Effects of Three Interventions on Children's Spatial Structuring and Coordination of Area Units

Amanda L. Cullen[1], Cheryl L. Eames[2], Craig J. Cullen[1], Jeffrey E. Barrett[1], Julie Sarama[3], Douglas H. Clements[3], and Douglas W. Van Dine[4]

[1]Illinois State University
[2]Southern Illinois University Edwardsville
[3]University of Denver
[4]Metropolitan State University of Denver

The research reported in this article was supported by the National Science Foundation under Grant No. DRL-1222944. Any opinions, findings, conclusions or recommendations expressed in this report are those of the authors and do not necessarily reflect the views of the National Science Foundation. A subset of the results discussed in this article was included in a presentation at the 2014 Annual Meeting of the American Educational Research Association. The authors would like to thank Drs. John Dossey, Esteban Lizano, Junvie Pailden, Jinfa Cai and his editorial team, and the anonymous reviewers for their contributions to our work.

We examine the effects of 3 interventions designed to support Grades 2–5 children's growth in measuring rectangular regions in different ways. We employed the microgenetic method to observe and describe conceptual transitions and investigate how they may have been prompted by the interventions. We compared the interventions with respect to children's learning and then examined patterns in observable behaviors before and after transitions to more sophisticated levels of thinking according to a learning trajectory for area measurement. Our findings indicate that creating a complete record of the structure of the 2-dimensional array—by drawing organized rows and columns of equal-sized unit squares—best supported children in conceptualizing how units were built, organized, and coordinated, leading to improved performance.

Keywords: Elementary grades; Geometric measurement; Learning trajectory; Spatial structuring

Geometric measurement is an important topic in school mathematics because it has practical applications to daily life, connects to multiple disciplines, and is a specific branch of mathematics that links number and space (Clements & Sarama, 2007). In this study, we examined and supported children's thinking and learning about area measurement in Grades 2 to 5. We take area measurement to mean the quantification of the amount of space within a two-dimensional, planar, closed surface or region (Sarama & Clements, 2009; Weisstein, 2016).

In this study, we extend the research on the learning of area measurement by investigating shifts in children's observable behaviors in response to one of three interventions designed to support children's growth in measuring rectangular regions but in different ways. We used a hypothetical learning trajectory for area measurement (Barrett, Clements, Sarama, Miller, et al., 2017) as a tool to track children's shifts and growth. To date, most learning trajectory-related research has been focused on building (e.g., Blanton, Brizuela, Gardiner, Sawrey, & Newman-Owens, 2015), revising (e.g., Sarama, Clements, Barrett, Van Dine, & McDonel, 2011), and extending (e.g., Barrett, Clements, & Sarama, 2017) learning trajectories. In 2010, Simon et al. asserted that "generally missing from the literature is research that examines the process by which students progress from one of these conceptual steps to a subsequent one" (p. 70). In this study, we incorporate the microgenetic method with a learning trajectory to identify level transitions, which can help teachers and researchers notice important shifts in observable behaviors, anticipate when level transitions are about to happen, and become more efficient at motivating these level transitions.

Review of the Related Literature

Although area measurement is a commonly taught form of geometric measurement (Curry, Mitchelmore, & Outhred, 2006), research indicates that the teaching and learning of area measurement has been inadequate for years (e.g., Bell, Hughes, & Rogers, 1975; National Assessment of Educational Progress [NAEP], 1983, 2016). In 2007, only 42% of fourth graders (9 year olds) taking the Trends in International Mathematics and Science Study (TIMSS, 2016) assessment chose the correct area of a fenced-in region when given the whole number length and width (e.g., a 4-meter by 3-meter rectangle). Although the fourth graders in the United States and England performed slightly better than this international average, only 48% and 44% of their participants selected the correct area, respectively. In contrast, only 28% of fourth graders in Australia made the correct selection. On the 2013 National Assessment of Educational Progress (NAEP, 2016), only 23% of fourth graders in the United States correctly chose the gym floor with the greatest area when given the whole number length and width of four gym floors (e.g., 95 feet by 40 feet). In the same year, only 47% of eighth graders correctly determined the area of a rectangle when given the length of one side and the perimeter of the rectangle: "One side of a rectangle is 14 meters. The perimeter of the rectangle is 44 meters. What is the area of this rectangle?" (NAEP, 2016). Most of the tasks on these large-scale assessments only required children to apply the area formula for rectangles.

Research indicates that assessing children's understanding of area measurement based on their application of a formula is insufficient because children in multiple countries are taught to apply area formulas without understanding (Battista, 2003; Clements & Sarama, 2007; Kamii & Kysh, 2006; Outhred, Mitchelmore, McPhail, & Gould, 2003). Zacharos (2006) asserted that this lack of understanding stems from teaching the formula prematurely. Battista (2003) posited that children's difficulty with understanding the area formula is related to a lack of understanding of the structure: Children multiply the length and width of a rectangle to produce a measure of area without realizing that this product produces an array of rows and columns of identical square units. Stephan and Clements (2003) argued that there is too much focus on procedures for measuring and not enough on the "*big ideas*" (p. 14) of measurement. These researchers agree that understanding area measurement requires the integration of experiences learning about unit concepts and spatial structuring of two-dimensional space.

Unit Concepts

Several researchers (e.g., Stephan and Clements, 2003) have investigated children's difficulties with area measurement in terms of unit concepts, such as unitizing, composing units to create units of units, iterating individual or groups of units, and coordinating units. Below we delineate each of these four unit concepts.

Unitizing. According to Steffe (1991), "Segmenting sensory experience into units is the result of a unitizing activity prior to measuring or to counting that makes these activities possible" (p. 63). In an area measurement context, unitizing requires the identification of a repeatable shape, piece, or object (i.e., the unit) that is part of the whole or region and segments or covers the two-dimensional space well.

Research indicates that children struggle to recognize area units. In their work with children in Grade 2, Lehrer, Jacobsen, et al. (1998) found that children initially selected objects that resembled the shape of the regions to be covered. For example, to cover the region within a traced hand outline, children selected beans, spaghetti, and rope as units of area measure. In another study, Lehrer, Jenkins, and Osana (1998) reported that 43% of the first-, second-, and third-grade participants selected circles as their unit of measure for covering the interior of a closed curve and that 73% of their participants were unperturbed by the suggested use of circles to cover a square region, even when directly asked about the gaps between the circles. Because the circles resembled the closed curve region, they posited that the children were again attending to resemblance rather than space-filling properties. These researchers also argued that even when children begin to recognize that some shapes or objects tessellate or cover regions better than others (i.e., without gaps or overlaps), they may not recognize the need for same size shapes or objects. Lehrer, Jenkins, and Osana (1998) also reported that when asked to measure the area of a square, 55% of their participants used a combination of squares and other shapes such as triangles to cover and thus “used manipulatives as a unit of cover” (p. 155). This is in contrast to consistently using one shape or object as a unit of measure.

In time and with experience or instruction, children can learn about space-covering and space-filling properties and begin to appreciate the square unit as a unit that segments, covers, fills, and tessellates rectilinear regions well (Lehrer, 2003; Lehrer, Jenkins, & Osana, 1998). However, Kamii and Kysh (2006) found that, given traditional instruction, it is not obvious to older children (Grades 4–8) that the square unit is the standard unit for area measurement. In other words, although older children may recognize the need for equal units of area that tessellate or cover space well, they may not recognize the square as the shape of the standard unit for area measurement.

Composing. Research indicates that children initially draw and count individual shapes that increasingly resemble squares when asked to copy an array (Sarama, Clements, Van Dine, et al., 2017), complete a partially-drawn array (Battista, Clements, Arnoff, Battista, & Barrow, 1998; Outhred & Mitchelmore, 2000; Sarama & Clements, 2009), or determine how many tiles (of the size and shape of the one provided) would be needed to cover a region and show how they fit (Miller, 2013; Outhred & Mitchelmore, 2000). With experience or instruction, children transition to thinking about individual squares as units and then to thinking about grouping units together to compose a composite unit (e.g., a row or column). This transition indicates that the child has “begun the coordinating action of seeing a square as both a unit and a component of a unit of units” (Sarama & Clements, 2009, p. 298), which Outhred and Mitchelmore (1992) claimed was a critical step. These composite units may or may not be rows or columns. Children may group units together to compose a nonrow unit of units, such as a partial row or column (Miller, 2013; Sarama, Clements, Van Dine, et al., 2017).

Iterating. Piaget, Inhelder, and Szeminska (1960) asserted that iteration is integral to measuring: “to measure is to take out of a whole one element, taken as a unit, and to transpose this unit on the remainder of the whole: measurement is therefore a synthesis of sub-division and change of position" (p. 3). In development, children's ability to iterate a single unit is an important conceptual advancement over a level of thinking in which they require enough physical tiles to cover a region completely to determine area. Children's early unit iterations may not be mathematically rigorous, but with experience, they modify their actions to minimize gaps and overlaps (Lehrer, 2003; Stephan & Clements, 2003). This development is difficult for children (Barrett, Clements, & Sarama, 2017) because unit iteration involves the repetition of an area unit, either an actual physical object (e.g., a square tile) or a mental image of a unit, which is geometrically translated repeatedly through two-dimensional space to occupy successive locations, always in an adjacent position with one concurrent edge. Later, children transition from iterating individual units to iterating units of units, a more sophisticated and efficient approach.

Coordinating. Moving from thinking about a square as an individual unit to thinking about a square as a component of a row and then to thinking about a square as a component of a row and a column requires a great deal of unit coordination. First, children must learn to coordinate area units within rows and columns (i.e., the child sees a unit as a member of both a row and a column, even as it covers only a single portion of space) by creating identical composite units and aligning composite units so that individual units are also aligned as that composite unit is repeated (Outhred & Mitchelmore, 1992, 2000; Sarama & Clements, 2009). Second, children must coordinate linear and area units to determine the number of area units that will fit along each side (Battista, 2003; Kara et al., 2011; Outhred & Mitchelmore, 1996, 2000; Sarama & Clements, 2009). The coordination of linear and area units involves using the linear dimensions to position area units, which can be taken for granted without consequence when the unit is a square unit (cf. Kara et al., 2011; Miller, 2013). Third, after creating a unit of units or row of units that fits along a side, children must learn to coordinate linear units—using the length and the width of both the unit and the region to be covered—to iterate that row or column of units in the orthogonal direction exhaustively (Outhred & Mitchelmore, 2000; Sarama & Clements, 2009). For example, when asked to cover an 8 cm by 9 cm rectangular region by drawing 2 cm by 3 cm rectangular area units, the child may use the length of the unit (3 cm) and length of the region (9 cm) to determine that three units would fit along that side, building a row of three rectangular units along the bottom side of the rectangle. The child may then repeat that row over the rectangular region, using the linear units of the width of both the unit (2 cm) and the region (8 cm) to determine that four rows would completely cover the region (cf. Barrett, Cullen, et al., 2017; Kara et al., 2011).

Spatial Structuring

Sarama and Clements (2009) defined spatial structuring as "the mental operation of constructing an organization or form for an object or set of objects in space, a form of abstraction, the process of selecting, coordinating, unifying, and registering in memory a set of mental objects and actions" (p. 296) and argued that the mental structuring of a two-dimensional array precedes the meaningful use of the mathematical structures of the array. To structure a two-dimensional array, a child may partition or cut the two-dimensional space into parts, or they may build the two-dimensional region from parts.

Spatial structuring may take a long time to develop, especially with a lack of guided experience (Smith, Males, & Gonulates, 2016). Young children may partition the space into unequal parts, such as by drawing shapes that vary in size and shape (e.g., Outhred & Mitchelmore, 2000; Sarama & Clements, 2009). Their organization of tiles or drawn objects is inadequate and unsystematic, causing them to have difficulty keeping track of what they have counted (e.g., Battista et al., 1998). Older children may partition the space into equal parts by drawing or iterating individual units (e.g., Sarama & Clements, 2009). According to Battista, Clements, Arnoff, Battista, and Barrow (1998), Outhred and Mitchelmore (1992, 1996, 2000), and Sarama and Clements (2009), with experience or explicit instruction, children begin to organize their units initially by utilizing a row structure (i.e., thinking about a unit of units), then a row-and-column structure (i.e., flexibly thinking about units of units—focusing on rows or columns but not both), and eventually an array structure (i.e., thinking about coordinated rows and columns—a unit of units of units). This array structure can be created by breaking down—subdividing the region into an array by drawing parallel row and column line segments—or building up—building an array by establishing a row of unit squares and iterating the row to fill the region. Taken together, these findings suggest that children's area measurement knowledge is supported by their ability to structure space.

Although arrays imply or suggest structure, children's ability to detect and use that structure is dependent on purposeful and repeated experiences (e.g., Battista et al., 1998; Outhred & Mitchelmore, 1992, 1996, 2000, 2004; Sarama & Clements, 2009). Initially, children may produce an array of squares by covering with physical tiles or by drawing parallel row and column line segments without visualizing the spatial structuring of rectilinear regions or the row and column array of squares. In other words, they may organize or structure a two-dimensional space before they can have a well-developed understanding of that structure (Sarama & Clements, 2009).

In their review, Outhred and Mitchelmore (2000) argued that using physical tiles to tile a rectangle "may conceal the very relations they are intended to illustrate" (p. 146). Some manipulatives, such as foam squares or grid-overlays, prestructure an array, allowing children to determine correctly the area of the region without attending to the structure (Lehrer, 2003; Outhred & Mitchelmore, 2000). In other words, children are often able to correctly create an array with square units and imitate the organization of an array by relying on discrete counting of objects or operating on rows or columns of individual units without conceptually understanding the array structure. Children need opportunities to mentally construct, organize, and integrate existing structures into new structures (Battista et al., 1998; Outhred & Mitchelmore, 1992; Sarama & Clements, 2009).

Goals and Research Questions

We sought to extend the research on area measurement by exploring interventions that were designed to support children's understanding of area measurement as a structuring process. We created and compared the effectiveness of three interventions, two of which were designed to support children's growth in measuring rectangular regions using spatial structuring. First, we describe two experimental interventions: (1) subdividing a region into an array by drawing parallel row and column line segments and (2) building an array by establishing a row of unit squares and iterating the row to fill a rectangular space. The third intervention was a comparison intervention in which the children were repeatedly exposed to an arithmetic approach of multiplication of two linear measures without direct appeal to an area unit.

