and colleagues used a variety of methods to collect data on teachers' enacted curriculum including daily logs, weekly surveys, classroom observations, and questionnaires (Porter, Kirst, Osthoff, Smithson, & Schneider, (1993). The agreement was consistently above .60 between classroom observations and teacher log data (calculated on each observation pair and averaged over all pairs) along the dimensions of broad content area (A), subskills within broad content area (AB), delivery of content (C), and cognitive demand (D). Porter and colleagues also noted significant correlations between log data and questionnaire data on content area of .50 to .93 in mathematics and of .61 to .88 in science (Smithson & Porter, 1994). Based on his own research and a number of studies investigating the validity of survey data, Porter (2002) stated "survey data is [*sic*] excellent for describing quantity—for example, what content is taught and for how long—but not as good for describing quality—for example, how well particular content is taught" (p. 9).

For purposes of validating teacher logs, Camburn and Barnes (2004) discussed the challenges of reaching (inter-rater) agreement as one of their validation strategies including rater background, type of instructional content, level of detail (e.g., subskills) associated with content, and frequency of occurrence. Agreement percentages across eight literacy topics between observers and teachers ranged between 37% and 75% (average agreement of 52%) using four levels of emphasis (i.e., primary focus, secondary focus, touched on only briefly, not a focus). The agreement percentages between two observers ranged between 52% and 90% (average agreement of 66%). On the basis of their results, Camburn and Barnes expressed confidence in teacher logs to measure instruction at grosser levels of detail and for activities that occurred more frequently. Rowan and Correnti (2009) concluded that teacher logs are (a) "far more trustworthy" than annual surveys to determine the frequency with which particular content and instructional practices are enacted; and (b) yield "nearly equivalent" data to what would be gathered via trained observers. Still, however, it seems that third-party observations are more highly regarded by a number of researchers for determining aspects of child-instruction or teacher-child interactions (Connor et al., 2009; Pianta & Hamre, 2009).

What should be measured? Based on our review of research, the answer to “what” should be measured for purposes examining OTL is: instructional time, content covered in the intended curriculum, and teacher actions that involve higher-order cognitive processes, evidence-based practices, and alternative grouping formats.

At what level should we measure classroom instruction? This question is also related to the nested nature of OTL within classrooms and challenges researchers to locate and sample for OTL. One of the first challenges is to decide the number of measurement occasions for purposes of documenting OTL at the enacted curriculum level. Rowan and Correnti (2009), who used daily teacher logs to measure different aspects of a teacher’s enacted curriculum, decomposed variance in time spent on reading/language arts instruction into three levels: time on instruction on a given *day* (Level 1), days nested within *teachers* (Level 2), and teachers nested within *schools* (Level 3). They found time on instruction varied from day to day with approximately 72% of the variance in instructional time among *days*, about 23% among *teachers* within schools, and about 5% among *schools*. Given their measurement system, Rowan and Correnti suggested 20 logs per year was optimal to reliably discriminate among teachers.

Connor et al. (2009), who used an observational measure, reported that different students nested within the same class may be experiencing different amounts and types of instruction. This finding highlights the need to think about the appropriate measurement level of OTL. That is, should OTL be documented at the student level or the classroom level? Most teacher self-report measures collect information on the enacted curriculum at the class level. Given evidence of significant variation along key instructional dimensions of OTL for students within the same class, data on classroom-level OTL cannot necessarily be generalized to individual students and, in the case of students with disabilities, cannot yield information on the extent to which students’ had the opportunity to learn their specific intended curriculum.

Croninger and Valli (2009) identified additional challenges related to the variability of instruction. Results from their 5-year longitudinal study of teaching reading in schools of poverty indicated that the majority of students received reading instruction from multiple sources in one or more locations. Croninger and Valli further noted that many students experienced more reading instruction outside their scheduled reading class than during their scheduled lesson. These findings underscore the challenge of accounting for all sources of instruction that contribute to a student’s opportunity to learn his or her intended curriculum.

Research Questions Addressed

The three questions that focused this research with MyiLGOS were:

1. Can teachers be trained to use MyiLOGS with high integrity to yield reliable OTL indices?

2. To what extent is there convergent and predictive validity evidence for the MyiLOGS indices?

3. What are the relations between student-based MyiLOGS indices and student achievement?

Method

Participants

Teachers in Arizona, Pennsylvania, and South Carolina were trained to report on OTL indicators via MyiLOGS for their 8th-grade Mathematics (MA) and English/Language Arts (ELA) classes and two students with disabilities nested within their classes during the 2010-2011 school year. The teacher participant sample featured 38 general and special education teachers from seven middle schools in Arizona (*n* = 15 teachers), five middle schools in Pennsylvania (*n* = 12 teachers), and five middle schools in South Carolina (*n* = 11 teachers). Each general and special education teacher had to provide MA and/or ELA instruction to two 8th-grade students with disabilities. In case multiple teachers were involved in the instructional provision for these target students, participation was further contingent on the voluntary consent of all involved teachers. This inclusion criterion was employed to ensure that every teacher who provided target students with the opportunity to learn the subject-specific intended curriculum was participating in the study. The final sample included three co-teaching pairs in Arizona. All three co-teaching pairs were asked to discuss their respective instructional provisions prior to the general education teacher reporting on OTL via MyiLOGS. The state- and subject-specific breakdowns of schools, teachers, classrooms, and target students are described in Table 2. Several teacher participants logged multiple classrooms within or across subjects, which featured some of the same target students. To highlight this overlap across content areas, Table 2 also lists unique teachers and target students. In South Carolina, two classrooms featured only one target student due to school transfers during the year. In sum, the final subject-specific samples across states were comprised as follows: (a) 19 teachers provided OTL data on 20 MA classes featuring 39 nested target students; and (b) 23 teachers provided OTL data on 26 ELA classes featuring 50 nested target students. All teacher participants were compensated for their time spent on study-related tasks. Each teacher received a $150 honorarium for participation in the MyiLOGS training, $100 per month for using MyiLOGS to report on daily classroom instruction, and $175 for the completion of the SEC at the end of study. It should be noted that all participants continued instructional logging through April, which resulted in data collection for five to eight months depending on start date.

Table 2
Breakdown of Schools, Teachers, Classrooms, and Target Students by State and Subject

Sample	Arizona			Pennsylvania			South Carolina		
	MA	ELA	Unique	MA	ELA	Unique	MA	ELA	Unique
Schools			7			5			5
Teachers	8	7	15*	5	8	12	6	8	11
Classes	9	7		5	8		6	11	
Target Students	18	14	22	10	16	19	11	20	15

Note. MA = Mathematics; ELA = English/Language Arts.
[a]Includes three special education co-teachers.

Table 3

Teacher Participant Characteristics by State

Characteristic	AZ		PA		SC		Total	
	n	***(%)***	***n***	***(%)***	***n***	***(%)***	***n***	***(%)***
Gender								
Female	9	*(75)*	11	*(92)*	9	*(82)*	29	*(83)*
Male	3	*(25)*	1	*(8)*	2	*(18)*	6	*(17)*
Ethnicity								
African American	1	*(8)*	0	*(0)*	0	*(0)*	1	*(3)*
Asian American	0	*(0)*	0	*(0)*	0	*(0)*	0	*(0)*
Caucasian	9	*(75)*	12	*(100)*	11	*(100)*	32	*(91)*
Hispanic	2	*(17)*	0	*(0)*	0	*(0)*	2	*(6)*
Other	0	*(0)*	0	*(0)*	0	*(0)*	0	*(0)*
Role								
General Education	12	*(100)*	7	*(58)*	5	*(45)*	24	*(69)*
Special Education	0	*(0)*	5	*(42)*	6	*(55)*	11	*(31)*
Degree								
Bachelor	5	*(42)*	2	*(17)*	6	*(55)*	13	*(37)*
Master	7	*(58)*	10	*(83)*	5	*(46)*	22	*(63)*
Doctorate	0	*(0)*	0	*(0)*	0	*(0)*	0	*(0)*
	M	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***
Years of Experience	7.6	7.4	10.1	8.4	10.7	8.5	9.4	8.0
PD Hours[a]	79	83	141	168	205	216	140	166

Note. AZ = Arizona; PA = Pennsylvania; South Carolina = SC; PD = Professional Development.
[a]Indicates PD hours on state- and district-specific academic standards during the past five years.

All teachers who reported on OTL via MyiLOGS completed a teacher profile. Table 3 displays teacher characteristics by state including years of experience and professional development hours on state- or district-specific academic standards during the past five years. The teacher sample was predominately female and Caucasian with a majority of teachers holding a graduate degree. The Arizona subsample was exclusively comprised of general education teachers because the three special education co-teachers did not complete a profile.

