Research Report

Elementary Students' Understanding of Geometrical Measurement in Three Dimensions

ETS RR–19-14

Eun Mi Kim
Leslie Nabors Oláh

December 2019

ETS

ETS Research Report Series

Since its 1947 founding, ETS has conducted and disseminated scientific research to support its products and services, and to advance the measurement and education fields. In keeping with these goals, ETS is committed to making its research freely available to the professional community and to the general public. Published accounts of ETS research, including papers in the ETS Research Report series, undergo a formal peer-review process by ETS staff to ensure that they meet established scientific and professional standards. All such ETS-conducted peer reviews are in addition to any reviews that outside organizations may provide as part of their own publication processes. Peer review notwithstanding, the positions expressed in the ETS Research Report series and other published accounts of ETS research are those of the authors and not necessarily those of the Officers and Trustees of Educational Testing Service.

The Daniel Eignor Editorship is named in honor of Dr. Daniel R. Eignor, who from 2001 until 2011 served the Research and Development division as Editor for the ETS Research Report series. The Eignor Editorship has been created to recognize the pivotal leadership role that Dr. Eignor played in the research publication process at ETS.

ETS Research Report Series ISSN 2330-8516

RESEARCH REPORT

Elementary Students' Understanding of Geometrical Measurement in Three Dimensions

Eun Mi Kim & Leslie Nabors Oláh

Educational Testing Service, Princeton, NJ

In this study, we investigated the potential of a hypothesized geometrical measurement learning progression (LP) to examine students' thinking and understanding in this domain. We interviewed 30 third to fifth graders using 3 LP-based cognitive tasks that asked the students to find the length, perimeter, area, surface area, and volume measurement of a given object. We analyzed the students' responses to the tasks to examine variation in levels of the students' geometrical measurement understanding and found evidence of understanding at 5 successive levels of a geometrical measurement LP in 1, 2, and 3 dimensions. From these findings, we concluded that an LP can be a practical tool for understanding students' existing thinking and understanding in a targeted domain and has the potential to support students' further learning in the domain.

Keywords Geometrical measurement; learning progressions; elementary-grades mathematics

doi:10.1002/ets2.12250

Principles and Standards for School Mathematics (National Council of Teachers of Mathematics [NCTM], 2000) described measurement as "a process that students in grades 3–5 use every day as they explore questions related to their school or home environment" (p. 171). Measurement knowledge, including its application for everyday life—both in and out of school contexts—is not only important in and of itself, but also for its relationship to other areas of mathematics and science (Clements, 2003). Yet, the most recent results from the U.S. National Assessment of Educational Progress (NAEP) indicated that performance in measurement (like performance in mathematics more broadly) is stagnant in both Grade 4 and Grade 8 (U.S. Department of Education, 2018). Furthermore, students' classroom experiences can be limited by the rote approaches to measurement found in many classrooms (Clements & Sarama, 2014). One way to better inform the teaching and learning of measurement is through the development of *learning progressions* (LPs)—descriptions of progressively more sophisticated ways of thinking about a concept or content that are hypothesized based on learning research (Smith, Wiser, Anderson, & Krajcik, 2006).

Descriptions of student learning in a domain can be valuable not only as foundational knowledge in this field but also as a way of informing instruction and assessment. In particular, Smith et al. (2006) suggested constructing hypothesized learning progressions within a particular content domain based on extant research on student thinking and learning, viewing them as useful tools to elaborate on national standards documents and to improve large-scale and classroom assessment. Additionally, according to Sztajn, Confrey, Wilson, and Edgington (2012), LPs grounded in learning research can support instruction when teachers use the LPs as "the basis for instructional decisions" (p. 147). Consistent with these views, we see research-based LPs as informing assessment design and as an instructional resource to improve student mathematics learning in school (see also Graf & van Rijn, 2016). In our previous work (Kim, Haberstroh, Peters, Howell, & Oláh, 2017), we described the development of a hypothesized LP for geometrical measurement in one, two, and three dimensions (hereafter referred to as the *geometrical measurement LP*) by synthesizing the findings of existing research on student learning of geometrical measurement (e.g., length, perimeter, area, surface area, and volume measurement). This LP was reviewed by experts in this field and was used to design a series of cognitive tasks to elicit evidence of understanding along the levels of the LP.

The purposes of the current study were to (a) gain early validation support for our hypothesized geometrical measurement LP (Kim et al., 2017) by analyzing variation in levels of students' geometrical measurement understanding within each dimension and (b) suggest the potential of using LP-based cognitive tasks in examining students' existing thinking

Corresponding author: E. M. Kim, E-mail: kimeun23@msu.edu

and understanding. To reach this goal, we examined third- to fifth-grade students' thinking and understanding of geometrical measurement presented in their responses to three LP-based cognitive tasks that asked students to find the length, perimeter, area, surface area, and volume measurement of a given object. The research question explored in this study was the following: Can variation in levels of students' geometrical measurement understanding be elicited by LP-based cognitive tasks measuring length (and perimeter), area (and surface area), and volume measures at the same time?

This report consists of five sections. In the opening section, we briefly describe how current elementary school (K–6) mathematics standards present geometrical measurement in terms of grade-level expectations. In the second section, we present the conceptual framework of this study related to the development of a hypothesized LP and its validation in a content domain, geometrical measurement. Next, we describe our research methods, including an inductively developed coding scheme. In the results section, we present and interpret observed variation in levels of geometrical measurement understanding elicited from the LP-based cognitive tasks, aligning student work with each level. Finally, we discuss the implications of the geometrical measurement LP as a basis for assessment design that would support teachers' instructional decisions by revealing students' existing thinking and understanding and interpreting it with regard to the levels of the LP.

Teaching and Learning of Geometrical Measurement in Elementary School (K–6) Mathematics Education

Participants in the Common Core State Standards Initiative (CCSSI, 2010) designed the *Common Core State Standards for Mathematics* (CCSS–M) to address the mathematical concepts and skills that U.S. students need to develop through mathematics education. The CCSSI (2010) stated that these standards documents draw on "research-based learning progressions detailing what is known today about how students' mathematical knowledge, skill, and understanding develop over time" (p. 4).

In the CCSS–M, the CCSSI (2010) outlined an instructional sequence of geometrical measurement that presents when and how elementary school students need to learn measurement of length, area, and volume—including perimeter and surface area. In kindergarten through Grade 2, students are expected to learn to measure length by "laying multiple copies of a shorter object" (p. 16) in a row from end to end with no gaps or overlaps. In this context, the shorter object is used as a length/linear measurement unit, and a count of the iterated units represents the length measure. In Grades 3 through 4, students are to develop understanding of area measurement with regard to covering the two-dimensional region of an object with squares with "side length 1 unit" (CCSSI, 2010, p. 25) or with same-sized rows or columns of the individual squares without gaps or overlaps and counting the iterated area units. In Grades 5 through 6, students need to develop understanding of measuring volume by filling the space of an object with cubes with side length one unit or with same-sized layers of cube arrays. In the mathematics standards, the teaching and learning of perimeter measurement is introduced immediately following area measurement in Grade 3 to show students the difference between one- and two-dimensional measurement as the "linear and area measures" (CCSSI, 2010, p. 25) of a given polygon, respectively. In a similar manner, surface area measurement is presented in Grade 6 after volume measurement is addressed in Grade 5. In sum, across the three dimensions, the use of standard units and their iteration for measurement are consistently emphasized.

According to the CCSS–M (CCSSI, 2010), the teaching and learning of geometrical measurement begins with length measurement, then continues to area and volume measurement; this may suppose a geometrical progression among the three dimensions (e.g., Barrett et al., 2011; Barrett, Clements, & Sarama, 2017; Curry, Mitchelmore, & Outhred, 2006). A similar instructional sequence is suggested in the *Principles and Standards for School Mathematics* (NCTM, 2000).

Conceptual Framework

Our study was informed by the literature in five areas: (a) conceptualizations and characteristics of LPs; (b) core cognitive constructs of geometrical measurement; (c) review and synthesis of existing LPs for length, area, and volume measurement; (d) a geometrical measurement LP hypothesized based on the extant research on measurement learning; and (e) the validation process for hypothesized LPs.