Based on our review of the related literature, we conjectured that both experimental interventions would provide greater support than the comparison intervention. The present study is part of a larger investigation in which we conducted one-on-one sessions with children in Grades 1–5 from the Rocky Mountain and Midwest regions of the United States to evaluate the three interventions. Elsewhere (Clements et al., 2017), we discuss the results from our work with 70 children in Grades 1–3 to illuminate the transition to operating on composite units. In this article, we present the results from our work with 54 children in Grades 2–5 to focus on the transition to using an array structure. The following research questions framed our work:

1. How are children's observable behaviors and numerical responses when measuring rectangular regions affected by repeated exposure to a video demonstration that focuses on either
    a. building an array by establishing a row of unit squares and iterating the row or
    b. subdividing a region into an array by drawing parallel line segments for rows and for columns?
2. What patterns emerge in children's observable behaviors just before and after they shift from measuring area by operating on individual or composite area units to using an array structure?

Theoretical Framework

To answer these questions, we employed a learning trajectory (LT) for area measurement (Barrett, Clements, Sarama, Miller, et al., 2017) to (a) inform the design of instructional tasks by focusing our attention on unit concepts and spatial structuring and (b) provide descriptions of the children's observable actions. According to Clements and Sarama (2007), an LT has three parts: an instructional goal in a mathematical domain, a likely path for learning through levels of increasing sophistication, and the instructional tasks specifically designed to engender the mental processes or actions that support children's progression through those levels.

The LT for area measurement (Barrett, Clements, Sarama, Miller, et al., 2017) utilized in this study is a revised and extended version of Sarama and Clements' (2009) LT for area measurement. The LT for area measurement produced by Sarama and Clements was based on a review of the literature as well as years of their cross-sectional research that included clinical interviews, individual teaching experiments, and classroom-based teaching experiments. The LT for area measurement produced by Barrett, Clements, Sarama, Miller, et al. was based on their retrospective analysis of a multi-site, 4-year longitudinal study of elementary children's developing measurement knowledge, which was funded by the National Science Foundation (for more information on the study, the revisions and extensions made to the initial LT, and justifications of those modifications, see Barrett, Clements, & Sarama, 2017).

We view the LT for area measurement employed in the present study from the hierarchic interactionalist perspective on learning and development (Clements & Sarama, 2007), which is a cognitive theoretical framework that synthesizes empiricism, nativism, and interactionalism. Specifically, this LT for area measurement is related to a key precept of hierarchic interactionalism, which postulates that children progress through domain-specific levels of understanding that build hierarchically out of the concepts and processes that constitute the previous levels. In the present study, the levels of this LT for area measurement served as a tool for measuring children's concept growth.

According to the hierarchic interactionalist perspective, as children progress through the levels of an LT, more than one level is within their reach on any given task. A child has a "dominant" level, yet more and less sophisticated levels need to be considered. This is a central aspect of the theory of hierarchical interactionalism (Clements and Sarama, 2007), and it is consistent with the zone of proximal development (ZPD) described by Vygotsky (1978) and the overlapping waves approach posited by Siegler (2002). On the one hand, behaviors indicative of one level more sophisticated than a child's current dominant level may be in reach in certain contexts or under certain conditions (e.g., with scaffolding and support from another). On the other hand, levels less sophisticated than a child's current dominant level are not abandoned. Children may fall back to make use of behaviors indicative of less sophisticated levels under conditions of increased stress, when confronted with more complex tasks, or when another process fails (Pirie & Kieren, 1994). Similarly, children may reach back (cf. Barrett, Clements, & Sarama, 2017) to make use of concepts and processes that constitute less sophisticated levels when a task can be efficiently and correctly resolved without making use of a more sophisticated level of thinking. Therefore, in the present study when we make a claim that a child is at a particular level of this LT for area measurement, we recognize that for a child to respond "at that level" on a given task or in a given context depends on additional factors. Thus, we are conservative in our claims regarding completed level transitions in the Results section.

The theory of hierarchical interactionalism posits that instructional practices that address the developmental progression are more effective, efficient, and generative for the child than those that do not (Clements & Sarama, 2007). Such instruction based on LTs builds on the hypothesized specific mental objects and actions that constitute children's thinking at a particular level by including the "external objects and actions that mirror the hypothesized mathematical activity of the children as closely as possible" (Clements & Sarama, 2007, p. 466). These characteristics are consistent with, but extend, other theories such as Siegler's (2002) overlapping waves approach, which focuses mainly on strategies and does not include instruction as an integrated component. In the following section we summarize the observable behaviors (hereafter referred to as behaviors) as well as the hypothesized mental actions on objects indicative of the levels of the LT for area measurement that we used to differentiate children's responses in the present study.

Levels of a Learning Trajectory for Area Measurement

The LT for area measurement (Barrett, Clements, Sarama, Miller, et al., 2017) employed in this study includes a developmental progression for how children develop area measurement concepts and spatial structuring schemes. Initially, children cannot or do not organize or structure the two-dimensional space. Instead, they draw approximations of rectangular shapes but leave gaps or overlaps to draw an incomplete covering (Physical Coverer and Counter). In time, children begin drawing complete coverings by drawing approximations of rectangular shapes without gaps or overlaps (Complete Coverer and Counter), which leads to the development of unitizing and iterating unit concepts (Area Unit Relater and Repeater, see Table 1 for this and other levels targeted in this study). This is followed by the transition to building, maintaining, and manipulating a unit of units or a composite unit (Initial Composite Structurer). Next, the coordination of area units within rows and columns emerges, which facilitates the development of a unit of units of units; in time and with experience or instruction, the coordination of linear and area units, the coordination of linear dimensions of the area unit and the region, and a global scheme for creating, organizing, and operating on an array are developed (Area Row and Column Structurer). The transition into the subsequent level (Array Structurer) marks a shift in the focus of the LT for area measurement from emphasizing spatial structuring and unit concepts to the development of increasingly sophisticated logical thought, reflection, explanation, and justification in geometric measurement situations (i.e., why multiplication creates a measure of area).

[Insert Table 1 here]

In our prior work in which we compiled longitudinal accounts of children's thinking and learning about geometric measurement from Grade 2 to Grade 5, we found that five out of seven children plateaued at the Initial Composite Structurer level for 12–30 months (Barrett, Cullen, et al., 2017). These participants were able to build, maintain, and manipulate a composite unit to structure an array (indicating that they were at least at the Initial Composite Structurer level) but struggled to coordinate linear and area units (thus not yet at the Area Row and Column Structurer level) during Grades 3–5. We considered this a plateau because the children transitioned into and out of other levels within a 6–12 month time span. This made us wonder if the transition from the Initial Composite Structurer level was more complicated than other transitions, if a level was missing, or if instruction needed to have more conceptual supports. We designed the present study, in part, to investigate ways of shortening this transition from the Initial Composite Structurer level to Area Row and Column Structurer level.

Method

Because we wanted to describe the transition from less sophisticated levels into the Area Row and Column Structurer level (and using an array structure), we designed a study that would allow us to observe children as they made this shift. To guide the design of these observations, we employed the microgenetic method (Siegler & Svetina, 2006). This method allows for the study of the circumstances preceding a conceptual change, the change itself, and the potential generalizability of the results beyond the context of the present study (Siegler & Crowley, 1991). It is historically rooted in the work of Heinz Werner and Lev Vygotsky who argued that change can be motivated through focused experiences and that change can be observed (Vygosky, 1978; Werner, 1925; as cited in Siegler & Crowley, 1991). There are three main characteristics of the microgenetic method: a) observations span the entire period from the beginning of the change to the time at which it reaches a relatively stable state; b) the density of observations within this period is high, relative to the rate of change; and c) observations of the changing performance are analyzed intensively to indicate the processes that give rise to them (Siegler, & Svetina, 2006).

Participants

To identify children not yet at the Area Row and Column Structurer (ARCS) level—specifically, children at the Area Unit Relater and Repeater (AURR) and Initial Composite Structurer (ICS) levels—we recruited children in 17 Grade 2 to 5 classes from two school districts in the Midwest (three classes each from Grades 2 and 3, six classes from Grade 4, and five classes from Grade 5). Both school districts had adopted the Common Core State Standards in Mathematics (CCSSM, National Governor's Association for Best Practices & Council of Chief State School Officers, 2010) and thus taught area during Grades 3–5. The Grade 2 children had not yet been formally exposed to the rectangular area formula, but the children in Grades 3–5 had. None of the teachers of participating classrooms were teaching measurement lessons during the study; therefore, the participants did not receive instruction about area measurement between our sessions.

Initial Screening Instrument

We administered a four-item LT-based screening instrument (see Appendix A) to all of the children (n = 240) in the participating classrooms. For Part 1 (Items 1 and 2), children did not have access to a ruler. After Part 1 was collected, we distributed Part 2

(Items 3 and 4) and rulers. This instrument included items from our earlier work that were designed to elicit different behaviors indicative of levels described in the LT for area measurement (Barrett, Clements, Sarama, Miller, et al., 2017).

To analyze the responses, we distributed the children's screening instruments among the researchers, each of whom had a minimum of 3 years (at the time) experience working with the LT levels and descriptions and identifying levels of thinking using that LT. We used Item 1 to analyze children's conceptions of area and area measurement through their definition of area. We used Items 2–4 to identify an LT level per item. On Item 2, we analyzed the children's drawings because the prompt elicited a range of responses that are describable by language in the LT.

For example, children exhibiting thinking at the Complete Coverer and Counter (CCC) level show an understanding that they must cover the entire region. John, a second-grade child, drew mostly closed individual squares using existing squares to guide his placement to cover the region (see Figure 1a). However, John had errors in the alignment of the squares and did not show that he recognized the need for equal sized units. Thus, we interpreted John's drawing as illustrating thinking at the CCC level but not yet at the AURR level.

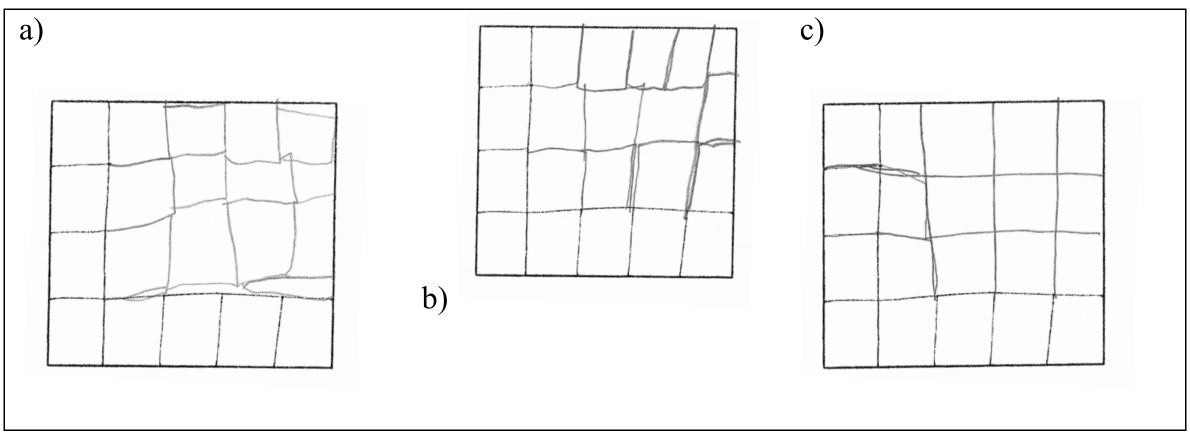


Figure 1. Three second grade children's drawings on Item 2: a) John's drawing indicating that he was not yet utilizing behaviors indicative of the AURR level, b) Elizabeth's drawing illustrating the AURR level, and c) Micah's drawing demonstrating thinking at the ICS level.

As children progress to the AURR level, they can still have some alignment errors but draw a complete covering. However, they attend to drawing equal-sized units, one at a time. Elizabeth, a second-grade child, illustrated thinking at the AURR level when she drew individual, approximately equal-sized units without gaps or overlaps (see Figure 1b). Because her placement of these units was strongly guided only by the previously drawn adjacent units, we take this as evidence that she was using the intuitive structure of a row as a marker to guide her drawing actions. However, Elizabeth did not use a row as a unit of units (e.g., she did not curtail the process to produce rows of individual units) nor did she show evidence that she understood that each row should have the same number of units, which is indicative of the next level.

Children operating at the ICS level can identify a square unit as both a unit and a component of a unit of units (e.g., a row, column). They can apply this unit of units repeatedly but not exhaustively. They may draw several rows by using line segments but then revert to drawing individual squares. Or they may begin drawing individual squares and then curtail this process to draw line segments to indicate rows. They also understand that each row must have the same number of units. Micah, a second-grade student, alternated between drawing individual units and completing rows and columns with line segments, as indicated by places in which he picked up his pencil and drew over existing marks (see Figure 1c).

The ARCS level is an advancement over the ICS level because children have moved from identifying squares as individual units to seeing the square also as a component of a row and a component of a column (i.e., coordinating area units). They also use the dimensions to constrain the unit size—the length of a side of a rectangle indicates the number of area units that will fit along that side. Because the dimensions of the rectangle in Item 2 are indicated by the printed row and column, we could not use this task to determine if children were operating at the ARCS level. At most, we could claim that they were operating at least at the ICS level. We used Items 3 and 4 to determine if they were already operating at the ARCS level. We wanted to know if they could use the dimensions (the lengths of two orthogonal sides) to determine how many units fit along those sides (i.e., coordinate linear and area units) as well as use the length and the width to constrain the unit size (i.e., coordinate linear units; see Coordination section in our Review of the Related Literature for more information on coordinating area units, coordinating linear units, and coordinating linear and area units).

Children not yet at the ARCS level are unable to produce a rectangle that has an area of 8 square inches and show how they fit when only given a ruler, writing utensil, and blank piece of paper for Item 3. They would also struggle to show how 10 square inches fit within a drawn rectangle when told the area is 10 square inches for Item 4. For example, John, who showed thinking at the CCC level on Item 2, did not draw 10 shapes (see Figure 2a). We interpret his line segments to indicate that he subdivided one side into centimeter length units. He did not produce a complete covering of units. However, because we used the word "area" in the prompt, this may indicate that he does not connect area with covering nor does he have an understanding of unit. Thus, we take this as evidence that he is not yet at the AURR level.