For purposes of establishing the target student sample, state personnel assisted teachers in randomly selecting two students with disabilities nested in each studied classroom. To be eligible for selection, a target student had to have a current IEP that indicated his or her participation in either the regular state assessment or the state's grade-level alternate assessment (i.e., AA-MAS). This selection criterion was used to ensure that all target students were within the legal and legislative framework that mandated their grade-level instruction in the academic content standards of the general curriculum. The target student sample (N = 56) was largely comprised of males and students with learning disabilities. The Arizona subsample was predominately Hispanic and the subsamples in Pennsylvania and South Carolina were predominately Caucasian and African American, respectively. The Arizona subsample further featured a very large proportion of students on free/reduced lunch.

Measures

All teacher participants completed MyiLOGS (Kurz, Elliott, & Shrago, 2009), the SEC (Porter & Smithson, 2001), and a pre-study evaluation survey of MyiLOGS training and a post-study MyiLOGS use survey. Teachers also administered their state-specific annual state achievement test used for accountability purposes. MyiLOGS served as the primary measure for determining students' opportunity to learn the intended curriculum. The SEC and achievement measure provided indices of similar and dissimilar constructs for initial validity analyses.

My instructional Learning Opportunities Guidance System (MyiLOGS). This online technology (www.myilogs.com) is designed to assist teachers with the planning and implementation of intended curricula at the class and student level. MyiLOGS features the state-specific academic standards of the general curriculum for various subjects and additional customizable skills that allow teachers to add student-specific objectives (e.g., IEP objectives). MyiLOGS provides teachers with a monthly instructional calendar that includes an expandable sidebar, which lists all intended objectives for a class (See Figure 1). Teachers drag-and-drop planned skills that are to be the focus of the lesson onto the respective calendar days and indicate the approximate number of minutes dedicated to each skill. After the lesson, teachers are required to confirm enacted skills, instructional time dedicated to each skill, and any time not available for instruction (due to transitions, class announcements, etc.) at the class level. In addition, two randomly selected days per week require further documentation. On these sample days, teachers report on additional time emphases related to the skills listed on the calendar including cognitive expectations, instructional practices, grouping formats, engagement, goal attainment, and time not available for instruction. This detailed reporting occurs at the class and student level along two two-dimensional matrices and two ratings. Teachers can further review a range of charts and tables that provide detailed information on their enacted curriculum and its relation to the intended curriculum (i.e., subject-specific academic standards and custom objectives). At the time of this initial study, the instructional reports were not available until after the project was completed. Screenshots of the MyiLOGS sample day matrices and ratings are displayed in Figure 2.

Figure 1. Screenshot of the MyiLOGS calendar interface.

Estimated Time Allocation Across Cognitive Process Dimensions for: Tunnell Gr. 8 Math

Skill	Attend	Remember	Understand/Apply	Analyze/Evaluate	Create	Sum	Calendar Minutes
S3C3PO3 Linear equations and inequalities	5	20	45	0	0	70	70
Concept Review Bell Work	0	5	5	0	0	10	10
Time Not Available for Instruction						0	0
				Update Totals	Total:	80	80

Estimated Time Allocation Across Instructional Practices for: Tunnell Gr. 8 Math

Teacher Actions	Individual	Small Group	Whole Class	Sum
Provided Direct Instruction	0	0	10	10
Provided Visual Representations	0	0	5	5
Asked Questions	0	5	5	10
Elicited Think Aloud	0	0	0	0
Used Independent Practice	0	0	0	0
Provided Guided Feedback	5	5	0	10
Provided Reinforcement	0	0	10	10
Assessed Student Knowledge	0	0	0	0
Other Instructional Practices	0	0	35	35
Time Not Available				0
		Update Totals	Calendar Total: 80	80

Engagement Matrix for: Tunnell Gr. 8 Math

Class Engagement	Learning Goal Attainment
○ Not Engaged (0%)	○ No effort or product observed (0%)
○ Low % of time (<50%)	○ Low effort or limited portion of work completed (<50%)
○ Moderate % of time (50% - 80%)	○ Moderate effort or moderate portion of work completed (50% - 80%)
◉ High % of time (>80%)	◉ High effort or substantial portion of work completed (>80%)

Figure 2. MyiLOGS sample day matrices and ratings.

For the first matrix, teachers report on the instructional minutes allocated per skill along five cognitive process expectations for student learning adapted from the revised version of Bloom's taxonomy (Anderson, 2002). For the second matrix, teachers report on the instructional minutes allocated per instructional practice along three grouping formats. Teachers further rate engagement and goal attainment along a 4-point scale. Student engagement and successful work completion are two previously discussed indicators for purposes of determining academic learning time. The definitions for the cognitive process expectations and instructional practices are provided in Tables 4 and 5. The grouping formats were defined as: (a) *Individual*: Instructional action is focused on a single individual; (b) *Small group*: Instructional action is focused on small groups; (c) *Whole Class*: Instructional action is focused on the whole class.

To minimize teachers' response burden, the related cognitive processes *Understand* and *Apply* as well as *Create* and *Analyze* were collapsed in the cognitive process matrix. In this study, the cognitive process matrix further included the *Attend* category, which is not part of the revised Bloom's taxonomy. The cognitive expectation of *Attend* allowed teachers to differentiate between the expectation of students (passively) listening to instructional tasks and related instructions and (actively) recalling information such as a fact, definition, term, or simple procedure.

Table 4

Cognitive Process Expectations for Student Learning and Definitions

Cognitive Process	Definition
Attend	Orient toward instructional task and related instructions. ▪ Synonyms include *listen, focus, pay attention*.
Remember[a]	Retrieve relevant knowledge from long-term memory. ▪ Synonyms include *recognize, identify, recall, retrieve*.
Understand[a]	Construct meaning from instructional messages. ▪ Synonyms include *interpret, exemplify, classify, summarize, infer, compare, explain*.
Apply[a]	Carry out or use a procedure in a given situation. ▪ Synonyms include *execute, implement, use*.
Analyze[a]	Break materials into its constituent parts and determine how the parts relate. ▪ Synonyms include *differentiate, organize, integrate, attribute*.
Evaluate[a]	Make judgments based on criteria and standards. ▪ Synonyms include *check, test, critique, judge*.
Create[a]	Put elements together to form a coherent whole or a new structure. ▪ Synonyms include *generate, hypothesize, plan, design, produce*.
[a]This cognitive process and definition is based on the revised Bloom's taxonomy (see Anderson et al., 2001).	

Table 5

Instructional Practices and Definitions

Instructional Practice	Definition
Provided Direct Instruction[a]	Teacher presents issue, discusses or models a solution approach, and engages students with approach in similar context.
Provided Visual Representations[a]	Teacher uses visual representations to organize information, communicate attributes, and explain relationships.
Asked Questions[a]	Teacher asks questions to engage students and focus attention on important information.
Elicited Think Aloud[a]	Teacher prompts students to think aloud about their approach to solving a problem.
Used Independent Practice	Teacher allows students to work independently to develop and refine knowledge and skills.
Provided Guided Feedback[a]	Teacher provides feedback to students on work quality, missing elements, and observed strengths.
Provided Reinforcement[a]	Teacher provides reinforcement contingent on previously established expectations for effort and/or work performance.
Assessed Student Knowledge[a]	Teacher uses quizzes, tests, student products, or other forms of assessment to determine student knowledge.
Other Instructional Practices	Any other instructional practices not captured by the aforementioned key instructional practices.
[a]This instructional practice has received empirical support across multiple studies.	

The second MyiLOGS matrix lists nine instructional practices and three grouping formats. The seven evidence-based instructional practices are supported by research syntheses and meta-analyses (e.g., Brophy & Good, 1986; Gersten et al., 2009; Swanson, 2000; Vaughn et al., 2000; Walberg, 1986). In addition, grouping formats other than whole class also have received empirical support for improving learning outcomes (see Elbaum et al., 2000). "Other instructional practices" represents a "catch-all" category to allow teachers to report on their entire allocated time per class using the available selection of instructional practices and/or "time not available for instruction." Teachers use the latter category to indicate any non-instructional minutes (e.g., transitions, announcements, fire drills), which together with instructional minutes should add up to the total allocated class time (e.g., 90-minute ELA class).

MyiLOGS Training Surveys. At the conclusion of the 4-hour MyiLOGS training session, all participants completed a 9-item survey to provide information on their satisfaction with the training and the software tool. Then 8 months after the completion of the data collection phase, all participants were asked to complete a follow-up survey regarding their assessment of the utility of MyiLOGS and the associated Teacher Reports that summarized their instructional efforts during the previous year. Participants were also asked to completed another OTL scenario like that used during their original MyiLOGS training. The purpose of this scenario re-test was to established some evidence regarding maintenance of the knowledge and skills needed to use MyiLOGS.

