



Contents lists available at ScienceDirect

Journal of Mathematical Behavior

journal homepage: www.elsevier.com/locate/jmathb

Evaluation of three interventions teaching area measurement as spatial structuring to young children[☆]

Douglas H. Clements^{a,*}, Julie Sarama^a, Douglas W. Van Dine^a, Jeffrey E. Barrett^b,
Craig J. Cullen^b, Aaron Hudyma^a, Ron Dolgin^a, Amanda L. Cullen^b, Cheryl L. Eames^c

^a University of Denver, United States

^b Illinois State University, United States

^c Southern Illinois at Edwardsville, United States

ARTICLE INFO

Keywords:

Area
Early childhood
Learning trajectories
Mathematics education
Measurement
Primary
Spatial structuring

ABSTRACT

We evaluated the effects of three instructional interventions designed to support young children's understanding of area measurement as a structuring process. Replicating microgenetic procedures we used in previous research with older children to ascertain whether we can build these competencies earlier, we also extended the previous focus on correctness to include analyses of children's use of procedural and conceptual knowledge and examined individual differences in strategy shifts before and after transitions, enabling a more detailed examination of the hypothesized necessity of development through each level of a learning trajectory. The two experimental interventions focused on a dynamic conception of area measurement while also emphasizing unit concepts, such as unit identification, iteration, and composition. The findings confirm and extend earlier results that seeing a complete record of the structure of the 2D array—in the form of a drawing of organized rows and columns—supported children's spatial structuring and performance.

1. Introduction

Measurement of area is not only an important mathematical topic, but also connects to children's experience with other mathematical topics and with the physical world. For example, developing concepts and procedures in area measurement builds a foundation for understanding topics such as multiplication, fractions, and composition of geometric figures, as well as for addressing real-world problems such as deforestation, navigation, designing of packing material, or measuring of molecules on a surface with a nanomeasuring machine. In this study, we replicated and substantially extended microgenetic procedures we used previously with students in Grades 2–5 to evaluate the effects of experimental interventions designed to support children's understanding of area measurement with spatial structuring on children from Grades 1–3.

2. Background

Students in the United States often use rote procedures for area measurement without demonstrating understanding of crucial

[☆] The research reported in this manuscript was supported by the National Science Foundation under Grant No. DRL-1222944. The findings and statements in this manuscript do not represent the views of the National Science Foundation.

* Corresponding author at: 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, United States.

E-mail address: Douglas.Clements@du.edu (D.H. Clements).

<https://doi.org/10.1016/j.jmathb.2017.12.004>

Received 4 May 2017; Received in revised form 24 December 2017; Accepted 25 December 2017

0732-3123/ © 2017 Elsevier Inc. All rights reserved.

area concepts and principles. Over 40 years of studies show that U.S. students, including young children, have performed poorly on measurement assessments, including area measurement (Battista & Clements, 1996; Carpenter, Coburn, Reys, & Wilson, 1975; Clements & Battista, 2001; Hirstein, 1981). As specific examples, on the fourth National Assessment for Educational Progress (NAEP) assessment, only 46% of seventh graders correctly identified the area of a 5 by 6 rectangle (Lindquist & Kouba, 1989). The trend in student performance on area does not seem to be improving. On the 2011 NAEP, only 24% of fourth graders were able to choose the correct value for the area of a square, and overall, measurement was the weakest of five substrands for fourth and eighth graders (National Assessment of Educational Progress, 2013).

Similar difficulties with area are revealed with younger children as well. A study of children from Grades 1, 2, and 3 revealed little understanding of area measurement (Lehrer, Jenkins, & Osana, 1998). When asked how much space a square (and a triangle) covers, 41% of children used a ruler to measure length and gave a numerical response, such as “9.” When asked what “9” would mean in that context, they said it would be “inches.” The second most frequent response, 22%, was “I don’t know.”