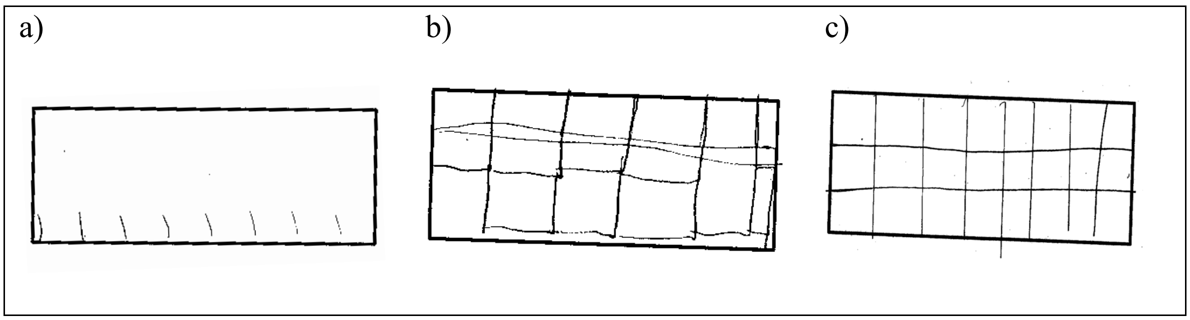


Figure 2. Three second grade children's drawings on Item 3: a) John's drawing showing that he is not yet at the AURR level, b) Elizabeth's drawing illustrating the AURR level, and c) Micah's drawing demonstrating thinking at the ICS level.

In contrast, Elizabeth, who provided evidence that she was at the AURR level on Item 2, drew 10 shapes individually but seemed to struggle to attend to unit size as well as collinearity of rows (see Figure 2b). We also do not have evidence that she used the length of a side to determine how many units would fit along that side. We take this as additional evidence that she was operating at the AURR level.

Micah continued to show evidence of thinking at the ICS level (see Figure 2c). He drew parallel row and column line segments to completely cover the rectangle with approximately equal sized units. However, he did not use the dimensions to determine how many units fit along those sides nor use the length and the width to constrain the unit size. We used the LT codes for Items 2–4 to determine each potential participant's dominant or summary LT level based on his or her responses across all of the items on instrument.

From the group of 240 children who took the screening instrument, we selected a total of 54 participants, all initially placed at the AURR (n = 24) and ICS levels (n = 30). Note that the purpose of identifying initial placements using the four-item LT-based screening instrument was to select participants who were at least at the AURR level but not yet at the ARCS level. We wanted to observe and describe conceptual transitions (from AURR into ICS, from AURR to ARCS, or from ICS to ARCS) and investigate how they may have been prompted by one of three interventions. John and other children who were not yet exhibiting thinking at the AURR level were not included among the 54 participants in the study. Children who exhibited thinking at the ARCS level on the initial screener were also not included in the study.

Using block random assignment, we divided the 54 participants into three groups of 18 participants and created three homogeneous groups in terms of the relevant attributes (i.e., grade and LT for area measurement levels). Each intervention group (n = 18) consisted of eight children initially placed at the AURR level, two from each grade (Grades 2–5) and 10 children initially placed at the ICS level, two per grade in Grades 2–4 and four in Grade 5.

Procedure

All of the 54 children participated in three 10- to 20-minute one-on-one sessions with a member of the research team. For each participant, these three sessions occurred on three separate days, and the mean time elapsed between the first and third sessions was 3.5 school days with a maximum of 6 school days. Every session was videotaped and was conducted during the school day in the child's school during the spring semester.

Each of the three sessions consisted of three trials, for a total of nine trials. Each trial consisted of a single task-intervention pair (Siegler & Crowley, 1991): finding the area of a given rectangle (task) and then watching a video that corresponded to their intervention group and the specific rectangle (intervention).[1]

[1] Simon et al. (2010) also utilized microgenetic methods by posing a sequence of instructional tasks that increased in complexity (from less to more sophisticated). This contrasts with our approach of repeated presentation of the task-intervention pair, varying only the rectangle's dimensions, to provide participants with opportunities to reflect on

Throughout each session, children were provided access to a standard 12-inch ruler, seven foam square-inch tiles, and a roll of transparent tape. Each of these tools was purposefully selected. First, the ruler was provided as a tool that would help the children identify linear units and therefore provided the children with an opportunity to coordinate linear and area units. Second, seven foam square-inch tiles were provided for two reasons: (1) none of the nine rectangles used in the study had an area less than 12 square inches, prohibiting the children from using a cover and count all strategy, and (2) none of the rectangles had a length or width of more than seven inches, allowing the children to build (and possibly iterate) a row for any given rectangle. Each trial began with the interviewer giving the child a pen and a rectangle (with side lengths of at least two and at most seven inches) printed on a piece of paper. Similar to the second condition of Nunes, Light, and Mason's (1993) study, each participant was provided with a ruler as well as the restricted number of area units (i.e., for the present study, seven foam square-inch tiles). In contrast to the Nunes et al. study, we did not provide them with a unit of units or row of area units glued together, yet we did provide them with the roll of transparent tape and the opportunity to create their own unit of units.

Our design decisions about the size of the nine rectangles used in this study as well as the order in which they were posed were also informed by prior research. To control for a possible effect of computational complexity on children's responses (cf. Vasilyeva, Ludlow, Casey, & St. Onge, 2009), we categorized the nine rectangles used in the study by area into three size groups: small, ranging from 10 to 12 square inches; medium, ranging from 15 to 24 square inches; and large, ranging from 24 to 30 square inches. Within each session we randomly assigned to each child one small, one medium, and one large rectangle. We randomly assigned the order in which we presented these rectangles with each of the rectangles presented to each of the 54 children exactly once.

The interviewer asked the child, "What is the area of this rectangle? You may use any of the tools on the table here to help you. Please write on the page while you think." If the child did not draw to produce her answer, the interviewer prompted the child to draw by stating, "Please show me how the <insert child's answer> fit" (e.g., If the child

---

repeated experience as well as provide researchers with the fine-grained detail needed to examine shifts in children's trial-by-trial behaviors.

provided a numeric answer of 15 without drawing, the interviewer responded with, "Please show me how the 15 fit"). We included this prompt because research indicates that a child's drawings of rectangular arrays, both their process and product (Miller, 2013), can provide insight into the child's thinking (e.g., Battista et al., 1998; Outhred & Mitchelmore, 2000; Stephen & Clements, 2003). After the child responded, the interviewer showed the child the instructional intervention video corresponding to the intervention group to which the child had been assigned (described below) and to the rectangle the child had just completed. Each child watched nine videos, one per rectangle, to provide repeated exposure to their assigned intervention.

Although each set of intervention videos was different (which we discuss below), there were some characteristics common to all of them. First, we delivered each intervention through short videos (25 s to 1 min 14 s) shown to individual participants on a laptop computer.[2] Second, in each intervention video the teacher used a ruler to measure the lengths of two adjacent sides of the rectangle and recorded those lengths outside of the rectangle. Third, in each video, the teacher provided the correct numerical answer (e.g., 10 square inches) to provide feedback in the form of knowledge of the correct response for the specific rectangle they were measuring in that trial.

Subdivision Intervention. We designed one of the interventions, hereafter referred to as the Subdivision Intervention, to emphasize coordinating linear and area units, coordinating linear dimensions of the area unit and the region, and subdividing (i.e., partitioning to establish the area unit). These intervention videos consisted mainly of a teacher using a ruler to draw parallel row and column line segments—in orthogonal directions, vertically then horizontally—to create a drawn array of square inches and reporting the correct area of the given rectangle. First, the teacher measured the length of one side of the rectangle, recording the length outside of the rectangle. Then she placed the ruler along that edge and used a pen to draw (freehandedly) a line segment extending across the region from each numbered tick mark on the ruler to the opposite side of the rectangle. While drawing this set of line segments, the teacher said, "This side is *a* inches

[2] In each video, the video camera was focused on the tabletop and depicted only the teacher's hands as she worked with the tools. Hence, her voice could be heard, but her face was not shown.

so that makes *a* rows [or *a* columns].” This process was repeated on an adjacent side thus creating an array of square inches. Finally, in the videos, the teacher skip counted along the rows. For example, in the case of the 3 × 7 rectangle, the teacher said, “So that makes 7, 14, 21 square inches.” Hence, the Subdivision Intervention videos illustrated big ideas of measurement, including spatial structuring and the unit concepts of unitizing and coordinating units.

Iteration Intervention. We designed a second intervention, hereafter referred to as the Iteration Intervention, to support the transition from covering with individual units to building, maintaining, and manipulating a unit of units or a composite unit as well as coordinating area units within rows and columns. The Iteration Intervention videos consisted of a teacher first measuring the length of a side, making tick marks at each of the inch markings from the ruler along the side, and iterating a single tile while saying, “So that makes 1, 2, 3… rows.” She then used the ruler to measure an adjacent side and placed a collection of tiles along that side while saying, “So that makes 1, 2, 3… in a row.” Next, the teacher taped the row of tiles together and iterated the taped row up through the tick marks on the adjacent side while skip-counting by the number of tiles in the row (e.g., “So that makes 7, 14, 21 square inches”). Hence, the Iteration Intervention videos also highlighted big ideas of measurement, including spatial structuring as well as unit concepts—unitizing, composing units to create units of units, iterating groups of units, and coordinating units.

Comparison Intervention. To help isolate the key features of the experimental interventions, we included a third intervention, which offered no visually supportive structural display and did not highlight any of the big ideas of measurement. We refer to this intervention as the Comparison Intervention. Similar to the experimental interventions, the teacher in the Comparison Intervention videos provided feedback to the child by reporting the correct linear measures of a pair of adjacent sides as well as the correct area of the rectangle measured in that trial. However, in contrast to the Subdivision and Iteration Intervention videos, the Comparison Intervention videos did not include a visual display of the structuring. For example, in the case of the 3 × 7 rectangle, the teacher demonstrated measuring two adjacent sides with a ruler, recording the length of each, and then reported and wrote the area. In the video the teacher stated, "This side is 3 inches, this side is 7 inches, so the area is 21 square inches." The inclusion of this intervention allowed us to examine the effects of repeated exposure to a rectangle area task and to check for the effect of modeling a procedure to multiply the measures found for the length and width to produce a measure of area. However, the teacher's use of an arithmetic approach of multiplying two linear measures was not explicitly identified.

Data Sources

The data set included the written LT-based screening instrument from the sample of 240 children, as well as researcher notes and participant written work taken from each of the nine trials with 54 participants. We used one video camera focused on the tabletop in front of the child to capture verbal responses, drawings, gestures, and use of tools. We organized the children's written work and researchers' notes per child, per session, and per trial. We present our data analysis and findings in two phases, one for each research question.

Phase 1

Research Question 1 asked: How are children's observable behaviors and numerical responses when measuring rectangular regions affected by repeated exposure to a video demonstration that focuses on either a) building an array by establishing a row of unit squares and iterating the row, or b) subdividing a region into an array by drawing parallel row and column line segments? To investigate the children's responses, we identified and examined shifts in their behaviors and numerical answers within and across the three intervention groups (including the Comparison Intervention, which offered no support for spatial structuring or unit concepts beyond the implicit procedure of multiplication for length and width).

Data Analysis

We coded the children's responses for each of the nine trials for both correctness and according to the levels of the LT for area measurement[3] (Barrett, Clements, Sarama, Miller, et al., 2017). Across all three interventions, we determined correctness based on the final numeric answer, which may or may not have included correct identification of the unit (square inches).[4] When coding for levels of the LT for area measurement, only three levels were germane to this study (see Table 1 for level descriptions). At times, the child's level of thinking was unclear, such as when the child did not produce a drawing and provided only a numeric answer. This is an example of when we made no level claim, hereafter referred to as "No claim."

To control for the effect of a potential covariate on the main dependent variable (i.e., the LT level exhibited on each trial by each child on correctness), we employed a mediation analysis adapted for categorical variables (Iacobucci, 2012). This mediation analysis technique allows for a combination of logistic and ordinary least squares regression models to test a hypothetical process or mechanism through which an independent variable, trial number (T), might elicit a dependent variable, correctness (C), through a mediating variable, the level of the LT for area measurement observed for each child for each trial (L). Our mediation analysis procedures involved the development, evaluation, and synthesis of three intermediate models for each intervention group (see Figure 3).

[3] Although it is possible for children to produce a correct numeric answer while using behaviors indicative of the AURR or ICS level (i.e., levels below the ARCS level), such behaviors are less sophisticated, in part because they are less efficient. Tracing one tile or drawing individual units to cover (AURR) takes much longer than building and repeating a composite unit with a mixed drawing strategy (ICS), which in turn takes longer than drawing parallel rows and column segments to produce an array of individual units (ARCS). Our design of repeated trials in a short amount of time has the potential to prompt children to adopt more efficient (and more sophisticated) behaviors due to the tediousness and the time and effort required by less sophisticated behaviors.

[4] We accepted numeric answers as evidence of children's claims for a measure of area given the children's tendencies to abbreviate their verbal report by neglecting to name a unit. We note this assumption as one limitation of the present study.