Survey of the Enacted Curriculum (SEC). This annual retrospective survey (www.seconline.org) is designed to provide information on the alignment between intended, enacted, and assessed curricula (Martone & Sireci, 2009; Porter, 2002; Roach, Niebling, & Kurz, 2008). The SEC alignment method hereby relies on content translations by teachers (for purposes of the enacted curriculum) and curriculum experts (for purposes of the intended and assessed curriculum) who code a particular curriculum into a content framework that features a comprehensive K-12 list of subject-specific topics. The SEC content frameworks in MA and ELA include 183 and 163 topics, respectively. All content translations occur along a two-dimensional matrix of topics (e.g., multiply fractions) and cognitive demands (e.g., memorize). Teachers report on their enacted curriculum at the end of the school year by describing different instructional emphases for each topic and any applicable cognitive expectations using a 4-point scale. As such, instructional time is not directly assessed via the SEC. To calculate alignment between two content matrices, the data in each matrix are reduced to cell-by-cell proportions with their sum across all rows and columns equaling 1.0. Porter's alignment index (*AI*) takes both dimensions (i.e., topics and cognitive demands) into consideration when calculating the content overlap between two matrices according to his formula: $AI = 1 - [(\Sigma|X - Y|)/2]$. The *AI* has been used as a proxy for OTL at the classroom level in studies (e.g., Kurz et al., 2010). Assuming accurate end of year reporting, the *AI* can provide information about the extent to which a teacher's enacted curriculum matches the content topics and cognitive expectations expressed in the academic content standards of the general curriculum.

State achievement tests. In three states, paper-and-pencil assessments designed to measure student achievement of state standards were used to provide summative data on the extent to which students have achieved the academic standards of the general curriculum for 8th-grade MA and ELA: (a) the *Arizona Instrument to Measure Standards* (AIMS); (b) the *Pennsylvania System of School Assessment* (PSSA), and South Carolina's *Palmetto Assessment of State Standards* (PASS).

Observation Materials. To estimate the extent to which teachers were using MyiLOGS reliably, each teacher participant was observed at least once during his/her logging period. In addition, a subset of three teachers per state was randomly chosen for two additional observations to determine the stability of the reliability estimates. To this end, the authors developed an observation form that mirrored the two two-dimensional MyiLOGS matrices. Trained observers used this form to code the dominant cognitive expectation for student learning and instructional practice observed during 1-minute intervals for an entire class. Inter-observer agreement (IOA) was collected on 30% of all observation sessions across states.

Procedures

State education personnel began the teacher recruitment process at the beginning of the 2010-2011 school year. During the recruitment months of August and early September, the authors trained two senior research professors and one doctoral graduate student in the use of MyiLOGS, its logging conventions, and the respective instructional scenarios for purposes of the professional development training. All three individuals had prior experience with conducting teacher workshops and experience as teachers.

Training to Use MyiLOGS. Each teacher participant received professional development training in the use of the online OTL measure MyiLOGS prior to starting data collection. The MyiLOGS competence-based training was comprised of four phases and required approximately 4 hours to complete. Materials developed to support teachers learning to use MyiLOGS with high integrity were: (a) the "MyiLOGS Teacher's Manual Part 1," which provided teachers with step-by-step instructions for completing the three main MyiLOGS tasks (i.e., daily calendar, sampled class details, sampled student details) as well as detailed answers to frequently asked questions; (b) the "MyiLOGS Teacher's Manual Part 2," which provided teachers with the first five instructional scenarios and their respective answer keys; (c) worksheets on the cognitive process expectations and instructional practices; (d) handouts of the agenda, presentation slides, the introductory worked example, a "cheat sheet" of important MyiLOGS conventions, and the performance assessment scenarios.

The first introductory phase of training featured a 30-minute video supported worked example, which provided a step-by-step demonstration of how to complete the three essential MyiLOGS tasks (i.e., daily calendar, sampled class details, sampled student details). The second phase was a guided practice session lasting about 2 hours. During that time, the lead trainer modeled the steps for completing each task followed by teacher pairs practicing these steps with the support of two additional trainers. To establish the definitions of the cognitive process expectations and instructional practices, teachers completed worksheets that asked them to define each category in their own words and provide examples. Subsequent discussion and modeling was used to resolve any questions and disagreements. The third phase featured a MyiLOGS performance assessment lasting about 1 hour. To ensure teachers had mastered the logging conventions of the technology to accurately represent their instruction, teachers had to pass a sequence of performance tests. These tests featured written instructional scenarios that summarized typical lessons along the calendar, class, and student level. Teachers had to correctly log the instructional scenario via the MyiLOGS software. The answers to two scenarios were modeled and discussed by a trainer. Subsequent instructional scenarios had to be completed independently by each teacher. Once completed, a trainer reviewed the accuracy of the logged scenario. Teachers had to independently pass two scenarios with 100% accuracy to be able to continue in the study. The fourth and final training phase was an independent practice session lasting about 1 hour. During that time, teachers were allowed to use their teaching materials, such as lesson plans and textbooks, to retrospectively log the previous month of instruction at the calendar level. For each of their registered classes, teachers were asked to indicate what academic and custom skills were taught and for how long.

In late September, the authors conducted the Arizona teacher training with 11 general and 5 special education teachers. In Pennsylvania, a trained senior research professor and graduate student conducted several smaller trainings during the month of November. A total of 12 general and special education teachers were trained to criterion. The procedures followed were the same as in Arizona. However, the retrospective logging for Pennsylvania teachers was limited to one month prior to the training thus leaving out the first month of the school year. Similarly, a large training was conducted in South Carolina in November. Out of 16 attendees, 13 general and special education teachers could be trained to criterion in the allotted time. South Carolina teachers' retrospective calendar-based logging was also limited to one month thus missing nearly all of the first two months of the school year.

For purposes of reporting OTL, all teacher participants were asked to log their daily classroom instruction at the calendar level (i.e., instructional time, content coverage) and twice a week in greater detail at the classroom and student level (i.e., instructional time, content coverage, cognitive expectations, instructional practices, grouping formats, engagement, goal attainment). In co-taught classes, the general education teacher was designated as the main reporter for purposes of MyiLOGS; teachers were asked to confer about their shared instruction, especially on any instructional differentiations that may have been implementing by the special education teacher. The use of MyiLOGS for planning purposes was optional. Teachers were asked to log their instruction shortly after having taught the lesson. The monthly compensation was contingent on timely completion of MyiLOGS, which was monitored remotely through the MyiLOGS server by the authors via bi-weekly procedural fidelity checks. To pass a fidelity check, each class had to have two weeks of daily logged skills and times as well as completed class and student details. Teacher participants who passed both fidelity checks during a four-week period were able to withdraw the monthly $100 compensation from a debit card that was loaded remotely by university personnel. The required logging period for all participants was four months with the option to continue through the month of April.

Observer Training and Data Collection. During the months of December and January, university and state department of education personnel were trained to use observation procedures and conducted inter-observer agreement (IOA) sessions in all three states. For training purposes, the MyiLOGS definitions and conventions were reviewed and an observation protocol that mirrored the content of the online MyiLOGS cognitive process and instructional quality matrices. Actual classroom observations were conducted as part of the training. Observers had to obtain an overall agreement percentage of 80% or higher on two consecutive 30-minute sessions to complete the training. After each classroom observation training session, the observers discussed any differences, revising definitions for instruction practices or cognitive processes. Across the three states, a total of six individuals were trained to criterion in approximately 3 hours each.

For observation purposes, all classroom observers (a) prerecorded the skills listed on the MyiLOGS calendar for the given day; (b) started the 1-minute interval with the bell or at the lesson's designated start time; (c) made a tally in both matrices according to the cognitive expectation and instructional practice that occupied the majority of the time during a 1-minute interval (by skill and grouping format); and (d) kept a frequency count of discreet events such as brief praise statements. At the conclusion of the observation, the observer was allowed to make time adjustments to reflect the summative duration of discreet events as well as the MyiLOGS convention of equal emphasis. The latter convention requires teachers to divide instructional minutes equally according to emphasis. For example, a teacher who allowed students to work independently for 10 minutes but concurrently provided students with individual guided feedback throughout the entire time could not log 10 minutes under each practice. Instead, the teacher must divide the instructional minutes accordingly (i.e., 5 minutes per practice). This convention constrains teachers to the allocated class time—the more skills and/or practices that are addressed, the less instructional time can be dedicated to each one. Accordingly, observers were allowed to make tally adjustment immediately following the observation.

For agreement purposes, cell-by-cell agreement was calculated for each matrix based on cell estimates within a 3-minute range or less. That is, two observer estimates of direct instruction at the whole class level of 20 minutes and 23 minutes respectively were counted as an agreement. Likewise, teacher and observer estimates of the Pythagorean Theorem at the Remember level of 4 minutes and 0 minutes respectively were counted as a disagreement. For each matrix, inter-rater agreement was calculated as the total number of agreements divided by the sum of agreements and disagreements. In addition, a combined inter-rater agreement percentage was calculated as the total number of agreements across both matrices divided by the sum of agreements and disagreements across both matrices. That latter index was used in establishing the training criterion (at or above 80%) and retraining criterion (below 80%) for observers.

SEC Administration. The study concluded with the administration of the SEC in early May. All teacher participants were asked to report on their annual enacted curriculum via the SEC for all classes logged via MyiLOGS. To this end, all teachers reviewed a 30-minute training video focused on the SEC coding conventions, a comparison between the cognitive process used in the SEC and MyiLOGS, and the SEC alignment reports. Teacher received $175 contingent on reviewing the training video and completing the SEC for one class. Participants who reported on two classes received an additional $75.