In investigating why children may struggle with area measurement, researchers have analyzed instructional textbooks and manipulative materials (Smith, Males, Dietiker, Lee, & Mosier, 2013), as well as common errors and misconceptions demonstrated by children. In presenting area tasks, textbooks often provide regions already subdivided or partitioned, so children simply need to count the number of squares, often counting one by one, to determine the area (Cavanagh, 2008). Representations such as this limit students’ experience with area measurement to simple counting of pre-formed units and later to using formulas without adequate understanding of constraints or adjustments to context. This results in incomplete understandings of fundamental area concepts, including understanding the attribute of area, unit concepts, accumulation and additivity, conservation, and the relation between number and space (Lehrer, 2003; Lehrer, Jacobson et al., 1998; Sarama & Clements, 2009).

A recent review supports the position that curricula have systemic deficits in teaching area measurement (Smith, Males, & Gonulates, 2016). Each of three elementary textbooks presented area measurement in almost exclusively procedural terms. Conceptual principles were infrequently provided and usually only after the procedures. A striking weakness was a lack of support for understanding how the multiplication of lengths produces area measures.

We postulate that addressing such weaknesses and prompting better understanding of area should be based on students’ development of spatial structuring competencies regarding organizing two- or three-dimensional space into orthogonal units. That is, spatial structuring is the mental operation of constructing an organization or form for an object or set of objects in space; it is a form of abstraction which involves the process of selecting, coordinating, unifying, and registering in memory a set of mental objects and actions. Based on Piaget and Inhelder’s (Piaget & Inhelder, 1967; Piaget, Inhelder, & Szeminska, 1948/1960) original formulation of coordinating dimensions, and observations by Outhred and Mitchelmore (1992), spatial structuring takes previously abstracted items as content and integrates them to form new structures, such as using a unit square to compose a row of units. It creates stable patterns of mental actions that an individual uses to link sensory experiences, rather than the sensory input of the experiences themselves, for example, selecting, coordinating, unifying, and registering in memory a set of mental objects and actions so that the row of units becomes a unit that can itself be iterated. Research indicates that such spatial structuring precedes meaningful mathematical use of the structures, such as determining area (Battista & Clements, 1996; Battista, Clements, Arnoff, Battista, & Borrow, 1998; Outhred & Mitchelmore, 1992). The scheme of spatial structuring in this context involves recognition of the goal of partitioning a region into parts, and the activity of organizing the region into a row-and-column structure, with the result of a fully partitioned and quantifiable region. Ultimately the structure is related to the linear dimensions of the rectangular region.

3. Theoretical framework

3.1. Hierarchic interactionalism

Learning trajectories (Simon, 1995) have served as the core of multiple research projects, curricula, and professional development projects (e.g., Clements & Sarama, 2014; Confrey, Maloney, Nguyen, & Rupp, 2014; Wilson, 2014). We define learning trajectories (LTs) as developmental progressions that include descriptions of children’s thinking and learning, as well as a related, conjectured route through a set of instructional activities to achieve a mathematical competence (Clements & Sarama, 2004).

We view LTs and children’s development of geometric measurement understanding through a theoretical lens termed *Hierarchic Interactionalism* (Sarama & Clements, 2009). A main tenet is that children progress through levels of understanding for measurement in ways that can be characterized by specific mental objects and actions (i.e., both concept and process) with the most visible progress through levels for domain-specific topics, such as area measurement. We postulate that various models and types of thinking grow in tandem to a degree, but 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, 1992; Clements, Battista, & Sarama, 2001). This often involves “fallback” to prior levels of thinking under increasingly complex demands. Further, intense “experiences can engender rapid change to a new level,” even progressing through several levels seamlessly (Sarama & Clements, 2009, p. 21). With experience over time, the level of thinking may become robust, and children’s learning follows a predictable pattern (Barrett, Clements, & Sarama, 2017).

