Confronting the Known Unknown: How the Concept of Opportunity to Learn Can Advance Tier 1 Instruction

Alexander Kurz

Even when students are taught in the same classroom by the same teacher, they are often offered very different opportunities to learn (Kurz, Elliott, Lemons et al., 2014; Rowan & Correnti, 2009; Schmidt, Burroughs, Zoido, & Houang, 2015). These differences in opportunity to learn (OTL) further correlate with differences in student achievement (Kurz, 2011; Schmidt & Burroughs, 2013). According to the *Standards for Educational and Psychological Testing*, OTL must be considered to ensure fair testing practices and prevent "misdiagnosis, inappropriate placement, and/or inappropriate assignment of services" (AERA, NCME, & APA, 2014, p. 57). These are important caveats whenever test score inferences are being made including in the context of response to intervention (RTI), which is predicated on the prevention of academic difficulties through the accurate identification of at-risk students and the timely delivery of effective interventions (Al Otaiba et al., 2014; Compton et al., 2012). Universal screening based on curriculum-based measurement is typically used to determine whether students are responding to high-quality Tier 1 instruction (Fuchs & Vaughn, 2012). This common screening approach, however, has shown unacceptably high rates of false positives (Fuchs, Fuchs, & Compton, 2012).

Differences in the provision and accessibility of said "high-quality" instruction represent one factor that could undermine classification accuracy.

Despite the correlation between OTL and student achievement, researchers and practitioners have largely ignored questions about the extent to which Tier 1 screening outcomes could be a function of "low-quality" instruction (Kurz, Elliott, & Roach, 2015; Reddy, Fabiano, & Jimerson, 2013). That is, positive screening results typically suggest two competing inferences: (a) a student's *inadequate response* to high-quality Tier 1 instruction or (b) a teacher's *inadequate implementation* of high-quality Tier 1 instruction. Ruling out the latter inference requires data about the implementation of Tier 1 instruction—a need often acknowledged but rarely addressed in research and practice (Gilbert et al., 2013; Kurz et al., 2015). Knowing that we don't know, of course, makes Tier 1 instructional data the proverbial "known unknown."

The rationale for confronting this known unknown is threefold. First, missing data on Tier 1 instruction represents a threat to fairness and thus potentially impacts the validity of test score inferences (AERA, NCME, & APA, 2014). In addition, reducing the number of false positives (i.e., students receiving Tier 2 who do not really need it) matters greatly in a resource-constrained environment such as public schools (Al Otaiba et al., 2014). Second, ensuring access to high-quality Tier 1 instruction is fundamental to the logic and

A. Kurz (✉)
Arizona State University, Tempe, AZ, USA
e-mail: alexander.kurz@asu.edu

efficacy of all RTI approaches (Fuchs & Vaughn, 2012). In fact, research findings suggest that students' unfettered access to primary prevention leads to fewer students requiring intervention initially and over time (Fuchs, Fuchs, Mathes, & Simmons, 1997; Vaughn et al., 2009), as well as a reduction in special education referrals and placements, with more proportionate representation of minorities, English language learners, and males (Torgesen, 2009; VanDerHeyden, Witt, & Gilbertson, 2007; Wanzek & Vaughn, 2010). Third, documenting the provision and accessibility of high-quality Tier 1 instruction can help educators make data-based decisions to improve "two important determinants of students success: opportunity to learn and quality of instruction" (Fuchs & Vaugh, 2012).

To ensure high-quality Tier 1 instruction as a form of primary prevention, researchers have focused on understanding and operationalizing evidence standards that can be used to identify so-called evidence-based practices (Cook & Odom, 2013; Gandhi, Holdheide, Zumeta, & Danielson, 2016). These laudable efforts, however, have emphasized secondary and tertiary levels of prevention. This omission is likely due to the many challenges of operationalizing and assessing generally effective instruction. Some researchers, however, have suggested that research about OTL holds important implications for addressing these challenges and supporting high-quality Tier 1 instruction (Holdheide, 2016; Kurz et al., 2015).

In this chapter, I argue that the establishment of Tier 1 quality standards and the conceptual expansion of OTL are interconnected, mutually beneficial activities important for truly accessible instruction. To this end, I focus on my conceptual synthesis of OTL (Kurz, 2011) due to its (a) operational definition (Kurz, Elliott, Lemons et al., 2014), (b) application in general and special education (Heafner & Fitchett, 2015; Roach, Kurz, & Elliott, 2015), and (c) measurement via an online teacher log (Kurz, Elliott, Kettler, & Yel, 2014). I begin by reviewing what is known about high-quality Tier 1 instruction and OTL, continue by synthesizing both literature bases to identify

potential sources of evidence for high-quality Tier 1, elaborate by highlighting possible measurement options via a case example, and conclude by setting a research and development agenda for Tier 1 OTL.

Tier 1 Instruction

As a multitiered approach to the early identification and support of students with learning and behavior needs, RTI's primary level of prevention (Tier 1) is focused on high-quality instruction and universal screening of all children. Across tiers or prevention levels, struggling learners are provided with interventions at increasing levels of intensity that make instruction more distinctive including students' grouping formats and teachers' skill level (Bradley, Danielson, & Doolittle, 2007; Chard, 2012). According to the high-quality Tier 1 instruction should be characterized by the use of (a) research-based core curriculum materials that are aligned to Common Core or other state standards, (b) data to make instructional decisions (e.g., selection of instructional practices, differentiation of learning activities, use of accommodations, use of problem-solving approaches to identify interventions), (c) teaching and learning objectives that are connected in progression within and across grades, and (d) enrichment opportunities for students who exceed benchmarks. Fuchs et al. (2012) argued that high-quality Tier 1 instruction should feature (a) core programs in reading and mathematics, (b) differentiated instruction, (c) accommodations to ensure accessibility to Tier 1 instruction for virtually all students, and (d) problem-solving strategies to address students' motivation and behavior. Fuchs et al. further acknowledged that the general effectiveness of Tier 1 instructional practices should be derived from research, which does not require the same empirical rigor of validation (via experimental or quasi-experimental studies) as is needed of evidence-based practices.

To determine the extent to which high-quality Tier 1 instruction is being delivered, Metcalf (2012) recommended educators and other

multidisciplinary team members review the scope and sequence of their day-to-day instruction and answer the following questions:

> What instructional routines are used? Are the routines consistent from classroom to classroom, general education to special education? Is there evidence of scaffolding and explicit instruction, especially when students are learning something new? Is there evidence of distributed practice of critical skills? Is cumulative review built in on a systematic basis? How much time is allocated? How is that time used (for example, whole group instruction, small group instruction, or independent practice)? Does the pace of the instruction match student needs? Do students have multiple opportunities for response and feedback? Are students actively engaged (that is, are they saying, writing, and doing)? (p. 5).

As such, Metcalf organized high-quality Tier 1 instruction around (a) use of instructional time, (b) emphasis of various instructional practices such as explicit instruction and scaffolding, and (c) student engagement with instructional materials. Moreover, he emphasized the review of ongoing, day-to-day information about classroom instruction.

Gandhi, Holdheide, Zumeta, and Danielson (2016) focused on the following characteristics for purposes of differentiating intervention tiers: (a) approach (i.e., comprehensive, standardized, individualized), (b) group size (i.e., whole class, small group, very small group), (c) progress monitoring (i.e., once per term, once per month, weekly), and (d) population (i.e., all students, at-risk students, students with significant and persistent needs). As evidence to document Tier 1 instruction, they noted comprehensive coverage of critical content and use of instructional practices derived from research. Holdheide (2016) further suggested use of instructional time and screening procedures for documenting the quality of Tier 1 instruction. She identified instructional time, content coverage, and quality of instruction as the defining dimensions of Tier 1 instruction. Holdheide and her colleagues did not suggest, however, what may constitute adequate use of instructional time (e.g., as a percentage of allocated class time) or comprehensive content coverage (e.g., as a percentage of available content standards).

In summary, researchers have identified a variety of characteristics indicative of high-quality Tier 1 instruction. As expected, most of these characteristics address the curricular and qualitative aspects of instruction—the "what" and "how" of a teacher's enacted curriculum. A few researchers also noted temporal considerations (i.e., use of instructional time). To date, these characteristics have lacked an organizing framework as well as operational definitions that address quality as a matter of degree. In other words, what is needed are definitions that go beyond denoting the mere presence of a characteristic (e.g., progress monitoring) by indicating the characteristic's degree of implementation necessary to be considered high quality (e.g., progress monitoring on a weekly basis versus quarterly basis). I argue that the concept of OTL be used to this end for two main reasons: (a) OTL is a well-researched concept with a solid conceptual and empirical basis (see Kurz, 2011); (b) recent advancements in OTL's measurement at the classroom level (e.g., Kurz, Elliott, Kettler et al., 2014; Rowan & Correnti, 2009) can address the next critical step after organizing and operationalizing key characteristics of high-quality Tier instruction, namely, measuring and providing feedback about the extent to which Tier 1 instruction is being implemented with high quality.