MyiLOGS Scoring Procedures. MyiLOGS, when completed as intended yields seven key calendar day indices (see Table 6). First, the Instructional Time (IT) index was specified according to instructional time spent on state-specific academic standards and instructional time spent on custom objectives. Second, the IT indices were calculated based on average minutes per day and as average percentages of allocated class time. The latter convention was used to allow for comparability between classes that differed in allocated class time. Third, time indices for non-instructional time collected via MyiLOGS were calculated separately. Fourth, the Content Coverage (CC) index was determined as previously described (i.e., percentage of addressed academic standards). Fifth, all indices related to instructional time and content coverage were calculated on the basis of calendar days and sample days with the former representing the largest set of measurement points. The three quality indices related to Cognitive Processes (CP), Instructional Practices (IP), and Grouping Formats (GF) were calculated on the basis of sample days only. Sixth, Engagement and Goal Attainment/Effort were two additional indices based on sample days. Lastly, all calendar-based indices reflected OTL at the class level, whereas indices based on sample days reflected OTL at the class and student levels.

Table 6
MyiLOGS OTL Indices and Operational Definitions

Index	Definition
Instructional Time on Standards (Min/Day)	Average amount of instructional minutes dedicated to the state-specific academic standards per day.
Instructional Time on Standards (%)	Average percentage of allocated class time used for instruction on the state-specific academic standards per day.
Instructional Time on Custom (Min/Day)	Average amount of instructional minutes dedicated to custom objectives per day.
Instructional Time on Custom (%)	Average percentage of allocated class time used for instruction on the custom objectives per day.
Non-Instructional Time (Min/Day)	Average amount of non-instructional minutes per day.
Non-Instructional Time (%)	Average percentage of allocated class time not used for instruction.
Content Coverage (%)	Percentage of state-specific academic standards addressed.
Cognitive Process Score	Sum of differentially weighted percentages of instructional time dedicated to each cognitive process expectation (*Attend* and *Remember* x1; *Understand/Apply*, *Analyze/Evaluate*, and *Create* x2).
Instructional Practice Score	Sum of differentially weighted percentages of instructional time dedicated to each instructional practice (*Used Independent Practice* and *Other Instructional Practices* x1; *Provided Direct Instruction*, *Provided Visual Representation*, *Asked Question*, *Elicited Think Aloud*, *Provided Guided Feedback*, and *Assessed Student Knowledge* x2).
Grouping Format Score	Sum of differentially weighted percentages of instructional time dedicated to each grouping format (*Whole Class* x1; *Individual* and *Small Group* x2)
Engagement	Average score based on "Not engaged (0%)" = 0; "Low % of time (<50%)" = 1; "Moderate % of time (50%-80%)" = 2; "High % of time (>80%)" = 3.
Goal Attainment/Effort	Average score based on No effort or product observed (0%) = 0; Low effort or limited portion of work completed (<50%) = 1; Moderate effort or moderate portion of work completed (50%-80%) = 2; High effort or substantial portion of work completed (>80%) = 3.

With respect to the differential weighting of instructional quality indicators, a weight of 1 was applied to all lower-order thinking skills, generic instructional practices, and whole class instruction for CP, IP, and GF scores, respectively. The weight of 2 was applied to all high-order thinking skills, empirically supported practices, and individual/small group instruction for CP, IP, and GF scores, respectively. As such, all cognitive expectations, instructional practices, and grouping formats received credit; yet those presumed to contribute more to enhance the quality of OTL received a greater weight. The CP, IP, and GF scores thus ranged between 1.00 and 2.00.

A CP, IP, and GF score of 1.00 indicates an exclusive focus on lower-order thinking skills (i.e., attend, remember), generic instructional practices (i.e., independent practice, other instructional practices), and whole class instruction, respectively. A CP, IP, and GF score of 2.00, on the other hand, indicates an exclusive focus on higher-order thinking skills (i.e., understand/apply, analyze/evaluate, create), generic instructional practices (i.e., direct instruction, visual representations, questions, think aloud, guided feedback, reinforcement, assessment), and individual/small group instruction, respectively. The teacher ratings for class and student engagement were based on a 4-point scale: "Not engaged (0%)" = 0; "Low % of time (<50%)" = 1; "Moderate % of time (50%-80%)" = 2; "High % of time (>80%)" = 3. The class and student ratings for goal attainment/effort were also based on a 4-point scale: "No effort or product observed (0%)" = 0; "Low effort or limited portion of work completed (<50%)" = 1; "Moderate effort or moderate portion of work completed (50%-80%)" = 2; "High effort or substantial portion of work completed (>80%)" = 3. Tables 6 lists all My OTL indices and their respective operational definitions.

Results

Teachers Use of MyiLOGS

We collected evidence along three steps of the training and data collection process to ensure that teachers understood how to use the MyiLOGS software correctly to record their daily classroom instruction reliably and with procedural fidelity and to estimate the extent to which teachers' log data represented a valid account of their classroom instruction. Specially, we used several methods to establish the quality of the OTL information collected. These methods were (a) teachers' performances on the final two tests during training, (b) teachers' survey responses following training and 8 months post-study; (c) bi-weekly procedural fidelity data and website user statistics across 30 weeks of instructional logging; and (d) agreement percentages between teachers and trained classroom observers.

Fidelity and reliability evidence. All participating teachers logged two instructional scenarios of the performance assessment with 100% accuracy, thus starting the project using MyiLOGS correctly to collect their instructional efforts accurately. For purposes of reporting daily OTL, all participants were asked to log their classroom instruction at the calendar level (i.e., instructional time, content coverage) and twice a week in greater detail at the classroom and student level (i.e., instructional time, content coverage, cognitive expectations, instructional practices, grouping formats, engagement, goal attainment/effort). Teachers' procedural fidelity (PF) based on completed calendar days and detailed sample days was monitored on a bi-weekly basis. Each check was scored dichotomously as either complete or incomplete. A missing calendar and/or sample day information was identified in a follow-up email along with a prompt to complete the missing information before the next check. Across states, a total of 15 PF checks were completed during 30 weeks of instructional logging. Across all checks, the completion rate ranged between 75% and 100% of classrooms. On average, 92% of classrooms were logged without any missing calendar or sample day information. Upon prompting, all teachers completed their missing data prior to the next PF check. The final instructional data set was 100% complete for all participating teachers.

All teachers were asked to report on their enacted curriculum concurrently with their daily instructional planning and implementation efforts. Although teachers were not required to log their instruction on a daily basis, the training materials recommended two to three logging times per week to minimize the burden on teacher recall. To determine the extent to which teachers followed this recommendation, the website was used to keep track of teachers' average number of log-ins per week (excluding Winter break) as well as their active log-in time per week. On average, participants logged into MyiLOGS 2.4 times per week (*SD* = 0.6) and clocked about 5.9 minutes per week (*SD* = 1.4) of active log-in time.

Finally, each teacher was observed at least once, and a randomly selected subsample of teachers was observed three times, to estimate the extent to which teachers' log data represented a valid account of their classroom instruction. Agreement percentages between teachers and independent observers were calculated on the basis of sample day details at the class level related to five cognitive process expectations per standard/objective and nine instructional practices per three grouping formats. Teachers and observers used the same matrix format to report on sample day details. Due to teacher attrition, South Carolina only featured two teachers with additional observations. Lastly, IOA was collected on 31% of all observation sessions between two trained observers. Table 7 shows the agreement percentages based on cognitive processes, instructional practices, and overall agreement. Across sessions, agreement between teachers and observers for cognitive processes per standard/objective ranged between 67% and 100% with an average of 93%. Across sessions, agreement for instructional practices per grouping format ranged between 89% and 100% with an average of 98%. Overall agreement between teachers and observers across sessions ranged between 85% and 100% with an average of 97%. Across states, agreement for cognitive processes per standard/objective ranged between 27% and 100% with an average of 63%. Across states, agreement for instructional practices per grouping format ranged between 64% and 100% with an average of 82%. Overall agreement across states ranged between 55% and 100% with an average of 77%.

Table 7

Percentage Agreement between Two Independent Observers

IOA Session	**Cognitive Processes**	**Instructional Practices**	**Overall Agreement**
1	100	96	**98**
2	100	96	**97**
3	100	100	**100**
4	100	100	**100**
5	88	100	**95**
6	82	100	**95**
7	100	100	**100**
8	100	100	**100**
9	100	96	**97**
10	100	100	**100**
11	100	100	**100**
12	100	100	**100**
13	91	100	**97**
14	67	100	**94**
15	100	96	**98**
16	67	89	**85**
M (SD)	93 (12)	98 (3)	**97 (4)**

In each state, agreement percentages for cognitive processes per standard/objective were consistently lower than agreement percentages for instructional practices per grouping format. In addition, the agreement percentages for cognitive processes per standard/objective were also more variable than agreement percentages for instructional practices per grouping format.

