Chapter 7
Commentary on Part II

Leanne R. Ketterlin-Geller

Abstract The purpose of this commentary is to summarize the information presented by Hunt and Tzur in Chap. 5 and Zhang, Maher, and Wilkinson in Chap. 6, identify common themes that bring these chapters together, and offer my reflection on these themes. Both chapters discuss ways in which assessment results can support decision-making for students who are experiencing difficulties in mathematics. Emerging from these chapters is the notion that when intentionally designed, assessment results can help teachers align instruction with students' learning needs to improve and accelerate student learning. Three themes that I build on are (a) the impact of teachers' attitudes toward students who are experiencing difficulty, (b) the need to carefully design assessments and tasks, and (c) the role of assessments in providing teachers with diagnostic information. Within each of these themes, I summarize the commonalities across chapters and extend the discussion to highlight additional perspectives or approaches.

Keywords Assessment · Students experiencing difficulty · Validity · Teacher decision-making · Item design · Asset-based assessment · Curriculum–instruction–assessment alignment · Fairness · Teachers' attitudes · Accommodations · Interview-based assessment · Diagnostic assessment · Learning progressions

Research conducted for this manuscript is based upon work supported by the National Science Foundation under Grant No. 1721100. Any opinions, findings, and conclusions or recommendations expressed in these materials are those of the authors and do not necessarily reflect the views of the National Science Foundation.

L. R. Ketterlin-Geller (✉)
Southern Methodist University, Dallas, TX, USA
e-mail: lkgeller@smu.edu

Y. P. Xin et al. (eds.), *Enabling Mathematics Learning of Struggling Students*, Research in Mathematics Education,
https://doi.org/10.1007/978-3-030-95216-7_7

7.1 Designing and Implementing Assessments to Support Learning for Students Who Are Experiencing Difficulties in Learning Mathematics

In this commentary, I reflect on the information presented by Hunt and Tzur in Chap. 5 and Zhang, Maher, and Wilkinson in Chap. 6. The overarching theme that ties these chapters together is the way in which assessment results can support decision-making for students who are experiencing difficulties in learning mathematics. In the chapter by Hunt and Tzur, the authors emphasize the role tasks play in connecting teaching and learning: Tasks are used to engage students' reflective thinking during the learning process, and teachers use prompting and gesturing to deepen students' reflection. In the chapter by Zhang, Maher, and Wilkinson, the authors broaden the discussion of assessments and focus on three major topics: (1) participation of students who are experiencing mathematics difficulties in traditional assessments for accountability and formative purposes, (2) dimensions of students' thinking that are assessed, and (3) diagnostic inferences to support instructional decisions. In this commentary, I provide a brief overview of the central points highlighted in these chapters. Then, I offer my observations about themes that cut across both chapters, calling attention to (a) the impact of teachers' attitudes toward students who are experiencing difficulty, (b) the need to carefully design assessments and tasks, and (c) the role of assessments in providing teachers with diagnostic information. Within each of these themes, I summarize the commonalities across chapters and extend the discussion to highlight additional perspectives or approaches.

To help situate my commentary, it is important to clarify what I mean when I reference students who are experiencing difficulties in learning mathematics. It is difficult to put bounds on the group of students for whom I am referring. To begin, I am not talking about students who may have temporary challenges understanding a new mathematical concept, applying a newly learned procedure to solve a problem, or providing a mathematically sound rationale to justify their procedures. Mathematics is a complex discipline for which many people will experience temporary difficulties or challenges at one time or another. Instead, I am referring to students who have persistent difficulties in learning or applying mathematical knowledge and skills, especially as concepts increase in complexity within and across grades. This population of students is heterogeneous; they have various strengths and learning needs. A small percentage of these students have been diagnosed with learning disabilities. Students with a diagnosed learning disability are receiving specialized instruction as outlined on their Individual Education Program (IEP), which may or may not include instructional goals related to mathematics.

For many students who experience difficulty, the underlying cause is not a diagnosed learning disability, but instead may emerge as a result of a misalignment between instruction and their learning needs (Tzur, 2013). The implications of this misalignment can be far-reaching for students. In a series of studies conducted by Jordan et al. (2007, 2009), they observed that over 50% of the children from low-income families had low or flat growth trajectories from kindergarten through grade

1. Many of the children entered kindergarten with less developed number sense than their middle- and high-income peers, and formal instruction did not accelerate their learning during this period. In their follow-up study, the authors noted that the low mathematics performance of the children from low-income families in grade 3 was mediated by their relatively weak number sense in kindergarten. In other words, formal mathematics instruction between kindergarten and grade 3 did not result in growth or meaningful learning. To explain these findings, the authors posit a mismatch between the instruction the children received and the targeted support they actually needed to reach the learning goals and/or the curricular expectations. Had the children who had less developed number sense received appropriately intensive supplemental instruction, they may have developed more robust early mathematics knowledge and skills and experienced steeper growth trajectories that put them on course for future success.

The misalignment between instruction and students' learning needs may be caused by multiple reasons. Common causes I have seen include instruction that does not activate or draw on students' prior knowledge or conceptualizations, instruction that is not sequenced and paced in accordance with students' current level of understanding, and instruction that does not address previous misconceptions or errors, thereby allowing them to linger. When instruction is not aligned with students' learning needs, their opportunities to learn mathematics are compromised. I explicitly raise these issues because they directly relate to the chapters by Hunt and Tzur and Zhang, Maher, and Wilkinson: When designed with intentionality, assessment results can help teachers align instruction with students' learning needs to improve and accelerate student learning. After describing the unique contributions made by each chapter, I highlight areas of convergence and extensions to consider.

7.2 Assessment Considerations Presented by the Authors

In the chapter by Hunt and Tzur (Chap. 5, this volume), the authors make an explicit connection between teaching and learning and the conditions that facilitate this connection. They define learning as changes in a student's conceptualization of the content via reflection on the effects of their activities and explain how teachers' instructional moves can directly impact this process.

Hunt and Tzur define teaching in terms of instructional moves teachers make to elicit a student's current ways of knowing or understanding and then facilitate reflection during tasks to optimize the student's learning. When a student is in a participatory stage of understanding, the teacher may need to prompt reflection to support the student's progress to the anticipatory stage of understanding. Once a student moves into the anticipatory stage, the teachers' actions may shift to focus on deepening or extending reflection to subsequently deepen understanding.