Another of the tenets of Hierarchic Interactionalism is the co-mutual development of concepts and procedures (tenet #5, Sarama & Clements, 2009, p. 22). We assume concepts can productively constrain procedures, and that concepts and procedures develop in mutual interaction (Baroody, Lai, & Mix, 2005; Greeno, Riley, & Gelman, 1984; Rittle-Johnson & Siegler, 1998). Thus, distinct from debates about which should be learned first (or focused on more intently), we posit that procedures can be learned meaningfully (connected with conceptual knowledge) or by rote—the latter we refer to as *mechanical procedures* (Hiebert & Wearne, 1993).

3.2. Learning Trajectory for Area Measurement

Thus, our goal is to help children develop and integrate conceptual and procedural knowledge of area measurement by employing instructional practices that build from 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). In Table 1, we summarize the observable behaviors as well as the hypothesized mental actions on objects indicative of the levels of the LT for area measurement we used to differentiate children's responses in the present study. The original LT (Sarama & Clements, 2009) was based on extant research on the implicit rules children appear to use in comparing or determining areas, spatial structuring, and the efficacy of various instructional approaches (see Clements & Sarama, 2009; Sarama & Clements, 2009). We also conducted Rasch analyses of tests designed to measure each level and revised the LT slightly based on these results (Clements, Sarama, & Liu, 2008).

In a subsequent project (Barrett, Clements, Sarama, Miller et al., 2017), we conducted teaching experiments with individual children as well as classroom-based teaching experiments with intensive qualitative analyses to clarify, revise, and expatiate the LT. Table 1 is the result.

4. Purpose

One purpose for this study is to replicate the structure of (replication being critical to educational science, Polanin, Tanner-Smith, & Hennessy, 2015), and also extend, a study completed with children in Grades 2 through 5 by Cullen et al. (in press) with a slightly younger age range (Grades 1–3). We hypothesized that children can learn substantial area competencies (concepts and processes) years earlier than the usual grade placement (e.g., Grade 3). We also extended the previous work in several ways. First, we included analyses of children's use of procedural and conceptual knowledge and individual differences in strategy shifts before and after level transitions. We examined if the experimental interventions would support co-mutual gains in procedural and conceptual knowledge. These analyses enabled us to achieve a second purpose: testing one assumption of the LT by determining whether children can learn advanced levels of thinking directly rather than developing through each level of thinking in the hypothesized developmental progression.

The previous study indicated that two experimental interventions promoted growth for children at the Area Unit Relater and Repeater (AURR) level as well as the Initial Composite Structurer (ICS) level: (a) constructing arrays by drawing parallel line segments in both dimensions across a rectangular region, subdividing it (here called the S intervention) and (b) building a row of square tiles (a unit of units) and iterating to fill a rectangular space (I intervention). The third intervention was a comparison condition (C Intervention), where we repeatedly exposed children to a verbal report of area that reflected the implicit (but never stated or shown) multiplication of length and width. We used these same conditions, extending the work to children who started at less sophisticated levels of the LT. We designed the study to address the following research questions:

1. How do the three treatment conditions affect children's production of accurate numerical measurements? Do effects differ for children of different grades, especially for those earlier than the typical introduction of instruction on area?
2. Are the patterns in children's observable behaviors before and after the shift from measuring area by operating on collections of individual units to coordinating rows and columns, including how they relate linear and area units consistent with the developmental progression?
3. In a related vein, do children operating at level $n - 2$ (or earlier, see Table 1) who experience the modeling of advanced-level strategies from level n learn these directly, or do they evince behaviors indicating that they develop through the sequence of levels ($n - 2$, $n - 1$, then n) as hypothesized in our theoretical framework (Sarama & Clements, 2009)?
4. How do the three treatment conditions affect children's use of strategies and their development of procedural and conceptual knowledge, including transfer of learning? Do those children who advance show increases in procedural knowledge only; do they progress from procedural to conceptual knowledge states; or do they develop intertwined procedural and conceptual knowledge in all phases of learning as hypothesized (Sarama & Clements, 2009, p. 22)?