User satisfaction evidence. Immediately following the 4-hour teacher training, all attending teachers completed a post-training survey using their anonymous identification names. The survey featured nine questions with a 6-point scale: *Strongly Disagree* = 1; *Disagree* = 2; *Somewhat Disagree* = 3; *Somewhat Agree* = 4; *Agree* = 5; *Strongly Agree* = 6.

Table 8

Post-Training and Follow-up Survey Results

Question Number	Question Stem	Post-Training (*n* = 41)		Follow-Up (*n* = 26)	
		M	***(SD)***	***M***	***(SD)***
1	Professional development related to the content standards is important for promoting effective instruction.	5.8	(0.4)	5.6	(0.6)
2	Comprehensive, high-quality coverage of the content standards is an important part of effective instruction.	5.8	(0.4)	5.6	(0.6)
3	The MyiLOGS training was helpful for understanding how to use the system.	5.9	(0.3)	5.4	(0.7)
4	Based on the MyiLOGS training, I was prepared to use the system reliably.	5.5	(0.5)	5.3	(0.8)
5	An online version of this training (e.g., webinar) could have been equally effective.	3.2	(1.5)	3.9	(1.4)
6	I think MyiLOGS can support my comprehensive, high-quality coverage of the content standards.	5.6	(0.6)	5.2	(0.7)
7*	The MyiLOGS training scenarios were helpful for understanding how to use the system.	5.9	(0.4)	--	--
8*	Overall, I think the trainers were well prepared.	5.9	(0.4)	--	--
9*	Overall, I think the training time was sufficient for understanding how to use the system.	5.7	(0.5)	--	--
10**	The charts and tables of the MyiLOGS Report provided meaningful information about my instruction.	--	--	5.3	(0.7)
11**	I would use the MyiLOGS Report feedback during the school year to improve my instruction.	--	--	5.2	(0.8)
12**	I think MyiLOGS Instructional Growth Plan could be helpful as a professional development tool.	--	--	5.2	(0.8)
13**	Using MyiLOGS substantially increase my self-reflection and awareness of how and what I was teaching.	--	--	5.3	(0.8)

Note. Strongly Disagree = 1; Disagree = 2; Somewhat Disagree = 3; Somewhat Agree = 4; Agree = 5; Strongly Agree = 6.
*Post-training only question. **Follow-up only question.

Descriptive data from using MyiLOGS. Teachers in each state reported on time and content indicators of OTL at the class level daily and on two random days per week, teachers also reported on additional quality indicators at the class and student level. With respect to the time dimension of OTL, teachers reported on three time-based indices: (a) instructional time on state-specific standards (*Time on Standards*), (b) instructional time on custom skills/activities (*Time on Custom*), and (c) non-instructional time (*Non-Instructional Time*). These class-specific time indices were calculated based on average minutes per day and as average percentages of allocated class time per day. With respect to the content dimension of OTL, teachers reported on the specific academic standards they covered during the course of the study. The calculated content-based index is the percentage of content standards addressed (*Content Coverage*). With respect to OTL indices related to instructional quality, teachers reported on time emphases along different cognitive processes, instructional practices, and grouping formats. These quality indices were calculated on the basis of summary scores and as total minute allocations and percentages for the different cognitive processes, instructional practices, and grouping formats. In addition, teachers rated class engagement and class goal attainment/effort.

For the 2010-2011 school year, all three states required 180 school days to be used for instruction. Across states, teachers logged between 85 and 178 school days via the calendar, which represented between 47% and 99% of the school year. On average, teachers logged calendar-based OTL indices for about 151 school days, or 84% of the school year. Across states, teachers' allocated class time (i.e., scheduled class length) ranged between 25 and 150 minutes with an average of 65 minutes per class. Table 9 lists all calendar-based OTL indices and the mean number of days on which the indices are based across the entire sample.

Table 9

Calendar-Based Class OTL Indices for Entire Sample

OTL Index	*n*	*M*	*(SD)*
Logged School Days	**46**	**151**	**(18)**
Instructional Time on Standards (Min/Day)	46	44	(23)
Instructional Time on Standards (%)	46	67	(18)
Instructional Time on Custom (Min/Day)	46	18	(11)
Instructional Time on Custom (%)	46	27	(17)
Non-Instructional Time (Min/Day)	46	3	(3)
Non-Instructional Time (%)	46	5	(4)
Number of Standards	**46**	**53**	**(28)**
Content Coverage of Standards (%)	46	68	(22)

Allocated class time was used to calculate all percentage-based indices. On average, teachers spent about 68% of allocated class time on teaching the state-specific standards per day. About 27% of allocated class time was spent on teaching custom skills/activities and an additional 5% was not available for instruction. The total across all percentage-based indices accounted for about 99% of allocated class time. Occasionally, the sum across percentage-based indices did not equal 100% due to time changes at the class or school level because some teachers had the flexibility to shorten or extend their class periods by a few minutes on a given day. In addition, the assignment of "half-days" due to inclement weather conditions or other administrative reasons effectively shortened all applicable class periods.

Table 10

Calendar-Based Class OTL Indices By Subject Area

OTL Index	MA			ELA		
	n	***M***	***SD***	***n***	***M***	***SD***
Arizona						
Logged School Days	**9**	**163**	**(5)**	**7**	**165**	**(7)**
Instructional Time on Standards (Min/Day)	9	52	(22)	7	71	(29)
Instructional Time on Standards (%)	9	73	(11)	7	72	(14)
Instructional Time on Custom (Min/Day)	9	17	(9)	7	24	(9)
Instructional Time on Custom (%)	9	25	(16)	7	27	(11)
Non-Instructional Time (Min/Day)	9	5	(4)	7	4	(4)
Non-Instructional Time (%)	9	6	(5)	7	3	(3)
Number of Standards	**9**	**61**	**(0)**	**7**	**115**	**(0)**
Content Coverage of Standards (%)	9	67	(11)	7	54	(16)
Pennsylvania						
Logged School Days	**5**	**142**	**(7)**	**8**	**128**	**(23)**
Instructional Time on Standards (Min/Day)	5	44	(11)	8	33	(14)
Instructional Time on Standards (%)	5	79	(13)	8	64	(14)
Instructional Time on Custom (Min/Day)	5	12	(12)	8	11	(5)
Instructional Time on Custom (%)	5	18	(15)	8	24	(14)
Non-Instructional Time (Min/Day)	5	2	(1)	8	4	(2)
Non-Instructional Time (%)	5	2	(2)	8	8	(6)
Number of Standards	**5**	**41**	**(0)**	**8**	**32**	**(0)**
Content Coverage of Standards (%)	5	69	(24)	8	87	(13)
South Carolina						
Logged School Days	**6**	**156**	**(11)**	**11**	**149**	**(13)**
Instructional Time on Standards (Min/Day)	6	30	(10)	11	37	(18)
Instructional Time on Standards (%)	6	55	(18)	11	63	(25)
Instructional Time on Custom (Min/Day)	6	23	(12)	11	18	(14)
Instructional Time on Custom (%)	6	38	(17)	11	31	(23)
Non-Instructional Time (Min/Day)	6	1	(1)	11	3	(3)
Non-Instructional Time (%)	6	2	(1)	11	4	(4)
Number of Standards	**6**	**33**	**(0)**	**11**	**40**	**(0)**
Content Coverage of Standards (%)	6	63	(29)	11	66	(25)
Across States						
Logged School Days	**20**	**156**	**(12)**	**26**	**147**	**(21)**
Instructional Time on Standards (Min/Day)	20	43	(19)	26	45	(25)
Instructional Time on Standards (%)	20	69	(16)	26	66	(19)
Instructional Time on Custom (Min/Day)	20	17	(11)	26	18	(11)
Instructional Time on Custom (%)	20	27	(17)	26	28	(17)
Non-Instructional Time (Min/Day)	20	3	(3)	26	3	(3)
Non-Instructional Time (%)	20	4	(4)	26	5	(5)
Number of Standards	**20**	**48**	**(13)**	**26**	**58**	**(36)**
Content Coverage of Standards (%)	20	66	(20)	26	69	(23)

Note. MA = Mathematics; ELA = English/Language Arts.

The general curriculum featured an average of 53 academic content standards, of which teachers' were able to address approximately 36 (about 68%) during their respective login period. Lastly, a review of the custom skills/activities indicated that the sample of students with disabilities did not receive additional IEP objectives beyond the academic standards of the general curriculum. Upon review, only one objective out of 554 was specifically identified as an IEP objective (logged by a special education teacher) related to fluency and comprehension. Table 10 shows the calendar-based OTL indices broken down for each state by subject area.