Conceptualizations and Characteristics of LPs

The idea of LPs has been developed as "an approach to research synthesis that could serve as the basis for a dialogue that includes researchers, assessment developers, policy makers, and curriculum developers" (National Research Council [NRC], 2007, p. 214). In the literature, conceptualizations of LPs share similar features with cognitive models in that they represent levels of sophistication of students' thinking and learning around core cognitive constructs in a particular content domain but are conceptualized and/or defined differently by scholars and researchers according to their purpose (e.g., Deane, Sabatini, & O'Reilly, 2012; Smith et al., 2006; Wilson, 2009; see also NRC, 2007). In addition, Clements and Sarama (2004) conceptualized learning trajectories as

> descriptions of children's thinking and learning in a specific mathematical domain and a related, conjectured route through a set of instructional tasks designed to engender those mental processes or actions hypothesized to move children through a developmental progression of levels of thinking, created with the intent of supporting children's achievement of specific goals in that mathematical domain. (p. 83)

Their conceptualization of learning trajectories can be distinguished from other LP conceptualizations in that "trajectories include descriptions of instruction, progressions do not" (Battista, 2011, p. 512). In following Smith et al. (2006), we conceptualize a hypothesized LP as a synthesis of existing research on how children learn in a targeted content domain.

Smith et al. (2006) defined a learning progression as "a sequence of successively more complex ways of thinking about an idea that might reasonably follow one another in a student's learning" (p. 5) within a specific content domain and promote its development around the big ideas of the content domain through syntheses of extant research on learning and conceptual analyses. Smith and colleagues described the basic characteristics of LPs as such: (a) successive levels of LPs do not represent a single correct sequence, but rather propose multiple pathways, which may be influenced by instructional intervention; (b) actual learning is viewed as ecological succession with simultaneous changes in multiple interrelated ways; and (c) with no "long-term longitudinal accounts of learning by individual students" (p. 6), the LPs are hypothetical to a certain extent.

Regarding levels of LPs, Battista (2004) argued that "the levels are compilations of empirical observations of the thinking of many students and because students' learning backgrounds and mental processing differ, a particular student might not pass through every level for a topic" (p. 187). Furthermore, LPs "are not developmentally inevitable" (Smith et al., 2006, p. 5) but "crucially dependent on instructional practices if they are to occur" (NRC, 2007, p. 214). As noted by Daro, Mosher, and Corcoran (2011), this last claim is likely truer of later-developing understanding than of earlier mathematical learning (e.g., quantity and shape), but it is certainly relevant for students' development of geometrical measurement.

For our research and development, we have followed the Smith et al. (2006) LP conceptualization. In framing an LP for geometrical measurement, we first considered the core cognitive constructs that underlie geometrical measurement in one, two, and three dimensions through the review of extant literature in the domain of geometry and measurement.

Core Cognitive Constructs of Geometrical Measurement

In their study on young students' conceptions of geometry in measurement, Piaget, Inhelder, and Szeminska (1960) stated that *to measure* is "to take out of a whole one element, taken as a unit, and to transpose this unit on the remainder of a whole" (p. 3); this conception of geometrical measurement associates the concept of unit, "identical" subdivisions of the object to be measured (Lehrer, Jaslow, & Curtis, 2003, p. 102), and two unit-related concepts: unit partition and unit iteration. The concept of *unit partition* refers to the mental and/or physical operation of subdividing an object by an identical, same-sized unit of measurement, and *unit iteration* refers to the operation of placing the taken unit end to end, with no gap or overlap, within the object being measured (Stephan & Clements, 2003, pp. 3–4).

In the domain of geometry and measurement, unit partition and unit iteration have been addressed by scholars and researchers as fundamental ideas needed to conduct the process of measuring meaningfully (Lehrer, 2003; Lehrer et al., 2003; Piaget et al., 1960; Stephan & Clements, 2003). These concepts are also emphasized in the geometrical measurement standards. In the CCSS–M, for instance, the CCSSI (2010) argued that in Grade 1, students need to develop the understanding of iteration of same-sized units in regard to the meaning and procedures of measuring length (see p. 13). The emphasis of this idea is also true for area and volume in later grades with respect to space covering with square units in Grade 3 (see p. 21) and space filling with cubic units in Grade 5 (see p. 33).

Yet we know that measuring is more than iterating single units. The conceptualization of unit partition and unit iteration for geometrical measurement also involves the understanding of partitioning by and iterating of a same-sized *composite unit*, "which is a unit consisting of more basic units" (Battista, 2003, p. 123), such as a set of individual squares for area and cubes for volume. According to Lehrer et al. (2003), iterating a composite unit as a "unit-of-units" (p. 105) can provide a natural context for constructing understanding of multiplication. For instance, the length of 12 linear units can be reconstructed as four iterations of the composites of three-linear-unit lengths (i.e., $12 = 4 \times 3$). As such, we propose that iterating composite units helps students understand and reason about the meaning of multiplying linear/length measures in formulas, such as volume = length × width× height.

Battista (2003) took this idea further by arguing for the importance of iterating *maximal composite units* such as "rows and columns of squares for area, layers for volume" (p. 127) with reference to spatial structuring of the maximal composites (row-by-column or layer structuring), arguing that "such structuring is more general and powerful than using standard area and volume formulas" (p. 129); see also Battista & Clements, 1996, for the idea of spatial structuring in detail. For instance, in measuring the area of a 4″-x-5″-inch rectangle, a student typically mentally integrates five squares in a row into a row composite and iterates the composite in the direction of a column to construct the entire rectangle (the actions of which are referred to as *maximal composite unit* and *row-by-column structuring*, respectively). Additionally, Battista and Clements (1998) pointed out that only students who visualize and construct such spatial structuring appropriately are ready to begin to formulate and abstract their enumeration process in terms of formulas.

As emphasized in the research literature in this domain, we deem the ideas of partition and iteration of (composite, maximal composite) units and spatial structuring of the iterated units as important cognitive features of geometrical measurement in one, two, and three dimensions. In considering measurement contexts for which no maximal composite unit exists, such as measuring the perimeter of a rectangle, we instead propose the idea of iterating efficiently sized composite units for that measurement context (e.g., iteration of the total length of two adjacent sides of a rectangle for perimeter).

Existing Learning Progressions for Length, Area, and Volume Measurement

After identifying the core cognitive constructs of geometrical measurement, we also reviewed and synthesized existing LPs for length, area, and volume measurement (e.g., for perimeter: Barrett, Clements, Klanderman, Pennisi, & Polaki, 2006; for area, Battista, Clements, Arnoff, Battista, & Borrow, 1998; for area and volume, Battista, 2004; for volume, Battista & Clements, 1996) to find a common developmental sequence of geometrical measurement within and across the dimensions. Here we briefly introduce the length measurement LP of Barrett et al. (2006) and the area and volume measurement LP of Battista (2004) that provided us with a theoretical grounding for hypothesizing levels of students' understanding in the three dimensions.

Barrett et al. (2006) interviewed 38 students in Grades 2 to 10, using two perimeter measurement tasks to examine the students' development of levels of thinking and reasoning about length measurement, focusing on use of units in measurement. Through their analysis of student responses to the tasks, the researchers found support for a developmental progression of five consecutive levels: (a) Level 1 assigns length measure by guessing visually, with no reference to linear units of measurement; (b) Level 2a makes inconsistent identification of linear units (e.g., partitioning with different-length segments) or coordinates iterated units improperly; (c) Level 2b makes consistent unit identification and coordinates iterated units properly; (d) Level 3a begins to iterate composite units for length by shifting measurement thinking "from partitioning to grouping and back and forth" (Barrett et al., 2006, p. 197); and (e) Level 3b iterates composite units efficiently with dynamic reasoning related to nested "part-whole relationships among units and groups of units" (Barrett et al., 2006, p. 209). This recognition that students use composite units with varying degrees of efficiently measuring length contributed to our understanding about iteration of composite units in measuring length and is featured in our LP.