Hunt and Tzur situate instructional tasks at the juncture of teaching and learning, while explicating the role tasks play in deepening understanding. Tasks connect teaching and learning through the teacher's intentional placement in the learning process and the teacher–student interactions that are facilitated during an instructional exchange. Hunt and Tzur dissect the various types of tasks and the level of interaction that encourages student reflection. In the procedures these authors describe, the teacher presents a student with tasks that are designed to bring forward the student's prior knowledge and available conceptualizations (labeled as "bridging tasks") from which they can build new conceptual understanding. Next, with student's prior knowledge primed, they engage in solving gradually more challenging tasks that include meaningful "variations" in the task design, but still focus on the key learning goal. Finally, deeper understanding is reached as the student moves into the "reinstating" phase where student learning is reinforced through reflection. Throughout this process, the teacher is interacting with the student to support their learning. The authors note the importance of the teacher's prompting and gesturing as a way of assessing student understanding and encouraging the student to clarify, justify, and critique their thinking and to deepen the student's reflection. I illustrate the connection between these components in Fig. 7.1.

Shifting to Chap. 6 by Zhang, Maher, and Wilkinson (this volume), these authors take on three major topics related to assessment. First, the authors provide an overview of various assessment approaches with an emphasis on how students with disabilities or students experiencing difficulties in mathematics participate in these programs. The authors reference the uses of large-scale assessments and curriculum-based measures (CBM) within accountability and instructional contexts, respectively. They provide examples of specific assessments and illustrate their strengths and limitations. In particular, they raise concerns about the content of many standardized mathematics assessments, noting that they may not be representative of the taught curriculum.

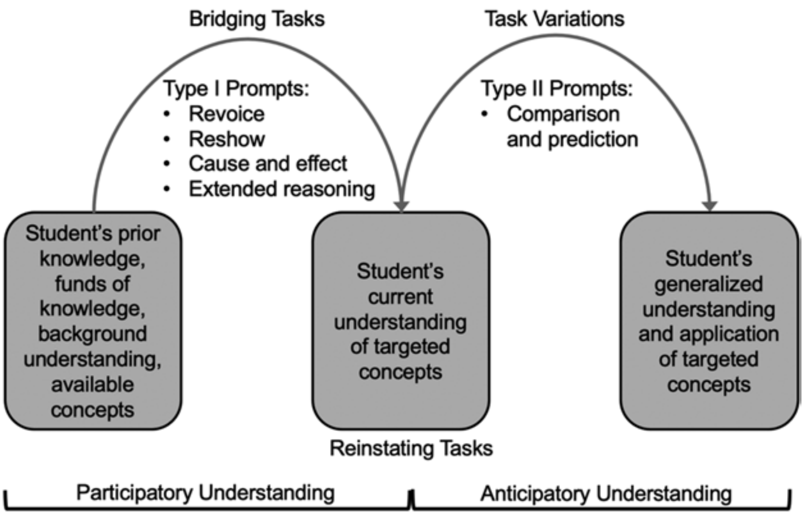


Fig. 7.1 Visual depiction of the components presented in Hunt and Tzur

To ensure that results from these assessment systems are reliable and support valid decision-making for students with disabilities or students experiencing difficulties in mathematics, Zhang, Maher, and Wilkinson discuss the provision of test accommodations and provide an important distinction from test modifications. The authors reference the lack of empirical evidence documenting the effectiveness of accommodations at supporting access and note that applying accommodations uniformly may not support greater access for students experiencing difficulty. Relatedly, when describing the role of language in communicating one's mathematical knowledge, the authors describe how language may cause an access barrier for some students, which may lead to teachers' underestimating their students' level of mathematical proficiency.

Next, Zhang, Maher, and Wilkinson consider the mathematical construct being measured and the value different constructs bring to understanding student thinking. The authors associate specific mathematical constructs with the application of theories of learning to mathematics education. For example, the authors reference the role of behaviorism in promoting instruction in fact fluency and retrieval along with the corresponding assessments designed to measure students' accuracy and computational fluency. The authors reference intentionally designing tasks to elicit students' reasoning and justification. They describe the importance of teachers attending to students' responding behaviors (e.g., how students represent and communicate their understanding) to deepen their understanding of students' mathematical thinking.

Lastly, Zhang, Maher, and Wilkinson describe the role of diagnostic assessments within the instructional design process. They provide an overview of some published diagnostic assessments and review several methods for gathering diagnostic information including clinical interviews, error pattern, and strategy analysis and integrating a microgenetic approach to strategy analysis. Through each of these approaches, the authors describe how information about students' thinking can be revealed to support future instruction and support student learning.

7.3 Convergence Across Chapters

Several themes emerged across the chapters by Hunt and Tzur and Zhang, Maher, and Wilkinson. First, both chapters emphasize the negative attribution teachers often place on students who are experiencing mathematics difficulty. Hunt and Tzur note how teachers' perceptions of their students' abilities may influence their instructional actions. Zhang, Maher, and Wilkinson similarly emphasize students' learned helplessness that sometimes accompanies low expectations. Second, both chapters describe the intentionality of assessment (or task) design as an essential component of using assessment results. Hunt and Tzur provide a detailed description of the types of tasks and accompanying prompting and gesturing that can be integrated into different phases of learning. Zhang, Maher, and Wilkinson note the important role accommodations play in administering assessments to overcome task design issues that may impact accessibility. Moreover, the authors of both chapters emphasize the importance of designing tasks to assess mathematical reasoning and justification. Third, and finally, both chapters emphasize the importance of implementing assessments and tasks to facilitate teachers' diagnostic inferences. In the remainder of this commentary, I will elaborate on each of these themes by offering some additional perspectives and extending the discussion in complementary ways.

7.3.1 Teachers' Attitudes Towards Students Experiencing Difficulties in Learning Mathematics

Highlighted in the chapters by Hunt and Tzur and Zhang, Maher, and Wilkinson are the consequences to student learning when teachers perceive students experiencing difficulties in learning mathematics as less capable of learning (e.g., deficit view). Hunt and Tzur describe how teachers' perceptions of their students' abilities may influence their instructional actions. They make a compelling case that teachers who perceive their students as lacking mathematical abilities will likely design instruction that sustains a deficit view. Similarly, Zhang, Maher, and Wilkinson recognize the role that standardized tests play in shaping teachers' perceptions of students' knowledge, skills, and abilities. They note that students may be judged as incapable of learning or applying advanced mathematics content based on their low scores.