Table 11

Sample-Day Based Class OTL Quality Indices By Subject Area

	MA			ELA		
	n	***M***	***SD***	***n***	***M***	***SD***
Arizona						
Logged Sample Days	**9**	**51**	**8**	**7**	**50**	**4**
Cognitive Process Score	9	1.69	0.16	7	1.82	0.10
Instructional Practice Score	9	1.67	0.08	7	1.57	0.11
Grouping Format Score	9	1.27	0.18	7	1.12	0.07
Engagement	9	2.60	0.30	7	2.63	0.27
Goal Attainment/Effort	9	2.59	0.29	7	2.60	0.29
Pennsylvania						
Logged Sample Days	**5**	**40**	**5**	**8**	**37**	**5**
Cognitive Process Score	5	1.71	0.17	8	1.79	0.13
Instructional Practice Score	5	1.70	0.09	8	1.69	0.18
Grouping Format Score	5	1.33	0.16	8	1.14	0.12
Engagement	5	2.42	0.22	8	2.71	0.19
Goal Attainment/Effort	5	2.36	0.28	8	2.69	0.21
South Carolina						
Logged Sample Days	**6**	**41**	**6**	**11**	**39**	**13**
Cognitive Process Score	6	1.67	0.13	11	1.74	0.11
Instructional Practice Score	6	1.68	0.18	11	1.49	0.25
Grouping Format Score	6	1.24	0.20	11	1.36	0.34
Engagement	6	2.52	0.32	11	2.43	0.40
Goal Attainment/Effort	6	2.50	0.31	11	2.43	0.40
Across States						
Logged Sample Days	**20**	**45**	**8**	**26**	**41**	**10**
Cognitive Process Score	20	1.69	0.14	26	1.78	0.11
Instructional Practice Score	20	1.68	0.12	26	1.57	0.21
Grouping Format Score	20	1.28	0.18	26	1.23	0.26
Engagement	20	2.53	0.28	26	2.57	0.33
Goal Attainment/Effort	20	2.50	0.29	26	2.56	0.33

Note. MA = Mathematics; ELA = English/Language Arts.

On sample days, teachers completed additional information beyond the calendar related to cognitive processes, instructional practices, grouping formats, class engagement, and goal attainment/effort. On average, teachers logged quality-related OTL indices for about 43 school days, or 24% of the school year. To complement the aforementioned calendar-based indices for time and content, Table 11 presents three quality-related summary indices—Cognitive Process Score, Instructional Practice Score, Grouping Format Score—each with a score range between 1.00 to 2.00 for each subject area by state. In addition, Table 11 provides a rating score for perceived class engagement and goal attainment/effort with a score range between 0 and 3. Across states, the summary indices indicate a greater emphasis of high-order thinking skills in ELA than in MA and a greater emphasis of evidence-based practices and *Individual* and *Small Group* grouping formats in MA than in ELA.

Convergent and Predictive Validity Evidence for the MyiLOGS OTL Indices

To provide initial validity evidence for the OTL measurement tool, we examined convergent validity values between the MyiLOGS OTL indices at the class level and the SEC AI index at the class level. In addition, we compared the predictive validity of both measures using their class-based indices to predict average class achievement on the state achievement for the state of Arizona—the only state that provided class-specific achievement data for students in participating classrooms.

Convergent Validity. It was hypothesized that the SEC AI index, which quantifies alignment based on a match of topic and cognitive demand between teacher instruction and state standards, should correlate differentially with the various OTL indices from MyiLOGS. Given that SEC does not account for instructional time, the correlations between the content- and quality-based OTL indices and the AI were hypothesized to exceed the correlations between the time-based OTL indices and the AI. In addition, the correlations between both measures were hypothesized to range between .15 and .30. The AI averages ranged between .14 and .16 with an average of .16 across states.

The difference in alignment between classroom instruction and state content standards for general and special education classrooms was statistically not significant ($p > .05$) with a medium effect size of $d = .44$. On average, alignment was lower in special education classrooms (.15) than in general education classrooms (.17).

None of the correlations between class-based OTL indices from MyiLOGS and the SEC AI were statistically significant ($p > .05$). We also examined scatterplots for the SEC AI and three OTL indices: *Time on Standards (Min/Day)*, *Content Coverage (%)*, and the *Cognitive Process Score*. None of the scatterplots displayed a clear relation between the SEC AI and the respective OTL indices. Based on these results, the two measurement tools provide indices that do not appear to be related.

Predictive Validity. For the Arizona subsample, state personnel provided class averages of the 2010-2011 AIMS state test for each class logged by a participating teacher. The subsample featured a total of 16 classes, which consisted exclusively of general education classrooms (three of which featured a general and special education co-teaching pair). The correlations between the SEC AI and time, content, and quality-related OTL indices are in Table 12. First, the bivariate correlation between the SEC AI and class achievement was statistically significant with $r = -.53$. As such, the AI accounted for about 28% of the variance in average class achievement. For this sample, the negative relation indicates that a higher AI corresponded with a lower average class achievement. With respect to the MyiLOGS measurement tool, one time-based and two quality-related OTL indices were related to average class achievement. Second, the bivariate correlation between the *Time on Standards* and class achievement was statistically significant with $r = .56$. As such, the *Time on Standards* index accounted for about 31% of the variance in average class achievement. For this sample, the positive relation indicates that more instructional time dedicated to the state-specific standards was associated with higher average class achievement. Third, the bivariate correlation between the *Cognitive Process Score* and class achievement was statistically significant with $r = .64$. As such, the AI accounted for about 41% of the variance in average class achievement. For this sample, the positive relation indicates that a greater emphasis of higher-order cognitive processes was associated with higher average class achievement. Fourth, the bivariate correlation between the *Grouping Format Score* and class achievement was statistically significant with $r = -.71$. As such, the AI accounted for about 50% of the variance in average class achievement. For this sample, the negative relation indicates that a greater emphasis of individual and small group formats was associated with lower average class achievement. Based on the current result, the expected convergent validity between the SEC AI and MyiLOGS OTL indices could not be corroborated.

Table 12

Correlations between SEC and MyiLOGS OTL Indices and Class Achievement Averages

Index	**2010-2011 Average Class Achievement**
SEC Alignment Index	$-.53^{*}$
Instructional Time on Standards (Min/Day)	$.56^{*}$
Instructional Time on Standards (%)	.06
Instructional Time on Custom (Min/Day)	.49
Non-Instructional Time (Min/Day)	-.04
Non-Instructional Time (%)	-.32
Content Coverage of Standards (%)	-.30
Cognitive Process Score	$.64^{**}$
Instructional Practice Score	-.34
Grouping Format Score	$-.71^{**}$
Note. N = 16. $^{*}p < .05$; $^{**}p < .01$.	

Relations between Student-Based OTL Indices and Student Achievement

The Arizona subsample was used to examine the extent to which student-based OTL indices were predictive of student achievement on the end-of year state test. Given that previous research has supported the relation between time, content, and quality-related OTL indices and student achievement, the following OTL indices were entered into the model: (a) *Time on Standards (Min/Day)*; (b) *Time on Custom (Min/Day)*; (c) *Non-Instructional Time (Min/Day)*, (d) *Content Coverage (%)*; (e) *Cognitive Process Score*; (f) *Instructional Practice Score*; and (g) *Grouping Format Score*. The time, content, and quality-related OTL indices were each entered as a set. Any non-significant predictors were removed prior to the next step. The order for the respective steps was based on prior research. The only student-based time index that showed a statistically significant relation with student achievement was *Time on Custom* (i.e., average amounts of minutes dedicated to custom skills/activities per day) with $R^2 = .24$.

Table 13

Hierarchical Regression Analysis Summary for Student-Based OTL Indices Predicting Student Achievement Controlling for Prior Achievement

Variable	*B*	*SEB*	*β*	R^2	ΔR^2
Step 1				**.62**	**.62**
Prior Achievement	0.76	0.11	0.79*		
Step 2				**.64**	**.02**
Prior Achievement	0.70	0.13	0.73*		
Time on Standards (Min/Day)	0.00	0.37	0.00		
Time on Custom (Min/Day)	0.46	0.51	0.13		
Non-Instructional Time (Min/Day)	0.20	0.40	0.06		
Step 3				**.63**	**-.01**
Prior Achievement	0.79	0.11	0.83*		
Content Coverage (%)	0.54	0.50	0.13		
Step 4				**.63**	**.00**
Prior Achievement	0.78	0.14	0.81*		
Cognitive Process Score	9.17	42.17	0.03		
Instructional Practice Score	36.75	55.37	0.09		
Grouping Format Score	2.26	37.30	0.01		
Final Model				**.62**	
Prior Achievement	0.76	0.11	0.79*		
Note. $p < .05$.					

The results for the same student-based OTL indices predicting student achievement while controlling for prior achievement for all three steps including the final model is presented in Table 13. The results indicated that none of the student-based OTL indices exhibited a statistically significant relation with student achievement controlling for students' prior achievement, which accounted for $R^2 = .62$.

Discussion

This study investigated initial validity evidence for a new tool for measuring the extent to which students have the opportunity to learn the intended curriculum as measured by instructional indicators of the enacted curriculum. The development of this teacher self-report measure advanced traditional teacher logging approaches by embedding logs into teachers' ongoing daily instructional practice. In combination with a sampling approach related to gathering additional details on aspects of instructional quality, the newly developed daily teacher log OTL measure permitted the establishment of a heretofore unavailable record of continuous teacher self-report data across the school year on OTL indices at both the class and student level.