In synthesizing previous LPs on area measurement (Battista et al., 1998) and volume measurement (Battista, 1999; Battista & Clements, 1996), Battista (2004) constructed a general model for the development of students' thinking and reasoning about area and volume measurement in terms of two cognitive processes in measurement: *units-locating* and *organizing-by-composites*. The process of units-locating refers to locating "squares and cubes by coordinating their locations along the dimensions that frame an array" (Battista, 2004, p. 192). The organizing-by-composites process refers to combining "an array's basic spatial units (squares or cubes) into more complicated *composite units* that can be repeated or iterated to generate the whole array" (Battista, 2004, p. 192). Battista suggested seven consecutive levels that progress in their levels of *abstraction,* or "the process by which the mind selects, coordinates, unifies, and registers in memory a collection of mental items or acts that appear in the attentional field" (Battista, 2004, p. 186): (a) Level 1 locates all squares and cubes in an array improperly (e.g., double-counting errors) and organizes no composites; (b) Level 2 begins to locate all squares and cubes in an array and organize equivalent-sized composites (e.g., for volume, after counting the number of cubes on the front of a cube building, see that there is the same number of cubes on the back); (c) Level 3 locates all squares and cubes in an array properly, thus eliminating double-counting errors; (d) Level 4 organizes maximal composite units for area and volume to iterate, but locates the iterated units in an array improperly; (e) Level 5 locates all squares and cubes in an array properly, but iterates less-than-maximal composite units for area and volume; (f) Level 6 employs both of the units-locating and organizing-by-composites processes sufficiently in structuring iterated maximal composites units for area and volume spatially (i.e., row-by-column or layer structuring) and enumeration of all squares and cubes in an array; and (g) Level 7 reaches "a level of abstraction" (Battista, 2004, p. 200) in units-locating and organizing-by-composites, thus relates row-by-column or layer structuring to numerical procedures, including application of area and volume formulas, for geometrical measurement and generalizes such reasoning to different measurement contexts (e.g., filling a rectangular box with rectangular prism-shaped packages, which are made from two identical cubes). The ideas of iteration of maximal composite units and spatial structuring of the iterated units influenced the distinction of levels in our LP with respect to area and volume.

Through the analysis and synthesis of previous work in this field across the LPs for length, area, and volume, we can see a common developmental sequence of geometrical measurement as "the progression from students' measurement thinking and reasoning from concrete and experiential to abstract with regard to unit iteration and spatial structuring of the iterated units, as well as with the use of efficient-sized composite units for presented measurement contexts" (Kim et al., 2017, p. 6). This LP is presented below.

Geometrical Measurement Learning Progression

Building upon previous empirical research on student learning of geometrical measurement in one, two, and three dimensions, we developed a hypothesized geometrical measurement LP. The geometrical measurement LP consists of five primary levels, with Level 3 having two sublevels (Kim et al., 2017). The description of levels is sequenced as

- Level 1 has no conception of unit and its iteration, compares size as a whole or counted parts of an object measured at the holistic level;
- Level 2 shows early unit conception, uses iterated units but improperly structures the iterated units;
- Levels 3 and 3.5 present sufficient conception of unit iteration by structuring the iterated (composite) units correctly;
- Level 4 formalizes the idea of iteration of an efficiently sized composite unit for an object being measured by visualizing the spatial structure of the iterated efficient composite unit; and
- Level 5 conducts measurement in the abstract, reasons about the multiplication of linear/length measures in formulas in terms of spatial structuring of iterated units.

With respect to Levels 3 and 3.5, students can coordinate and structure iterated units properly (i.e., with no gap or overlap) to fill the space of an object. Level 3.5 is characterized by iteration of composite units, yet this level of understanding does not involve reasoning about efficiently sized composite units for an object being measured (i.e., iteration of maximal composite units) (Battista, 2003, 2004; Battista & Clements, 1996). In developing this LP, we considered iteration of efficient composite units in the measuring process to be a conceptual shift necessary to form a more general model of geometrical measurement.

In proposing one LP for geometrical measurement, we put three LPs for length, area, and volume together within the larger LP (see Wilson, 2009) with propositions of vertical progression along levels within each dimension and horizontal progression across the three dimensions (illustrated in Figure 1). With respect to progressing within each dimension, we hypothesized the transition between levels based on the research of Gutiérrez, Jaime, and Fortuny (1991) on evaluating students' acquisition of the van Hiele levels of geometrical reasoning. From their analysis of students' responses to a test evaluating reasoning ability in three-dimensional geometry, Gutiérrez and colleagues identified students "who are in transition between levels" (Gutiérrez et al., 1991, p. 237) because they showed "two consecutive levels of reasoning at the same time, although what usually happens is that the acquisition on the lower level is more complete than the acquisition of the upper level" (Gutiérrez et al., 1991, p. 250). We also propose that growth across the three dimensions is staggered,

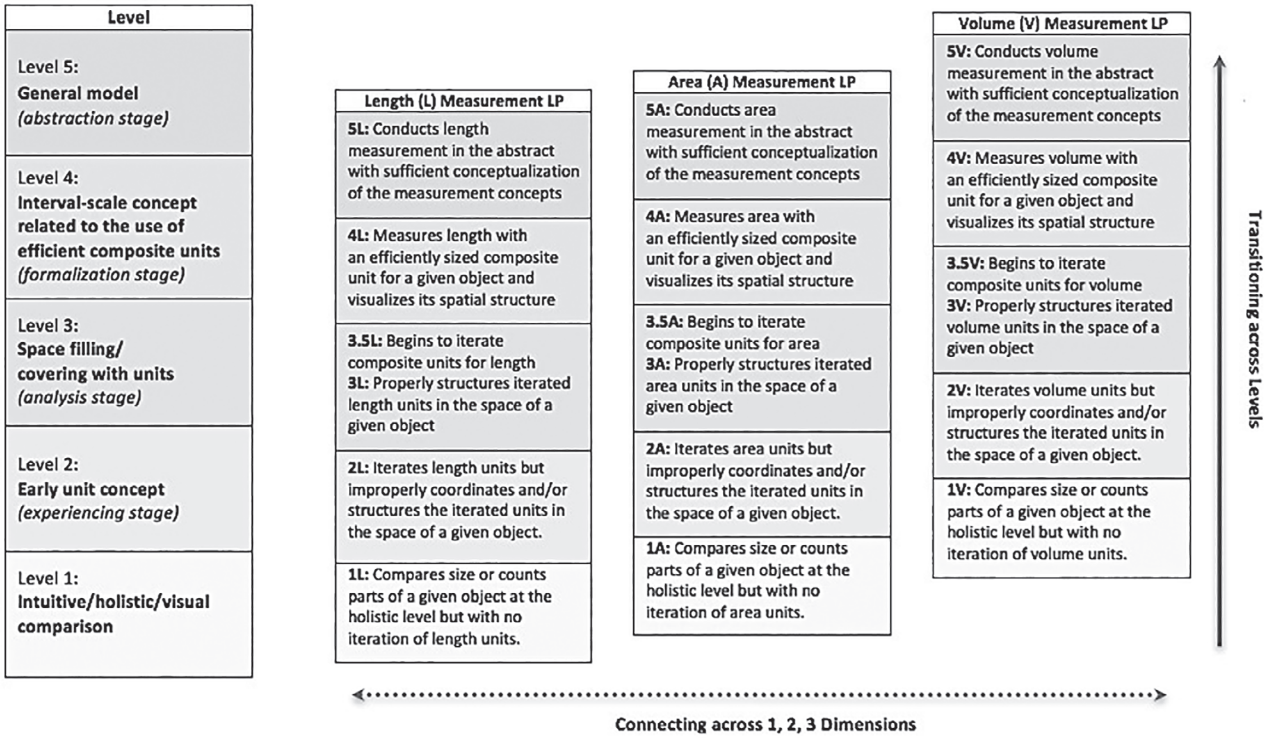

Figure 1 Geometrical measurement learning progression for length, area, and volume measurement. Adapted from "A learning progression for geometric measurement in one, two, and three dimensions" by E. M. Kim, J. Haberstroh, S. Peters, H. Howell, and L. N. Oláh, 2017, Research Report No. RR-17-55, p. 7. Copyright 2017 by Educational Testing Service.

such that understanding of linear measurement precedes that of area and volume measurement but that it is not necessary to have completely surpassed a given level in length (or area) to show understanding at that level in area (or volume).