I would like to broaden this discussion to focus on the systems in which teachers and students engage and highlight the role a school's culture plays in shaping teachers' attitudes and ultimately their actions and/or beliefs. A school's culture can have a significant impact on how students who are experiencing difficulties in learning mathematics are perceived and supported. A school's culture can be thought of as the shared values, norms, and beliefs that explicitly or implicitly guide the actions and decisions of teachers, school leaders, and the broader school community (Kilgore & Reynolds, 2011). A school's culture impacts the ways in which students experiencing difficulties in mathematics are perceived and the perception of whose responsibility it is to support their learning. As evidenced by their actions and decisions, some schools may espouse underlying beliefs about these students' mathematical abilities. For instance, these students may be thought of as "lacking" or "deficient" in their mathematical knowledge. In turn, those deficiencies may be perceived to be associated with disadvantages that are rooted in cultural, social, economic, or political differences from the normative group in the school (Healy & Powell, 2013).

Schools with a deficit view of students often hold the perception that low performance on standardized tests is caused by problems inherent to the student. Such problems may include limited parental support or involvement, insufficient preparation from their home environment or previous schooling, underlying lack of motivation, and students' insufficient background knowledge or even capacity to learn. In some instances, this deficit view permeates so deeply that it becomes a self-fulfilling prophecy. Beliefs that students have deficits or deficiencies that limit their ability to learn mathematics may fuel underlying biases that impact teachers' and/or school-level actions and decisions that harm student learning. An outcome of this perspective is where the blame for poor performance lies: In a deficit model, the blame lies with the student, their parents, or possible background or demographic characteristics, but not with teaching that takes place within the school or educational system through which the poor performance originated.

Conversely, an asset-based approach to supporting students assumes that all students have both knowledge and the capacity to learn—and deserve to learn. Accordingly, the school and related stakeholders are responsible for providing appropriately designed instruction (including assessment) that can address each student's needs. Schools in which the perception of students is grounded in an asset model believe that students' backgrounds provide a strong basis from which they can connect new knowledge and view families and communities as mathematically rich resources that can be harnessed to engage students. Actions that characterize teachers with an asset orientation include designing instruction to build new knowledge based on what students know, recognizing and honoring students' reasoning and sense making, and extending what they recognize and value as mathematical practices (Healy & Powell, 2013). At the system level, schools can examine how they incorporate families and communities into their practices, ways in which current operations may alienate people, and what additional resources are needed to engage the student, their families, and the community. Based on the actions and beliefs that are espoused and enacted, the school is establishing their culture. Each school has the authority and responsibility to establish a positive culture that values all students and views all students as capable of learning.

Intertwined with schools' and teachers' perceptions of students who are experiencing difficulties in learning mathematics is the notion of struggle. In mathematics education, this term is often used to denote how students wrestle with and persevere through problem-solving as they apply their mathematical knowledge. The field of mathematics education has coined the term "productive struggle" to indicate the experiences students face as they tackle challenging mathematical content, integrate and apply their knowledge, and persevere through finding a solution. This type of struggle is deemed productive because the content is accessible yet also challenging, and the problem-solving experience generally results in positive cognitive and/or non-cognitive outcomes (e.g., persistence, grit). However, there is an opposite experience had by some students: unproductive struggle. When students are unproductively struggling, it is often because they do not yet have an entry point for approaching the content or problem, they have not yet gained sufficient knowledge to apply to the problem situation, or they have persistent errors or naive conceptions that impact their current thinking about the problem space. Instead of resulting in positive cognitive and/or non-cognitive outcomes, the result of unproductive struggle is often negative (e.g., unwillingness to persevere, diminished feelings of self-worth). At this point, and with repeated experiences with unproductive struggle, students may lose interest in mathematics, perceive themselves as not successful in mathematics, and cement a negative identity. Students with a negative mathematical identity who define themselves as “not a math person” may act in ways that perpetuate this self-impression, such as deferring to others who are perceived of as more mathematically capable, engaging with mathematics in low-level ways, and talking and acting with peers in a way that propagates this identity (Bishop, 2012). As such, without intervention to change the underlying causes of unproductive struggle, the lasting consequences could be dire.

Recognizing that some students experience unproductive struggle sometimes and with some mathematics content does not mean that they will continuously and persistently struggle with mathematics. Moreover, saying that someone is struggling is not akin to saying that they are failing or incapable of learning mathematics. Instead, it recognizes that all students learn differently, and some students may need additional instructional support to build and demonstrate understanding in specific mathematics concepts or procedures.

In summary, both chapters recognize the negative consequences to student learning when teachers perceive students experiencing difficulties in learning mathematics from a deficit view. I extend this discussion to recognize the role a school’s culture plays in shaping teachers’ attitudes, actions, and beliefs. I note that a deficit view of students may cause a misalignment between students’ learning needs and their instructional opportunities. By not providing access to proximally relevant instruction and tasks, students miss the opportunity to build on their prior knowledge as a basis to deepen their understanding of mathematics concepts and may develop poor dispositions toward mathematics.

7.3.2 Intentionally Designed Assessments and Tasks

Another theme that emerged in the chapters by Hunt and Tzur and Zhang, Maher, and Wilkinson is the importance of intentionally designing assessments or tasks. Hunt and Tzur provide detailed guidance on the types of tasks and prompts that may facilitate learning across different stages of students’ understanding. Zhang, Maher, and Wilkinson approach the topic from two perspectives: (1) the importance of providing accommodations to improve accessibility and (2) the need to extend the range of constructs that are measured so as to cover the breadth of mathematical practices. In this section, I will expand upon the authors’ discussions by first emphasizing the importance of alignment and then underscoring the importance of fairness and equity in assessment or task design.

Alignment of Curriculum, Instruction, and Assessment: Implications for Assessment and Task Design In 1999, the National Research Council (Bransford et al., 1999) recognized the alignment between curriculum, instruction, and assessment as a key issue impacting students’ opportunities to learn. Within an aligned system, students have access to important curricular expectations during instruction, and their progress toward reaching these expectations can be monitored through an integrated and ongoing process of formative and summative assessment.

All students are provided with instructional opportunities that facilitate their learning of the curricular expectations. Assessment results are used to guide teachers' decisions about how best to facilitate learning.

For assessment results to be relevant and informative for supporting learning, formative and summative assessment practices must be clearly and carefully aligned with curriculum and instruction along multiple dimensions. Such an alignment includes the content domains, levels of cognitive engagement, and strategies and thinking processes through which students interact with the content (Ketterlin-Geller, 2016). No single test can sample the full range of curricular expectations; data combined from multiple tests should provide a comprehensive picture of student proficiency.