We provided evidence that teachers could be trained to use MyiLOGS with a high degree of procedural fidelity and good reliability and provided descriptive statistics on the time, content, and quality instruction indices for the teachers across five or more months of classroom instruction. With respect to time, teachers reported on three time-based indices: (a) instructional time on state-specific standards (*Time on Standards*), (b) instructional time on custom skills/activities (*Time on Custom*), and (c) non-instructional time (*Non-Instructional Time*). With respect to content, teachers reported on the specific academic standards they covered during the course of the study. The calculated content-based index is the percentage of content standards addressed (*Content Coverage*). With respect to instructional quality, teachers reported on time emphases along different cognitive processes, instructional practices, and grouping formats. In addition, teachers rated class engagement and class goal attainment/effort. All OTL indices used to address our first validity evidence question were based on the class level. This represents the traditional view of OTL, which treats the teacher's instructional provision of the enacted curriculum as universal and undifferentiated.

Based on website user statistics, teachers applied the concurrent logging approach as instructed, logging their daily classroom instruction, on average, 2.4 times a week covering, on average, about 151 school days, or 84% of the school year. With respect to basic time and content frameworks, teachers within and between states demonstrated a great deal of variation both in terms of allocated class time and the number of academic standards for each subject area. Across states and subject areas, the allocated class time ranged between 25 and 150 minutes and the number of academic content standards covered ranged between 32 and 115. Within these basic frameworks of allocated class time and number of content standards, teachers further varied in the extent to which they dedicated instructional time to the content standards and different custom skills, as well as the extent to which allocated time was non-instructional (e.g., transitions, announcements). Irrespective of the large standard deviations, the average percentage-based indices across states were similar for MA and ELA with 69% and 66% for *Time on Standards*, 27% and 28% for *Time on Custom*, 4% and 5% for *Non-Instructional Time*, as well as 66% and 69% for *Content Coverage*, respectively.

With respect to OTL indices for instructional quality, teachers completed additional information on cognitive processes, instructional practices, grouping formats, class engagement, and goal attainment/effort. Based on summary data across states, subject-specific differences in OTL indices were noted along the *Cognitive Process*, *Instructional Practice*, and *Grouping Format* scores. These summary indices indicated a greater emphasis of high-order thinking skills in ELA than in MA, a greater emphasis of evidence-based practices in MA than in ELA, and a greater emphasis of alternative grouping formats in MA than in ELA. None of these general trends, however, represented statistically significant differences based on this sample.

Subsequent descriptions of total time allocations across the different cognitive process, instructional practices, and grouping formats indicated that across states, the most emphasized cognitive processes were *Understand/Apply*. The *Remember* process was more prevalent in MA than in ELA, and the *Create* process more prevalent in ELA than in MA. Both findings appear reasonable given the large number of memorizable MA facts and the ability for ELA teachers to utilize the *Create* process during composition tasks. With respect to instructional practices, *Independent Practice* represented the most commonly emphasized practice among available choice across both subject areas. Moreover, *Direct Instruction* and *Assessed Student Knowledge* followed *Independent Practice* as the second and third order of emphasis across subject areas. Lastly, *Whole Class* was the most commonly emphasized grouping format across subject areas. Conversely, *Small Group* represented the least commonly emphasized grouping format across subjects.

To provide initial validity evidence for the OTL measurement tool, we examined convergent validity values between the MyiLOGS OTL indices at the class level and the SEC AI index at the class level. The SEC AI has been previously identified as an OTL proxy (e.g., Kurz et al., 2010; Porter, 2002) although it is first and foremost an alignment measure. With it, low alignment can be function of misalignment among topics covered, cognitive demands emphasized, or both. The results of this study indicated that MyiLOGS and the SEC did not measure the same construct, thus reinforcing that the SEC is primarily a measure of alignment, while MyiLOGS appears to measure aspects of instruction that are different from alignment.

In addition, we compared the predictive validity of both MyiLOGS and the SEC using their class-based indices to predict average class achievement on the state achievement for the state of Arizona—the only state that provided class-specific achievement data for students in participating classrooms. Despite low power, the results indicated several statistically significant correlations with medium effect sizes. Three class-based MyiLOGS indices showed statistically significant relations with class achievement: *Time on Standards*, the *Cognitive Process Score*, and the *Grouping Format Score*. First, the average amount of minutes per day dedicated to the state-specific standards had a positive relation with class achievement with a medium effect size. Second, a greater emphasis on high-order thinking skills correlated positively with class achievement also with a medium effect size. Third, a greater emphasis on small group and individual grouping formats correlated negatively with class achievement with a medium negative effect size. The latter finding is also surprising given prior research indicating a positive relation between achievement and grouping formats other than whole class (e.g., Elbaum et al., 2000). In addition, this finding cannot be attributed to class type—the prevalence of alternative grouping formats in special education classroom, which may further coincide with lower academic achievement—because the Arizona subsample was entirely comprised of general education classrooms. For the SEC, the AI was negatively correlated with class achievement. This finding is surprising given prior research findings, which have supported a positive relation between the AI and student achievement (e.g., Kurz et al., 2010; Smithson & Collares, 2007). An important difference between this subsample and samples in other predictive studies such as the ones in Kurz et al. (2010) is the sample's sensitization to their daily instructional practices. That is, teachers in this study reviewed their daily instruction several times a week for up to eight months prior to taking the SEC's annual survey.

Further examination of the relation between student-based OTL indices and individual student achievement for the Arizona subsample revealed that without controlling for prior achievement, instructional time on custom skill/activities (*Time on Custom*) was the only student-based OTL index that exhibited a positive relation with student achievement accounting for about 24% of the variance. This finding is surprising in the context of a non-significant finding for *Time on Standards*. That is, one would expect that more instructional time on the state-specific standards be related to higher achievement based on an assessment that covers those standards—rather than an index related to instructional time on objectives/activities outside the standards. However, as noted previously, many teachers logged review activities and technology-based elements of their lesson under *Time on Custom*. As such, it is very likely that *Time on Custom* reflected *additional* time on standards-based instruction rather than instructional time unrelated to the general curriculum standards.

Limitations

In general, the study's results were based on a relatively small volunteer sample across states, subject areas, and class types. As such, these initial OTL results lack generalizability. In addition, the missing achievement data from the states of Pennsylvania and South Carolina significantly limited the predictive findings related to student achievement.

The study findings are also subject to limitations due to several unconfirmed assumptions and methodological challenges. With respect to assumptions, the following ones remain unconfirmed: (a) the state tests used for determining the relation between OTL and achievement were aligned with the state-specific standards and exhibited instructional sensitivity; and (b) the intended curriculum for students with disabilities was congruent with the general curriculum standards applicable to students without disabilities. A violation of the first assumption related to alignment could have led to underestimation of the relation between the various OTL indices and student achievement. Given that most OTL indices in this study were based on the state-specific general curriculum standards, a strong relation between these indices and achievement cannot be expected, if the respective state tests are not well aligned with the standards used to determine OTL. In addition, we have no evidence of instructional sensitivity for the respective state tests. That is, the extent to which the state assessments were sensitive to differences in instruction remains unclear. Low instructional sensitivity could result in test scores that cannot fully reflect differences in OTL. Consequently, the presumed relation between OTL and achievement could be underestimated.

A violation of the second assumption could limit the extent to which the conclusions are related to students' opportunity to learn the *intended curriculum*. As noted at the outset of this article, the intended curriculum for students with disabilities is dually determined by both the general curriculum and additional IEP objectives. The current conclusion based on students' opportunity to learn the intended curriculum assumes that teachers accurately logged all applicable IEP objectives. Based on the current results, it appears that students' intended curriculum overlapped entirely with the general curriculum standards. Given that students in the participating states were expected to have standards-based IEPs this assumption is logical, but was never directly confirmed through an actual review of the target students' IEPs. The findings therefore may under-represent students' intended curricula. In other words, the current findings may be a more accurate description of students' opportunity to learn the general curriculum.

A final limitation stems for two methodological challenges related to the observation system. Given the possibility that a teacher can address all cognitive processes and instructional practices in one lesson, the observation protocol allowed any categories that were neither reported by the teacher nor observed by the observer to be counted as an agreement. This convention may have contributed to inflated agreement percentages in certain cases. A second methodological challenge of the observation system was the varying cell sizes by which agreement percentages were calculated. Depending on the number of standards/objectives per lesson, the possible number of agreements/disagreements varied from teacher to teacher. This prevented the application of alternative agreement statistics such a Kappa, which could have accounted for chance agreement.

Implications for Practice and Future Research

A major implication for both practice and research lies in the development of the applied OTL measurement tool, MyiLOGS, which was used successfully to collect data on a range of instructional indices related to time, content, and quality. Specifically, we provided evidence to support the feasibility, usability, and promise of MyiLOGS and its training and follow-up procedures for measuring OTL at class and student level. As such, large-scale research on OTL including normative studies as well as subgroup-specific investigations can be launched.