5. Method

5.1. Microgenetic method

The goal of a microgenetic study is to ascertain how children learn (Siegler, 2006). Microgenetic methods have three main characteristics: (a) observations span the period of rapidly changing competence; (b) within this period, the density of observations is high, relative to the rate of change; and (c) observations are analyzed intensively, with the goal of inferring the representations and processes that gave rise to them. With this in mind, finding a balance between the number of children and number of observations is key in order to capture the change point (the shift described previously), allowing exploration of what happens right before the change, during the change, and after the change has stabilized.

For this study, we implemented the interventions with children whose dominant level of thinking was identified ranging from Area Unit Relater and Repeater (AURR, level n) and Physical Coverer and Counter (PCC, level $n - 2$) to isolate specific instances of growth (defined here as a change to a higher LT level), including better descriptions of behaviors and changes in behaviors indicating non-observable mental actions-on-objects, and to connect those instances to specific aspects of area problems and teaching strategies

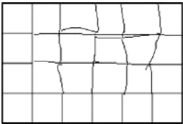
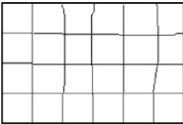
Table 1

Select levels from learning trajectory for area measurement.

Developmental Progression	Mental Actions on Objects	Anticipated Misconceptions or Partial Conceptions
Physical Coverer and Counter (PCC) Attends to some aspects of the structure. <i>Tiling.</i> Completely covers a rectangular space with physical tiles.	With perceptual support, can visualize that regions can be covered by other regions. With strong guidance provided by pre-structured materials, can direct the covering of that space and recognize that covering as complete. Can represent figural units as unanchored, approximately rectangular shapes, aligning them (applied concept of collinearity), but often only intuitively and in one dimension (Mullet & Paques, 1991) or using height + width rules (Cuneo, 1980; Rulence-Paques & Mullet, 1998).	May not organize, coordinate, and structure two-dimensional space without perceptual support. <i>Quantifying.</i> When counting to a total, often relies on existing guides to direct counting but may count unsystematically in the interior. <i>Drawing.</i> When representing a rectangular tiling task, may initially draw approximately rectangular shapes, often leaving overlaps and gaps and often aligning only in one dimension.
Complete Coverer and Counter (CCC) <i>Drawing.</i> Draws a complete covering of a specific region without gaps or overlaps and in approximations of rows (see example on right). <i>Producing.</i> When provided with more than the total number of physical tiles needed, builds a region of specified area. <i>Builds a rectangle with an area of 12 from a pile of 20 tiles.</i>	Applies explicit understanding that entire region must be covered with approximately rectangular shapes (in most rectilinear contexts). Implicit visual patterning of multiple concatenations of rectangles (and constraints of physical materials in physical tiling tasks) guides placement of anchored units in rows, approximately aligned with parallel rows. The counting-all action schemes (without use of rows and columns) are used to enumerate the covering of a region. <i>Note on psychological foundation:</i> Not yet able to understand the basis for the multiplicative transformation of lengths into area measurements.	<i>Quantifying.</i> In counting to a total, may count around the border and then systematically on the inside (possibly in a spiral, in an S pattern, or in approximations of rows). Initially, may not accurately count shapes, losing track or double counting. Later, may not lose track, counting more accurately, but without row or column organization. <i>Drawing.</i> When drawing approximations of rows, often has errors in alignment of the shapes and does not yet recognize the need for equal-sized units.
Area Unit Relater and Repeater (AURR) <i>Quantifying.</i> Counts individual units to measure, guided by rows. <i>Drawing.</i> Draws a complete covering based on an intuitive row or column structure. Children attend to drawing equal sized units, one at a time. <i>Iterating.</i> When provided with fewer than the total number of physical tiles needed, iterates individual tiles to measure and eventually cover a given region. <i>Producing.</i> Builds a region of area from an insufficient number of unit tiles through individual unit iteration (e.g., “leap frogging,” when given a set of n tiles, the child may translate the first tile to represent both tile 1 and tile $n + 1$, the second tile to represent both tile 2 and tile $n + 2$...).	Stronger constraints on object counting (e.g., counts all objects once and only once), and use of rows as an intuitive structure or explicit application of labeling as marker, allows accurate keeping track. <i>Notes on the psychological foundation:</i> This level is foundational to the later development of the multiplicative transformation of lengths into area measurement. This type of multiplicative structure is analogous to n times 1 square unit (e.g., for 12 square units to be drawn one at a time and counted one at a time to be thought of as 12 times 1 square unit). But at this level, a child would not be expected to notice and represent collections of rows along a column to find an array.	<i>Tiling.</i> When asked to cover a region with physical tiles, may cover the whole region but use tiles of varying dimensions. <i>Iterating.</i> When asked to cover without a complete set of physical tiles, may attempt to iterate but unable to maintain equal iterations. <i>Drawing.</i> When drawing, complete coverings may have some errors of alignment.
Initial Composite Structurer (ICS)	Builds, maintains, and manipulates mental images of composite units (e.g., conceptualizing and anticipating their existence or construction before they are	Organizes counting, drawing, or moving objects in composite units (unit of units), yet without a consistent application of that structure. Early on, may be unaware of the congruence of rows, yet use a variety of rows to determine a measure of