Opportunity to Learn

Since the 1960s, researchers have used the OTL acronym to examine a variety of schooling variables and their relationships to student achievement. Due to its continued relevance and applicability, OTL is widely considered a "generative concept" (McDonnell, 1995). That is, OTL has been relevant and applicable for identifying the normative assumptions of policy goals, focusing empirical research on strategies for reaching these goals, and informing ways to measure progress toward these goals. For example, OTL has been used to identify several normative assumptions of standards-based reform such as alignment (Porter, 2002) and access to the general curriculum (Kurz & Elliott, 2011). It has also

Fig. 9.1 Three dimensions of the enacted curriculum that define OTL
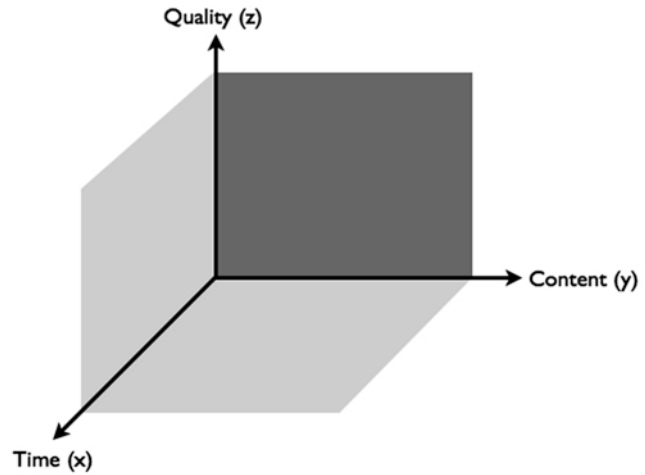
been used to focus empirical research on examining relations between different schooling variables and student outcomes (e.g., Herman, Klein, & Abedi, 2000; Wang, 1998). In addition, several measurement tools of OTL currently hold implications for measuring progress toward certain policy goals (see Kurz, 2011).

I previously reviewed the large variety of OTL variables found in the extant literature focusing on those that showed empirical relations to student achievement (Kurz, 2011) and grouping them into three broad categories related to the time, content, and quality of classroom instruction (e.g., Borg, 1980; Brophy & Good, 1986; Porter, 2002). Based on a review of these three distinct OTL research strands, I provided a conceptual synthesis of OTL that focused on three key dimensions of the enacted curriculum—time, content, and quality—all of which co-occur during instruction (see Fig. 9.1). That is, teachers distribute OTL of what they want students to know and be able to do by allocating instructional time and content coverage to intended objectives using a variety of pedagogical approaches. As such, OTL is considered to be a teacher effect. Teachers provide OTL through their instruction, which is part of the enacted curriculum. The extent to which students engage in that opportunity, of course, is a separate matter. What follows is a summary of the three major OTL research strands I discussed in my earlier work (Kurz, 2011).

Time

The first research strand emerged with John Carroll (1963), who introduced the concept of OTL as part of his model of school learning: "Opportunity to learn is defined as the amount of time allowed for learning, for example by a school schedule or program" (Carroll, 1989, p. 26). Subsequent research on time and school learning (see Borg, 1980; Gettinger & Seibert, 2002) began to empirically examine this OTL conceptualization using general indicators such as *allocated time* (i.e., scheduled time to be allocated to instruction) or more instructionally sensitive indicators such as *instructional time* (i.e., proportion of allocated time actually used for instruction), *engaged time* (i.e., proportion of instructional time during which students are engaged in learning), and *academic learning time* (i.e., proportion of engaged time during which students are experiencing a high success rate).

The amount of time dedicated to instruction has received substantial empirical support in predicting student achievement (e.g., Carroll, 1989; Denham & Lieberman, 1980; Fisher & Berliner, 1985; Walberg, 1988). In a research synthesis on teaching, Walberg (1986) identified 31 studies that examined the "quantity of instruction" and its relation to student achievement. Walberg reported a median (partial) correlation of 0.35 controlling for other variables such as student ability and socioeconomic status. In a meta-analysis on

educational effectiveness, Scheerens and Bosker (1997) examined the effect of (allocated) time on student achievement using 21 studies with a total of 56 replications across studies. The average Cohen's *d* effect size for time was 0.39 (as cited in Marzano, 2000). Considering that time usage related to instruction represents one of the best documented predictors of student achievement across schools, classes, student abilities, grade levels, and subject areas (Vannest & Parker, 2010), it is not surprising that research regarding time on instruction continues to this date.

Content

The second research strand emerged with studies that focused on the *content overlap* between the enacted and assessed curriculum (e.g., Comber & Keeves, 1973; Husén, 1967). Husén, one of the key investigators for several international studies of student achievement, developed an item-based OTL measure that required teachers to report on the instructional content coverage for each assessment item via a 3-point Likert scale: "Thus opportunity to learn from the Husén perspective is best understood as the match between what is taught and what is tested" (Anderson, 1986, p. 3682). The International Association for the Evaluation of Educational Achievement (IEA) has conducted several comparative studies of international student achievement, the results of which have supported "students' opportunity to learn the assessed curriculum" as a significant predictor of systematic differences in student performance. This content overlap conceptualization of OTL remained dominant in several other mostly descriptive research studies during the 1970s and 1980s (e.g., Borg, 1979; Winfield, 1993) and continues to date (e.g., Schmidt et al., 1997, 2001). The latter findings based on international studies such as the Second and Third International Mathematics Studies (SIMS, TIMS) as well as the Programme for International Student Assessment (PISA) have established consistent evidence that greater OTL is related to higher student achievement (Schmidt et al., 2015).

Another line of research on content overlap focused on students' opportunity to learn important content objectives (e.g., Armbuster, Stevens, & Rosenshine, 1977; Jenkins & Pany, 1978; Porter et al., 1978). Porter et al., for instance, developed a basic taxonomy for classifying content included in mathematics curricula and measured whether different standardized mathematics achievement tests covered the same objectives delineated in the taxonomy. Porter continued his research on measuring the content of the enacted curriculum (e.g., Gamoran, Porter, Smithson, & White, 1997; Porter, Kirst, Osthoff, Smithson, & Schneider, 1993) and developed a survey-based measure that examined the content of instruction along two dimensions: topics and categories of cognitive demand (Porter & Smithson, 2001). He subsequently developed an *alignment* index (Porter, 2002), which qualified content overlap based on these two dimensions. The findings of Gamoran et al. indicated that alignment between instruction and a test of student achievement in high school mathematics accounted for 25% of the variance among teachers.

Quality

The third and most diverse research strand related to an instructional dimension of OTL can be traced back to several models of school learning (e.g., Bloom, 1976; Carroll, 1963; Gagné, 1977; Harnischfeger & Wiley, 1976). Both Carroll's model of school learning and Walberg's (1980) model of educational productivity, for example, featured quality of instruction alongside quantity of instruction. The operationalization of instructional quality for purposes of measurement, however, resulted in a much larger set of independent variables than instructional time. Most these variables were focused on *instructional practices* related to student achievement. In his research synthesis on teaching, Walberg (1986) reviewed 91 studies that examined the effect of quality indicators on student achievement, such as frequency of praise statements, corrective feedback, classroom climate, and instructional groupings. Walberg reported the highest mean effect sizes for

(positive) reinforcement and corrective feedback with 1.17 and 0.97, respectively. Brophy and Good's (1986) seminal review of the process-product literature identified aspects of giving information (e.g., pacing), questioning students (e.g., cognitive level), and providing feedback as important instructional quality variables with consistent empirical support. Additional meta-analyses focusing on specific subjects and student subgroups are also available (e.g., Gersten et al., 2009; Vaughn, Gersten, & Chard, 2000). Gersten et al. (2009), for example, examined various instructional components that enhanced the mathematics proficiency of students with learning disabilities. Gersten and colleagues hereby identified two instructional components that provided practically and statistically important increases in effect size: teaching students the use of heuristics (i.e., general problem-solving strategy) and explicit instruction. In addition, researchers have identified *grouping formats* other than whole class (e.g., Elbaum, Vaughn, Hughes, Moody, & Schumm, 2000) and cognitive expectations for learning, so-called cognitive demands (e.g., Porter, 2002), as important qualitative aspects of instruction. With respect to cognitive expectations, 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).