Within an aligned system, the curricular expectations should form a foundation of all instruction and assessment efforts. The curricular expectations to which I am referring are the state or national content standards in mathematics. In the United States, these content standards are often informed by research on how children learn mathematics (National Research Council [NRC], 2001) and may be aligned with the *Common Core State Standards in Mathematics* (CCSS-M; National Governors Association & Council of Chief State School Officers, 2010). A key shift in the CCSS-M that was informed by research in mathematics education (National Mathematics Advisory Panel [NMAP], 2008; NRC, 2001) was the focus on a balanced approach to integrating ways of knowing mathematics. These ways of knowing, emphasized in the CCSS-M, include conceptual understanding, procedural fluency, and application through problem-solving. In *Adding It Up,* published by the NRC (2001), additional specification was provided to call out strategic competence, adaptive reasoning, and productive disposition as important dimensions of mathematical proficiency (these dimensions of mathematical proficiency are defined elsewhere in this volume). In addition to these dimensions of knowing, mathematical practices were embedded in the CCSS-M and in most states' content standards. Mathematical practices represent ways in which students engage with the mathematical content through solving meaningful real-world problems, reasoning abstractly and quantitatively, constructing viable arguments about mathematical conjectures, modeling, appropriately using mathematical tools, engaging with precision, discerning patterns and structures in mathematics, and searching for regularity when solving like problems.

I intentionally define and call out these curricular expectations for several reasons. First, the definition of mathematical proficiency just described represents an integrated understanding of what mathematical competency means and looks like, and by extension, how we can evaluate performance. Hunt and Tzur contend that mathematical proficiency and performance are separate constructs that lead to different interpretations of students' knowing and thinking in mathematics. However, in an aligned system, formative and summative assessment practices lead to performance that directly indicates and/or illustrates the depth and breadth of students' proficiency across these dimensions of mathematics. When assessments are not aligned to the curricular expectations, performance may not be indicative of mathematics proficiency. I recommend we focus on building aligned systems of instruction and assessment to facilitate teachers' decision-making and, ultimately, students' development of mathematical proficiency.

Second, as the guidepost for an aligned system, the curricular expectations should be at the forefront of all instructional opportunities and assessment practices. It is essential that all students—whether or not they are experiencing difficulties in learning mathematics—have opportunities to learn the depth and breadth of the curricular expectations and demonstrate their understanding. Zhang, Maher, and Wilkinson pointed out an over-emphasis on procedural fluency during instruction and on assessments for students experiencing difficulties in learning mathematics. I contend that if instruction does not emphasize the range of learning expectations specified in the content standards, then the curriculum is narrowed, and students are denied the opportunity to learn essential mathematics content that has been deemed important and necessary for them to learn in each grade. A growing number of interventions and intervention frameworks are available that effectively represent the depth and breadth of the curricular expectations (c.f., Fuchs et al., 2017; Powell et al., 2020; Zhang et al., 2014). An exception to alignment with grade-level content standards may exist for students with disabilities whose IEPs specify that they receive a modified curriculum and are assessed based on alternate content and performance standards. For these students, the IEP team determined that the most suitable instruction and assessment should focus on modified content standards.

Relatedly, for assessments and tasks that are intended to inform teachers' decision-making about students' development of mathematical proficiency across the range of curricular expectations, the content should be representative of the depth and breadth of the content standards. In other words, items or tasks should align with the content and the strand(s) of mathematical proficiency that is intended by the standard. For example, consider the Common Core State Standard in Mathematics (2010):

> 7.EE.4. Use variables to represent quantities in a real-world or mathematical problem, and construct simple equations and inequalities to solve problems by reasoning about the quantities.

This standard focuses on expressions and equations, in particular the use of variables. Students engage their conceptual understanding of variables to represent problem situations, procedural knowledge to solve problems, and strategic competence to construct equations and inequalities, and quantitative reasoning to consider the real-world outcomes. To adequately assess this standard, the range of proficiency should be elicited and will require multiple items. If, instead, this standard was only assessed with items that elicited students' procedural fluency, interpretations about students' level of understanding and reasoning related to this standard may not be valid.

To support some decisions, teachers may need data about specific aspects of students' mathematical proficiency that are inclusive but not exhaustive of the curricular expectations. For example, Zhang, Maher, and Wilkinson referenced the need to understand students' reasoning to make decisions about their abilities to develop arguments and justify their solution frameworks when solving problems. An assessment aligned with this purpose may not include items that assess other aspects of mathematical proficiency such as conceptual understanding, procedural fluency, or strategic competence. Similarly, as described by Zhang, Maher, and Wilkinson, CBMs in mathematics often focus on computations. Results from these assessments can provide teachers with valuable insights into students' procedural fluency and can guide instructional decisions relevant to the accuracy and efficiency in which students execute procedures, but not their conceptual understanding, strategic competence, or reasoning. In both of these examples, the intended construct of the tests is narrower than the range of proficiencies specified in the curricular expectations. Alone, results from these tests do not provide a comprehensive view of students' mathematical proficiency. However, because the decisions resulting from these assessments are specific to the assessed content (and, therefore, not the depth and breadth of the curricular expectations), these assessments may continue to have value for teachers. As I hope this discussion illustrates, it is important to keep an assessment's intended uses and interpretations in the forefront when examining alignment and evaluating validity.

Various item or task formats can adequately assess a range of students' knowledge, skills, and abilities. Although student interviews or individually administered tasks as described in the chapters by Hunt and Tzur and Zhang, Maher, and Wilkinson may provide direct evidence of student thinking, this item or task format may not be feasible within all situations due to the time needed for administration and interpretation. Selected response items (e.g., multiple choice, true–false, matching) may aid in efficiency of administration and scoring, while still assessing multiple dimensions of student proficiency. For example, well-designed multiple-choice items can assess students' conceptual understanding, reasoning, and other higher-order thinking skills (Downing, 2006; Haladyna, 2004; Schneider et al., 2013) and, when intentionally designed to do so, may provide insights into the students' thinking based on their selection of distractors (Briggs et al., 2006; Luecht, 2007).

Third and lastly, it is important to remind ourselves that we—public school teachers, administrators, curriculum developers, assessment designers, and education researchers—are facilitators of the learning process and should engage in activities that support student learning of the depth and breadth of the curricular expectations. We are responsible for providing each student with carefully constructed opportunities to meet these expectations. In most instances, a state board of education or a national ministry of education has decided what students are required to know and be able to do; we cannot and should not change these expectations. As such, our efforts to facilitate mathematical proficiency should be in service of the ratified content standards. A teacher may have concerns about the nature of their state or country's content standards, particularly as they represent ways of knowing of the dominant culture and themselves are socially, politically, and historically bound. These concerns should be communicated with the school and district leadership, and/or state or national legislators. However, while these content standards remain the legislated mandate, it is imperative that educators provide each and every student with opportunities to learn the breadth and depth of these curricular expectations.