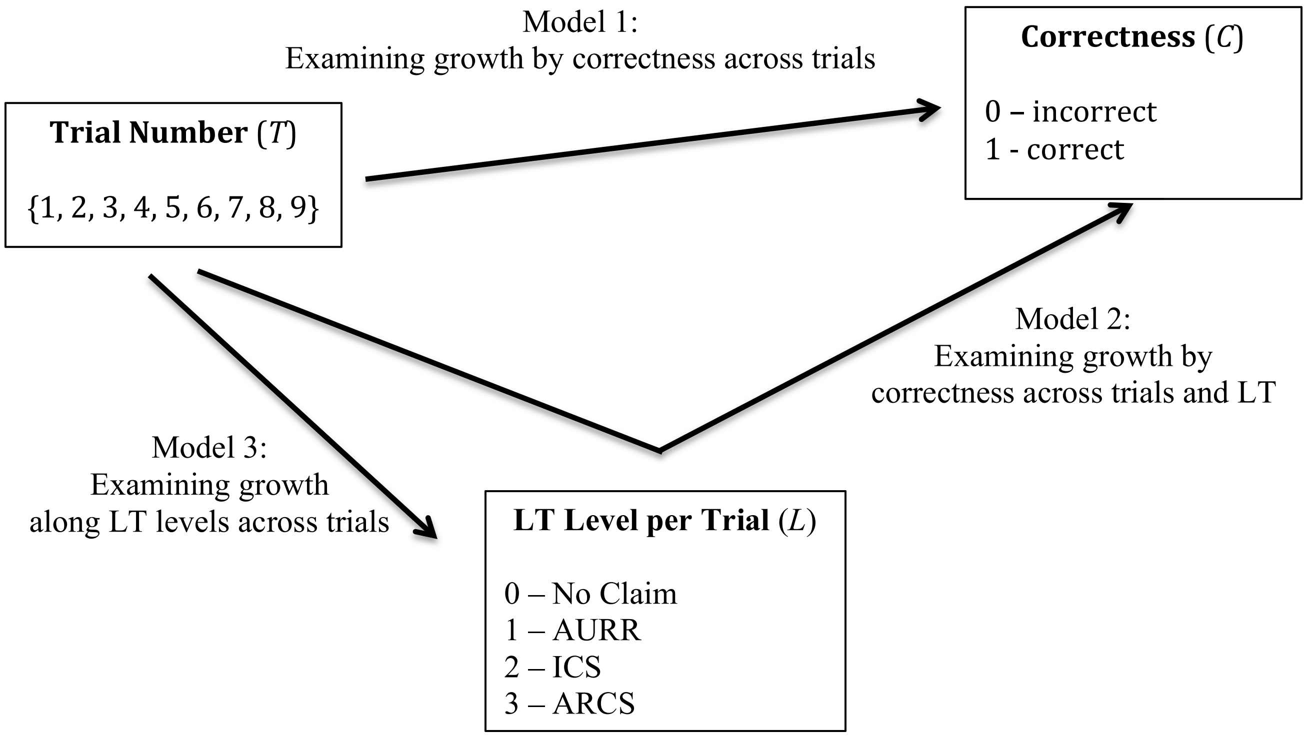


Figure 3. Illustrating model development procedures.

Additionally, we conducted a series of two-proportion *z* tests to investigate differences in the cumulative number of each LT level observed for each intervention group. Note that data from all 54 participants were included in Phase 1.

Findings for Phase 1

Models 1 and 2: Examining growth by correctness across trials or LT levels. A logistic regression was conducted to examine the direct relationship between trial number (T) and correctness (C) for each of the three intervention groups, and results are presented under Model 1 of Table 2. For each of the three intervention groups, a significant association between trial number (T) and correctness (C) was observed. Increases of 41%, 36%, and 18% in the odds of providing a correct answer with each subsequent trial (T) were observed for the Subdivision, Iteration, and Comparison Groups, respectively. Model 2 adds the levels of the LT for area measurement (L). The levels of the LT for area measurement (L) had significant associations with correctness

(C) for the Iteration and Comparison Groups, but this association was not significant for the Subdivision Group.

Table 2

Logistic Regression of Trial Number or LT Level on Correctness for Each Intervention Group

	Independent variable	Model 1 Odds ratio	Model 1 Standard error	Model 2 Odds ratio	Model 2 Standard error
Subdivision Group	Constant	0.95	.39	0.70	.31
	Trial Number (T_S)	1.41**	.13	1.32**	.13
	Trial LT Level (L_S)	-		1.41	.27
	Pseudo R^2	.10		.13	
	Model χ^2	16.49, $df = 1, p < .0001$		19.75, $df = 2, p < .001$	
Iteration Group	Constant	1.16	.47	0.49	.25
	Trial Number (T_I)	1.36**	.12	1.28**	.12
	Trial LT Level (L_I)	-		2.0**	.49
	Pseudo R^2	.08		.13	
	Model χ^2	13.02, $df = 1, p < .001$		20.94, $df = 2, p < .0001$	
Comparison Group	Constant	0.88	.31	0.44*	.18
	Trial Number (T_C)	1.18*	.08	1.19*	.08
	Trial LT Level (L_C)	-		2.00***	.36
	Pseudo R^2	.03		.12	
	Model χ^2	6.37, $df = 1, p < .05$		24.41, $df = 2, p < .0001$	

* significant at a level of $p < .05$
** significant at a level of $p < .01$
*** significant at a level of $p < .001$

A key assumption underpinning Model 2 is that, within each intervention group, trial number (T) and trial LT level (L) are both independent variables that affect the probability that a child will respond correctly. However, with each subsequent trial, children in the Subdivision and Iteration Groups were exposed to instruction designed to support their growth along the LT for area measurement. Thus, we conjectured that a child's predominant LT level (L) is dependent upon the trial number (T) and intervention group the child was assigned to.

Model 3: Examining growth along the LT for area measurement across trials. To test this conjecture, we calculated a simple linear regression to predict the level of the LT for area measurement (*L*) exhibited by each child based on trial number (*T*) for each intervention group. For the Subdivision and Iteration Groups, significant regression equations were found ($F(1, 160) = 26.91$, $p < .0001$ and $F(1, 160) = 11.25$, $p < .01$), with R^2 values of .14 and .06, respectively. A significant regression equation was not found for the Comparison Group. Children's LT level (*L*) is predicted by the equations $L_S = 1.09 + .17T$ and $L_I = 1.33 + .08T$ for the Subdivision and Iteration Groups, respectively. A *t* test used to compare the slopes of the regression equations for the Subdivision and Iteration Groups revealed a significant difference with $t = 2.16$, $p < .05$. This suggests that, for the Subdivision Group, children's LT level increased by .17 for each subsequent trial number[5]. For the Iteration Group, children's LT level increased by .08 levels per trial. These findings suggest that the children in the Subdivision and Iteration Groups exhibited concept growth along the LT for area measurement, with children in the Subdivision Group growing significantly faster and demonstrating an overall higher gain along the LT than the children in the Iteration Group. Children in the Comparison Group did not exhibit significant growth along the LT for area measurement across the nine trials.

A synthesis of models: Measuring the mediation effect. Because children in the Comparison Group did not exhibit significant growth along the LT for area measurement, we examined the mediation effect of the LT (*L*) on correctness (*C*) only for the Subdivision and Iteration Groups. To examine the mediating effect of the LT for area measurement (*L*), we calculated $z_{\text{mediation}}$ (Iacobucci, 2012)[6]. These calculations yielded $z_{\text{mediation}} = 3.65$, $p < .001$ and $z_{\text{mediation}} = 2.55$, $p < .05$ for the Subdivision and Iteration Groups, respectively. These findings suggest that the independent variable of trial

---

[5] The levels of the LT for area measurement are discrete and hierarchically ordered. However, we treated the LT as a continuous scale in Model 3 for the purpose of comparing the effects of the interventions on children's growth along the LT. Some may argue that this is a potential limitation of our model development approach, but we see this as analogous to Iacobucci's (2012) treatment of rating scales as continuous.

[6] This involves computing the standardized elements for the trial number (*T*) parameter in Model 3, z_{T3}, and LT level parameter (*L*) in Model 2, z_{L2}, by using the parameter estimates and their standard errors. Determining $z_{\text{mediation}}$ then consists of calculating the ratio of the product of these standardized elements, $z_{\text{T3}}\, z_{\text{L2}}$, to their collected standard error, $\sqrt{z_{\text{T3}}^2 + z_{\text{L2}}^2 + 1}$.

number (*T*) elicits correctness (*C*) indirectly through a mediating variable, which is the trial LT level (*L*) for the Subdivision and Iteration Groups.

Comparing by concept growth. Pairwise two-tailed, two-proportion *z* tests[7] revealed that the Iteration Intervention prompted significantly more instances of ICS-level behaviors over the nine trials than the Subdivision and Comparison Interventions, with $z = 6.80$, $p < 0.001$ and $z = 7.49$, $p < 0.001$, respectively. Because our sampling method produced intervention groups with the same number of AURR and ICS level children at the beginning of the study, this finding suggests that the Iteration Intervention may be effective in promoting growth into the ICS level of the LT for area measurement. The Subdivision Intervention supported significantly more ARCS-level behaviors over the nine trials than the Iteration and Comparison Groups, with $z = 7.44$, $p < 0.001$ and $z = 6.10$, $p < 0.001$, respectively. We also observed significantly fewer instances of no level claim in the Subdivision and Iteration Groups than in the Comparison Group, with $z = 5.68$, $p < 0.001$ and $z = 7.03$, $p < 0.001$, respectively. Furthermore, we observed significantly more instances of no level claim than AURR (with $z = 7.41$, $p < 0.001$), ICS ($z = 4.63$, $p < 0.001$), or ARCS ($z = 6.44$, $p < 0.001$) level claims for the Comparison Group. These findings suggest that the Comparison Intervention was not effective in eliciting behaviors that could be associated with the concepts or mental actions on objects described in the LT for area measurement.

Phase 2

Research Question 2 asked: What patterns emerge in children's observable behaviors just before and after they shift from measuring area by operating on individual or composite area units to using an array structure? Thus, during Phase 2 we studied our set of dense observations "before the change began and…continue[d] until a point of relative stability was reached" (Siegler & Crowley, 1991, p. 607), which is a tenet of the microgenetic method.

[7] To control for a potential Type I error, a Bonferroni adjusted *p*-value of $p = 0.005$ was used as the significance criteria for the sequence of pairwise two-tailed, two-proportion *z* tests reported here.

Data Analysis

To account for and make sense of these changes observed in Phase 1, we performed a qualitative, comparative analysis during Phase 2 (Corbin & Strauss, 2008). The comparative analysis consisted of constant comparisons, a process of watching one child's videos from the three sessions, describing the behaviors we observed, creating phrases (codes) to describe these behaviors, and creating categories to group codes.[8] We repeated this process with two researchers independently coding each child's videos, measuring interrater reliability, discussing code discrepancies, and modifying codes and code descriptors until the interrater reliability between all pairs exceeded 80%.[9] We then used this initial list of codes to independently code three children, with each child being coded by two researchers. After coding these children independently, we met to discuss code discrepancies, modifying code descriptors to increase consistency and clarity when necessary. At that point, we deemed the code list as final and deleted the coding sheets for the previously coded children. These children remained in the list of 54 children to be coded.

We distributed the 54 participants among three researchers, maintaining efforts to distribute children from each intervention equally among the researchers. At least two researchers coded each child's data. We triple coded a child's data (n = 6) when interrater reliability was less than 70% between any two coders. Hence, all three researchers agreed upon each coding decision when there was a discrepancy.

Relevant Findings for Phase 2: Behavior Shifts and Level Transitions

Although the LT was applicable to all of the participants because their behaviors, behavior shifts, and level transitions exhibited during the study were consistent, 25

---

[8] These codes reflected observable behaviors and inferred "mental actions on objects" that spanned multiple levels of the LT for area measurement. Assignment of these codes occurred simultaneously with assignment of the correctness and level placement codes.

[9] We evaluated reliability through interrater reliability measures using percent agreement: R = number of agreements / (number of agreements + number of disagreements) × 100. In examining our interrater reliability, we did not correct for chance agreement for two reasons. First, there were a large number of codes, and the likelihood for two coders to agree by chance was low. Second, three researchers met during each code discrepancy meeting, even if one researcher did not participate in the coding of a particular child's data.

children did not exhibit relatively stable growth—placed at a specific level at least three times on subsequent trials (Siegler & Svetina, 2006)—during the time of the study (see Figure 4).

Subdivision Intervention

	T1	T2	T3	T4	T5	T6	T7	T8	T9
S1						C	C	C	C
S2		C	C	C	C	C	C		C
S3			C	C	C		C	C	C
S4	C	C	C	C		C		C	C
S5	C	C	C			C		C	C
S6	C		C	C	C	C	C	C	C
S7	C		C	C	C	C	C	C	C
S8	C	C	C	C	C	C	C	C	C
S9	C				C	C	C	C	C
S10		C	C	C	C	C			C
S11		C	C	C	C	C	C	C	C
S12			C	C	C	C	C	C	C
S13	C		C	C	C	C	C	C	C
S14			C		C	C	C	C	C
S15	C	C	C	C	C	C	C	C	C
S16		C	C	C	C	C	C	C	C
S17	C	C	C	C	C	C	C	C	C
S18	C	C	C	C	C	C	C	C	C
	10	10	16	14	15	17	15	16	18

Iteration Intervention

	T1	T2	T3	T4	T5	T6	T7	T8	T9
I1			C		C	C	C	C	C
I2			C	C	C	C		C	C
I3			C		C	C	C	C	C
I4	C	C	C	C	C	C	C	C	C
I5	C	C	C	C	C	C	C	C	C
I6		C	C	C		C	C	C	C
I7	C			C	C	C	C	C	C
I8	C	C	C	C	C	C	C	C	C
I9	C	C	C	C	C	C	C	C	C
I10	C	C	C	C	C	C	C	C	C
I11	C	C	C	C	C	C	C		C
I12						C		C	C
I13							C	C	
I14	C	C	C	C	C		C	C	C
I15	C	C	C	C	C	C	C	C	C
I16	C	C	C	C	C	C	C	C	C
I17	C	C	C	C	C	C	C	C	C
I18		C	C	C	C	C	C	C	C
	11	12	15	14	15	16	16	17	17

Comparison Intervention

	T1	T2	T3	T4	T5	T6	T7	T8	T9
C1					C	C		C	C
C2			C	C	C	C	C	C	C
C3	C	C	C	C		C	C	C	C
C4									
C5	C	C	C	C	C	C	C	C	C
C6									
C7		C	C	C	C	C		C	C
C8	C	C	C	C	C	C	C	C	C
C9						C	C	C	C
C10	C	C	C	C	C	C	C	C	C
C11	C								
C12	C	C	C	C	C	C	C	C	C
C13	C	C	C	C	C	C	C	C	C
C14									
C15		C	C	C	C	C	C	C	C
C16		C	C	C	C	C	C	C	C
C17	C	C	C	C	C	C	C	C	C
C18		C	C	C	C	C	C	C	C
	8	11	12	12	12	14	12	14	14

Figure 4. LT level placement and correctness by trial per intervention group. Within each subfigure, participants are represented by subrows: S1–S18 for Subdivision Group; I1–I18 for Iteration Group; and C1–C18 for Comparison Group. The nine trials are indicated by each of the nine subcolumns within each subfigure as T1–T9. LT level placement per trial is indicated by shading: darkest shading is for ARCS, second darkest shading for ICS, third darkest shading for AURR, lightest shading for levels less sophisticated than AURR, and no shading for No claim. Correctness per trial is indicated by text in the cell: C for Correct or no text for Incorrect numeric response.