Figure 9.2 compares three classification categories of cognitive process expectations: (a) Webb's Depth-of-Knowledge (DOK) levels (see Webb, 2006), the categories of cognitive demand used as part of the SEC (see Porter, 2002), and the six categories of the cognitive process dimension from the revised Bloom's taxonomy (see Anderson et al., 2001). It should be noted that the latter taxonomy situates all educational objectives within a two-dimensional framework that includes both a knowledge dimension and a cognitive process dimension.
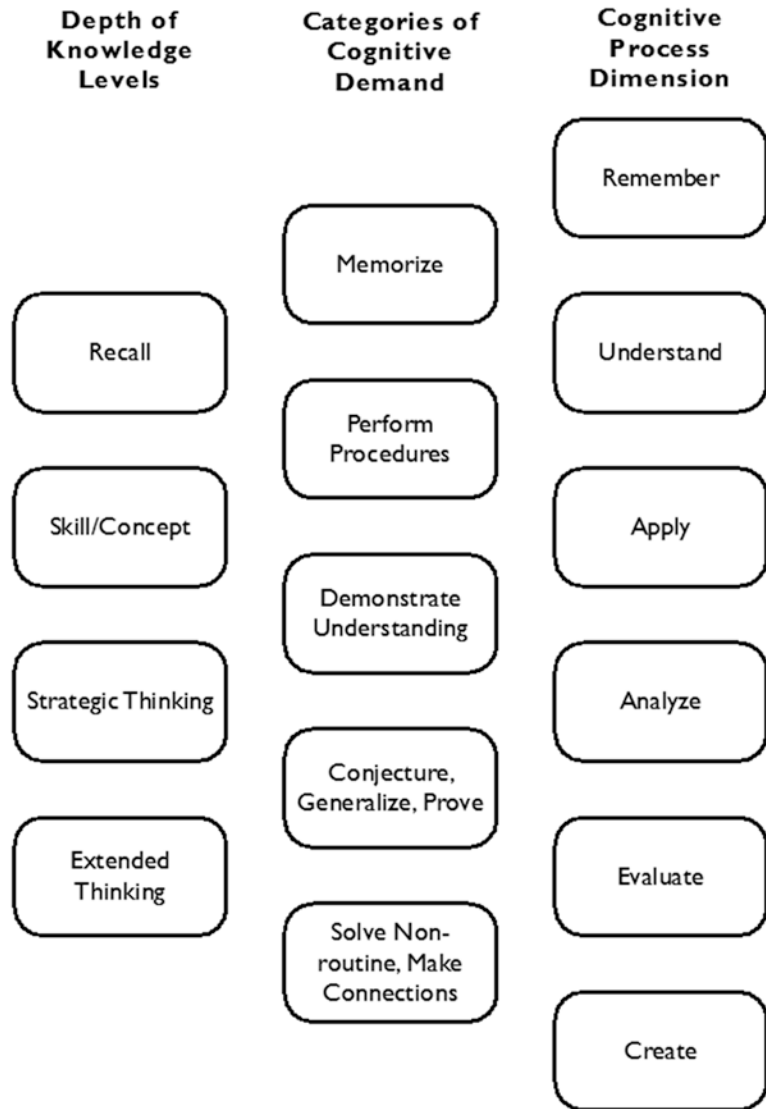
Operational Definition and Measurement

As noted earlier, all three dimensions of the enacted curriculum co-occur during classroom instruction. As such, a comprehensive definition of OTL should include previously discussed OTL indices along each of the three enacted curriculum dimensions: time, content, and quality. Anderson first articulated a merger of the various OTL conceptualizations in 1986: "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). Following Anderson's suggestion, I developed a conceptual synthesis of OTL, which was subsequently operationalized by defining OTL as "the degree to which a teacher dedicates instructional time and content coverage to the intended curriculum objectives emphasizing higher-order cognitive processes, evidenced-based instructional practices, and alternative grouping formats" (Kurz, Elliott, Lemons et al., 2014, p. 27). Building on the work of Rowan and colleagues (e.g., Rowan, Camburn, & Correnti, 2004; Rowan & Correnti, 2009), Kurz and colleagues conducted several studies measuring OTL via an online teacher log that can collect daily self-report data based on several of the previously discussed OTL indices (see Kurz et al., 2015). The online teacher log specifically incorporates indices related to instructional time, content coverage, cognitive processes, instructional practices, and grouping formats into one assessment.

Historically, researchers have relied on classroom observations and teacher surveys to measure OTL. The variability of classroom instruction, however, presents unique challenges for both options (Rowan & Correnti, 2009). To ensure generalizability of classroom observations, researchers must sample a large number of lessons to make valid inferences about OTL for the entire school year. Due to the high costs associated with this approach, most assessments of OTL rely on teacher surveys, typically conducted at the end of the school year (Porter, 2002). Teacher recall, especially across longer time periods, is subject to recall error (Mayer, 1999; Rowan et al., 2004). Teacher logs represent an alternative approach that is intended to (a) reduce a teacher's response burden by focusing on a discreet set of behaviors, (b) increase accuracy of teacher recall by focusing on a recent time period,

Fig. 9.2 Comparison of several classification categories for cognitive process expectations

Depth of Knowledge Levels	Categories of Cognitive Demand	Cognitive Process Dimension
		Remember
	Memorize	
Recall		Understand
	Perform Procedures	
Skill/Concept		Apply
	Demonstrate Understanding	
Strategic Thinking		Analyze
	Conjecture, Generalize, Prove	
Extended Thinking		Evaluate
	Solve Non-routine, Make Connections	
		Create

and (c) increase generalizability through frequent administrations across the school year (Kurz, Elliott, Lemons et al., 2014).

The online teacher log used by Kurz and colleagues requires teachers to undergo proficiency-based knowledge and skill tests prior to use. The log aggregates daily OTL data into a variety of scores, which have been examined for technical adequacy. Kurz, Elliott, Kettler et al. (2014) provided a summary of initial evidence supporting intended score interpretations for assessing OTL. The summary included multiple sources of evidence: usability, reliability, and validity evidence based on content, responses processes, internal structure, relations to other variables, and consequences of using the measure. More recently, Berkovits, Kurz, and Reddy (2017) provided evidence for the measure's convergent validity with a classroom observational assessment. Table 9.1 adapted from Kurz, Elliott, Kettler et al. (2014) provides details on five major scores provided by the online teacher log called the *Instructional Learning Opportunities Guidance System* (MyiLOGS; Kurz & Elliott, 2012).

Data collection using MyiLOGS has occurred in the context of research studies with incentivized participants as well as part of teachers' regular professional development. Collectively, these

Table 9.1 Enacted curriculum dimensions, OTL indices, and score definitions

Dimension	OTL index	Score definition[a]
Time	Instructional time (IT)	Score between 0.00 and 1.00, which represents the percentage of allocated time used for instruction on standards and custom objectives
Content	Content coverage (CC)	Score between 0.00 and 1.00, which represents the cumulative percentage of academic standards covered for 1 min or more
Quality	Cognitive processes (CP)	Score between 1.00 and 2.00, which represents the percentage of time spent on higher-order cognitive processes +1.00[b]
	Instructional practices (IP)	Score between 1.00 and 2.00, which represents the percentage of time spent on evidence-based instructional practices +1.00[b]
	Grouping formats (GF)	Score between 1.00 and 2.00, which represents the percentage of time spent on individual and/or small group formats +1.00[b]

[a]*Note*: Score definitions are specific to the Instructional Learning Opportunities Guidance System (MyiLOGS; Kurz & Elliott, 2012)

[b]*Note*: The addition of +1.00 is intended to prevent negative user associations with a score of 0

efforts have yielded three important findings: (a) teachers can be trained to provide reliable self-report data when used for self-reflection and professional development; (b) these data can yield valid inferences about their provision of OTL; and (c) their OTL score profiles can be used to establish normative standards (Berkovits, Kurz, & Reddy, 2017; Kurz, Elliott, Kettler et al., 2014; Kurz et al., 2015; Kurz, Reichenberg, & Yel, 2017). Moreover, teachers' OTL scores have shown to be sensitive to job-embedded professional development (i.e., instructional coaching). Controlling for several teacher characteristics and class size, Kurz (2016) reported on coaching status predicting OTL score increases for instructional time (IT), content coverage (CC),

instructional practices (IP), and grouping formats (GF). These OTL scores further maintained moderate correlations with median growth scores on benchmark assessments in reading and mathematics ranging from 0.31 to 0.49 (Kurz, 2017).

In summary, researchers have used the concept, definition, and measurement of OTL to describe classroom instruction and its relationship to student achievement. Measurement tools that can be completed on a daily basis such as online teacher logs are capable of capturing the scope and sequence of teachers' day-to-day instruction—an important criterion—if Tier 1 data are to be used formatively (Metcalf, 2012). Several research studies have already provided evidence that OTL data can be used in the context of professional development to drive measurable instructional changes. Research on OTL is thus positioned to confront the known unknown of Tier 1 instruction in several ways: (a) provide data on teachers' use of instructional time, content coverage, and instructional quality, (b) collect these data formatively to drive improvements in Tier 1 instruction, and (c) set quality standards based on these data. As a first step, however, we need to synthesize the discussed RTI and OTL literature relevant to Tier 1 OTL to identify potential sources of evidence for high-quality Tier 1.