Designing Assessments and Tasks to Promote Fairness, Accessibility, and Validity Another consideration when intentionally designing assessments and tasks is fairness and accessibility, because they directly impact the validity of the intended uses and interpretations of assessment and task results. Zhang, Maher, and Wilkinson emphasize the role accommodations play in supporting access to the assessed construct for students with disabilities. They also note that standardized accommodations may not live up to this goal because the variety of students' needs that may not be well matched with standardized accommodations. I would like to build on this discussion by recentering the purpose of accommodations within the perspectives of fairness and validity.

To support fair and valid decision-making, assessments must be free from bias, minimize the impact of sources of construct-irrelevant variance, and be based on statistical analyses (e.g., item calibrations, percentile and cutoff scores, validity evidence) that were obtained from a representative sample of students. Items, and the assembled test, should yield results that can fairly inform decision-making regardless of the examinee's race or ethnicity, gender, socioeconomic status, educational classification, or any other demographic variable. To verify comparability, sufficient evidence is needed that documents the absence of differential item functioning (DIF) and differential classification outcomes.

Regardless of the type of assessment, some students with disabilities will need testing accommodations to accurately demonstrate their knowledge, skills, and abilities. Zhang, Maher, and Wilkinson introduced the purpose of testing accommodations and provided an important distinction from test modifications. To further elaborate, testing accommodations are designed to mitigate the impact of students' personal characteristics that negatively interact with item design features and subsequently lead to construct-irrelevant variance (CIV) in students' scores. As an example of this interaction, consider a student who may have difficulty retaining information in short-term memory. This student may experience additional difficulty when solving multi-step word problems that require the student to use information temporarily stored in working memory to solve subsequent components of the problem. In these instances, the student may benefit from using a graphic organizer as an accommodation to record intermediary steps and solutions within the problem. However, when the same student is responding to items that elicit conceptual understanding or application of single-step procedures that do not require extensive use of working memory, the student may not encounter additional challenges that result in CIV. In those instances, the student may not benefit from the accommodation. Comparatively, a student who has difficulty decoding the text used in multi-step word problems may benefit from a reading-based accommodation (e.g., read aloud, simplified language) and may not benefit from using a graphic organizer.

The primary purpose of administering tests is to capture variability in student performance that can be attributable to variability in construct-relevant knowledge, skills, and abilities. However, the examples above illustrate how students' personal characteristics influence how they engage with test items in construct-irrelevant ways. Moreover, the examples also highlight how variations in item design features may support or hinder students' engagement depending upon their personal characteristics. It is because of this interaction that the evidence about the effectiveness of accommodations is inconclusive. I have long argued that accommodations should be assigned at the item level where these interactions impact students' ability to accurately demonstrate their construct-relevant knowledge, skills, and abilities (c.f., Ketterlin-Geller, 2008; Ketterlin-Geller, 2016). Accommodations can be effective only when they are applied at the juncture where CIV occurs.

As this discussion illustrates and is noted by Zhang, Maher, and Wilkinson, assigning test accommodations should not be conceptualized as a one-size-fits-all approach. IEP teams need to consider the student's personal characteristics and item and test design features to determine the probable sources of CIV. IEP teams can then assign allowable accommodations that may mitigate the negative interaction. Our previous research (c.f., Ketterlin-Geller et al., 2014) provides guidance on this process.

I want to explicitly note that accommodations are not intended to promote a deficit view of students with disabilities. Differences in students' cognitive processing, attention, language or linguistic processing, and physical characteristics are not deficiencies and may promote mathematical sense making in other ways (Healy & Powell, 2013). Insomuch as the test design features can support accurate elicitation of students' mathematical knowledge and skills, there may be little or no CIV introduced into students' scores. However, in instances where the test design features may hinder students' ability to demonstrate their mathematical understanding, accommodations will be needed to provide equitable opportunities for students with disabilities to demonstrate their mathematical sense making. Alternatively, tests can be designed to minimize the reliance on construct-irrelevant skills by applying the principles of universal design (Ketterlin-Geller, 2008; Ketterlin-Geller, 2016). For example, reading is often a leading cause of CIV on mathematics tests for many students with and without disabilities. Tests can support variability in students' reading skills by providing options to have the test items read aloud to anyone. Although it is not possible to mitigate all sources of CIV through intentional test design, universally designed tests may reduce the need for externally applied accommodations.

In summary, building on the theme introduced by Hunt and Tzur and Zhang, Maher, and Wilkinson, it is important to carefully consider the design of assessments and tasks when using data to guide instructional decisions for students who are experiencing difficulties in learning mathematics. Valid uses and interpretations of results are incumbent on teachers' access to data that are aligned with curricular expectations and instructional opportunities and that accurately reflect students' knowledge, skills, and abilities without bias or CIV.

7.3.3 Assessments and Tasks to Support Diagnostic Inferences

Quite often, mathematics classrooms do not comprise a homogeneous grouping of students. Instead, many of today's classrooms are characterized by considerable heterogeneity across a variety of factors, including students' background knowledge or experiences with formal and informal mathematics, prior conceptualizations of mathematical concepts, and facility with different mathematical representations. As a result, students within a typical mathematics class may benefit from different approaches to instruction, such as representing mathematical concepts in multiple ways, altered amounts of practice opportunities, and different levels of scaffolding. Students may also need varying levels of supplemental instructional support to reach their goals, with some benefiting from minimal additional support and others needing considerably more to meet the curricular expectations. Some schools use a system-level framework (e.g., multi-tiered system of support) to coordinate instructional delivery, while other schools use a more diffuse model. Within either approach, diagnostic data may be needed to efficiently and effectively design supplemental instructional support to address these students' learning needs.

Diagnostic assessment practices form an important part of an integrated and aligned system of assessments. As described in Chaps. 5 and 6 of this volume, teachers often seek diagnostic information to generate hypotheses about the nature of the instructional support from which a student who is experiencing difficulties in mathematics may benefit. These hypotheses can be based on information such as the student's understanding of specific mathematical concepts or skills, persistent errors, or naive conceptions the student makes while completing their work, the strategies the student uses to arrive at a solution and/or the efficiency with which the student executes the strategy, and/or the representations they use to demonstrate understanding. This information recognizes the student's prior knowledge and their ways of knowing and representing mathematics—and helps identify areas in which additional instruction could support future learning. With this information, the teacher can design a supplemental instructional plan that builds on the student's current understanding and facilitates the next steps in the student's learning.