No level transitions and no behavior shifts. Ten of the 25 children who did not exhibit a change in their level of thinking across trials were also consistent in their trial-by-trial behaviors. Nine consistently used a ruler or tiles to measure side lengths and then reported a numeric answer (correct or incorrect) without drawing, even when asked to show how they fit. We were unable to make a level claim based on the children's written, verbal, or nonverbal responses for most of the trials (eight or nine) for these nine children. One of the 10 children consistently drew individual units by tracing one tile, providing a correct numeric answer and correct drawing for all trials—although there was some alignment and spacing issues. This child demonstrated behaviors indicative of the AURR level on all trials.

No level transitions but small behavior shifts. Four of the 25 children who did not show stable growth during the study did not exhibit a change in their level of thinking but did show a shift in their trial-by-trial behaviors. They still exhibited ICS-level behaviors but in different ways across the trials. They built composite units for one or two trials before transitioning to building and repeating composite units for the remaining trials.

No overall level transitions but some behavior shifts. Two of the 25 children who did not show stable growth during the study did not exhibit an overall change in their levels but did shift in their trial-by-trial behaviors by exhibiting at least one reach back. For example, these children built and repeated composite units on Trials 2 and 3 but then operated on individual units on Trial 4. One child used tiles to cover, placing one by one until they ran out of tiles then iterated the previously used tiles to finish covering. The other child traced tiles to cover, creating a correct drawing and counting one by one to report an answer. For both of these children, their level of thinking on previous and subsequent trials was coded at the ICS level but on Trial 4, they exhibited thinking at the AURR level.

Incomplete level transitions and behavior shifts. The remaining nine of the 25 children exhibited some growth, but that growth did not meet our requirements for relatively stable growth. We identified two of these nine children as operating at the AURR level with the screening instrument and determined that they had demonstrated behaviors indicative of the ICS level on one or two (nonconsecutive) trials during the study. That is, they used tiles to cover, through tracing or iterating individual tiles, on some trials but used a mixed drawing strategy—alternating between drawing individual units and drawing rows or columns of individual units—on other trials. Hence, these two children had started to build and operate on a composite but not consistently.

Seven of the nine children started the study at the ICS level and exhibited ARCS level thinking at some point during the study but not consistently. They started to coordinate linear and area units in one dimension or in two dimensions but not consistently. Many seemed to have other misconceptions to overcome. Two of the children confounded area and perimeter on some of the trials and a third child made errors while using the ruler. We posit that these nine children may have been transitioning and exhibiting shifts in their trial-by-trial behaviors, but they had not completed that transition by the end of the study. We removed these nine children, as well as the previously discussed 16 children, from our Phase 2 analysis because we wanted to examine the level shifts, and these 25 children did not exhibit a complete level transition.

Complete level transitions that may have begun prior to Trial 1. Eleven children started the study at the AURR level and demonstrated behaviors described at the ICS level by Trial 1 or 2, and five children started the study at AURR or ICS and exhibited thinking indicative of the ARCS level by Trial 2. We removed these 16 children from further analysis because they exhibited growth so quickly that their transition to the next level may have begun prior to Trial 1. Thus, we could not document the entire period from the beginning of the change until reaching stability (Siegler & Svetina, 2006)..

Complete level transitions that occurred after Trial 2. The remaining 13 children met the criteria for relatively stable growth: Three demonstrated growth into the ICS level and 10 demonstrated growth into the ARCS level at least three times on subsequent trials. Nine of these 13 children were in the Subdivision Group, three were in the Iteration Group, and one was in the Comparison Group. Background information for these 13 children[10] is presented in Table 3.

Table 3

Intervention Group, Initial Placement, and Grade per Selected Child

Child	Intervention Group	Initial Placement	Final Placement	Grade	Figure 4[a]
Sadie_2	Subdivision	AURR	ICS	2	S11
Ian_4	Iteration	AURR	ICS	4	I15
Ianto_5	Iteration	AURR	ICS	5	I7
Samuel_4	Subdivision	AURR	ARCS	4	S6
Saul_4	Subdivision	AURR	ARCS	4	S15
Sarah_5	Subdivision	AURR	ARCS	5	S9
Sierra_2	Subdivision	ICS	ARCS	2	S1
Sidney_3	Subdivision	ICS	ARCS	3	S3
Sidra_5	Subdivision	ICS	ARCS	5	S7
Sibley_5	Subdivision	ICS	ARCS	5	S17
Simon_5	Subdivision	ICS	ARCS	5	S18
Iiago_4	Iteration	ICS	ARCS	4	I16

[10] All names are pseudonyms. The first letter of the name indicates intervention group, the second letter indicates initial level placement, and the number subscript indicates grade level. For example, Sadie_2 was a Grade 2 child (subscript of a 2) in the subdivision group (first letter of her name is an s) who started the study at the AURR level (second letter of her name is an a).

Cameron$_2$	Comparison	AURR	ARCS	2	C13

[a]This is the identifier from Figure 4 to show how children identified in Table 2 correspond to children in Figure 4.

Sadie$_2$, Ian$_4$, and Ianto$_5$ demonstrated relatively stable growth from AURR into the ICS level without then transitioning into the ARCS level. For the first three to five trials of the study, these children traced one tile to draw complete or incomplete arrays. On the first five trials, Sadie$_2$ traced one tile to complete an array of individual squares, and although they were aligned in rows and columns, she did not provide evidence that she was coordinating area units within rows and columns (e.g., demonstrated that she expected rows to have the same number of units). On the first three trials, Ian$_4$ and Ianto$_5$ both showed evidence that they were building a composite unit and coordinating area units within rows and columns by either using a mixed drawing strategy—alternating between drawing row line segments and tracing individual tiles—or by indicating one row and one column with tiles. Then on Sadie$_2$'s sixth trial and Ian$_4$'s and Ianto$_5$'s fourth trials, a shift occurred.

Sadie$_2$ started by tracing one tile but then curtailed her drawing actions by using the ruler as a straight edge to finish some but not all row and column segments to produce a correct and complete drawing of an array. We take her curtailment as evidence that she was coordinating area units within rows and columns and beginning to think about a column of two squares as a repeatable unit. Ianto$_5$ used a sequence of behaviors similar to what was displayed in the Iteration Intervention video by drawing tick marks along the vertical side, but he did this without using the ruler or tiles. He then placed six tiles along the horizontal side of the rectangle, taped the tiles, and then iterated this composite unit five times, once per tick mark, skip counting "6, 12, 18, 24, 29 [sic]" as he went. In contrast, Ian$_4$ iterated a single tile to determine that four square inches would fit along the vertical side of the rectangle and six along the horizontal side. He asserted that he made "four rows of six" and that "6 times 4 is 24." Because all three children built and repeated a composite area unit as well as coordinated area units within rows and columns, we claimed that they were exhibiting behaviors at the ICS level.

Throughout the remaining trials, these three children continued to operate on composite area units. However, they did not provide evidence that they had begun (much less completed) the transition from ICS to ARCS. Specifically, they did not yet demonstrate that they were applying the concept that the side length determines the number of units that will fit along the side; using the length and the width, coordinating linear dimensions of both the area unit and the region to be covered to iterate a row or column of units in the orthogonal direction exhaustively; or drawing parallel row and column line segments during the study. Hence, they did not exhibit any of the behaviors indicative of the ARCS level by the end of the study.

Main Findings for Phase 2: Complete Level Transition to ARCS.

To describe the typical patterns in children's behaviors as they shifted from measuring area by operating on individual or composite area units to utilizing an array structure, we examined 10 children who made this transition. For each of the 10 children who exhibited growth to the ARCS level during this study (see Table 3), we isolated the first trial on which we assigned an ARCS level code. To identify patterns during level transitions, we highlighted the behaviors present surrounding the change by reassigning the trial on which we first coded the child at the ARCS level as Trial 0. We then identified the two trials preceding and the three trials following Trial 0 as Trials -2, -1, +1, +2, and +3. We made comparisons among these 10 children per behavior and per LT level. For seven of these 10 children, Trial 0 was Trial 3 (Sidney$_3$, Samuel$_4$, Sidra$_5$, Sarah$_5$, Saul$_4$, Sibley$_5$, and Cameron$_2$). For Simon$_5$, Trial 0 was Trial 4; for Iiago$_4$, Trial 0 was Trial 5; and for Sierra$_2$, Trial 0 was Trial 6 (see Table 4).

Table 4

Learning Trajectory for Area Measurement Placement per Selected Child per Trial

Child	T1	T2	T3	T4	T5	T6	T7	T8	T9
Samuel$_4$	ICS	NC	***ARCS***	ARCS	ARCS	ARCS	ARCS	ARCS	ARCS
Saul$_4$	NC	NC	***ARCS***	ARCS	ICS	ARCS	ARCS	ARCS	ARCS
Sarah$_5$	AURR	AURR	***ARCS***	ARCS	ARCS	ARCS	ARCS	ARCS	ARCS
Sierra$_2$	PCC	NC	ICS	NC	ICS	***ARCS***	ARCS	ARCS	ARCS
Sidney$_3$	NC	NC	***ARCS***	ARCS	ARCS	ARCS	ARCS	ARCS	ARCS
Sidra$_5$	ICS	ICS	***ARCS***	ARCS	ARCS	ARCS	ARCS	ARCS	ARCS
Sibley$_5$	NC	NC	***ARCS***	ARCS	ARCS	ARCS	ARCS	ARCS	ARCS
Simon$_5$	ICS	ICS	ICS	***ARCS***	ARCS	ARCS	ARCS	ARCS	ARCS
Iiago$_4$	ICS	ICS	ICS	ICS	***ARCS***	ARCS	ARCS	ICS	ARCS
Cameron$_2$	ICS	ICS	***ARCS***	ICS	ARCS	ARCS	ARCS	ARCS	ARCS

Note. NC = No level claim was made for that particular child on that trial; Trial 0 identified with italicized and bolded ARCS per child. PCC = Physical Coverer and Counter.

Behavior patterns surrounding the change. To identify patterns surrounding the change for the 10 children who demonstrated growth into the ARCS level at least three times on subsequent trials, we examined their behaviors on Trials -2, -1, +1, +2, and +3. Six of these children (Samuel$_4$, Sierra$_2$, Sidney$_3$, Sidra$_5$, Sibley$_5$, and Simon$_5$) exhibited similar patterns of growth into and then within ARCS, despite some instances of subtle within-child variability between trials before the shift for four of them (Samuel$_4$, Sierra$_2$, Sidney$_3$, and Sidra$_5$). In contrast, three of the 10 children exhibited within-child variability after the shift: Cameron$_2$, Saul$_4$, and Iiago$_4$ exhibited fallback between sessions. Only one child (Sarah$_5$) transitioned from AURR to ARCS without going through the ICS level. We next discuss the typical patterns of growth demonstrated by the six children before considering the more divergent cases.

Four of the children (Samuel$_4$, Sidra$_5$, Sibley$_5$, and Simon$_5$) provided a correct numeric response on Trial -2. Although Sierra$_2$ and Sidney$_3$ used the ruler to measure adjacent sides, Sierra$_2$ and Sidney$_3$ reported numeric answers reflective of finding semiperimeter and perimeter, respectively. Neither produced a drawing, even when asked to show how their units fit. In comparison, Sibley$_5$ also did not produce a drawing, but she did provide a correct numeric response. When asked to show how they fit, she responded that she multiplied.

The other three children produced correct arrays. Initially, Samuel$_4$ placed six tiles along the horizontal side of the rectangle, removed these tiles, placed five tiles along the vertical side, and then wrote $5 \times 6 = 30$. It was only after he was asked to show how the 30 fit that he created a drawing: He extended horizontal parallel line segments from the tiles still along the vertical side of the rectangle and then drew vertical parallel line segments freehandedly. Sidra$_5$ and Simon$_5$ also produced correct arrays by drawing parallel row and column line segments. Sidra$_5$ used the ruler as a straightedge only, using it to draw one horizontal line segment and four vertical line segments (left to right), effectively and correctly partitioning it into 10 equal sized pieces. However, she did not use the ruler to measure the sides first; hence, we do not have evidence that she knew how many vertical segments to draw before she started drawing them in. In contrast, Simon$_5$ measured the vertical and horizontal sides of the rectangle, recording their measures. Then he drew three parallel but not equally spaced line segments to indicate four rows. Before drawing four parallel but not equally-spaced line segments to subdivide the region into five columns, he wrote, "4 colums [sic] and 5 colums [sic] would equal 20 square inches" above the rectangle (see Figure 5).

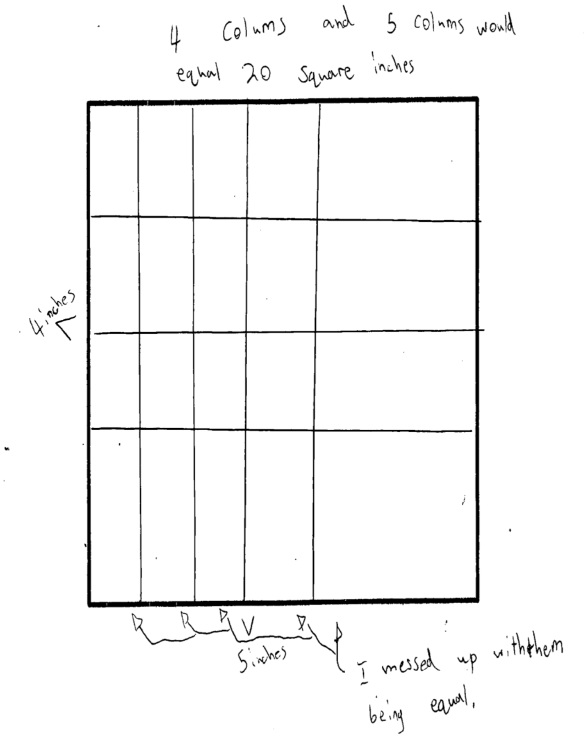


Figure 5. Simon$_5$'s imprecise coordination of linear and area units on Trial -2.