Tier 1 OTL

Expectations for what students should know and be able to do must be articulated across all levels of the educational environment. To delineate these levels, researchers have developed curriculum frameworks, which typically emphasize how different curricula relate to the intended curriculum (e.g., Anderson, 2002; Porter, 2002; Webb, 1997). That is, the intended curriculum represents the normative target for all other curricula by defining students' learning objectives (usually by way of subject- and grade-specific standards). Based on this premise, I developed a curriculum framework delineating key curricula at the system, teacher, and student level (see Fig. 9.3).

As shown in the figure, students access the intended curriculum through the teacher's enacted

Fig. 9.3 Curriculum framework that delineates normative function of intended curriculum
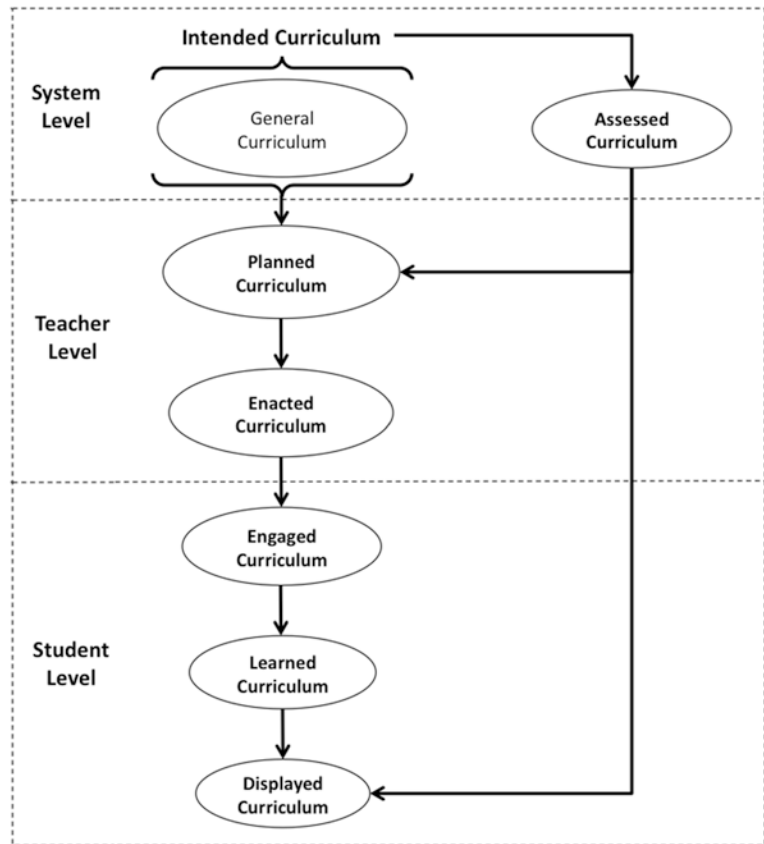
curriculum. The extent to which this happens is captured by OTL along the three dimensions of the enacted curriculum: time, content, and quality. By definition, multitiered support systems are predicated on positing different intended curricula depending on the needs of students. A teacher's enacted curriculum must consequently differ at each level of prevention: primary, secondary, and tertiary. If OTL is to serve as a generative concept for RTI, then OTL must be defined and refined according to each tier: Tier 1 OTL, Tier 2 OTL, and Tier 3 OTL.

The comprehensive definition of OTL provided by Kurz, Elliott, Lemons et al. (2014) was based on empirically supported OTL indices that came out of the three major OTL research strands, which have focused almost exclusively on general instruction provided to the vast majority of students. As such, their OTL definition is most directly applicable to Tier 1 instruction. The OTL scores based on this definition (see Table 9.1) are

designed to address several important features of general instruction for students with and without disabilities. Although students with disabilities served in general education under a full inclusion model typically exhibit less severe disabilities, they are still served via an Individualized Education Program (IEP). Both reauthorizations of the IDEA in 1997 and 2004 emphasize the IEP as the central mechanism for detailing a student's access, involvement, and progress in the general curriculum (Karger, 2005). The IEP also documents educational objectives relevant to a student's present levels of performance as well as accommodations and modifications that facilitate the student's access to enacted and assessed curricula (Ketterlin-Geller & Jamgochian, 2011). The IEP thus augments the general curriculum, which is the reason I qualified the curriculum framework displayed in Fig. 9.3 for students with disabilities. That is, I noted that the intended curriculum for students with disabilities is dually

determined by both the IEP and the general curriculum (Kurz, 2011). Therefore, measurement of OTL at the enacted curriculum must be able to account for a teacher's time spent teaching the students' IEP objectives in addition to the academic standards that define the general curriculum.

For purposes of Tier 1, assessment of OTL along the time dimension of the enacted curriculum can occur based on any of the aforementioned indices: allocated time, instructional time, engaged time, and academic learning time. While these indices exhibit stronger relations to achievement the more directly they measure student engagement and success, Metcalf's (2012) imperative to measure the scope and sequence of teachers' day-to-day instruction for purposes of high-quality Tier 1 instruction limits the practicality of measuring these more student-centric indices on a daily basis. The instructional time (IT) index, I contend, strikes a reasonable balance between empirical evidence (i.e., the extent to which time indices are predictive of student achievement) and feasible measurement. Moreover, we have evidence that teachers can be trained to estimate their time used for instruction reasonably well by subtracting non-instructional time such as transitions and other interruptions from their allocated time (Kurz, Elliott, Kettler et al., 2014). In addition, the operationalization of IT for measurement purposes via an online teacher log (see Table 9.1) also addresses the previous concern of capturing time spent on both academic standards and custom objectives (i.e., IEP objectives). The score developed by Kurz and Elliott (2012)—IT as the percentage of allocated time used for instruction on standards and custom objectives—further can be interpreted as a measure of efficiency. The IT score does not capture total minutes of instruction but rather how efficiently a teacher can use the allocated time for a particular class (e.g., 60-min mathematics class) for instruction on standards and custom objectives (e.g., 53 min) resulting in a percentage (e.g., IT = 88%).

In the context of RTI Tier 1, researchers have argued for the implementation of a research-based reading or mathematics program that provides instructional objectives and materials

aligned to Common Core or other state standards (Fuchs et al., 2012). While IT can capture the amount of time spent teaching intended standards as well as any custom objectives specific to a reading or mathematics program, the time index does not address whether the number of intended standards taught is adequate. In fact, Holdheide (2016) specifically argued for adequate coverage of Tier 1 content. As such, assessment of OTL along the content dimension of the enacted curriculum can be used to address this concern. The content coverage (CC) index of OTL provides a basic measure of the breadth of coverage. MyiLOGS, for example, simply provides a cumulative percentage of academic standards covered (i.e., Common Core State Standards) for at least 1 min or more. This threshold can be adjusted, but it is purposefully set low to avoid construct overlap with the time index.

Thus far, we have identified at least two sources of evidence for high-quality Tier 1 instruction related to the time and content dimensions of the enacted curriculum. Given these are fully aligned with previously established OTL indices, we can use prior research based on these indices to draft operational definitions—short of setting actual standards. For example, large-scale research based on hundreds of teachers across multiple states teaching a variety of subjects and grade levels provided averages for IT ranging between 84% and 94% (Kurz, 2017). At a minimum, we can state that teachers should be able to spend the majority of allocated time on instruction for proposes of high-quality Tier 1 instruction. Looking at content coverage using the same large-scale research, we were able to calculate averages for CC ranging between 54% and 68% (Kurz, 2017). At a minimum, we can make a similar statement, namely, that teachers should be able to cover the majority of subject- and grade-specific content standards for proposes of high-quality Tier 1 instruction.

Assessment of OTL along the quality dimension of the enacted curriculum suggests three sources of evidence of primary prevention: (a) emphasis of higher-order cognitive processes, (b) emphasis of research-based instructional

practices, and (c) emphasis of grouping formats other than whole class such as small groups (Kurz, Elliott, Lemons et al., 2014). To develop summative scores, Kurz and colleagues operationalized emphases based on time spent in one of two categories (low-order vs. higher-order cognitive process, non-research-based vs. research-based instructional practices, whole class vs. alternative grouping formats). For the cognitive processes, they used the work by Anderson et al. (2001). For the instructional practices and grouping formats, they focused on practices and formats with empirical support based on meta-analyses (e.g., Elbaum et al., 2000; Gersten et al., 2009). The use of two categories for all three quality-related scores was based on two operating assumptions: (a) teachers who address a range of cognitive processes, instructional practices, and grouping formats during the course of their instruction, and (b) teachers who emphasize higher-order cognitive processes, research-based instructional practices, and alternative grouping formats that improve the quality of students' OTL. Given that the empirical basis for these assumptions is insufficient to single out specific processes, practices, or formats, they argued for a dichotomous grouping. Disaggregated information by each cognitive process, instructional practice, and grouping format is also available.