(continued on next page)

Table 1 (continued)

Developmental Progression	Mental Actions on Objects	Anticipated Misconceptions or Partial Conceptions
Identifies a square unit as both a unit and a component of a unit of units (a row, column, or group); however, the child needs figural support to structure the space (this may include physical motions of some of the tiles or drawing some collections of units rather than from using the dimensions.) Reasonably estimates areas of regions. May apply an upper (counting all wholes and partials as wholes) or lower bound (counting only wholes as wholes) strategy. Tacitly recognizes usefulness of dimension displays as an indicator of the number of units along one dimension (row or column). <i>When asked to draw a rectangle of specified area, may succeed in identifying dimensions of a region without correctly drawing the array of units.</i> Organizes counting, drawing, or moving of objects in composites units (unit of units). 	perceptually present), structuring them as composites of individual shapes as a single entity – a row (a unit of unit recognized as such). Applies this composite unit repeatedly (built and repeated at least once), but not necessarily exhaustively, as its application remains guided by intuition. <i>Notes on the psychological foundation: This level anticipates later development of the multiplicative transformation of lengths into area measurement. This type of multiplicative structure can be analogous to n rows times m square units per row (e.g., for 12 square units drawn and/or counted, in part, as 3 rows with 4 square units per row).</i>	area. Later, structures the rectangle as a set of rows, understanding the collinearity of rows and the constraint that each row must have the same number of units. <i>Drawing.</i> May draw several rows using parallel line segments but revert to drawing individual squares. May begin drawing by repeating individual square units and curtail this process to draw parallel line segments to demarcate rows. May ignore number labels or other dimension demarcations while drawing collections of units May coordinate only one dimension, the width or the height of the unit and the region.
Area Row and Column Structurer (ARCS) Decomposes and recomposes partial units to make whole units. <i>Drawing.</i> Uses the dimensions to constrain the placement of parallel row and column line segments. Measures both dimensions to determine the size of the iterated squares or rows of squares, as well as determine the number of rows needed in drawing. May not need to complete the drawing to determine the area by counting systematically while attending to rows (most younger children) or computation (repeated addition or multiplication).  <i>Comparing.</i> Explicitly relates size and number of area units (conversion). <i>Producing.</i> When asked to draw a rectangle of specified area, connects the measure of area to a count of area units without connecting the count of area units to the dimensions of the region.	Builds, maintains, and manipulates mental images, as supported by drawings, of composite units; has moved from identifying squares are seen as individual units and as component of a unit of unit (e.g., a row and a column). Applies this composite unit repeatedly and exhaustively to fill the array (or instantiating this operation symbolically)-coordinating this movement in 1-1 correspondence with the elements of the orthogonal column. In a measurement context, applies the concept that the length of a line specifies the number of unit lengths that will fit along that line, but the child may need to create a perceptual array to support such reasoning. This scheme may not be reversible. May connect the meaning of area measurement to a count of area units without relating those area units to linear units.	<i>Drawing.</i> May succeed in identifying dimensions but need to draw the array of square units to check. For example, may realize that a rectangle having an area of 16 square inches has 16 squares in it but struggles to determine the linear dimensions of that rectangle. <i>Quantifying.</i> May form an overgeneralization of the rectangular area formula, such as referring to this operation as n rows times m columns.