This provided us with evidence that Simon$_5$ was starting to attempt to coordinate linear and area units. Without tools, he coordinated linear and area units imprecisely (i.e., numerically but less precisely spatially) along the vertical and horizontal sides of the rectangle (see Appendix B for definitions). For this trial, Samuel$_4$, Sidra$_5$, and Simon$_5$, were placed at the ICS level because of the combination of behaviors.[11] No level claim

[11] Recall that level placement is based on a combination of observed behaviors. To be placed at the ARCS level, a child needs to apply the concept that the side length determines the number of linear and area units that will fit along the side and coordinate the linear dimensions of both the area unit and the region to be covered to iterate a row or column of units in the orthogonal direction exhaustively. Thus, it is possible for a child to

was made for $Sierra_2$ and $Sibley_5$ because neither made any marks to show how the units would fit, nor for $Sidney_3$ because her drawing did not match her numeric answer.

On Trial -1, only two of the children ($Sibley_5$ and $Simon_5$) provided a correct numeric response, and there was less consistency in their behaviors. Two of the children who provided incorrect numeric responses, $Sierra_2$ and $Sidra_5$, drew parallel row and column line segments to produce incorrect arrays. $Sierra_2$ first measured the vertical side of the rectangle to be 3 inches and the horizontal side to be 7 inches. Next, she drew three row line segments (instead of two) and six column line segments freehandedly to produce a 4 × 7 array (instead of a 3 × 7 array). Given a 4-inch by 5-inch rectangle, $Sidra_5$ again used the ruler as a straightedge only, using it to draw one horizontal line segment and then three vertical line segments (left to right), partitioning the rectangle into eight somewhat equal sized pieces (i.e., producing a 2 × 4 array instead of a 4 × 5 array).

In contrast, $Sidney_3$ and $Samuel_4$ used behaviors similar to the one displayed in the Subdivision Intervention video. They measured adjacent sides of the rectangle and used the tick marks on the ruler to draw parallel row and column line segments and produce correct arrays. However, $Sidney_3$ and $Samuel_4$ reported numeric answers reflective of finding semi-perimeter and perimeter, respectively. Similar to his behaviors in Trial -2, $Simon_5$ measured adjacent sides of the rectangle and recorded their linear measures. Then he drew five parallel but not equally spaced, vertical line segments and then three parallel but not equally spaced, horizontal line segments to produce a numerically correct array.

Because of their use of the ruler to draw the line segments, $Sidney_3$ and $Samuel_4$ coordinated linear and area units precisely with tools in both dimensions. Whereas, $Simon_5$ did not use the ruler, causing him to coordinate linear and area units imprecisely (numerically but not spatially) without tools in both dimensions. Although $Sibley_5$ provided a correct numeric response, she was the only child who did not create a drawing on Trial -1. Again, when asked to show how they fit, she responded that she multiplied. $Sierra_2$, $Sidra_5$, and $Simon_5$, were placed at the ICS level for Trial -1 because of their

---

be placed at the ARCS level without drawing parallel row and column segments and for a child to draw parallel row and column segments without being placed at the ARCS level. The same is true for any other observed behavior in isolation.

combination of behaviors, whereas no level claim was made for three children: Sidney_3 and Samuel_4 because their drawings did not match their numeric answers, and Sibley_5 because she did not show how the units would fit.

On Trial 0, all six of the children provided a correct numeric response. They also produced correct drawings of the array, drew parallel row and column segments, and coordinated linear and area units precisely with tools in both dimensions (vertical and horizontal). Their sequence of actions was similar to those they watched in the Subdivision Intervention video. This was a noticeable decrease in variability in collections of behaviors from Trials -2 and -1. All six of the children continued to produce correct drawings of arrays, draw parallel row and column line segments, and coordinate linear and area units precisely with tools in both dimensions on the remaining trials (Trials +1, +2, and +3). All of them also produced correct numeric responses on Trials +1 and +2, and most of them did so on Trial +3. (Sidney_3 provided an incorrect numeric response on Trial +3 when she asserted that 4 times 6 is 32.) This indicates that for these six children, their behaviors were relatively consistent after transitioning into the ARCS level.

Examining fallback from Area Row and Column Structurer. On Trial 0, all 10 of the children who exhibited growth to the ARCS level during this study had been placed at the ARCS level (by definition). On each of Trials +1, +2, and +3, nine of these children were at the ARCS level and one child was at the ICS level. However, the child at the ICS level was different on each of those three trials (Cameron_2 on Trial +1, Saul_4 on Trial +2, and Iiago_4 on Trial +3). Interestingly, Cameron_2 and Iiago_4 both fell back to ICS when presented with Rectangle A (4×3), but Saul_4 fell back to ICS when presented with Rectangle H (4×6). Iiago_4 and Cameron_2 still produced correct arrays for Rectangle A but with less sophisticated behaviors than ARCS. Therefore, we claim that Iiago_4 and Cameron_2 reached back to use ICS-level behaviors. However, Saul_4 did not.

To investigate how Saul_4 fell back, we returned to his work. Figure 6 illustrates Saul_4's drawings before, when, and after he fell back to ICS-level behaviors on Trial +2 by drawing a 6×4 array of rectangular units rather than a 4×6 array of square units. That is, Saul_4 partitioned a length of 6 inches into four sections and a length of 4 inches into six sections. Interestingly, the units within the region are equivalent in area to square inches, and his answer was numerically correct; however, he exhibited a disconnect between the length of a side and the number of area units fitting along that side in both dimensions. Because Saul$_4$ used a collection of behaviors that exhibited a linear and area unit disconnect to produce an incorrect array, we posit that he fell back to use an ICS-level behavior (cf. Barrett, Clements, & Sarama, 2017; Pirie & Kieren, 1994).

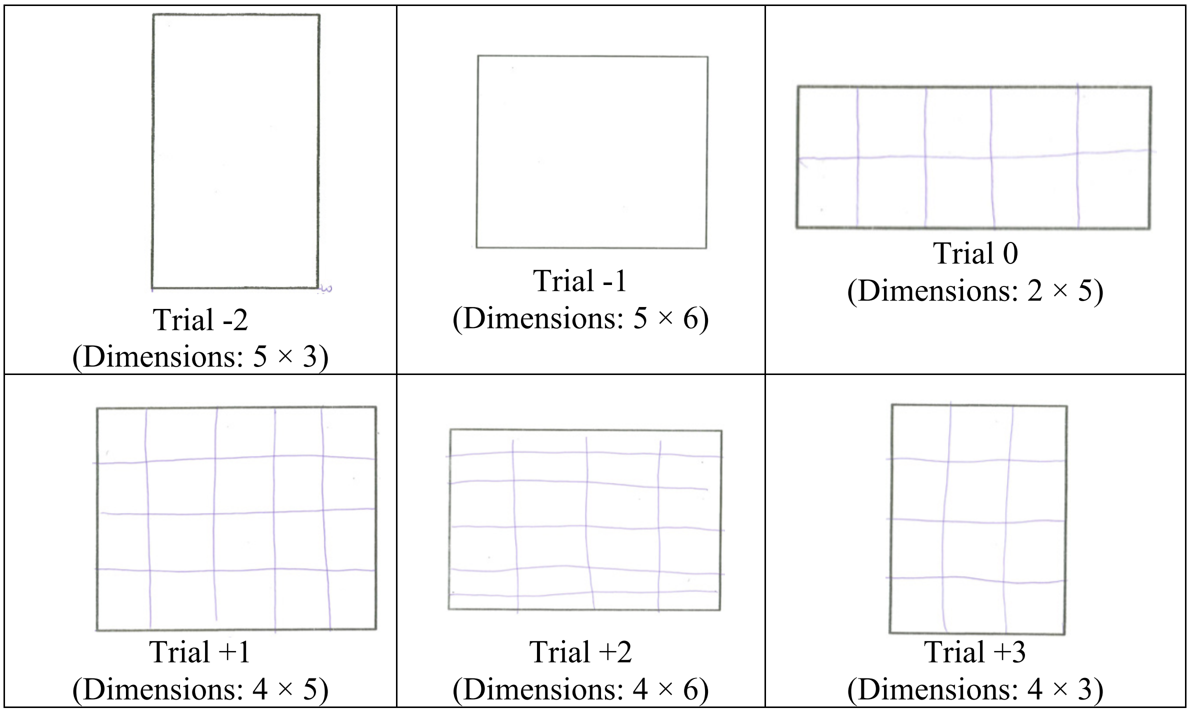


Figure 6. Saul$_4$'s drawings, indicating fall back to ICS on Trial +2.

Saul$_4$ then returned to using ARCS-level behaviors on Trial +3 by coordinating linear and area units precisely with tools. Saul$_4$ continued to use ARCS-level behaviors for the remainder of the trials in the study (see Trial +4, Trial +5, and Trial +6 in Table 4).

Examining growth from Area Unit Relater and Repeater. Sarah$_5$ was the only child to exhibit a transition from AURR to ARCS without demonstrating behaviors indicative of ICS. We did not see evidence of Sarah$_5$ curtailing her tracing actions to build a composite unit (i.e., a row of individual units) and then repeat that composite unit, nor did we see evidence that she was coordinating linear units or coordinating linear and area units in one dimension. Instead, Sarah$_5$ shifted from thinking about individual area units to thinking about an array, as demonstrated by her coordination of linear and area units in both dimensions when she subdivided the region into an array by drawing parallel row and column line segments. We next examine the trials immediately preceding and following this transition in terms of her drawings, operations on units, and ability to coordinate linear and area units.

On Trials -2 and -1, Sarah$_5$ traced individual square tiles to structure the rectangles (see Figure 7). Her tedious (over)attention to tracing square tiles indicated that she was reliant on an individual area unit as a marker to help her keep track of where she had previously iterated square tiles. These behaviors indicated that she was thinking about an individual unit to cover. Sarah$_5$ may have used a row as an intuitive structure because her individual units appear to be aligned in rows, but she did not indicate (e.g., verbally or with motion) that this was a row to her. Also, although her units within "rows" also appear to be aligned within "columns," Sarah$_5$ did not provide evidence that she was coordinating area units within rows and columns (e.g., indicating that she expected rows to have the same number of units). Thus, Sarah$_5$ was coded as exhibiting the AURR level on those two trials.

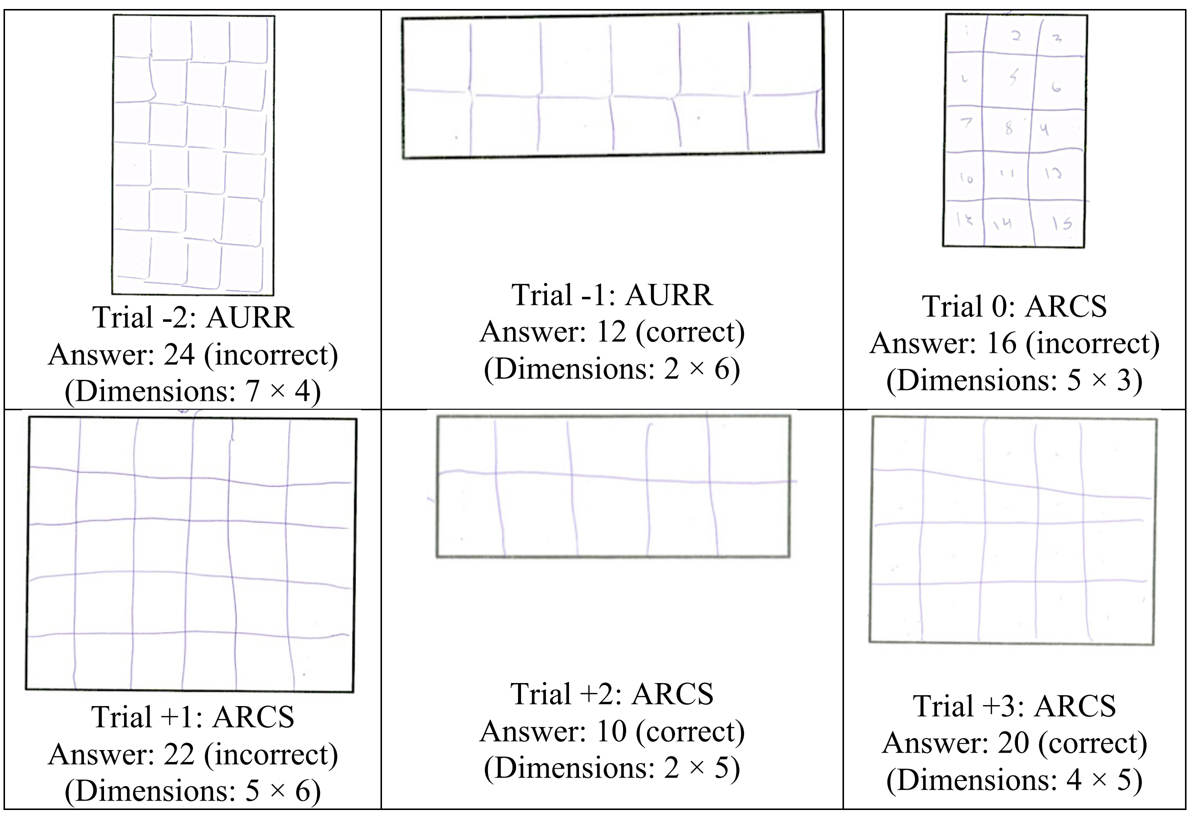


Figure 7. Sarah$_5$'s numeric responses and drawings from Trials -2 through +3.