I contend that the quality index-based cognitive process, instructional practices, and grouping formats remain relevant for purposes of high-quality Tier 1 instruction. But rather than establishing the normative goal of "emphasis" (i.e., majority of time spent in a preferred category), I recommend these quality indices (not their summative scores) be used to operationalize a key characteristic of high-quality Tier 1 instruction noted in the RTI literature: differentiated instruction (e.g., Fuchs et al., 2012). In other words, the use of differentiated instruction represents evidence of primary prevention and is operationalized by the use of different cognitive processes, instructional practices, and grouping formats. The next section provides a brief case example to illustrate some measurement options related to differentiated instruction.

The next critical characteristic of high-quality Tier 1 discussed in the RTI literature that falls under the quality dimensions of the enacted curriculum is the use of instructional accommodations (e.g., Fuchs & Vaughn, 2012; RTI Action Network, n.d.). Instructional accommodations are intended to increase students' access to the enacted curriculum allowing them to learn the same material to the same level of performance as other students in the general education classroom (Ketterlin-Geller & Jamgochian, 2011). More specifically, the teacher should make instructional adaptations to the design and deliver of instruction and associated materials based on presentation, setting, timing/scheduling, or response mode (see Chap. 14 of this volume). This source of evidence of primary prevention could be operationalized based on these adaptation categories and measured via frequency counts using a simple checklist. In fact, Elliott, Kratochwill, and Schulte (1999) developed a detailed assessment accommodations checklist that could be adapted for gathering instructional accommodation evidence.

The final two Tier 1 characteristics that fall under the quality dimensions of the enacted curriculum are the use of universal screening and progress monitoring (Fuchs et al., 2012). As such, high-quality Tier 1 instruction includes brief screening assessments for all students—ideally followed by additional or short-term progress monitoring to confirm risk status and movement in subsequent tiers of prevention (Fuchs & Vaughn, 2012). The Center for Response to Intervention (2014) put forth three criteria for high-quality universal screening:

> (1) screening is conducted for all students (i.e., is universal); (2) procedures are in place to ensure implementation accuracy (i.e., all students are tested, scores are accurate, cut points/decisions are accurate); and (3) a process to screen all students occurs more than once per year (e.g., fall, winter, spring) (p. 1).

For purposes of progress monitoring, teachers should conduct regular assessments to monitor students' academic performance, quantify student rate of improvement or responsiveness to instruction,

Table 9.2 Enacted curriculum dimensions, evidence sources, and operational definitions for high-quality Tier 1 instruction

Enacted curriculum dimension	Evidence of primary prevention	Operational definition
Time	Efficient use of instructional time	Teacher uses the vast majority of allocated class time for instruction
Content	Implementation of a research-based and aligned core program	Teacher fully implements a research-based reading or mathematics program that provides instructional objectives and materials aligned to common core or other state standards
	Comprehensive content coverage	Teacher covers the majority of subject- and grade-specific content standards
Quality	Use of differentiated instruction	Teacher regularly differentiates cognitive process expectations, instructional practices, and grouping formats to address individual needs based on students' background knowledge, readiness for the instructional objective, language skills and abilities, preferences, and interests
	Use of instructional accommodations	Teachers regularly make adaptations to the design and deliver of instruction and associated materials based on presentation, setting, timing or scheduling, and response mode
	Use of progress monitoring	Teachers regularly assess students to monitor academic performance, quantify student rate of improvement or responsiveness to instruction, and evaluate the effectiveness of instruction
	Use of universal screening	Teacher regularly conducts brief screening assessments with all students followed by additional testing or short-term progress monitoring to confirm students' risk status and need for additional interventions

and evaluate the effectiveness of instruction (Fuchs & Vaughn, 2012). The Center for Response to Intervention (2014) also developed two criteria for high-quality progress monitoring:

> (1) progress monitoring occurs at least monthly for students receiving secondary-level intervention and at least weekly for students receiving intensive intervention; and (2) procedures are in place to ensure implementation accuracy (i.e., appropriate students are tested, scores are accurate, decision-making rules are applied consistently). (p. 2).

At this point, I have identified seven sources of evidence for primary prevention that fall along the three dimensions of the enacted curriculum: (a) efficient use of instructional time, (b) implementation of a research-based and aligned core program, (c) comprehensive content coverage, (d) use of differentiated instruction, (e) use of instructional accommodations, (f) use of progress monitoring, and (g) use of universal screening. Table 9.2 includes these evidence sources by

each enacted curriculum dimension along with their respective operational definitions.

The operational definitions listed in the table, however, still fall short of setting actual quality standards for high-quality Tier 1 instruction. That is, the operational definitions currently contain qualitative qualifiers (e.g., majority, fully, regularly) rather than quantitative ones. For two dimensions—time and content—I provided quantitative ranges, because I was able to draw from several descriptive OTL studies that can be used to establish some initial base rates. More research is needed, specifically, on base rates regarding the use of differentiated instruction and instructional accommodations. Some suggested frequencies that can quantitatively define "regular" progress monitoring and universal screening are actually available. The Center for Response to Intervention (2014) suggested three universal screenings per year (i.e., fall, winter, spring), and Fuchs and Fuchs (2006) suggested an ideal progress moni-

toring frequency of every 2 weeks. Next, I provide a brief case example using MyiLOGS and its newly developed lesson planner, easyCBM, and the *Checklist of Learning and Assessment Adjustments for Students* (CLASS; Davies, Elliott, & Cumming, 2016) to show how data on the seven evidence sources can be collected.

Case Example

To illustrate how Tier 1 data collection can occur, I discuss a hypothetical fifth-grade English Language Arts (ELA) classroom. The focus of this brief example is *how* these data can be collected using existing tools. The ELA classroom consists of 28 fifth-grade students who are currently in their third quarter working on a unit that explores examining two texts with similar themes. The units are based on a core reading curriculum. The K-8 curriculum was adopted by the district based on its research-based elements for instructional content (i.e., phonemic awareness, phonics, fluency, vocabulary, comprehension) and instructional design (i.e., explicit instructional strategies, coordinated instructional sequences, ample practice opportunities, aligned student materials). In addition, all curriculum units are aligned to the Common Core State Standards (CCSS). The current unit, which features five mini lessons, addresses four CCSS standards (i.e., RL5.2, RL.5.2, RL.5.3, RL.5.9). The first two standards, RL.5.2 and RL.5.3, focus on summarizing texts based on their themes and comparing/contrasting parts of a story such as characters, settings, and events. The teacher intends to cover the first two standards during the first half of the week. The

ELA block typically lasts 75 min with Wednesday being an early release day, which shortens the lesson to 45 min. To document evidence of high-quality Tier 1 instruction, the teacher plans and monitors her lessons via MyiLOGS (Kurz & Elliott, 2012), an online teacher log and its newly developed lesson planning feature.

The MyiLOGS lesson planner provides teachers with a monthly instructional calendar that includes an expandable sidebar, which lists all intended objectives for a class. Teachers drag and drop planned objectives that are to be the focus of upcoming lessons onto the respective calendar days. After implementing their lessons, teachers are required to confirm enacted objectives, instructional time dedicated to each objective, and any time not available for instruction at the class level. Figure 9.4 shows an excerpt from the teacher's instructional calendar in MyiLOGS.

Her calendar shows that she intends to cover the current unit on parallel texts and, more importantly, the two content standards RL.5.2 and RL5.3. MyiLOGS also allows her to enter any intended custom objectives under the green "+" sign. She anticipates about 8 min of non-instructional time on Monday and Tuesday and about 5 min of non-instructional time for her early release day. Each day also features a notepad icon, which brings the user to the actual lesson plan. Figure 9.5 shows an excerpt from the teacher's lesson planner in MyiLOGS.

The figure shows the content to be covered on the top left under *Standards* and then several options for three additional lesson elements: *Activities*, *Practices*, and the *Class Roster*. The content standards, student activities, and instructional practices can be dragged and dropped on
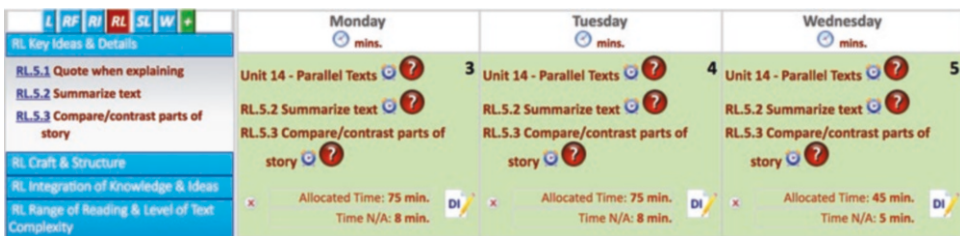

Fig. 9.4 MyiLOGS instructional calendar excerpt

Fig. 9.5 MyiLOGS
lesson planner excerpt

the right-hand side under either *General Instruction* or *Differentiated Instruction*. Several tabs are available for various lesson segments such as *Opening* or *Tell/Show*. These tabs can be customized to whatever lesson plan formats and segments are desired by different school leaders or district requirements. Both spaces allow standards, activities, and practices boxes to be sized according to the intended time frame. In this case, she begins the lesson with 15 min of independent reading on Unit 14, which covers excerpts from "White Socks Only" and "The Story of Ruby Bridges." The Standards box further contains a DOK level indicator, which is focused on recall and reproduction (i.e., Level 1). The Activity box indicates "W" for whole class

instruction. All three boxes further contain a text editor, where the details of the lesson plan can be captured. Three students have advanced in their readings during the previous lesson. She decides to move these students directly into a small group assessing their knowledge of theme based on details from both parallel texts. Given that the questions on the quiz require students to think deeply about text details including analysis and judgment, she assigns a Level 3 DOK (i.e., short-term strategic thinking). As can be seen, this online teacher log can be used to capture the scope and sequence of day-to-day instruction.