A second implication for practice lies in the remediation of potential OTL gaps through the development of teacher-level interventions. The findings of this study have demonstrated feasibility, usability, and promise of using an online technology such as MyiLOGS for purposes of concurrent teacher logging of OTL indices at the class and student level. Therefore, the collected data 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. Future research on the formative aspects of OTL, especially in conjunction with student outcomes data, appears to be particularly salient, because it would allow teachers to use data on instructional inputs, processes, and outcomes for informing instruction.

Conclusions

The majority of findings of this study are unique, because no investigator has previously reported a study where OTL data were continuously collected and analyzed along all three instructional dimensions—time, content, and quality—at the class and student level for a large portion of the school year. The evidence collected with MyiLOGS by teachers substantiated that: (a) teachers can be trained to criterion within 4-hour to report reliably on various OTL indices based on instructional scenarios at the class and student level; (b) teachers can be supported to maintain high procedural fidelity logging various OTL indices at the class and student level across the duration of a school year; and (c) teachers' concurrent log data provided a valid account of their classroom instruction based on agreement percentages between teachers and independent observers. The results of the classroom observations indicated that two independent observers were able to achieve high agreements across both observation categories and teachers and observers generally had lower agreements for cognitive processes than instructional practices. The current findings do support the conclusion that the teacher self-report data from MyiLOGS provides a rich picture and reliable account of opportunities to learn in middle school classrooms across several states.

Footnote

[1]The Modified Alternate Assessment Participation Screening (MAAPS) project addresses federal regulations, which note that participation in alternate assessments based on modified achievement standards (AA-MAS) is, in part, dependent on a student's failure to reach grade-level proficiency despite access to "appropriate instruction" (U.S. Department of Education, 2007; Award # S368A090006). In the context of MAAPS, the concept of OTL is used to circumscribe appropriate instruction and its measurement is intended to support IEP teams in a data-driven placement decision.

References

Abedi, J., Courtney, Leon, & Azzam (2006).

Aguirre-Munoz (2006)

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Anderson, L. W. (1986). Opportunity to learn. In T. Husén & T. Postlethwaite (Eds.), *International encyclopedia of education: Research and studies*. Oxford, UK: Pergamon.

Anderson, L. W. (2002). Curricular alignment: A re-examination. *Theory into Practice*, *41*(4), 255-260.

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: Macmillian.

Burns, M. K., & Ysseldyke, J. E. (2009). Reported prevalence of evidence-based instructional practices in special education. *Journal of Special Education*, *43*(1), 3-11.

Camburn, E., & Barnes, C. A. (2004). Assessing the validity of a language arts instruction log through triangulation. *Elementary School Journal*, *105*(1), 49-73.

Connor, C. M., Morrison, F. J., Fishman, B. J., Ponitz, C. C., Glasney, S., Underwood, P. S., Piasta, S. B., et al. (2009). The ISI classroom observation system: Examining the literacy instruction provided to individual students. *Educational Researcher*, *38*(2), 85-99.

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.

D'Agostino, J. V., Welsh, M. E., & Corson, N. M. (2007). Instructional sensitivity of a state's standards-based assessment. *Educational Assessment*, *12*(1), 1-22.

Elbaum, B., Vaughn, S., Hughes, M., T. Moody, S. W., & Schumm, J. S. (2000). How reading outcomes for students with learning disabilities are related to instructional grouping formats: A meta-analytic review. In R. Gersten, E. P. Schiller, & S. Vaughn (Eds.), *Contemporary special education research: Syntheses of the knowledge base on critical instructional issues* (pp. 105-135). Mahwah, NJ: Lawrence Erlbaum.

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

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.

Herman, J. L., & Abedi, J. (2004). *Issues in assessing English language learners’ opportunity to learn mathematics* (CSE Report No. 633). Los Angeles, CA: Center for the Study of Evaluation, National Center for Research on Evaluation, Standards, and Student Testing.

Herman, J. L., Klein, D. C., & Abedi, J. (2000). Assessing students’ opportunity to learn: Teacher and student perspectives. *Educational Measurement: Issues and Practice*, *19*(4), 16-24.

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

Individuals with Disabilities Education Improvement Act of 2004, (amending 20 U.S.C. §§ 1400 *et seq.*)

Jackson, P. W. (1990). *Life in Classrooms*. Teachers College Press.

Karger, J. (2005). *Access to the general education curriculum for students with disabilities: A discussion of the interrelationship between IDEA and NCLB*. Wakefield, MA: National Center on Accessing the General Curriculum.

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.), *The handbook of accessible achievement tests for all students: Bridging the gaps between research, practice, and policy*. New York: Springer.

Kurz, A., & Elliott, S. N. (2011). Overcoming barriers to access for students with disabilities: Testing accommodations and beyond. In M. Russell (Ed.), *Assessing students in the margins: Challenges, strategies, and techniques* (pp. 31-58). Charlotte, NC: Information Age Publishing.

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.

Martone, A., & Sireci, S. G. (2009). Evaluating alignment between curriculum, assessment, and instruction. *Review of Educational Research*, *79*(4), 1332-1361.

McDonnell, L. M. (1995). Opportunity to learn as a research concept and a policy instrument. *Educational Evaluation and Policy Analysis*, *17*(3), 305-322.

No Child Left Behind Act of 2001, 20 U.S.C. §§ 6301 *et seq*.

O'Sullivan, P. J., Ysseldyke, J. E., Christenson, S. L., & Thurlow, M. L. (1990). Mildly handicapped elementary students' opportunity to learn during reading instruction in mainstream and special education settings. *Reading Research Quarterly*, *25*(2), 131-146.

Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher*, *38*(2), 109-119.

Porter, A. C. (1991). Creating a system of school process indicators. *Educational Evaluation and Policy Analysis*, *13*(1), 13-29.

Porter, A. C. (1995). The uses and misuses of opportunity-to-learn standards. *Educational Researcher*, *24*(1), 21-27.

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

Porter, A. C., & Smithson, J. L. (2001). Are content standards being implemented in the classroom? A methodology and some tentative answers. In S. Fuhrman (Ed.), *From the Capitol to the classroom: Standards-based reform in the states*. *One Hundredth Yearbook of the National Society for the Study of Education* (pp. 60-80). Chicago: University of Chicago Press.

Porter, A. C., Kirst, M. W., Osthoff, E. J., Smithson, J. L., & Schneider, S. A. (1993). *Reform up close: An analysis of high school mathematics and science classrooms* (Final Report). Madison, WI: University of Wisconsin, Wisconsin Center for Education Research.

Roach, A. T., Niebling, B. C., & Kurz, A. (2008). Evaluating the alignment among curriculum, instruction, and assessments: Implications and applications for research and practice. *Psychology in the Schools*, *45*(2), 158-176.

Rogosa, D., Floden, R., & Willett, J. B. (1984). Assessing the stability of teacher behavior. *Journal of Educational Psychology*, *76*(6), 1000-1027.

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.

Smithson, J. L., & Collares, A. C. (2007, April). *Alignment as a predictor of student achievement gains*. Paper presented at the meeting of the American Educational Research Association, Chicago, IL.

Smithson, J. L., & Porter, A. C. (1994). *Measuring classroom practice: Lessons learned from efforts to describe the enacted curriculum* (CPRE Research Report No. 31). New Brunswick, NJ: Consortium for Policy Research in Education.

Stevens, F. I., & Grymes, J. (1993). *Opportunity to learn: Issues of equity for poor and minority students* (NCES No. 93-232). Washington, DC: National Center for Education Statistics.

Swanson, H. L. (2000). What instruction works for students with learning disabilities? Summarizing the results from a meta-analysis of intervention studies. In R. Gersten, E. P. Schiller, & S. Vaughn (Eds.), *Contemporary special education research: Syntheses of the knowledge base on critical instructional issues* (pp. 1-30). Mahwah, NJ: Lawrence Erlbaum.

Vannest, K. J., & Hagan-Burke, S. (2010). Teacher time use in special education. *Remedial and Special Education*, *31*(2), 126-142.

Vaughn, S., Gersten, R., & Chard, D. J. (2000). The underlying message in LD intervention research: Findings from research syntheses. *Exceptional Children*, *67*(1), 99-114.

Vaughn, S., Levy, S., Coleman, M., & Bos, C. S. (2002). Reading instruction for students with LD and EBD: A synthesis of observation studies. *Journal of Special Education*, *36*(1), 2-13.

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.

Walberg, H. J. (1988). Synthesis of research on time and learning. *Educational Leadership*, *45*(6), 76-85.

Wang, J. (1998). Opportunity to learn: The impacts and policy implications. *Educational Evaluation and Policy Analysis*, *20*(3), 137-156.

Wehby, J. H., Symons, F. J., & Canale, J. A. (1998). Teaching practices in classrooms for students with emotional and behavioral disorders: Discrepancies between recommendations and observations. *Behavioral Disorders*, *24*(1), 51-56.

Wehmeyer, M. L., Sands, D. J., Knowlton, H. E., & Kozleski, E. B. (2002). *Teaching students with mental retardation: Providing access to the general curriculum*. Baltimore: Brookes.