In contrast, on Trial 0, when asked to find the area of a 3-inch by 5-inch rectangle, $Sarah_5$ used the numbered tick marks on the ruler to constrain the placement of parallel row and column line segments to produce a correct array. In other words, she used both dimensions to constrain the placement of parallel row and column line segments, behaviors indicative of the ARCS level. Although $Sarah_5$ produced a correct drawing, she gave a numeric answer that reflected an attention to perimeter. When she was asked to show how the 16 would fit, $Sarah_5$ numbered individual squares within the correct array as she counted 15 squares (see Figure 7) but did not change her final answer from 16 to 15. On Trial +1 $Sarah_5$ repeated her drawing strategy to draw a correct array but reported a numeric answer of 22 (i.e., the perimeter) as her numeric answer. It was not until Trial +2 that $Sarah_5$'s numeric answer matched the number of area units in the array. We have two interpretations for $Sarah_5$'s disconnect between space and number (i.e., coordinating linear and area units in both dimensions to draw parallel row and line column segments and produce a correct array but reporting a numeric answer that did not correspond to the number of units drawn in the array). On the one hand, her shift from using just tiles to using a ruler may have prompted her to confound perimeter and area. On the other hand, $Sarah_5$ may be demonstrating an initial onset of ARCS-level thinking that is messy and complicated because of her large conceptual leap from AURR.

Discussion

We evaluated three interventions designed to support Grades 2–5 children's growth in measuring the area of rectangular regions in different ways. The Subdivision Intervention privileged the subdivision of a region into an array by drawing parallel row and column line segments, the Iteration Intervention emphasized the building of an array by establishing a row of unit squares and iterating the row, and the Comparison Intervention reflected the implicit multiplication of length and width to obtain a numeric answer for area.

Phase 1 Conclusions

We found that children's observable behaviors and numerical responses when measuring areas of rectangular regions were more affected by the Subdivision and

Iteration Interventions than the Comparison Intervention. Specifically, although all three groups exhibited an increased likelihood of answering correctly across the nine trials, the children in the Subdivision and Iteration Groups were more likely to provide a correct numeric response than the children in the Comparison Group. Furthermore, only the children in the Subdivision and Iteration Groups exhibited significant growth along the LT for area measurement, with the growth being the most rapid for the Subdivision Group. These findings suggest that interventions designed to support concept growth (e.g., big ideas of measurement) are more efficient and effective than an intervention that does not.

In addition, for those children in the Subdivision and Iteration Groups, the LT level exhibited on each trial served as a significant mediating variable for correctness. This finding suggests that, although all of the children exhibited an increased likelihood of responding correctly with each subsequent trial, only the Subdivision and Iteration Groups' increases were shown to be the result of more sophisticated conceptualizations. In other words, the Comparison Intervention may help children produce more correct answers; however, it does not help children improve their understanding of area measurement concepts.

The Subdivision Intervention also prompted significantly more instances of ARCS-level behaviors than the other interventions. Comparably, the Iteration Intervention prompted significantly more instances of ICS-level behaviors than the other interventions. It is our conjecture that a key characteristic contributing to the effectiveness of the Subdivision Intervention was seeing a complete record of the structure of the two-dimensional array. We believe that the teacher's process of drawing an array (as observed in the Subdivision Intervention video) supported children in the Subdivision Group in conceptualizing how linear and area units were coordinated to partition the region into an array by drawing parallel row and column line segments, regardless of whether the children exhibited global structuring techniques themselves (cf. Battista et al., 1998). This is in contrast to children in the Iteration Group who were shown how to build an array by establishing a row of unit squares and iterating the taped row to fill the region because there was no imprint, stamping, or other record of the structure of the two-dimensional array after the teacher finished the process of iterating the row of square tiles.

Phase 2 Conclusions

We also examined the children's observable behaviors before and after they shifted from measuring area by operating on individual or composite area units to using an array structure. Because we utilized the microgenetic method in this study, we were able to investigate this variability within and among trials and children. Twenty-five children were removed from Phase 2 analysis because they did not fit our conservative criteria of relatively stable growth, and 16 children were removed from Phase 2 analysis because they exhibited relatively stable growth so quickly that we could not document and analyze the period of change. However, this does not mean that our interventions were ineffective for 76% of our participants. As noted in our Phase 1 analysis, children in all three groups exhibited an increased likelihood of responding correctly with each subsequent trial. Thus, most of these 41 children made shifts in their ability to produce a correct numeric answer.

The nine children who appeared to begin—but not complete—a level transition during the study exhibited behavior shifts. These children exhibited variability in their behaviors. Some had other misconceptions that may have been interfering with their growth, such as confounding area and perimeter or making errors when using a ruler.

The 13 children who completed a level transition during the study, and especially the 10 children who completed their transition into the ARCS level, also exhibited behavior shifts but before the level transition. We noticed subtle variability in behaviors among the 10 children who transitioned into the ARCS level before they transitioned. We also found variability in behaviors within children from trial to trial: Four exhibited slightly more variability before the shift (Samuel$_4$, Sierra$_2$, Sidney$_3$, and Sidra$_5$), and three exhibited slightly more variability after the shift (Cameron$_2$, Saul$_4$, and Iiago$_4$). Yet, this slight variability after the shift was in the form of fall back or reach back on a single trial, indicating that the 10 children's behaviors were relatively stable after they transitioned to the ARCS level. The before-shift variability has a different explanation. Siegler (2006) demonstrated that variability within individuals can be a prelude for learning, which in our study was indicated by a transition to a more sophisticated level in the LT for area measurement. In 2002, Siegler argued, "Just prior to discoveries, children show increased solution times (Sielger & Jenkins, 1989), increased verbal disfluencies (Perry & Lewis, 1999), increased gesture-speech mismatches (Alibali & Goldin-Meadow, 1993), and increased cognitive conflict (Piaget, 1952)" (p. 52). We concur that variability within a child from trial to trial is not an anomaly, but instead a harbinger or early phase of substantial change, and thus a component of the "path of change" (Siegler, 2006).

Our findings regarding variability are also consistent with our theoretical perspective of hierarchic interactionalism. Because growth into the next level depends on the concepts and processes that constitute the previous levels, "a critical mass of ideas from each level must be constructed before thinking characteristic of the subsequent level becomes ascendant in the child's thinking and behavior" (Clements & Sarama, 2007, p. 465). We posit that the within-child variability that we observed on Trials -2 and -1 was indicative that a proper subset of the requisite concepts and processes was emerging as the children modified their mental models for coordinating and structuring each individual unit and its relation to the group or groups of units. However, the full set had not yet congealed in a cohesive, efficient, and dominant mental structure.

The six children who exhibited similar patterns of growth into and then within ARCS were all in the Subdivision Group. A skeptic may argue that these children did not learn about area measurement; instead they learned to mimic a sequence of behaviors exhibited in the Subdivision Intervention video. Although the Subdivision Intervention video could be interpreted as a form of demonstration, these children still required repeated experience or exposure to that demonstration before exhibiting changes in their own drawings or notations. It took the eight of 18 children in the Subdivision Group who completed their transition into the ARCS level between two and five viewings of the Subdivision Intervention video (see Table 4) to reflect upon (and perhaps use) what they were observing to exhibit a sequence of behaviors that was similar to what was displayed in the Subdivision Intervention video.

When a child observes another person (directly present or shown in a video recording) constructing an array by drawing two orthogonal sets of line segments onto a rectangle, the child must incorporate that example into their own activity to provide a basis for advancing their knowledge of the array as an organizing principle to guide the measure of the area of that rectangle. In this study, other than asking the participants to "show how the area units fit," there was no prompt to compel them to use the collection of behaviors illustrated in the videos, nor a way to ensure they used them meaningfully. For example, $Saul_4$ did not meaningfully use the linear scale when he partitioned a length of 6 inches into four sections and a length of 4 inches into six sections. He produced approximately equal-sized rectangles to produce an array of units equivalent in area to square inches. His numeric answer was correct, but he did not have an array of 24 *square* inches. In other words, the meaningful use of a linear scale (i.e., a 12-inch ruler) to guide the placement of area units and draw parallel row and column line segments and the coordination of linear and area units in both dimensions are nontrivial behaviors and cannot be incorporated prior to making sense of those actions. Therefore, we contend that some of the participants learned that the numerical response to the question "What is the area of this rectangle?" should indicate that that number of area units should cover the space (i.e., they learned to associate the word area with structuring space); some learned how to coordinate linear and area units; and some learned a more sophisticated way to show how the area units fit. This implies more than mimicry; we take this as evidence of cognitive restructuring. As for the participants who may have "just" learned to mimic some of the behaviors illustrated in the videos, we argue that mimicry is still learning—they still learned how parts fit into a whole and applied this to new situations.

Limitations

Although this study was designed to extend the research on the learning of area measurement by investigating shifts in children's behaviors in response to one of three interventions, it is not without limitations. Some of the limitations of this study can be attributed to the nature of the tasks posed in the study. In the present study, children had access to seven square-inch foam tiles, a ruler, and a pen, and they were asked to determine the area of a given rectangle. The growth exhibited here was observed in a clinical setting with a researcher. It is unknown whether children would maintain the growth observed in the study or transfer their knowledge when confronted with new situations involving other aspects of area measurement, more complex tasks, or with different tools. In addition, a delayed posttest may have given additional credibility to our findings.

Another limitation of the present study is related to our decision to determine correctness based on children's numeric answers. Our participants may not have realized that they were often reporting the number of square units. Previous research indicates that children do not think of a square unit as the standard unit of area (Kamii & Kysh, 2006). However, our participants were not always counting unit squares. Sometimes they were counting rectangular units equivalent in area to square units (e.g., $Saul_4$). Other times they were counting approximately rectangular shapes (e.g., $Simon_5$). Although we argue that for most of the children in this study the unit is implied because the teacher in each of the intervention videos reports both the numeric answer and the unit (e.g., "21 square inches"), additional research is needed to investigate the impact of directly identifying the unit for the child through instruction or demonstration.

Implications and Suggestions for Future Research

Our findings have several implications for teaching. First, the effectiveness of the Subdivision and Iteration Interventions indicates that children benefit from experiences learning about unit concepts (e.g., unitizing, composing units to create units of units, iterating individual or groups of units, and coordinating units) and spatial structuring of two-dimensional space. We recommend that teachers make these big ideas explicit and help children connect these big ideas (e.g., how multiplying the length and width of a rectangle to produce a measure of area is related to drawing an array of rows and columns of identical square units). Second, the rate of change for children in the Subdivision Group suggests that having a complete record of how the area units fit led to improved performance. Thus, it is important for children to not only have opportunities to see a complete record, but also to reflect on its creation and organization, such as by having multiple children share their records of how the area units fit and comparing them in a whole class discussion. Third, the relative ineffectiveness of the Comparison Intervention indicates that a focus on procedures may contribute to children's difficulties when learning about area measurement (Smith et al., 2016).

This study also has several implications for research. In this article, we report on our findings from a multisite, cross-sectional study that integrated multiple methodologies. We also utilized the microgenetic method with an LT for area measurement as a lens to investigate children's behavior shifts and level transitions. Siegler (2002) argued that such analysis is important: "Examining the way that children learn under various instructional procedures, contrasting the characteristics of more and less successful learners, and identifying where learning goes awry when it goes awry—all can contribute to improving instructional procedures" (p. 36). Hence, this research has the potential to help teachers and researchers notice important shifts in behaviors, anticipate level transitions, and provide meaningful experiences at important times.

Of the 10 children who transitioned from not yet ARCS into ARCS during this study, only one transitioned from AURR to ARCS without providing evidence of behaviors indicative of the ICS level. Did $Sarah_5$ skip a level? Or, did our session protocol limit our ability to observe the intermediate transition? The tenets of hierarchical interactionalism indicate that these levels are not only sequential, but also that growth into the next level depends on the concepts and processes that constitute the previous levels (although the theory does not prohibit advances on several contiguous levels simultaneously nor a fast acquisition of multiple levels that occurs without observation of each level in sequence, see Clements & Sarama, 2014; Sarama & Clements, 2009). We do not consider $Sarah_5$'s growth to be disconfirming evidence for the LT, but rather we consider her a novel case. Additional research is needed. We wonder—is "skipping" a level efficient or is it problematic in the long run for children like $Sarah_5$?

Another suggestion for future research is the scale up of this study to a wider range of children's abilities and ages across various settings. This line of research should examine the effectiveness of the Subdivision Intervention and Iteration Intervention both with small groups of children and large groups of children in a classroom setting. We also anticipate the extension of the study presented in this article to children at levels less sophisticated than AURR. Thus far, we have studied how the Subdivision and Iteration Interventions promoted growth and caused a change in behaviors of children beginning at the AURR or ICS levels. One natural extension would be to investigate if or how the Subdivision and Iteration Interventions promoted growth or caused changes in behaviors of children not yet at the AURR level. Another extension would be to examine modifications of the existing Subdivision and Iteration Interventions, such as synthesizing the Subdivision and Iteration Interventions (e.g., modifying the process of iterating rows of taped tiles so that an imprint, stamp, or record of the structure of the two-dimensional array is left behind; L. Steffe, personal communication, February 18, 2016) or privileging continuous motion (e.g., dragging or sweeping one length through another; Kobiela, Lehrer, & Pfaff, 2010). We look forward to extensions of this research.

References

Barrett, J. E., Clements, D. H., & Sarama, J. (Eds.). (2017). Children's measurement: A longitudinal study of children's knowledge and learning of length, area, and volume. *Journal for Research in Mathematics Education* monograph series (Vol. 16). Reston, VA: National Council of Teachers of Mathematics.

Barrett, J. E., Clements, D. H., Sarama, J., Miller, A. L., Cullen, C. J., Van Dine, D. W., . . . Klanderman, D. (2017). Integration of results: A new learning trajectory for area. In J. E. Barrett, D. H. Clements, & J. Sarama (Eds.), Children's measurement: A longitudinal study of children's knowledge and learning of length, area, and volume. *Journal for Research in Mathematics Education* monograph series (Vol. 16, pp. 129–148). Reston, VA: National Council of Teachers of Mathematics.

Barrett, J. E., Cullen, C. J., Miller, A. L., Eames, C. L., Kara, M., & Klanderman, D. (2017). Area in the middle and later elementary grades. In J. E. Barrett, D. H. Clements, & J. Sarama (Eds.), Children's measurement: A longitudinal study of children's knowledge and learning of length, area, and volume. *Journal for Research in Mathematics Education* monograph series (Vol. 16, pp. 105–127). Reston, VA: National Council of Teachers of Mathematics.