In closing the review of existing literature in the domains of geometry and measurement in his chapter in the *Second Handbook of Research on Mathematics Teaching and Learning*, Battista (2007) called for further development of an iterated model for student learning of length, area, and volume measurement, because in the field of geometrical measurement "individual research studies usually have focused on only one type of geometric measurement at a time (length, area, or volume), research has not yet produced a comprehensive theory of geometric measurement" (p. 902). In this aspect, our geometrical measurement LP contributes to the field of study by integrating the LPs of length (and perimeter), area (and surface area), and volume within the larger LP construct for geometrical measurement. We are not the only ones who have aimed to integrate the three LPs (see Barrett et al., 2017; Battista, 2007; Battista, 2012); however, our propositions about vertical transitions between consecutive levels of the length and area and volume LPs and horizontal progression among the three LPs to integrate the individual LPs into a single LP for geometrical measurement distinguishes our model from those of other researchers (Kim et al., 2017).

Validation Process for Hypothesized Learning Progressions

Because a hypothesized LP is the result of a synthesis of literature in a particular topic or domain, "each of which focuses on more specific content and a narrower time frame than the learning progression itself" (Graf & van Rijn, 2016, p. 167), it is partially inferential (Smith et al., 2006). Thus, there is a need to conduct subsequent empirical research to provide evidence for the validation or refinement of the LP. Perhaps first among research priorities is to establish whether or not the levels of the LP provide a reasonable description of the order in which the given concept is acquired. At the same time, information can be collected on the development of important ideas within a larger domain. In the case of geometrical measurement, there is the use of (composite, efficient composite) units in unit partition, unit iteration, and spatial structuring of the iterated units.

In considering the provisional nature of hypothesized LPs, which are "subject to empirical verification and theoretical challenge" (Deane et al., 2012, para. 1), Graf and van Rijn (2016) proposed a validation cycle for LPs consisting of four steps: (a) the development of a hypothesized LP through research synthesis and domain analysis, as well as expert opinion; (b) the design of tasks associated with the hypothesized LP to examine whether the ordering of the levels in the hypothesized LP can be recovered empirically (e.g., mapping students' task performance to the levels of a hypothesized LP); (c) the comparison of the LP to other competing models to seek its disconfirmation; and (d) the evaluation of the instructional efficacy of the LP, which refers to "the degree to which it is used successfully in the classroom" (p. 167). According to Graf and van Rijn, the development of an LP should go through multiple iterations of its evaluation and revision processes. In this study, we evaluate our hypothesized geometrical measurement LP in light of empirical evidence of student thinking in response to LP-based tasks.

Methods

As a first step in gathering evidence of student performance with respect to the geometrical measurement LP, we examined individual students' understanding of geometrical measurement in one, two, and three dimensions through cognitive interviews (Ginsburg, 2005). This method was chosen because it would provide us with preliminary evidence of student performance along the hypothesized LP while allowing us to gain in-depth information on students' thinking that could be used to modify the LP, if needed. For the goal of our study, a purposeful sample of 30 students was sufficient, in consideration of both available resources and expected variation in participant responses (see Blair & Conrad, 2011).

Participants

Students were recruited through an online posting to the ETS site as well as through an e-mail list of families who had indicated interest in participating in ETS research studies. The first students who volunteered were chosen. The study involved 30 participants who had just finished third, fourth, or fifth grades in the summer of 2016 (specifically, 12 third graders, 9 fourth graders, and 9 fifth graders). This grade span was chosen to reflect the grade-level expectations of geometrical measurement learning for length (kindergarten through Grade 2), area (Grades 3 through 4), and volume (Grades 5 through 6) in the CCSS–M (CCSSI, 2010). Thus, it was expected that much of the initial learning of geometrical measurement in the three dimensions takes place across this grade span. It was expected that all students would be able to complete the simplest task of measuring height and that we would see variation in performance among all of the tasks.

The vast majority of the students attended local public schools (25 of 30). Three students attended independent schools, one attended a charter school, and one attended a parochial school. Thirteen students were male, and 17 were female. A slight plurality of students identified as Caucasian (12 of 30), 10 identified as Asian, three as African American, and two as multiracial (three students did not report a race or ethnicity). As a group, the students reported using at least eight different mathematics curricula, and perhaps more, as 12 students reported that they did not know what curriculum they used in class, or if they used a text at all, or were not sure what materials were used in class. On average, students reported having average to strong math performance in school; 10 students reported receiving the highest possible grade in mathematics (whether it be a grade of A, or a 3 on a scale of 1 to 3, or "exceeds expectations."); 10 reported receiving a B, or "proficient"; nine did not report their most recent grade in math; and one student reported that his school does not assign tests or grades.

Interview Tasks and Procedure

Because this study aimed to investigate levels of understanding about one-, two-, and three-dimensional measurement, we designed three interview tasks to target different levels of the LP (Kim et al., 2017). Each task consisted of three main components: (a) a drawn three-dimensional figure, (b) a set of written questions for students to respond to, and (c) a set of interview probes given in the presence of a physical model of the drawn figure presented in (a). All three tasks asked students to provide measurements of the dimensions of a drawn three-dimensional figure; however, the three tasks presented increasingly challenging stimuli. The stimulus for Task 1 consisted of a four-by-four cuboid on which the 1-unit cubes were marked. Task 2 provided a similar stimulus, except that it featured an irregular shape (with the 1-unit cubes clearly marked). Task 3 showed a 3-by-3 cuboid; however, unlike the Task 1 cuboid and the Task 2 shape, only some 1-unit

Table 1 Inductively Developed Coding Scheme From Participant Responses to the Geometrical Measurement Tasks

Level	Defining characteristic of geometrical measurement understanding using units
Level 5	Determines measurement by applying formula(e) and makes visual inference in measurement
Level 4	Iterates efficient composite units in measuring
Level 3.5	Iterates composite units in measuring
Level 3	Iterates consistent and appropriate dimensional units
Level 2	Iterates inconsistent or inappropriate dimensional units
Level 1	Counts at a holistic level with no unit iteration; computes measurement through inappropriate application of formula(e)
Level 0	No conception or presence of misconceptions of the attributes to be measured; no response

cubes were marked on the cube. A set of parallel questions was asked for each task, focusing on measurement of height, perimeter, area, surface area, and volume (in that order). Tasks 1 and 2 were designed to elicit understanding at Levels 1 to 3 of the LP, whereas Task 3 was designed to elicit evidence at all five levels. (See Kim et al., 2017, for task design in detail.)

Students began by working independently to respond to the set of questions. Once students had responded to the questions about the drawn figure, they were provided the three-dimensional object on which the drawing was based and were allowed to change their responses if needed. At this time, the interviewer posed follow-up questions to explore their measurement strategies as well as their thinking of and reasoning about the measurement processes (e.g., “How could you find the height of this cuboid?”). These cognitive interviews took, on average, 20–30 minutes. Interviews were audio- and video-recorded. Following the interviews, student work was scanned as PDFs, and the audio recordings were professionally transcribed. These two data sources were linked by a participant ID number to the video recordings. All three data sources were used in this analysis.

Coding and Analysis

Table 1 presents an overview of our inductively developed coding scheme used for classifying participant responses to questions of height, perimeter, area, surface area, and volume measurement in Tasks 1, 2, and 3. In development of this coding scheme, we began by grouping common and similar participant responses given to each measurement task for height, perimeter, area, surface area, and volume and defined each level description to capture the characteristics of each group of the responses. In this grouping, we paid particular attention to how students used units to measure.

In classifying participant responses at levels of the geometrical measurement LP, we examined participants’ written responses to questions about the height, perimeter, area, surface area, volume measurements of a given object in each of the three tasks, and their verbal accounts about the measures and strategies for measuring each attribute of the given object in response to the interviewer’s questions (e.g., “What does perimeter mean?”, “When we look at this front face, where is the perimeter?”, or “How could you figure out the perimeter of the front face of this object?”). For each participating student, their responses to each question in each task were coded at one level only, considering their response holistically by taking into account their written and verbal responses as well as their gestures.

The coding scheme is a result of continuous classification of participant responses according to shared characteristics of geometrical measurement understandings across the three attributes of length, area, and volume as they appear in the three tasks. We used this coding scheme to determine whether the three tasks elicited variation in levels of student understanding in one-, two-, and three-dimensional measurement; thus, it can be considered as evidence in support of the geometrical measurement LP among the three dimensions.

In establishing inter-rater reliability, the second author double-coded 10 transcriptions selected from the 30 transcriptions that the first author had initially completed coding. Few discrepancies were found. All the discrepancies between first and second coding were resolved through discussion, and any changes made to the coding were then applied to all 30 transcriptions.