MyiLOGS further calculates a variety of OTL scores and descriptive information based on the data logged via the lesson planner. The teacher

can review over a dozen figures that detail her use of instructional time, implementation of the core program, content coverage, as well as details about her differentiated instruction. Given that allocated time is logged on a daily basis, IT can be calculated accurately. For her class, her average IT score across 107 school days is 89%. Logged time per standard, student activity, and instructional practice further permits the calculation of several additional OTL scores. During 107 school days, for example, she has covered 14 out of 21 units (66%) in the core curriculum with a CC score of 57% for her cumulative coverage of the CCSS reading and writing standards. The current evidence on time and content thus indicates that she is (a) using the majority of her allocated class time for instruction, (b) on track for covering most core curriculum units, and (c) on track for covering the majority of subject- and grade-specific content standards. To better target the remaining CCSS standards, she reviews the MyiLOGS content coverage bar chart that delineates time emphases along all intended standards for the subject and grade in question. Figure 9.6 shows an excerpt of the bar chart, which reveals several content standards yet to be covered.

The newly developed MyiLOGS lesson planner further permits several scores related to differentiated instruction. In addition to MyiLOGS's previous CP, IP, and GF (summary) scores, the lesson planner scores provide information about, and comparisons between, general instruction and any instructional changes logged under the differentiated instruction track. Based on her 107 logged days, for example, she has implemented differentiated instruction (for at least part of her lesson) on 58 days (54%). Based on her logged general instruction time (GET), she specifically provided differentiated instruction during 38% of her GET. Given that the lesson planner requires the assignment of differentiated instruction based on students using the class roster, she further knows that 100% of her students have received some differentiated instruction during the past 107 school days. MyiLOGS provides additional charts that detail the type of instructional differences based on the following categories: (a) content standards, (b) DOK levels, (c) student activities, (d) grouping formats, and (e) instructional practices. The charts are descriptive in as so far as the only document instructional differences. To determine the adequacy of the match between these instructional differences and students' individual needs would require detailed information about students' background knowledge, readiness for the instructional objective, language skills and abilities, preferences, and interests. At a minimum, however, the teacher has

RL Key Ideas & Details	**60.9%**	**17 hrs 30 mins**
RL.5.1 Quote when explaining		
RL.5.2 Summarize text	29.6%	8 hrs 30 mins
RL.5.3 Compare/contrast parts of story	31.3%	9 hrs 0 mins
RL Craft & Structure	**13.0%**	**3 hrs 45 mins**
RL.5.4 Similies and methaphors	5.2%	1 hrs 30 mins
RL.5.5 Explain structure		
RL.5.6 POV influences descriptions	7.8%	2 hrs 15 mins
RL Integration of Knowledge & Ideas	**11.3%**	**3 hrs 15 mins**
RL.5.7 Visual elements & meaning		
RL.5.9 Compare/contrast stories	11.3%	3 hrs 15 mins
RL Range of Reading & Level of Text Complexity	**14.8%**	**4 hrs 15 mins**
RL.5.10 Read/comprehend literature	14.8%	4 hrs 15 mins
SL Comprehension & Collaboration	**0%**	
SL.5.1 Engage in discussions		
SL.5.1a Come to discussions prepared		
SL.5.1b Follow discussion rules		
SL.5.1c Ask questions to contribute to discussion		
SL.5.1d Review key ideas		
SL.5.2 Summarize		
SL.5.3 Summarize points with support		

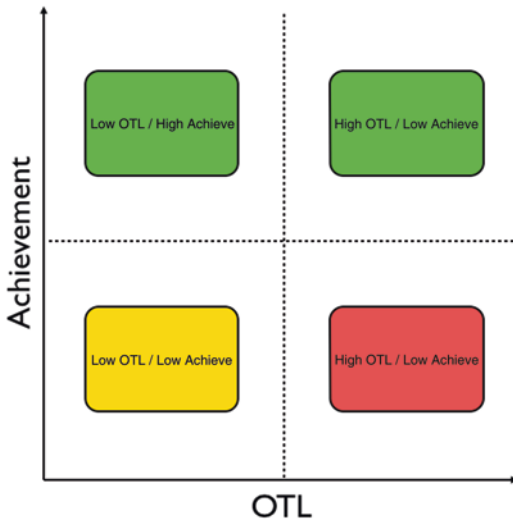**Fig. 9.6** Excerpt from the MyiLOGS content coverage bar chart

Fig. 9.7 Achievement/OTL chart for instructional priority setting

evidence that she regularly differentiates content standards, DOK levels, student activities, grouping formats, and instructional practices.

MyiLOGS can further provide integration with a curriculum-based measurement called easyCBM to schedule regular progress monitoring and review instructional data alongside student achievement. As such, she knows that she schedules and administers easyCBM probes every 2 weeks to monitor students' academic performance, quantify their rate of improvement or responsiveness to instruction, and evaluate the effectiveness of instruction. In addition, she regularly reviews easyCBM data alongside her content coverage. After each probe, for example, she examines students' achievement data by content domains such as *Key Ideas & Details*, *Craft & Structure*, and *Integration of Knowledge & Ideas* alongside her content coverage as displayed in Fig. 9.6. She therefore knows what content standards were emphasized and the extent to which her students answered items in that domain correctly. Based on the excerpt shown in Fig. 9.6, she knows that she did not yet cover the standards under *Comprehension & Collaboration*. To adjust her scope and sequence of instruction, she follows a simple decision-making chart.

Figure 9.7 shows how she prioritizes content areas. The y-axis refers to student achievement on her curriculum-based measurement. The higher the class or student performance is, the further along the placement on the y-axis. The x-axis refers to OTL for a class or student. The figure highlights two quadrants in green. Both quadrants indicate high student achievement either in the context of high OTL (e.g., extensive content coverage, high use of instructional time) or in the absence of OTL (i.e., student prior knowledge). Either way, these content domains for which we have evidence of high student achievement are not a priority for immediate teaching purposes. Low student achievement, however, prompts her to review OTL data, especially in the context of high OTL.

To gather evidence on instructional accommodations, the MyiLOGS lesson planner further provides the option to complete the CLASS (Davies et al., 2016). The CLASS allows her to log 67 accommodations grouped into eight categories: (a) motivation, (b) scheduling, (c) setting, (d) assistance with directions, (e) assistance prior to testing, (f) assistance during learning or assessment, (g) equipment or assistive technology, and (h) changes in format. These adjustments have their basis in educational instruction, testing standards, and accessible educational practices. Figure 9.8 shows an excerpt from the CLASS, which allows her to record instructional accommodations on an individual basis. While current technology integration does not permit data aggregation such as the percentage of students receiving accommodations or information about the types of accommodations used, the teacher is at least able to document her efforts along key accommodation categories.

As illustrated throughout this case, it is currently possible to document all necessary evidence sources for high-quality Tier 1 instruction with relatively few tools. Further technology integration has the potential to make these efforts even more efficient. As a concluding step, I will summarize my argument and provide some key points for a future research agenda.

Fig. 9.8 Excerpt from the checklist of learning and assessment adjustments for students

Assistance During Learning or Assessment
33 Arrange for a special education teacher or other qualified person to manage learning or assessment
34 Read learning or assessment expectations and content to student
35 Sign learning or assessment expectations and content to student
36 Restate the learning or assessment task with more appropriate vocabulary or define unknown vocabulary
37 Turn pages for the student in learning or assessment
38 Record student's responses (in writing or by audio recording)
39 Provide spelling assistance, where appropriate, during learning or assessment
40 Have teacher sit near student during learning or assessment
41 Use test form with vertically arranged multiple-choice items that have an answer circle to the left of each choice
42 Provide cues such as stop signs or arrows on learning activities or the test form
43 Allow responses to be marked in the learning or test book rather than on a separate answer document
44 Assist the student in tracking learning or test items by pointing to or by placing student's finger on item

Tier 1 OTL: Why, How, and What For?