There are various approaches to designing diagnostic assessments to inform these instructional hypotheses and, ultimately, guide the design of supplemental instructional opportunities. Zhang, Maher, and Wilkinson provide an overview of some published diagnostic assessments, as well as various approaches that can be applied without the use of standardized assessments, such as clinical interviews, error pattern and strategy analysis, and strategy analysis with a microgenetic approach. Hunt and Tzur describe an approach to gathering diagnostic information through interviews that scaffold tasks through the learning process. Each chapter documents the procedures of these approaches and offers examples in practice.

Another approach to gathering information that may be particularly useful for guiding the design of supplemental instructional opportunities for students who are experiencing difficulties in learning mathematics is the use of learning progression-based diagnostic assessments. In the remainder of this section, I provide an overview of learning progressions in mathematics, discuss ongoing development work to design diagnostic assessments based on learning progressions, and explain how these assessments may provide unique information that can accelerate the learning process.

Learning progressions (or learning trajectories) are theoretical models of learning that describe the development of sophistication in students' thinking within a discipline. While different disciplines tend to use one term over the other (e.g., science education tends toward learning progressions, while mathematics education often uses learning trajectories), in her exhaustive review, Confrey (2018) noted that there are minimal differences in the propositions underlying these terms. As such, I will use the term learning progressions with the intention of being inclusive of both learning progressions and learning trajectories.

Learning progressions illustrate a sequential progression of understanding within a discipline that leads from foundational knowledge to more advanced thinking (Bennett, 2015). Knowledge and skills are integrated together in a meaningful way to build deeper understanding. Inherently, learning progressions present an asset-based model of learning by framing students' development of understanding as continuing from their available conceptualizations through to more sophisticated and comprehensive understandings (Alonzo, 2018). Confrey (2018) likened this development to a climbing wall in which students deepen their understanding (move up the wall) by combining knowledge and skills (using each hand and foot hold) in a thoughtful, purposeful, and sometimes iterative way. As what happens with a climbing wall, the specific starting point and the pathway each child takes may look different, but the outcome is comparable (reaching the top of the wall, or completing a course).

When most learning progressions are described, they begin with a set of foundational skills that mark the entry point or lower boundary. Similarly, they have an upper boundary that identifies the targeted learning goals for that progression (Confrey, 2018; Corcoran et al., 2009). The lower and upper boundaries are connected by a series of intermediary phases of learning that are akin to the hand and foot holdings on the climbing wall mentioned earlier. These phases of learning make up the network of knowledge, skills, and processes that build in sophistication and complexity as they progress from the lower toward the upper bounds. Each student's pathway through this network may be different and will be guided by their prior experiences and exposure to the content, as well as their instructional opportunities.

The intermediary phases of learning can exist with varying levels of specificity from coarse-grained representations of concepts to fine-grained micro-conceptualizations of student thinking. The granularity in which the learning progression is specified depends on the learning outcomes. For example, the Common Core State Standards in Mathematics are based on the principles of learning progressions and are specified at a relatively coarse grain size. Conversely, the learning progression explicating the process by which students learn to divide fractions articulated in Ketterlin-Geller et al. (2013) is specified at a relatively fine-grained size.

The level of specificity has important implications for informing teaching and learning, notably the design of instruction and assessment. For the purposes of this commentary, I discuss the design of diagnostic assessments to inform supplemental instruction for students who are experiencing difficulties in learning mathematics.

Diagnostic assessments based on learning progressions provide teachers with information about the nature of students' learning, so as to directly inform instruction. Teachers are able to interpret student performance from a lens of students' knowing and understanding and move away from dichotomizing interpretations of performance as "they got it" or "they didn't get it" (Alonzo, 2018). For example, teachers can identify student understanding in relation to their foundational knowledge and skills (e.g., where they started on the climbing wall). This information can inform where instruction may intersect with students' prior conceptualizations. Relatedly, teachers can monitor students' movement through the intermediary phases of learning to determine the knowledge, skills, and processes students have developed and what instructional actions should come next to support learning. Also, teachers can examine fine-grained information within the intermediary phases of learning to better understand the students' conceptualizations, integration of prior learning, facility with mathematical representations, and other aspects of students' mathematical thinking. Understanding students' partial, emerging, or naive conceptions may help teachers identify prior knowledge that can be integrated with future learning to create a more complete representation of knowing in the domain (Confrey, 2018). This information can directly inform the design of instruction.

To illustrate the decisions teachers can make using learning progression-based classroom assessments, I briefly describe a learning progression that illustrates the knowledge, skills, and reasoning that students in kindergarten through grade 2 develop when reasoning about numeric relations. In the project *Measuring Early Mathematics Reasoning Skills* (NSF #1721100), we articulated a learning progression through an iterative and systematic process specified by Ketterlin-Geller et al. (2013). This process integrated evidence from (a) theoretical propositions underlying learning in the domains, (b) detailed reviews by nationally recognized mathematicians and mathematics educators, (c) in-depth analyses of students' thinking as elicited by cognitive interviews conducted with children in kindergarten through grade 2, and (d) input from teachers about their understanding of child development cultivated through extensive observations. The resulting learning progression for numeric relational reasoning has three targeted learning goals (relations, composition and decomposition, and properties of operations), each with three- or four-core concepts (see Kuehnert et al., 2020 for more information). The overall structure is displayed in Fig. 7.2.

Each core concept is further explicated into finer-grained subcomponents that describe the intermediary phases of learning, including the conceptualizations, reasoning, and strategic processes underlying learning.