Findings

In this section, we illustrate how LP-based cognitive tasks in measuring length, area, and volume concurrently can be used to elicit variation in levels of students’ geometrical measurement understanding. In doing so, we provide student responses

Table 2 Frequency of Participant Responses Classified at Each Level of Geometrical Understanding in Response to Height Measurement Tasks

	Grade 3 ($n = 12$)			Grade 4 ($n = 9$)			Grade 5 ($n = 9$)		
Level	Task 1[a]	Task 2[a]	Task 3	Task 1[a]	Task 2[a]	Task 3	Task 1[a]	Task 2[a]	Task 3
Level 5			8			9			9
Level 4			0			0			0
Level 3.5			0			0			0
Level 3	10	9	0	9	9	0	9	9	0
Level 2	0	0	2	0	0	0	0	0	0
Level 1	0	0	0	0	0	0	0	0	0
Level 0	2	3	2	0	0	0	0	0	0

[a]Height measurements in Tasks 1 and 2 are designed to target Levels 1 to 3 in length measurement LP only; in other words, Level 3 is the highest level of understanding that is expected to be shown in these height measurement tasks.

Table 3 Frequency of Participant Responses Classified at Each Level of Geometrical Understanding in Response to Perimeter Measurement Tasks

	Grade 3 ($n = 12$)			Grade 4 ($n = 9$)			Grade 5 ($n = 9$)		
Level	Task 1	Task 2	Task 3	Task 1	Task 2	Task 3	Task 1	Task 2	Task 3
Level 5	1	0	1	2	1	2	0	0	1
Level 4	0	0	0	1	0	2	0	0	0
Level 3.5	0	1	0	0	1	0	0	0	0
Level 3	2	1	2	3	3	2	1	1	1
Level 2	1	2	2	2	3	2	2	2	3
Level 1	1	1	0	0	0	0	0	0	0
Level 0	7	7	7	1	1	1	6	6	4

representing different levels of geometrical understanding in response to the tasks. To address our research question, we examined student performance along the three LPs separately—height and perimeter measurement, area and surface area measurement, and volume measurement—to determine whether variation in levels of student understanding within each dimension could be elicited with these tasks. We observed variation not only across students but also within students, by task. For example, some participants showed different levels of understanding according to the measurement context; thus, we do not claim that students are "at" a given level, but rather that their performance during a particular task provides evidence of understanding at a particular level.

Height and Perimeter Measurement

Tables 2 and 3 display the frequency of student responses classified at different levels of understanding in height and perimeter measurement, respectively, for Tasks 1, 2, and 3. In addition to Levels 1 to 5, we included a Level 0, indicating no conception of height. For example, in response to Task 1, Participant 9 (Grade 3) added up the length measures of the four vertical sides of the cuboid to find the height of the cuboid (Level 0).

The data in Table 2 indicate that, with the exception of a few students in Grade 3 classified at Level 0, most or all students across all grades show sufficient understanding of iteration of consistent and appropriate linear units with respect to height measurement, as indicated by their responses at Level 3 to Tasks 1 and 2. In other words, most students can consistently and appropriately iterate units to solve problems of height measurement. In addition, all students in Grades 4 and 5 and most students in Grade 3 can make proper inferences about length measurement, as shown by the frequency of responses at Level 5 to Task 3. For example, as her response to Task 3 asking for the height of a stack of rectangular blocks of three different sizes, Participant 19 (Grade 5) wrote down the height measurement as "3 cm" by making visual inferences with the given length measure of 1 cm (Level 5). See Figure 2 for her drawing of the lines representing her visual inference that she explained as: "I separated them [blocks] into one-size centimeter."

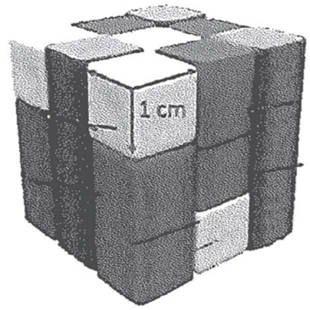

Figure 2 Work produced from Participant 19 in response to Task 3.

Two students' performances were classified at Level 2 because they counted one 2-cm block and one 1-cm block together as two "whole" blocks in response to Task 3.

The data in Table 3 indicate that, with respect to perimeter measurement, the three tasks elicited responses that were more varied in levels of understanding than those about height measurement. Responses were observed along all levels of the LP and included counting parts of a given object at a holistic level with no unit iteration (Level 1), (efficient) composite unit iteration and its spatial structuring (Levels 3.5 to 4), and perimeter formula application in measurement of a path length (Level 5).

We present four student responses representing thinking and reasoning at Levels 1, Levels 3.5 to 4, and Level 5 of the length measurement LP that were observed in response to perimeter measurement questions. In a response to Task 1, finding the perimeter of the front face of a cuboid, Participant 4 (Grade 3) incorrectly wrote down "4 lines" in his written work by counting the four sides of the front face, indicating that he was counting at a holistic level with no unit iteration (Level 1). A Level 3.5 response to Task 2 was shown by Participant 8 (Grade 4), who attempted to find the perimeter of the front face of an irregular shape made of cubes by counting by "twos" (i.e., iterating sets of two equal-interval sized line segments to form the perimeter of the given shape), arriving at the correct answer. Participant 12 (Grade 4) showed a Level 4 response to Task 1 by first counting the four equal-interval sized line segments along the top side of the front face and multiplied by 4, explaining that "… each of them [sides] has four" (i.e., iterating efficient composite units to form the perimeter of the given shape). In response to the same measurement task, Participant 25 (Grade 4) computed the perimeter of the given shape, describing his process as "Um, to get the perimeter, you would do length plus width plus length plus width. So length is 4 cubes and width's 4 cubes. So 4 plus 4 plus 4 plus 4, 16." This application of a perimeter formula $W + L + W + L$ with regard to the spatial structure of the given shape was classified at Level 5.

As displayed in Table 3, each task elicited responses at all targeted levels of the LP, with the exception of Level 3.5 in response to Task 1, Level 4 in response to Task 2, and Levels 1 and 3.5 with respect to Task 3. Still, taken as a set, these three tasks were successful at targeting all levels of the LP. It is important to note, however, that half of the third-grade responses and nearly half of the fifth-grade responses were classified at Level 0, revealing that students either had no conception, or evidenced misconceptions of, perimeter (e.g., confusing perimeter with area). In sum, Tasks 1, 2, and 3 measuring height and perimeter were successful at eliciting evidence of student understanding at all targeted levels of the length measurement LP, including the highest level of understanding that was expected to be examined in these tasks (see Tables 2 and 3).

Area and Surface Area Measurement

The three tasks also targeted student understanding of area measurement, including surface area measurement. Tables 4 and 5 display the frequency of the responses classified at each level of geometrical understanding in area and surface area measurement for Tasks 1, 2, and 3; these three tasks were developed to target all five levels of the LP. As with the previous analysis, Level 0 indicates no conception, or presence of misconceptions, of area measurement.

Although some students showed either no conception of, or misconceptions with, area measurement (e.g., confusing area with perimeter, surface area, or volume), the data in Table 4 show that across grades, most students had at least some understanding of area measurement and that the level of understanding evident in student responses varied widely, both within and across tasks. Task 1 elicited evidence of understanding at three levels only (Levels 0, 3, and 5), whereas Tasks 2

Table 4 Frequency of Participant Responses Classified at Each Level of Geometrical Understanding in Area Measurement

	Grade 3 ($n=12$)			Grade 4 ($n=9$)			Grade 5 ($n=9$)		
Level	Task 1	Task 2	Task 3	Task 1	Task 2	Task 3	Task 1	Task 2	Task 3
Level 5	7	2	2	6	1	6	6	2	4
Level 4	0	3	2	0	4	0	0	2	1
Level 3.5	0	0	0	0	1	0	0	0	0
Level 3	2	3	3	2	2	2	1	1	2
Level 2	0	0	1	0	0	0	0	0	0
Level 1	0	0	0	0	0	0	0	1	0
Level 0	3	4	4	1	1	1	2	3	2

Table 5 Frequency of Participant Responses Classified at Each Level of Geometrical Understanding in Surface Area Measurement

	Grade 3 ($n=12$)			Grade 4 ($n=9$)			Grade 5 ($n=9$)		
Level	Task 1	Task 2	Task 3	Task 1	Task 2	Task 3	Task 1	Task 2	Task 3
Level 5	1	0	1	2	0	2	0	0	0
Level 4	0	1	0	0	0	0	0	0	0
Level 3.5	0	0	0	0	0	0	0	0	0
Level 3	0	0	0	0	0	0	0	0	0
Level 2	0	0	0	1	1	1	0	0	0
Level 1	0	0	0	0	0	0	0	0	0
Level 0	11	11	11	6	8	6	9	9	9

and 3 elicited responses at six and seven of the seven levels, respectively. The number of responses classified at the higher levels indicate that more than half of the students demonstrated understanding of (composite, efficient composite) unit iteration (i.e., Levels 3 to 4) and area formula and visual inference (i.e., Level 5).