In most cases, primary prevention at Tier 1 is simply assumed to occur. That is, positive screening results are assumed to represent a student's inadequate response to high-quality Tier 1 instruction. Instructional data about a teacher's enacted curriculum are typically not available and therefore not considered in the interpretation of screening results or subsequent progress monitoring based on formative assessments. In other words, we do not know about the extent to which Tier 1 instruction was actually implemented with high quality. This lack of instructional data undermines the validity of critical test score inferences (i.e., a student's inadequate response to high-quality Tier 1 instruction) and subsequent RTI decisions (i.e., a student's need for additional intervention). Confronting this known unknown is thus essential to maintaining the integrity of RTI at its most fundamental level, namely, ensuring high-quality Tier 1 instruction that can support the vast majority of students.

In recent years, researchers have provided more nuanced descriptions of Tier 1 instruction with the intent to better define characteristics of high quality. The concept of OTL has been used by researchers to examine classroom instruction based on instructional indices along all three dimensions of enacted curriculum—time, content, and quality. In this chapter, I argued that these empirically validated indices and their respective measurement through teacher logs are key to confronting the known unknown of Tier 1 instruction. As such, I used the enacted curriculum (see Fig. 9.1) as an organizing framework to highlight how OTL addresses several, previously unaccounted characteristics of high-quality Tier 1 instruction (i.e., instructional time, content coverage) and to reveal several shortcomings of OTL for capturing high-quality Tier 1 instruction (i.e., core program, differentiated instruction, instructional accommodations, universal screening, progress monitoring). These efforts culminated in seven sources of evidence for high-quality Tier 1 instruction (see Table 9.2) that collectively represent Tier 1 OTL. I further provided operational definitions for each evidence source drawing from previously defined OTL scores. In the case of differentiated instruction, I used multiple OTL scores (i.e., cognitive processes, instructional

practices, and grouping formats) to better operationalize measurable differences. I also provided a case example to illustrate how currently available tools can be used to gather evidence for high-quality Tier 1 instruction.

Despite this progress, one may still wonder as to why the concept of OTL should be expanded and refined to represent Tier 1 OTL. First, we need to remember that the concept of OTL originated with the intent to better understand student learning, specifically, student learning as a function of classroom instruction, which is largely under the control of the teacher. Over the years, OTL was used to define and measure the "what" and "how" of classroom instruction. This led to the understanding of OTL as a teacher effect, which provides students access to the intended curriculum through the teacher's enacted curriculum. Much of the current volume, however, argues for a shift from *access to* toward *accessibility of* instruction and testing. As such, we should ensure that any conceptual definition of OTL should not only address access to intended curriculum but also accessibility of the enacted curriculum. The inclusion of differentiated instruction and instructional accommodations thus clarifies that students' opportunity to learn also depends on the extent to which the enacted curriculum accounts for their background knowledge, readiness for the instructional objective, language skills, abilities, preferences, and interest. This necessitates not only changes to general instruction using different practices, cognitive processes, grouping formats, presentations, materials, settings, schedules, or response modes but also changes based on assessment results via screening and progress monitoring.

Evidence for high-quality Tier 1 instruction, while necessary for accurate test score interpretations, nonetheless holds its greatest potential for purposes of instructional improvement of the enacted curriculum. Teachers, especially early career teachers, are often asked to become self-reflective practitioners. Self-reflection, however, requires a framework. We need to know what to reflect upon, and, ideally, we have quantitative information to support our subsequent decisions. Throughout this chapter, I have used the concept of OTL to develop such a framework.

In addition, I have incorporated what is known about high-quality Tier 1 instruction. The proposed evidence sources can guide teachers' instructional improvement efforts and lead to a more seamless integration of data about instruction and student achievement.

Future research and development efforts must address the feasibility and adequacy challenges of Tier 1 OTL. The feasibility challenge stems for the additional time and effort needed to document the various evidence sources. While measurement options are indeed available, they are not well integrated and, in some instances, may even require duplicate efforts (i.e., submitting a paper lesson plan and then logging the lesson plan via the teacher log). Without offering teachers a value-added proposition such as saving time on activities they already do without adding new tasks, it is unlikely that teachers will find the extra time and effort to engage meaningfully with Tier 1 data. Integration of the various tools and clear ways to replace already required tasks (i.e., print options for the lesson planner that provides the format of a traditional lesson plan) is the first step to overcoming the feasibility challenge. The second step is to support teachers in their self-reflection and instructional decision-making. Instructional coaching represents a viable strategy for assisting teachers in these efforts (Joyce & Showers, 2002; Kurz, Reddy, & Glover, 2017). Tier 1 OTL data are complex and require careful decision-making, especially when student background characteristics come into play for purposes of instructional accommodations. Instructional coaches can provide the intensive, job-embedded professional development necessary to use Tier 1 OTL data in the context of goal setting and Tier 1 instructional changes.

To address the adequacy challenge, future research and development efforts must deploy the integrated measurement tools at scale to establish base rates for the various evidence source of Tier 1 OTL across a variety of subjects, grades, as well as additional teacher, student, and school characteristics. Standard-setting methods such as the iterative judgmental policy capturing (JPC) performance-standard-setting procedure could then be used for the multidimensional score profiles (Kurz, Reichenberg, & Yel, 2017)

to establish overall standards of quality provision (i.e., ineffective, partially effective, effective, highly effective). At the very least, base rates could be used to translate the qualitative qualifiers (e.g., majority, fully, regularly) of the current Tier 1 OTL definitions into quantitative ones.

Given that high-quality Tier 1 instruction is intended to serve as the primary prevention for the vast majority of the over 50 million K-12 students, the importance of the outlined Tier 1 OTL efforts can hardly be overstated. The current framework provides an initial step. And last but not least, the newly proposed concept of Tier 1 OTL also initiates a fundamental shift in the conceptualization of OTL from addressing students' access to the enacted curriculum toward OTL ensuring the accessibility of the enacted curriculum.

References

Al Otaiba, S., Connor, C. M., Folsom, J. S., Wanzek, J., Greulich, L., Schatschneider, C., & Wagner, R. K. (2014). To wait in tier 1 or intervene immediately. *Exceptional Children*, *81*(1), 11–27. https://doi.org/10.1177/0014402914532234

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

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.

Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., … Wittrock, M. C. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. New York, NY: Longman.

Armbuster, B. B., Stevens, R. J., & Rosenshine, B. (1977). *Analyzing content coverage and emphasis: A study of three curricula and two tests (Technical Report No. 26)*. Urbana, IL: Center for the Study of Reading, University of Illinois.

Berkovits, I., Reddy, L. A., & Kurz, A. (2017). *Teacher log of students' opportunity to learn and classroom observation: A preliminary investigation of convergence*. Manuscript submitted for publication.

Bloom, B. S. (1976). *Human characteristics and school learning*. New York, NY: McGraw-Hill.

Borg, W. R. (1979). Teacher coverage of academic content and pupil achievement. *Journal of Educational Psychology*, *71*(5), 635–645.

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

Bradley, R., Danielson, L., & Doolittle, J. (2007). Responsiveness to intervention: 1997 to 2007. *Teaching Exceptional Children*, *39*(5), 8–12.

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

Carroll, J. B. (1963). A model of school learning. *Teachers College Record*, *64*(8), 723–733.

Carroll, J. B. (1989). The Carroll model: A 25-year retrospective and prospective view. *Educational Researcher*, *18*(1), 26–31.

Chard, D. J. (2012). Systems impact: Issues and trends in improving school outcomes for all learners through multitier instructional models. *Intervention in School and Clinic*, *48*(4), 198–202. https://doi.org/10.1177/1053451212462876

Comber, L. C., & Keeves, J. P. (1973). *Science education in nineteen countries*. New York, NY: Halsted Press.

Compton, D. L., Gilbert, J. K., Jenkins, J. R., Fuchs, D., Fuchs, L. S., Cho, E., et al. (2012). Accelerating chronically unresponsive children to tier 3 instruction: What level of data is necessary to ensure selection accuracy? *Journal of Learning Disabilities*, *45*(3), 204–216. https://doi.org/10.1177/0022219412442151

Cook, B. G., & Odom, S. L. (2013). Evidence-based practices and implementation science in special education. *Exceptional Children*, *79*(2), 135–144.

Davies, M. D., Elliott, S. N., & Cumming, J. (2016). Documenting support needs and adjustment gaps for students with disabilities: Teacher practices in Australian classrooms and on national tests. *International Journal of Inclusive Education*, *20*(12), 1252–1269. https://doi.org/10.1080/13603116.2016.1159256

Denham, C., & Lieberman, A. (Eds.). (1980). *Time to learn*. Washington, DC: National Institute for Education.

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., Kratochwill, T. R., & Schulte, A. G. (1999). *Assessment accommodations guide*. Monterey, CA: CTB/McGraw Hill.

Fisher, C. W., & Berliner, D. C. (Eds.). (1985). *Perspectives on instructional time*. New York, NY: Longman.

Fuchs, D., & Fuchs, L. S. (2006). Introduction to response to intervention: What, why, and how valid is it? *Reading Research Quarterly*. https://doi.org/10.1598/RRQ.41.1.4

Fuchs, D., Fuchs, L. S., & Compton, D. L. (2012). Smart RTI: A next-generation approach to multilevel prevention. *Exceptional Children*, *78*(3), 263–279.