To illustrate the way in which learning progressions can inform instructional decision-making, I will elaborate on the first core concept in relations, labeled "*Comparison of Unequal Numbers*." The intermediary phases of learning include the following:

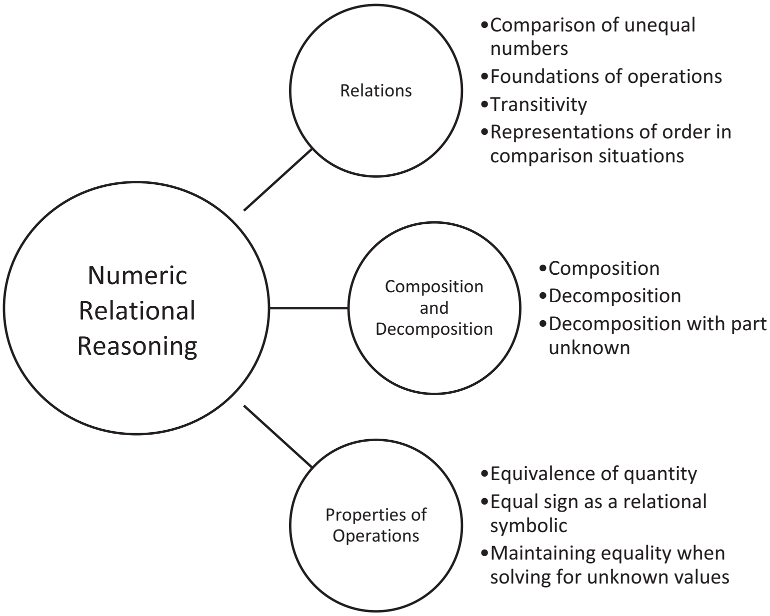
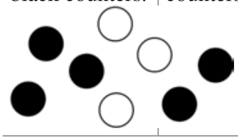


Fig. 7.2 Visual display of the learning progression for numeric relational reasoning

- Compare two quantities to find which has more or fewer items using one-to-one matching and counting strategies without a specific arrangement.
- Compare two quantities to find which has more or fewer items when given a specific arrangement.
- Compare two quantities grouped in tens and ones to find which has more or fewer items using one-to-one matching and counting strategies.
- Compare two quantities to find which has more or fewer items when given a specific arrangement with place value (e.g., place value blocks).
- Compare two numbers using the understanding of the number sequence to determine which is larger or smaller.
- Compare two numbers using written number lines to determine which is larger or smaller.
- Compare two numbers using open number lines to determine which is larger or smaller.
- Compare two numbers using symbolic notation.

These intermediary phases of learning occur iteratively as students build fluency and flexibility working with increasingly larger number ranges (e.g., from 1–5 to 1–19). Through each intermediary phase, students deepen their understanding and ability to apply number relations concepts. Moreover, they are introduced to the number line as an important tool for modeling mathematics. Ultimately, these intermediary phases culminate in a strong conceptual understanding of the number sequence, magnitude, place value, and the usefulness of communicating mathematical concepts through visual representations. These concepts lay the foundation for future work with operations.

Classroom assessments aligned to this learning progression can provide teachers with diagnostically useful information about students' conceptualizations, reasoning, and strategic processing at each intermediary phase as the student progresses to larger number ranges. For example, consider this sample item for the first subcomponent, *compare two quantities to find which has more or fewer items using one-to-one matching and counting strategies without a specific arrangement*. This item would be presented to a student who is working within the number range of 0–10 in an interview format with manipulatives (e.g., counters).

Stimulus	Prompt to elicit conceptualizations	Prompt to elicit reasoning	Prompt to extend students' thinking
Look at these white and black counters. 	Without touching these counters, are there more black counters or white counters?	Show me how you got your answer. You may touch the counters.	Show me how you would make the groups so there are more white counters.

Because the purpose of administering diagnostic assessments is to gather data to generate hypotheses about student learning to inform the design of supplemental instruction, students' responses to these types of questions do not need to be scored in the traditional sense (e.g., scored for correctness, evaluated against a rubric). Regardless of whether the assessment includes constructed or selected response items, students' responses provide windows into their thinking. Teachers can interpret students' responses in relation to the learning progression to identify the students' prior conceptualizations. Moreover, teachers may get a glimpse into the strategies students use when solving problems and the varied ways of knowing and making sense of mathematics.

Using the item above as an example, a student may be able to count the number of counters by color and be able to state which group has more counters only when allowed to rearrange the counters so as to employ a one-to-one matching strategy. In this instance, the teacher gains insights into the student's understanding of relationships between quantities and the strategies they use to understand these relationships. The teacher may infer that the student is progressing in their understanding of the number sequence, has a grasp of cardinality, and has a strategy for comparing quantities using physical manipulatives. In turn, the teacher can use this information to understand

- what content to teach next and to what level of intensity;
- student's conceptualizations underlying their responses, which may illuminate varying levels of sophistication and efficiency in the strategies; and
- student's mental models of mathematics concepts.

As these interpretations inform the teacher's inferences about the student's thinking, they serve to guide instructional decisions to support future learning.

In sum, within an integrated and aligned system of assessments, diagnostic assessments serve an important role in promoting equitable outcomes in mathematics classrooms by providing teachers with valuable information to support inferences about student learning. These inferences may include an understanding of students' current conceptualizations, persistent errors, or naive conceptions, strategies the student uses and/or the efficiency with which the student executes the strategy, and the representations the student uses to demonstrate understanding. From this information, teachers can formulate hypotheses about the nature of the instructional support from which students may benefit. In the example described above, the teacher may facilitate the student's future learning by emphasizing the meaning of the number sequence, extending the range of strategic approaches to comparing quantities (including introducing visual and abstract representations of quantities), and expanding the number range in which the student is making comparisons—all of which align with the intermediary phases of the learning progression. As this example illustrates, because learning progressions are based on theoretical models of learning, they provide a unique lens through which students' knowing and thinking can be interpreted to directly guide instructional design decisions.

7.4 Conclusions

Students who are experiencing difficulties in learning mathematics may be supported in a variety of classroom settings within a school. It is essential that the values, norms, and beliefs that define a school's culture recognize the assets all students bring to the learning environment and take actions that signify these beliefs. Importantly, all students should view themselves as valued contributors to the learning environment who are capable of learning complex mathematical concepts and have unique and important perspectives that shape their learning.

When intentionally designed, assessments can help teachers align instruction with students' learning needs to improve and accelerate student learning. Designing assessments to provide meaningful, trustworthy, and reliable data to inform classroom decisions requires careful consideration of several factors, such as the depth of understanding elicited by the items, content representation at the item and test levels, and the procedures for administration, scoring, and interpretation. In particular, although there are various approaches to designing diagnostic assessments, aligning the content with learning progressions may help teachers make inferences about students' knowing, which can be translated into actionable steps for designing learning opportunities.

On a final note, I want to acknowledge and thank the editors for inviting me to contribute this commentary. We often point to differences in the ways in which special education and mathematics education researchers approach supporting students who are experiencing difficulties in learning mathematics (c.f., Woodward &

Tzur, 2017). In this commentary, I have tried to emphasize the commonalities both between the chapters and also among the broader community of educators who are committed to improving outcomes for students who are experiencing difficulties. I urge us—the community of educators, researchers, and other stakeholders—to build on our shared goals and vision of all students realizing positive outcomes in mathematics. The more we can work toward understanding and away from emphasizing difference, the better equipped our community will be to support all students, especially those who are experiencing difficulties in learning mathematics.