so as to extend our knowledge of the critical third component of a LT, instruction that helps children progress.

5.2. Participants

Participants who participated in all sessions were 70 children from 12 classrooms in four elementary schools in a large, suburban school district in the Rocky Mountain region: 29 children from five Grade 1, 23 from four Grade 2, and 18 from three Grade 3 classrooms.

5.3. Procedure

Within each grade level, we randomly assigned children to one of three interventions with weighting because our previous study and our piloting indicated that the control intervention (the C intervention) was unlikely to change children's behaviors. Thus, we included the C intervention but limited the number of children at each grade level in this intervention to five. We balanced the remaining children between the experimental interventions at each grade level, beginning with 12 in each of these interventions at Grade 1, 10 in each at Grade 2, and 7 in each at Grade 3. Due to attrition, we lost one child in the C intervention group at each of Grade 2 and 3, as well as one child at Grade 1. The final count was 27 and 30 in the two experimental interventions (S and I, respectively) and 13 in the control intervention. All children were pretested before the intervention sessions and posttested within 2 weeks after these sessions. All children participated in three 15- to 20-min sessions conducted as one-on-one interviews, each on a different day. Over the three sessions, children were presented a total of nine area measurement tasks, herein called "trials" (we use the term "tasks" to include both the test items and these session trials). Each session occurred during the school day and was videotaped. During each of the interviews, children were provided a standard 12-inch ruler, 20 1-inch tiles, a pencil, and an eraser.

For the pre- and posttests, we first asked the children "What is area?" and then presented them with three area measurement items. In the first item, we presented a 4-by-5 rectangle with the bottom row and left column of squares drawn in. We told the child, "I wanted to cover this rectangle with these squares. I started drawing them in. Please finish the drawing by completely covering the rectangle." In the second item, we presented a 2-inch by 5-inch, blank rectangle and said, "The area of this rectangle is 10 square inches. Draw how each of the 10 square inches fit." In the final area item, we asked children to "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."

Each intervention session consisted of three rounds of presenting the child with a blank rectangle and asking, "What is the area of this rectangle?" We allowed the child time and flexibility to determine the area of the rectangle via any method they desired. Next, we showed the child a video of someone finding the area of the same rectangle. The video shown varied depending on the intervention. Indicating the video, we asked the child, "What did you notice?" We did not question the children further (as one would in a clinical interview), however, for two related reasons. First, this is replication of the earlier study; therefore, the interventions needed to be the same. Extension questioning could fundamentally change the intervention. Secondly, such questioning would also introduce variation as the interviewer/child interactions would not be consistent, adding variability that would make clear warranted statements difficult or impossible.

All children, regardless of intervention group, received Rectangle A on the first trial and were asked to determine the area. They were then shown a video of someone finding the area of Rectangle A using the method identified for their intervention group. In all, we presented children with nine rectangles (Table 2) over the 3-day period.