Fuchs, D., Fuchs, L. S., Mathes, P., & Simmons, D. (1997). Peer-assisted learning strategies: Making classrooms more responsive to student diversity. *American Educational Research Journal*, *34*, 174–206.

Fuchs, L. S., & Vaughn, S. (2012). Responsiveness-to-intervention: A decade later. *Journal of Learning Disabilities*, *45*(3), 195–203. https://doi.org/10.1177/0022219412442150

Gagné, R. M. (1977). *The conditions of learning*. Chicago, IL: Holt, Rinehart & Winston.

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

Gandhi, A. G., Holdheide, L., Zumeta, R., & Danielson, L. (2016, February). *Understanding and operationalizing evidence-based practices within multi-tiered systems of support*. Paper presented at the annual Pacific Research Coast Conference, San Diego, CA.

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.

Gettinger, M., & Seibert, J. K. (2002). Best practices in increasing academic learning time. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology IV* (Vol. 1, pp. 773–787). Bethesda, MD: National Association of School Psychologists.

Gilbert, J. K., Compton, D. L., Fuchs, D., Fuchs, L. S., Bouton, B., Barquero, L. A., & Cho, E. (2013). Efficacy of a first-grade responsiveness-to-intervention prevention model for struggling readers. *Reading Research Quarterly*, *48*(2), 135–154. https://doi.org/10.1002/rrq.45

Harnischfeger, A., & Wiley, D. E. (1976). The teaching–learning process in elementary schools: A synoptic view. *Curriculum Inquiry*, *6*(1), 5–43.

Heafner, T. L., & Fitchett, P. G. (2015). An opportunity to learn US history: What NAEP data suggest regarding the opportunity gap. *The High School Journal*, *98*(3), 226–249. https://doi.org/10.1353/hsj.2015.0006

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.

Holdheide, L. (2016, February). *Tier 1 instructional practice: Mixed messages and missed opportunities*. Paper presented at the annual Pacific Research Coast Conference, San Diego, CA.

Husén, T. (1967). *International study of achievement in mathematics: A comparison of twelve countries*. New York, NY: John Wiley & Sons.

Jenkins, J. R., & Pany, D. (1978). Curriculum biases in reading achievement tests. *Journal of Reading Behavior*, *10*(4), 345–357.

Joyce, B., & Showers, B. (2002). *Student achievement through staff development*. Alexandria, VA: Association for Supervision and Curriculum Development.

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.

Ketterlin-Geller, L. R., & Jamgochian, E. M. (2011). Accommodations and modifications that support accessible instruction. 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, NY: Springer.

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

Kurz, A. (2016, February). *Measuring Opportunity to learn through a teacher log*. Paper presented at the annual Pacific Research Coast Conference, San Diego, CA.

Kurz, A. (2017, February). *Educational redemption and instructional coaching*. Paper presented at the annual Pacific Research Coast Conference, San Diego, CA.

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

Kurz, A., & Elliott, S. N. (2012). *MyiLOGS: My instructional learning opportunities guidance system (Version 2) [Software and training videos]*. Tempe, AZ: Arizona State University.

Kurz, A., Elliott, S. N., Kettler, R. J., & Yel, N. (2014). Assessing students' opportunity to learn the intended curriculum using an online teacher log: Initial validity evidence. *Educational Assessment*, *19*(3), 159–184. https://doi.org/10.1080/10627197.2014.934606

Kurz, A., Elliott, S. N., Lemons, C. J., Zigmond, N., Kloo, A., & Kettler, R. J. (2014). Assessing opportunity-to-learn for students with and without disabilities. *Assessment for Effective Intervention*, *40*(1), 24–39. https://doi.org/10.1177/1534508414522685

Kurz, A., Elliott, S. N., & Roach, A. T. (2015). Addressing the missing instructional data problem: Using a teacher log to document tier 1 instruction. *Remedial and Special Education*, *36*(6), 361–373. https://doi.org/10.1177/0741932514567365

Kurz, A., Reddy, L. A., & Glover, T. A. (2017). A multidisciplinary framework of instructional coaching. *Theory Into Practice*, *56*, 66–77. https://doi.org/10.1080/00405841.2016.1260404

Kurz, A., Reichenberg, R., & Yel, N. (2017). *Setting opportunity-to-learn standards for effective teaching*. Manuscript submitted for publication.

Marzano, R. J. (2000). *A new era of school reform: Going where the research takes us (REL no. #RJ96006101)*.

Aurora, CO: Mid-continent Research for Education and Learning.

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

Mayer, R. E. (2008). *Learning and instruction* (2nd ed.). Upper Saddle River, NJ: Pearson.

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

Metcalf, T. (2012). *What's your plan? Accurate decision making within a multi-tier system of supports: Critical areas in Tier 1*. Retrieved from http://www.rtinetwork. org/essential/tieredinstruction/tier1/accurate-decision-making-within-a-multi-tier-system-of-supports-critical-areas-in-tier-1

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

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.

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, IL: University of Chicago Press.

Porter, A. C., Schmidt, W. H., Floden, R. E., & Freeman, D. J. (1978). *Impact on what? The importance of content covered (Research Series No. 2)*. East Lansing, MI: Michigan State University, Institute for Research on Teaching.

Reddy, L. A., Fabiano, G. A., & Jimerson, S. R. (2013). Assessment of general education teachers' Tier 1 classroom practices: Contemporary science, practice, and policy. *School Psychology Quarterly*, *28*(4), 273–276. https://doi.org/10.1037/spq0000047

Roach, A. T., Kurz, A., & Elliott, S. N. (2015). Facilitating opportunity to learn for students with disabilities with instructional feedback data. *Preventing School Failure*, *59*(3), 168–178. https://doi.org/10.1080/1045988X.2014.901288

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

Rowan, B., & Correnti, R. (2009). Studying reading instruction with teacher logs: Lessons from the study of instructional improvement. *Educational Researcher*, *38*(2), 120–131.

RTI Action Network, http://www.rtinetwork.org.

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

Schmidt, W. H., & Burroughs, N. A. (2013). Opening the black box: Prospects for using international large-scale assessments to explore classroom effects. *Research in Comparative and International Education*, *8*(3), 236–212. https://doi.org/10.2304/rcie.2013.8.3.236

Schmidt, W. H., Burroughs, N. A., Zoido, P., & Houang, R. T. (2015). The role of schooling in perpetuating educational inequality: An international perspective. *Educational Researcher*, *44*(7), 371–386. https://doi.org/10.3102/0013189X15603982

Schmidt, W. H., McKnight, C. C., Houang, R. T., Wang, H. A., Wiley, D. E., Cogan, L. S., & Wolfe, R. G. (2001). *Why Schools Matter: a cross-national comparison of curriculum and learning*. San Francisco, CA: Jossey-Bass.

Schmidt, W. H., McKnight, C. C., Valverde, G. A., Houang, R. T., & Wiley, D. E. (1997). *Many visions, many aims volume 1: A cross- national investigation of curricular intentions in school mathematics*. Dordrecht, Netherlands: Kluwer Academic.

Torgesen, J. K. (2009). The response to intervention instructional model: Some outcomes from a large-scale implementation in Reading First schools. *Child Development Perspectives*, *3*, 38–40.

VanDerHeyden, A. M., Witt, J. C., & Gilbertson, D. (2007). A multi-year evaluation of the effects of a response to intervention (RTI) model on identification of children for special education. *Journal of School Psychology*, *45*, 225–256.

Vannest, K. J., & Parker, R. I. (2010). Measuring time: The stability of special education teacher time use. *Journal of Special Education*, *44*(2), 94–106.

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., Wanzek, J., Murray, C. S., Scammacca, N., Linan-Thompson, S., & Woodruff, A. L. (2009). Response to early reading interventions: Examining higher responders and lower responders. *Exceptional Children*, *75*, 165–183.

Walberg, H. J. (1980). A psychological theory of educational productivity. In F. H. Farley & N. Gordon (Eds.), *Psychology and education* (pp. 81–110). Berkeley, CA: McCutchan.

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, NY: 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.

Wanzek, J., & Vaughn, S. (2010). Is a three-tier reading intervention model associated with reduced placement in special education? *Remedial and Special Education*, *32*(2), 167–175.

Webb, N. L. (1997). *Criteria for alignment of expectations and assessments in mathematics and science education (NISE Research Monograph No. 6)*. Madison, WI: University of Wisconsin-Madison, National Institute for Science Education.

Webb, N. L. (2006). Identifying content for student achievement tests. In S. M. Downing & T. M. Haladyna (Eds.), *Handbook of test development* (pp. 155–180). Mahwah, NJ: Lawrence Erlbaum.

Winfield, L. F. (1993). Investigating test content and curriculum content overlap to assess opportunity to learn. *The Journal of Negro Education*, *62*(3), 288–310.