References

Alonzo, A. C. (2018). An argument for formative assessment with science learning progressions. *Applied Measurement in Education, 31*(2), 104–112.

Bennett, R. E. (2015). The changing nature of educational assessment. *Review of Research in Education, 39*, 370–407. https://doi.org/10.3102/0091732X14554179

Bishop, J. P. (2012). "She's always been the smart one. I've always been the dumb one": Identities in the mathematics classroom. *Journal for Research in Mathematics Education, 43*(1), 34–74. https://doi.org/10.5951/jresematheduc.43.1.0034

Bransford, J., Brown, L. A., & Cocking, R. R. (1999). *How people learn: Brain, mind, experience, and school*. National Academy Press.

Briggs, D. C., Alonzo, A. C., Schwab, C., & Wilson, M. (2006). Diagnostic assessment with ordered multiple-choice items. *Educational Assessment, 11*(1), 33–63.

Confrey, J. (2018). *Future of education and skills 2030: Curriculum analysis – A synthesis of research on learning trajectories/progressions in mathematics* (EDU/EDPC[2018]44/ANN3). Organisation for Economic Co-operation and Development.

Corcoran, T., Mogat, F. A., & Rosher, A. (2009). *Learning progressions in science: An evidence-based approach to reform* (CPRE Research Report #RR-63). Consortium for Policy Research in Education.

Downing, S. M. (2006). Selected-response item formats in test development. In S. M. Downing & T. M. Haladyna (Eds.), *Handbook of test development* (pp. 287–301). Lawrence Erlbaum.

Fuchs, L. S., Malone, A. S., Schumacher, R. F., Namkung, J., & Wang, A. (2017). Fraction intervention for students with mathematics difficulties: Lessons learned from five randomized controlled trails. *Journal of Learning Disabilities, 50*(6), 631–639.

Haladyna, T. M. (2004). *Developing and validating multiple-choice test items* (3rd ed.). Lawrence Erlbaum.

Healy, L., & Powell, A. B. (2013). Understanding and overcoming "disadvantage" in mathematics learning. In M. A. Clements, A. J. Bishop, C. Keitel, J. Kilpatrick, & F. K. S. Leung (Eds.), *Third international handbook of mathematics education* (pp. 69–100). Springer.

Jordan, N. C., Kaplan, D., Locuniak, M. N., & Ramineni, C. (2007). Predicting first-grade math achievement from developmental number sense trajectories. *Learning Disabilities Research and Practice, 22*(1), 36–46.

Jordan, N. C., Kaplan, D., Ramineni, C., & Locuniak, M. (2009). Early math matters: Kindergarten number competence and later mathematics outcomes. *Developmental Psychology, 45*(3), 850–867. https://doi.org/10.1037/a0014939

Ketterlin-Geller, L. R. (2008). Testing students with special needs: A model for understanding the interaction between assessment and student characteristics in a universally designed environment. *Educational Measurement: Issues and Practice, 27*(3), 3–16. https://doi.org/10.1111/j.1745-3992.2008.00124.x

Ketterlin-Geller, L. R. (2016). Understanding and improving accessibility for special populations. In A. Rupp & J. P. Leighton (Eds.), *Handbook of cognition and assessment* (pp. 198–225). Wiley-Blackwell.

Ketterlin-Geller, L. R., Crawford, L., & Huscroft-D'Angelo, J. N. (2014). Screening to assign accommodations: Using data to make decisions. *Learning Disabilities: A Multidisciplinary Journal, 20*(2), 61–74.

Ketterlin-Geller, L. R., Yovanoff, P., Jung, E., Liu, K., & Geller, J. (2013). Construct definition using cognitive-based evidence: A framework for practice. *Educational Assessment, 18*, 122–146. https://doi.org/10.1080/10627197.2013.790207

Kilgore, S. B., & Reynolds, K. J. (2011). *From Silos to Systems: Reframing schools for success. Thousand Oaks, CA: Corwin.*

Kuehnert, E. A., Geller, J., Perry, L., Hatfield, C., & Ketterlin-Geller, L. R. (2020). *Numerical relational reasoning (NRR): Learning progressions development* (Tech. Rep. No. 20-02). Southern Methodist University, Research in Mathematics Education. Available online at: https://blog.smu.edu/MMaRS/technical-reports/

Luecht, R. M. (2007). Using information from multiple-choice distractors to enhance cognitive-diagnostic score reporting. In J. P. Leighton & M. J. Gierl (Eds.), *Cognitive diagnostic assessment for education: Theory and applications* (pp. 319–340). Cambridge University Press.

National Governers Association & Council of Chief State School Officers. (2010). *Common Core state standards for mathematics*. Authors.

National Mathematics Advisory Panel. (2008). *Foundations for success: The final report of the National Mathematics Advisory Panel*. US Department of Education.

National Research Council. (2001). Adding it up: Helping children learn mathematics. In J. Kilpatrick, J. Swafford, & B. Findell (Eds.), *Mathematics learning study committee, center for education, division of behavioral and social sciences and education*. National Academy Press.

Powell, S. R., Berry, K. A., Fall, A.-M., Roberts, G., Fuchs, L. S., & Barnes, M. A. (2020). Alternative paths to improved word-problem performance: An advantage for embedding pre-algebraic reasoning instruction within word-problem intervention. *Journal of Educational Psychology, 113*(5), 898–910. https://doi.org/10.1037/edu0000513

Schneider, C. M., Huff, K. L., Egan, K. L., Gaines, M. L., & Ferrara, S. (2013). Relationships among item cognitive complexity, contextual demands, and item difficulty: Implications for achievement-level descriptors. *Educational Assessment, 18*, 99–121.

Tzur, R. (2013). Too often, these children are teaching-disabled, not learning-disabled. In *Proceedings of the 11th annual Hawaii international conference on education*. Author (DVD).

Woodward, J., & Tzur, R. (2017). Final commentary on the cross-disciplinary thematic special series: Special education and mathematics education. *Learning Disability Quarterly, 40*(3), 146–151.

Zhang, D., Xin, Y. P., Harris, K., & Ding, Y. (2014). Improving multiplication strategic development in children with math difficulties. *Learning Disability Quarterly, 37*(1), 15–30.