Barrett, J. E., Sarama, J., Clements, D. H., Cullen, C. J., & Van Dine, D. W. (2017). Reflections on the design and outcomes of the research. In J. E. Barrett, D. H. Clements, & J. Sarama (Eds.), Children's measurement: A longitudinal study of children's knowledge and learning of length, area, and volume. *Journal for*

Research in Mathematics Education monograph series (Vol. 16, pp. 223–243). Reston, VA: National Council of Teachers of Mathematics.

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., Clements, D. H., Arnoff, J., Battista, K., & Borrow, C. (1998). Students' spatial structuring of 2D arrays of squares. *Journal for Research in Mathematics Education, 29*(5), 503–532. doi:10.2307/749731

Bell, A., Hughes, E. R., & Rogers, J. (1975). *Area, weight, and volume*. London, Great Britain: Nelson for the Schools Council.

Blanton, M., Brizuela, B. M., Gardiner, A. M., Sawrey, K., & Newman-Owens, A. (2015). A learning trajectory in 6-year-olds' thinking about generalizing functional relationships. *Journal for Research in Mathematics Education, 46*(5), 511–558. doi:10.5951/jresematheduc.46.5.0511

Clements, D. H., & Sarama, J. (2007). Early childhood mathematics learning. In F. K. Lester, Jr. (Ed.), *Second handbook of research on mathematics teaching and learning* (Vol. 1, pp. 461–555). Charlotte, NC: Information Age Publishing.

Clements, D. H., & Sarama, J. (2014). Learning trajectories: Foundations for effective, research-based education. In A. P. Maloney, J. Confrey, & K. H. Nguyen (Eds.), *Learning over time: Learning trajectories in mathematics education* (pp. 1–30). Charlotte, NC: Information Age Publishing.

Clements, D. H., Sarama, J., Van Dine, D. W., Barrett, J. E., Cullen, C. J., Hudyma, A.,…Eames, C. L. (2017). Evaluation of three interventions teaching area measurement as spatial structuring to young children. (Manuscript in preparation).

Corbin, J., & Strauss, A. (2008). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (3rd ed.). Thousand Oaks, CA: Sage. doi:10.4135/9781452230153

Curry, M., & Outhred, L. (2005). Conceptual understanding of spatial measurement. In P. Clarkson, A. Downton, D. Gronn, M. Horne, A. McDonough, R. Pierce, & A. Roche (Eds.), *Building connections: Theory, research and practice* (Proceedings of the 27th annual conference of the Mathematics Education Research Group of Australasia, Melbourne, pp. 265–272). Sydney: MERGA. Retrieved from http://www.merga.net.au/documents/RP262005.pdf

Curry, M., Mitchelmore, M., & Outhred, L. (2006). Development of children's understanding of length, area, and volume measurement 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* (Vol. 2, pp. 377–384). Prague, Czech Republic: Psychology of Mathematics Education.

Iacobucci, D. (2012). Mediation analysis and categorical variables: The final frontier. *Journal of Consumer Psychology*, *22*(4), 582–594. doi:10.1016/j.jcps.2012.03.006

Kamii, C., & Kysh, J. (2006). The difficulty of "length x width": Is a square the unit of measurement? *Journal of Mathematical Behavior, 25*(2), 105–115. doi:10.1016/j.jmathb.2006.02.001

Kara, M., Eames, C. L., Miller, A. L., Cullen, C. J., & Barrett, J. E. (2011). Developing an understanding of area measurement concepts with triangular units. In L. R. Wiest & T. Lamberg (Eds.), *Proceedings of the 33rd Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 1015–1023). Reno, NV: North American Chapter of the International Group for the Psychology of Mathematics Education.

Kobiela, M. A., Lehrer, R., & Pfaff, E. (2010, April). *Students developing conceptions of area via partitioning and sweeping.* Paper presented at the Annual Meeting of the American Educational Research Association, Denver, CO.

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., Jacobson, C., Thoyre, G., Kemeny, V., Strom, D., Horvath, J., et al. (1998). Developing understanding of geometry and space in the primary grades. In R. Lehrer & D. Chazan (Eds.), *Designing learning environments for developing* *understanding of geometry and space* (pp. 169–200). Mahwah, NJ: Lawrence Erlbaum Associates.

Lehrer, R., Jenkins, M., & Osana, H. (1998). Longitudinal study of children's reasoning about space and geometry. In R. Lehrer & D. Chazan (Eds.), *Designing learning environments for developing understanding of geometry and space* (pp. 137–167). Mahwah, NJ: Lawrence Erlbaum Associates.

Miller, A. L., (2013). *Investigating conceptual, procedural, and intuitive aspects of area measurement with non-square area units* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (Order No. 3572989)

National Assessment of Educational Progress. (1983, April). *The third national mathematics assessment: Results, trends, and issues* (Report No. 13-MA-01). Denver, CO: Education Commission of the States. Retrieved from ERIC database. (ED228049)

National Assessment of Educational Progress. (2016). *NAEP data explorer*. Retrieved from http://nces.ed.gov/nationsreportcard/naepdata/dataset.aspx

National Governor's Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common core state standards for mathematics.* Washington, DC: Authors. Retrieved from http://www.corestandards.org/assets/CCSSI_Math%20Standards.pdf

Nunes, T., Light, P., & Mason, J. H. (1993). Tools for thought: The measurement of length and area. *Learning and Instruction, 3*(1), 39–54. doi:10.1016/S0959-4752(09)80004-2

Outhred, L., & Mitchelmore, M. (1992). Representation of area: A pictorial perspective. In W. Geeslin & K. Graham (Eds.), *Proceedings of the 16th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 194–201). Durham, NH.

Outhred, L., & Mitchelmore, M. (1996). Children's intuitive understanding of area measurement. In L. Puig & A. Gutierrez (Eds.), *Proceedings of the Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 91–98). Valencia, Spain: Psychology of Mathematics Education.

Outhred, L. N., & Mitchelmore, M. C. (2000). Young children's intuitive understanding of rectangular area measurement. *Journal for Research in Mathematics Education, 31*(2), 144–167. doi:10.2307/749749

Outhred, L., & Mitchelmore, M. (2004). Students' structuring of rectangular array. In M. J. Hoines & A. B. Fuglestad (Eds.), *Proceedings of the 28th Annual Meeting of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 465–472). Bergen, Norway: Psychology of Mathematics Education.

Outhred, L., Mitchelmore, M. C., McPhail, D., & Gould, P. (2003). Count me into measurement: A program for the early elementary school. In D. H. Clements & G. Bright (Eds.), *Learning and teaching measurement: 2003 yearbook* (pp. 81–99). Reston, VA: National Council of Teachers of Mathematics.

Piaget, J., Inhelder, B., & Szeminska, A. (1960). *The child's conception of geometry*. New York, NY: W. W. Norton & Company.

Pirie, S., & Kieren, T. (1994). Growth in mathematical understanding: how can we characterise it and how can we represent it? *Educational Studies in Mathematics, 26*(2–3), 165–190. doi:10.1007/BF01273662

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

Sarama, J., Clements, D. H., Barrett, J. E., Van Dine, D. W., & McDonel, J. S. (2011). Evaluation of a learning trajectory for length in the early years. *ZDM: The International Journal on Mathematics Education, 43*(5), 667–680. doi:10.1007/s11858-011-0326-5

Sarama, J., Clements, D. H., Van Dine, D. W., Vukovich, M., Newburgh, K., McDonel, J. S., & Napora, L. (2017). Area in the early childhood and primary grades. In J. E. Barrett, D. H. Clements, & J. Sarama (Eds.), Children's measurement: A longitudinal study of children's knowledge and learning of length, area, and volume. *Journal for Research in Mathematics Education* monograph series (Vol. 16, pp. 82–104). Reston, VA: National Council of Teachers of Mathematics.

Siegler, R. S. (2002). Microgenetic studies of self-explanation. In N. Granott & J. Parziale (Eds.), *Microdevelpment: Transition processes in development and* *learning* (pp. 31–58). Boston, MA: Cambridge University Press. doi:10.1017/CBO9780511489709.002

Siegler, R. S. (2006). Microgenetic analyses of learning. In W. Damon & R. M. Lerner (Series Eds.) & D. Kuhn & R. S. Siegler (Vol. Eds.), *Handbook of child psychology: Volume 2: Cognition, perception, and language* (6th ed., pp. 464–510). Hoboken, NJ: Wiley.

Siegler, R. S. & Crowley, K. (1991). The microgenetic method: A direct means for studying cognitive development. *American Psychologist, 46*(6)*,* 606–620. doi:10.1037/0003-066X.46.6.606

Siegler, R. S., & Svetina, M. (2006). What leads children to adopt new strategies? A microgenetic/cross sectional study of class inclusion. *Child Development, 77*(4)*,* 997–1015. doi:10.1111/j.1467-8624.2006.00915.x

Smith, J. P., III, Males, L. M., & Gonulates, F. (2016). Conceptual limitations in curricular presentations of area measurement: One nation's challenges. *Mathematical Thinking and Learning, 18*(4), 239–270. doi:10.1080/10986065.2016.1219930

Steffe, L. P. (1991). Operations that generate quantity. *Learning and Individual Differences, 3*(1), 61–82. doi:10.1016/1041-6080(91)90004-K

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

Trends in International Mathematics and Science Study. (2016). *International data explorer.* Retrieved from https://nces.ed.gov/timss/idetimss/

Vasilyeva, M., Ludlow, L. H., Casey, B. M., & St. Onge, C. (2009). Examination of the psychometric properties of the measurement skills assessment. *Educational and Psychological Measurement*, *69*(1), 106–130. doi:10.1177/0013164408318774

Weisstein, E. W. (2016). *Area*. Retrieved from http://mathworld.wolfram.com/Area.html

Zacharos, K. (2006). Prevailing educational practices for area measurement and students' failure in measuring areas. *Journal of Mathematical Behavior*, *25*(3)*,* 224–239. doi:10.1016/j.jmathb.2006.09.003

Authors

Amanda L. Cullen, Department of Mathematics, Illinois State University, Campus Box 4520, Normal, IL 61790; almille@ilstu.edu

Cheryl L. Eames, Department of Mathematics and Statistics, Southern Illinois University Edwardsville, Edwardsville, IL 62026; ceames@siue.edu

Craig J. Cullen, Department of Mathematics, Illinois State University, Campus Box 4520, Normal, IL 61790; cjculle@ilstu.edu

Jeffrey E. Barrett, Department of Mathematics, Illinois State University, Campus Box 4520, Normal, IL 61790; jbarrett@ilstu.edu

Julie Sarama, Morgridge College of Education, Marsico Institute for Early Learning and Literacy, Katherine A. Ruffatto Hall 154, 1999 East Evans Avenue, University of Denver, Denver, CO 80208-1700; Julie.Sarama@du.edu

Douglas H. Clements, Morgridge College of Education, Marsico Institute for Early Learning and Literacy, Katherine A. Ruffatto Hall 154, 1999 East Evans Avenue, University of Denver, Denver, CO 80208-1700; Douglas.Clements@du.edu

Douglas W. Van Dine, Department of Mathematical and Computer Sciences, Metropolitan State University of Denver, Denver, CO 80217-3362; dvandine@msudenver.edu

Submitted November 11, 2016

Accepted March 20, 2017

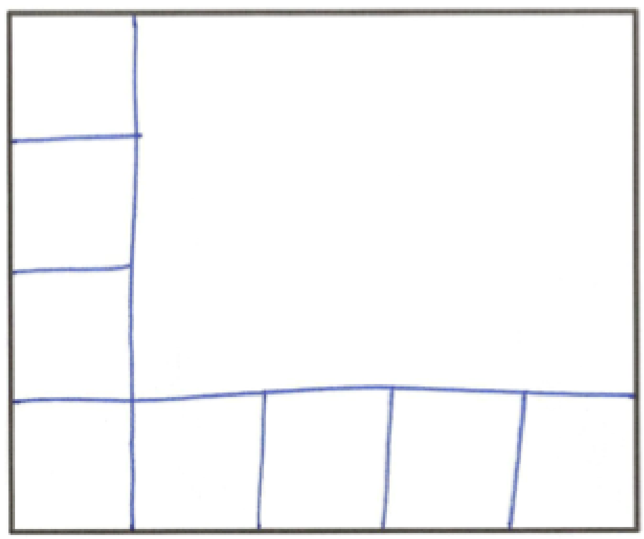
Appendix A

Four-item Screening Instrument

Part 1: Without a ruler

1. What is area?

2. I wanted to cover this rectangle with these squares. I started drawing them in. Please finish the drawing by completely covering the rectangle.



Part 2: With a ruler

3. The area of this rectangle is 10 square inches. Draw how each of the 10 square inches fit.

4. Draw a rectangle that has an area of 8 square inches in the space below. You may use a ruler to help you. Show on your rectangle how the 8 square inches fit.

Appendix B

Codes	Descriptor
Unit coordination codes	
No claim	Unable to make a claim about unit coordination based on the child's written, verbal, or nonverbal responses.
Linear and area unit disconnect	Did not apply the concept that the length of a side indicates the number of area units that should fit along that side (e.g., created a 6 by 4 array for a 4-in. by 6-in. rectangle, see $Saul_4$'s drawing for Trial +2 in Figure 6).
Coordinated linear and area units imprecisely (numerically but not spatially) without tools	Placed tick marks or line segments on the rectangle, resulting in the correct number of rows or columns, but the rows or columns were not close to the same size (i.e., unequally spaced).
Coordinated linear and area units precisely (numerically and spatially) without tools	Without tools applies the concept that the side length determines the number of area units that will fit along the side (i.e., did not use the ruler to guide the placement of tick marks or line segments; may have used a mental image of an inch or square inch to place tick marks appropriately). For example, demonstrated or verbalized that a side length of 4 inches would necessarily have 4 square units that would fit along that side.
Coordinated linear and area units precisely with tools	With tools applies the concept that the side length determines the number of area units that will fit along the side (e.g., used the numbered tick marks on the ruler to guide the placement of tick marks or line segments). For example, used the linear units on a ruler to measure the length of the side and then drew line segments per inch to show where each square inch would fit along that side.