As previously stated, we included three different treatments in this study: an active control ("C"), and two experimental interventions, one emphasizing iteration of units-of-units ("I") and the other emphasizing subdivision of the entire region into rows and columns ("S"). All three take a worked-examples approach; that is, seeing example problems along with their worked-out solutions (Bokosmaty, Sweller, & Kalyuga, 2014). Although this approach emphasizes procedures, the two interventions were devised to embody the critical concept of spatial structuring. It is important to note that although our usual approach is to match instruction to children's developmental level, in this replication of a study conducted with older students, the instruction was at a high level (which had the added advantage of allowing us to answer the third research question). Videos of the interventions can be viewed here: <http://www.childrensmeasurement.org/resources.html>.

Table 2
Rectangles in order of presentation in the area interventions.

Day	Rectangle	Dimensions (width $\times$ length)
1	A	4" $\times$ 3"
	B	3" $\times$ 7"
	C	7" $\times$ 4"
2	D	2" $\times$ 6"
	E	5" $\times$ 3"
	F	5" $\times$ 6"
3	G	2" $\times$ 5"
	H	4" $\times$ 5"
	I	4" $\times$ 6"

5.3.1. The C intervention

The C intervention was an active control for the effects of repeated exposure to the general features of a rectangular area measurement task. In this limited intervention, children were shown videos of a person using a ruler to measure the length (bottom side) and width (left side) of a rectangle, labeling these dimensions including inch units on the corresponding sides, and then writing the area including square inch units at the top of the paper. The calculations for the area were not shown. Instead, the teacher in the video simply said, for example, “This side is 5 in. and this side is 3 in., so the area of the rectangle is 15 square in.”

Thus, the C intervention did not attend to geometric structure, which was the focus of both the experimental interventions. These both used gestures (Cook, Duffy, & Fenn, 2013; Cook & Goldin-Meadow, 2006; Hu, Ginns, & Bobis, 2015) and verbalizations in a dual-modality presentation technique (i.e., one is visual, one auditory, reducing cognitive load, see Mousavi, Low, & Sweller, 1995; Tindall-Ford, Chandler, & Sweller, 1997), albeit in distinct ways, to model area measurement as spatial structuring.

5.3.2. The I intervention

The I intervention focused on development and iteration of a composite unit (Cullen et al., in press). Children in this intervention group saw videos of a person measuring the width of a rectangle in inches with a ruler, labeling the dimension including inch units, marking tick marks at each inch along that side, then iterating a single tile along that side while saying, “So that makes 1, 2, 3... rows.” Next in the video, she measured and labeled the length side of the rectangle, then laid out a row of inch tiles along the length while saying, “So that makes 1, 2, 3... in a row.” She then taped this row of tiles together, iterated it along the first dimension, and used skip counting to determine the area, marking the area at the top of the paper in square inches.

5.3.3. The S intervention

The S intervention focused on subdividing a rectangle into rows and columns (Cullen et al., in press). Children in this intervention group were shown a video of a person measuring the vertical side of a rectangle in inches with a standard ruler, labeling the length of this side including inch units, then laying the ruler at each mark and drawing horizontal line segments across the rectangle. While drawing these line segments she said, “This side is a inches so that makes a rows [or a columns].” She next measured and labeled the length of the horizontal side, then similarly used the ruler as a guide to draw vertical segments across the rectangle at each inch mark and described these actions similarly. This resulted in an array of inch squares filling the rectangle.

5.4. Coding criteria and procedure

Although we had access to the coding scheme used in the previous study, we decided to use emergent coding in our analysis. We randomly selected two videos from each intervention group and three researchers independently watched the videos, allowing important behaviors and verbalizations to emerge as codes from each. We then met to discuss the independent ways of coding and compile a summative list. This list was then utilized by the three researchers to independently code one more randomly-selected video from each group, while allowing for other codes to emerge. We conferred again to finalize a list of codes. No codes were changed during analysis following the conclusion of the pilot stage. Next, we tagged each video clip using NVivo data analysis software using the qualitative codes through a categorical method. Two raters independently coded each trial and later discussed any disagreement (less than 10%) with each other and a third coder until consensus was reached. Our coding system included four categories that emerged from behaviors and verbalizations children exhibited as they worked on area trials. We abbreviated the categories through use of interrogatives reflecting the category being examined. Specific behaviors of interest were coded with abbreviations placed within their corresponding category.