Here we present four student responses representing thinking and reasoning at Levels 3 to 5; these are all in response to Task 2, asking to find the area of the front face of a stack of cubes of the same size. To this area measurement task of a nonrectangular shape, Participant 6 (Grade 4) correctly answered "16" by counting "all the squares" on the shape, thereby showing Level 3 reasoning by iterating equal-sized squares and correctly structuring the iterated units to form the whole shape. In the same measurement context, Participant 29 (Grade 4) showed Level 3.5 understanding by adding "12 plus 4 to get 16" by counting a set of four squares and two rows of six squares on the shape (i.e., iterating the composites, a set of four squares (i.e., the top) and two rows of six squares (i.e., the base), correctly structuring the composites to form the whole shape. This is distinguished from Level 4 understanding shown by Participant 8 (Grade 4), who iterated four sets of four squares on the face (i.e., iterating efficient composite units to form the whole shape). Level 5 understanding was also seen in response to this task by Participant 7 (Grade 5) who "divided" the face into two rectangular-shaped parts, the top and the base (see Figure 3 for his marking to divide the shape into two parts) to do "2 times 6 and then 2 times 2" and then "added them" (i.e., applying an area formula, W × L, twice with reference to the spatial structure of the given nonrectangular shape).

The data in Table 5 show that, as with perimeter measurement, students across Grades 3–5 have difficulty with the understanding of surface area measurement, as most of the student responses were classified at Level 0, revealing no conception, or substantial misconceptions, of surface area (e.g., confusing surface area with volume or thinking of surface area as the surface/top side of a given object). However, the tasks were also able to elicit evidence of student understanding at Levels 2, 4, and 5, indicating that some students demonstrated understanding of surface area (but used inappropriate dimensional units for area; Level 2), efficient composite unit iteration (Level 4), and formula and inference (Level 5).

Here we present three student responses representing thinking and reasoning at Levels 2, 4, and 5. In response to Task 1, finding the surface area of a given cuboid, Participant 8 (Grade 4) responded at Level 2 by counting "56 cubes" shown on the six faces of the given object (i.e., iterating cubic units for area measurement; Level 2). In response to Task 2, finding the surface area of a stack of cubes of the same size, Participant 17 (Grade 3) showed evidence of Level 4 understanding in that he added "the 16 and 16" for the areas of "the front face and the back face" and then added "8 and 8" for the left-side

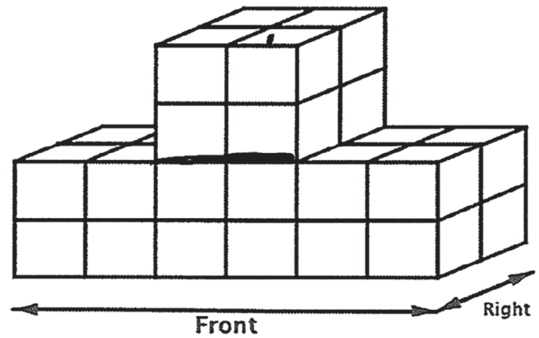

Figure 3 Participant 7's work on Task 2 separating the top from the base.

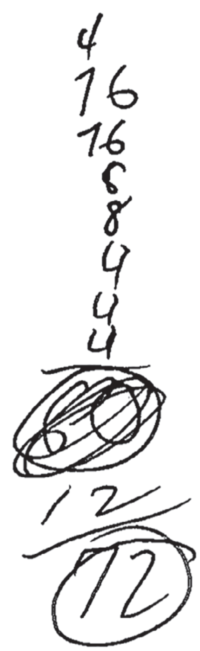

Figure 4 Work produced from Participant 17 in response to surface area measurement in Task 2.

and right-side faces and 12 for the bottom after finding the area of top parts by adding 4 three times as "There's four ... four, four." See Figure 4 for his computation that reveals his reasoning about iteration of efficient composite units to form the surface area of the given object. In response to Task 3, Participant 17 showed Level 5 understanding when finding the surface area of a stack of rectangular blocks of three different sizes. He computed "3 times 3 times 6," applying a surface area formula for a cube, $S \times S \times 6$, with regard to the spatial structure of the given cube-shaped object.

In sum, Tasks 1, 2, and 3 measuring area and surface area were successful at eliciting evidence of student understanding with regard to unit iteration and its spatial structuring at almost all targeted levels of area measurement LP (see Tables 4 and 5).

Volume Measurement

Table 6 displays the frequency of the responses classified at each level of geometrical understanding in volume measurement for Tasks 1, 2, and 3 that were developed to target all levels of the LP. Level 0 indicates no conception or misconceptions of volume measurement.

As with other topics in measurement, many students, particularly third graders, showed either no conception of, or had misconceptions with, volume measurement (e.g., confusing volume with height or surface area). Even with this in mind, the tasks still successfully elicited evidence of student understanding across most levels of the LP, except Level 2

Table 6 Frequency of Participant Responses Classified at Each Level of Geometrical Understanding in Volume Measurement

	Grade 3 ($n = 12$)			Grade 4 ($n = 9$)			Grade 5 ($n = 9$)		
Level	Task 1	Task 2	Task 3	Task 1	Task 2	Task 3	Task 1	Task 2	Task 3
Level 5	3	0	3	3	1	3	3	1	2
Level 4	0	2	0	1	3	2	2	2	2
Level 3.5	0	0	0	0	1	0	0	0	0
Level 3	0	1	0	1	0	0	0	1	1
Level 2	0	0	0	0	0	0	0	0	0
Level 1	1	1	1	0	0	0	0	1	0
Level 0	8	8	8	4	4	4	4	4	4

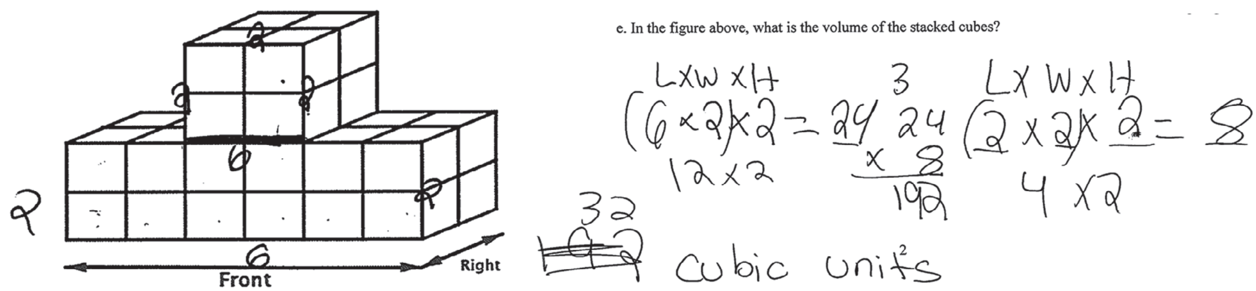

Figure 5 Participant 30's work on Task 2.

(iterating inconsistent or inappropriate dimensional units in measuring volume), which had no responses coded at this level.

Below we present five student responses representing thinking and reasoning at Levels 1 and 3 to 5. In response to Task 2, finding the volume of a stack of cubes of the same size, Participant 13 (Grade 5) showed understanding at Level 1 when he applied the volume formula "length times width times height" in measuring the volume of the non-rectangular-shaped given object. He showed inappropriate application of the volume formula with no reasoning about the length, width, and height measures. Participant 19 (Grade 5) showed evidence of Level 3 understanding in response to Task 2 by counting "all of the cubes" and correctly finding "32." This student iterated equal-sized cubes and correctly structured the iterated units to form the whole object. In response to the same measurement task, Participant 26 (Grade 4) showed Level 3.5 understanding when he counted the eight cubes on the top part of the given object and multiplied this number by the number of sets of eight cubes, thereby iterating efficient composite units to form the whole shape. This same student showed Level 4 understanding in responding to Task 1, an easier task asking to find the surface area of a given cuboid (built from small cubes of same size). He counted the "16" cubes on the front face of the cuboid and counted by 16s to reach the correct answer (i.e., iterating efficient composite units to form the whole object in terms of layer structuring). Participant 30 (Grade 4) demonstrated Level 5 understanding in response to Task 2 by dividing the given stack of cubes into two rectangular-prism-shaped parts (see Figure 5) to do "6 times 2 times 2" and "2 times 2 times 2" and then "add them together because the volume is the whole thing." She applied a volume formula, $L \times W \times H$, twice with regard to the spatial structure of the given irregular-shaped object.

In measuring volume, therefore, Tasks 1, 2, and 3 were successful at eliciting evidence of student understanding at almost all targeted levels of volume measurement LP, excluding Level 2 (see Table 6). These tasks revealed that surface area is by far the most challenging for students, followed by perimeter and volume.

Discussion and Conclusion

In this study, we examined variation in levels of students' geometrical measurement understanding by using the three LP-based cognitive tasks that each measure length, area, and volume understanding. We observed some task effects in the above data, but we also observed that these tasks, as a set, were able to successfully elicit evidence of student understanding across most of the levels within each of length, area, and volume measurement. Through this study, we found evidence of understanding at five successive levels of a geometrical measurement LP within each dimension and provide early validation support for our hypothesized geometrical measurement LP (Kim et al., 2017). Our discussion focuses on three key points inferred from the analyses of the 30 individual students' concurrent responses to the three LP-based cognitive tasks.

First, from third to fifth grade, participating students evidenced variation in levels of thinking and reasoning around the use of (composite, efficient composite) units within each dimension in terms of unit partition, unit iteration, and spatial structuring of the iterated units (see Tables 2–5, and 6). The range of the variation of the levels of geometrical measurement discerned from this study (across the three dimensions) spans from no unit iteration (Level 1) and early unit iteration conception (Level 2, with the exception of volume), to iteration of individual units (Level 3) and composite units (Level 3.5), to iteration of efficient composite units (Level 4), to the application of formulas with reference to the spatial structure of an object being measured (Level 5). The observed level variation may reflect that more abstraction is involved in iteration of composites of units and structuring the iterated composites spatially than iteration of single units. Although we did not observe evidence of student thinking at Level 2 in response to questions about volume, our research design does not allow us to conclusively rule it out. The viability of this level will need to be tested and addressed through a further confirmatory study.

Second, as evidenced by responses representing thinking and reasoning at Level 4 across the three dimensions, some participants showed their understanding about the use of efficient composite units and their perception of the spatial structure of the iterated efficient composite units for perimeter, area, surface area, and volume (see Tables 3–5, and 6). In the previous research on area and volume measurement (Battista, 2003, 2004; Battista & Clements, 1996), iteration of maximal composite units and their spatial structuring have been conceptualized in the contexts of regular-shaped objects (e.g., for the area of a rectangle, iterating of rows or columns of squares regarding its row-by-column structuring, and for the volume of a rectangular prism, iterating of layers of cubes that gives layer structuring). The responses given to Task 2 in which a given object features an irregular shape demonstrate that students in Grades 3–5 can apply the iteration of efficient composite units to irregular-shaped objects. Additionally, the responses given to perimeter measurement tasks in Tasks 1 and 3 allow us to expand the application of the idea of efficient composite units to length measurement, as had previously been done for area and volume (Battista, 2003, 2004; Battista & Clements, 1996).

Third, some of the participants revealed confusion about perimeter, surface area, and volume. Students' confusion between surface area and volume has been reported in earlier literature (e.g., Ben-Haim, Lappan, & Houang, 1985; Hirstein, 1981; Tan Sisman & Aksu, 2016). However, we also observed some participants' confusion between perimeter and area (e.g., counting all the squares on the front face of a given object to find the perimeter of the face); confusion between area and surface area (e.g., multiplying the area of a face by the number of faces of a given object to find the area of the face); confusion between area and volume (e.g., counting the cubes on the front face of a given object and multiplying the number of vertical layers to find the area of the face); and confusion between volume and height/length (e.g., counting all the squares/cubes at a corner of a given object to find the volume of the object). We classified these responses at Level 0 of the LP because they featured no conception, or substantial misconceptions, of the attributes to be measured within each dimension.

We would like to mention three main limitations of the study. First, note that the LP-based tasks were all given at one sitting; therefore, one could say that the students had the opportunity to learn from their performance on the previous tasks and questions, influencing performance on later tasks. However, because successive parts of each task targeted understanding of measurement in other dimensions, any learning from the tasks themselves would have been minimal. Second, using the current method of data analysis, we did not observe consistency in participant performance across the three dimensions; in other words, a number of participants exhibited different levels of understanding among the three dimensions. This observation may reflect the aspects of hierarchic development (the third tenet of *hierarchic interactionalism*, Sarama & Clements, 2009) that describe development as "an interactive interplay among specific existing components of knowledge and processes" (p. 21), where "each level builds hierarchically on the concepts and processes of the previous levels" (p. 21) and with probability fall back to earlier levels in the contexts of "increased task complexity, stress, or failure" (p. 21). Finally, we should recognize that we did not collect data on student opportunities to learn geometric measurement, including instruction in their classrooms. Thus, we cannot speak to potential effects of certain kinds of learning activities on development of geometrical measurement.

The goal of this study was to examine elicited variation among levels of geometrical measurement understanding in each dimension, not to verify a horizontal progression across the three dimensions. From the findings of this study, we inferred only that among length, area, and volume measurement, volume is by far the most challenging for students, a finding that has been observed in prior research (e.g., Curry et al., 2006).

This study was the first occasion in which we collected empirical data on student understanding of geometrical measurement through the lens of this LP. As such, it should be recognized as the first step in a larger research agenda. For example, cross-sectional research often precedes longitudinal research so that instruments can be developed and tested. The fact that student variation in responses was observed consistent with the levels of the LP points to the importance of further longitudinal study to provide an account of the relationships among the LPs for the three dimensions by individual students over time (see Barrett et al., 2017).

From the findings of the current study, we suggest that the geometrical measurement LP can be a practical tool for designing and assessing students' thinking and understanding in geometrical measurement. We also note the potential of our geometrical measurement LP as an instructional resource when the LP is presented to teachers along with the associated cognitive tasks and annotated examples of student response given to the tasks (see also Battista, 2012). In particular, the annotated examples of student response regarding levels of the LP might assist teachers in making sense of differences in performance across levels in the LP and reflecting on how those differences may indicate a need for instruction on particular aspects of geometrical measurement (e.g., use of composite units). Therefore, we argue that the LP can inform assessment of student understanding and communicate important learning research in this domain to teachers (Graf & van Rijn, 2016; Smith et al., 2006; Sztajn et al., 2012) by presenting a set of cognitive tasks with selected student work associated with LPs (Battista, 2012) and thus foster student movement to more sophisticated understanding of measurement for their further learning and everyday life.

Acknowledgments

The authors wish to acknowledge Edith Aurora Graf, Madeleine Keehner, and E. Caroline Wylie for their comments on earlier versions of this manuscript. We are grateful to Dawn Leusner and Stephanie Peters for their assistance with planning and conducting the cognitive laboratory study.

References

Barrett, J. E., Clements, D. H., Klanderman, D. B., Pennisi, S. J., & Polaki, M. V. (2006). Students' coordination of geometric reasoning and measuring strategies on a fixed perimeter task: Developing mathematical understanding of linear measurement. *Journal for Research in Mathematics Education*, *37*, 187–221.

Barrett, J. E., Clements, D. H., & Sarama, J. (2017). Children's measurement: A longitudinal study of children's knowledge and learning of length, area, and volume [Monograph]. Journal for Research in Mathematics Education*, 16*. Retrieved from https://www.nctm.org/Store/Products/JRME-Monograph-16--Children-s-Measurement--A-Longitudinal-Study-of-Children-s-Knowledge-and-Learning-of-Length,-Area,-and-Volume/

Barrett, J. E., Cullen, C. J., Sarama, J., Clements, D. H., Klanderman, D., Miller, A. L., & Rumsey, C. (2011). Children's unit concepts in measurement: A teaching experiment spanning grades 2 through 5. *ZDM: The International Journal on Mathematics Education, 43*, 637–650. https://doi.org/10.1007/s11858-011-0368-8

Battista, M. T. (1999). Fifth graders' enumeration of cubes in 3D arrays: Conceptual progress in an inquiry-based classroom. *Journal for Research in Mathematics Education, 30*, 417–448. https://doi.org/10.2307/749708

Battista, M. T. (2003). Understanding students' thinking about area and volume measurement. In D. H. Clements & G. Bright (Eds.), Learning and teaching measurement: 2003 *Yearbook* (pp. 122–142). Reston, VA: National Council of Teachers of Mathematics.

Battista, M. T. (2004). Applying cognition-based assessment to elementary school students' development of understanding of area and volume measurement. *Mathematical Thinking and Learning, 6*, 185–204. https://doi.org/10.1207/s15327833mtl0602_6

Battista, M. T. (2007). The development of geometric and spatial thinking. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 843–908). Charlotte, NC: Information Age.

Battista, M. T. (2011). Conceptualizations and issues related to learning progressions, learning trajectories, and levels of sophistication. *The Mathematics Enthusiast, 8*(3), 507–569.

Battista, M. T. (2012). Cognition-based assessment and teaching of geometric measurement. *Building on students' reasoning*. Portsmouth, NH: Heinemann.

Battista, M. T., & Clements, D. H. (1996). Students' understanding of three-dimensional rectangular arrays of cubes. *Journal for Research in Mathematics Education*, *27*, 258 – 292. https://doi.org/10.2307/749365

Battista, M. T., & Clements, D. H. (1998). Finding the number of cubes in rectangular cube buildings. *Teaching Children Mathematics, 4*(5), 258 – 264.

Battista, M. T., Clements, D. H., Arnoff, J., Battista, K., & Borrow, C. V. A. (1998). Students' spatial structuring and enumeration of 2D arrays of squares. *Journal for Research in Mathematics Education*, *29*, 503 – 532. https://doi.org/10.2307/749731

Ben-Haim, D., Lappan, G., & Houang, R. T. (1985). Visualizing rectangular solids made of small cubes: Analyzing and affecting students' performance. *Educational Studies in Mathematics*, *16*(4), 389 – 409. https://doi.org/10.1007/BF00417194

Blair, J., & Conrad, F. (2011). Sample size for cognitive interview pretesting. *Public Opinion Quarterly, 75*(4), 636 – 658. https://doi.org/10.1093/poq/nfr035

Clements, D. H. (Ed.). (2003). Learning and teaching measurement (*2003* yearbook). Reston, VA: National Council of Teachers of Mathematics.

Clements, D. H., & Sarama, J. (2004). Learning trajectories in mathematics education. *Mathematical Thinking and Learning, 6*, 81 – 89. https://doi.org/10.1207/s15327833mtl0602_1

Clements, D. H., & Sarama, J. (2014). *Learning and teaching early math: The learning trajectories approach* (2nd ed.). New York, NY: Routledge. https://doi.org/10.4324/9780203520574

Common Core State Standards Initiative. (2010). *Common Core State Standards for Mathematics*. Retrieved from www.corestandards.org

Curry, M., Mitchelmore, M., & Outhred, L. (2006). Development of children's understanding of length, area, and volume principles. In J. Novotná, H. Moraová, M. Krátká, & N. Stehlíková (Eds.), *Proceedings of the 30th Conference of the International Group for the Psychology of Mathematics Education* (pp. 377 – 384). Prague: PME.

Daro, P., Mosher, F. A., & Corcoran, T. B. (2011). *Learning trajectories in mathematics: A foundation for standards, curriculum, and instruction* (CPRE Research Report #RR-68). Philadelphia, PA: Consortium for Policy Research in Education. https://doi.org/10.12698/cpre.2011.rr68

Deane, P., Sabatini, J., & O'Reilly, T. (2012). *The* CBAL™ *English language arts (ELA) competency model and provisional learning progressions*. Retrieved from http://elalp.cbalwiki.ets.org/Outline+of+Provisional+Learning+Progressions

Ginsburg, H. P. (2005). *Entering the child's mind: The clinical interview in psychological research and practice*. Cambridge, UK: Cambridge University Press.

Graf, E. A., & van Rijn, P. (2016). Learning progressions as a guide for design: Recommendations based on observations from a mathematics assessment. In S. Lane, M. R. Raymond & T. M. Haladyna (Eds.), *Handbook of test development* (2nd ed., pp. 165 – 189). New York, NY: Routledge.

Gutiérrez, A., Jaime, A., & Fortuny, J. M. (1991). An alternative paradigm to evaluate the acquisition of the van Hiele levels. *Journal for Research in Mathematics Education*, *22*, 237 – 251. https://doi.org/10.2307/749076

Hirstein, J. J. (1981). The second national assessment in mathematics: Area and volume. *Mathematics Teacher, 74*, 704 – 708.

Kim, E. M., Haberstroh, J., Peters, S., Howell, H., & Oláh, L. N. (2017). *A learning progression for geometric measurement in one, two, and three dimensions* (Research Report No. RR-17-55). Princeton, NJ: Educational Testing Service. https://doi.org/10.1002/ets2.12189

Lehrer, R. (2003). Developing understanding of measurement. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *A research companion to principles and standards for school mathematics* (pp. 179 – 192). Reston, VA: National Council of Teachers of Mathematics.

Lehrer, R., Jaslow, L., & Curtis, C. L. (2003). Developing understanding of measurement in the elementary grades. In D. H. Clements & G. Bright (Eds.), *Learning and teaching measurement* (pp. 100 – 121). Reston, VA: National Council of Teachers of Mathematics.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: Author.

National Research Council. (2007). *Taking science to school: Learning and teaching science in grades K-8*. Washington, DC: The National Academies Press. https://doi.org/10.17226/11625

Piaget, J., Inhelder, B., & Szeminska, A. (1960). *The child's conception of geometry*. London, England: Routledge & Kegan Paul.

Sarama, J., & Clements, D. H. (2009). *Early childhood mathematics education research: Learning trajectories for young children*. New York, NY: Routledge.

Smith, C. L., Wiser, M., Anderson, C. W., & Krajcik, J. (2006). Implications of research on children's learning for standards and assessment: A proposed learning progression for matter and the atomic-molecular theory. *Measurement: Interdisciplinary Research and Perspectives, 4*(1 – 2 1 – 98. https://doi.org/10.1080/15366367.2006.9678570

Stephan, M., & Clements, D. H. (2003). Linear and area measurement in prekindergarten to grade 2. In D. H. Clements & G. Bright (Eds.), Learning and teaching measurement*, 2003* yearbook of the National Council of Teachers of Mathematics (pp. 3 – 16). Reston, VA: NCTM.

Sztajn, P., Confrey, J., Wilson, P. H., & Edgington, C. (2012). Learning trajectory based instruction: Toward a theory of teaching. *Educational Researcher, 41*(5), 147 – 156. https://doi.org/10.3102/0013189X12442801

Tan Sisman, G., & Aksu, M. (2016). A study on sixth grade students' misconceptions and errors in spatial measurement: Length, area, and volume. *International Journal of Science and Mathematics Education 14*, 1293–1319. https://doi.org/10.1007/s10763-015-9642-5

U.S. Department of Education. (2018). *Explore assessment data* [NAEP data tools]. Retrieved from https://nces.ed.gov/nationsreportcard/about/naeptools.aspx

Wilson, M. (2009). Measuring progressions: Assessment structures underlying a learning progression. *Journal of Research in Science Teaching, 46*(6), 716–730. https://doi.org/10.1002/tea.20318

Suggested citation:

Kim, E. M., & Oláh, L. N. (2019). *Elementary students' understanding of geometrical measurement in three dimensions* (Research Report No. RR-19-14). Princeton, NJ: Educational Testing Service. https://doi.org/10.1002/ets2.12250

Action Editor: Heather Buzick

Reviewers: Edith Aurora Graf and Madeleine Keehner