5.4.1. Broad observable behaviors and strategies

We identified 10 broad, observable behaviors used by children in attempts to solve area tasks (Table 3).

5.4.2. Specific, observable behaviors

This category specifies the various ways in which a child could perform the broad behavior identified. For example, if a child

Table 3
Broad, observable behaviors and strategies used.

Code	Observed Behavior/Strategy
Draws	Draws individual shapes, rows/columns, tick marks, boundary
Demarcates	Demarcates lines
Traces tile	Traces provided square inch tiles
Counts	Performs any counting
Tiles	Tiles area using square inch tiles
Ruler	Uses provided ruler in any way
Composite	Creates a composite unit taping (sometimes holding) together the square inch tiles
Labels	Labels sides of rectangle (e.g., 5 in.)
Gives unit	Gives unit when providing final answer (e.g., 12 square inches)
Uses arithmetical reasoning	Uses arithmetical reasoning at any point to find the area

Table 4

Specific, observable behaviors that reflect how the children implemented strategies and used tools.

Code	Observed Behavior/Strategy
Errors	Committed some type of error (e.g., left gaps while drawing, errors in alignment, miscounts, etc.) that impacted their ability to determine the correct area.
Anchoring	Does or does not anchor when drawing units. Anchoring refers to a drawing strategy in which each new line segment begins from a point on an existing line (Thomassen & Tibosch, 1991). For example, when drawing individual units, rather than drawing next to or over a side of an adjacent unit, a child may share the side of the shape with the one being drawn.
Iterates	Uses an iterating strategy to find answer.
Additive reasoning/Skip counts	Either uses addition or skip counts when finding area from individual units.
Multiplicative reasoning	Uses multiplication when finding their answer.
Attempts to maintain equal size unit	Appears to implement strategy with an attempt to keep unit sizes the same.
Recognizes need for equal size unit	Either clearly recognizes the need for equal sized units or clearly does not recognize the need for equal sized units (often discerned through verbalization).
Recognizes row or column	Clearly recognizes either rows or columns in implementing their strategy.
Labels sides	Labels one, two, or all sides of the rectangle when solving the area problem.
Gives units prompted/unprompted & correct/incorrect	Gives unit (i.e., square inches) with final answer either prompted or unprompted by the facilitator and either correctly or incorrectly.

labeled sides of a rectangle during an area trial and we coded this with “Labels” from above, we would also assign one of the codes specifying how many sides the child labeled (“Labels one/two/all sides”). Overall, we coded 13 specific, observable behaviors and strategies (Table 4).

5.4.3. General approach to determining area

This category contains three approaches, mimic, mechanical, or conceptual (Table 5). A mimic response was one in which a child appeared to attempt elements of the procedure shown in the intervention video, correctly or incorrectly, but with no solid indication of conceptual understanding (“going through the motions”); that is, they employed a behavior demonstrated in the intervention video without any demonstration of grasping the relationship between the behavior and its intended effect (Simon & Tzur, 2004). For example, creating a composite unit and then attempting to iterate it with arbitrary and inaccurate intervals, or subdividing a rectangle in an unguided fashion with an arbitrary number and spacing of demarcations.

Responses coded as mechanical were those in which a child appeared to perform a procedure by rote that was different in at least some way from the intervention’s procedure but did not demonstrate an understanding of area concepts such as those detailed in the background section. These responses could encompass basic approaches as well as relatively advanced approaches to finding area, such as the use of the $a = lw$ formula.

Responses coded as conceptual were those in which a child performed procedures while explicitly demonstrating that their actions were informed by an understanding of area concepts. Verbal reports, self-corrections after mistakes, novel approaches that deviated from replication of previous procedures (from the video or their previous behaviors), and children’s responses to the “What did you notice?” query after being shown each video were often useful in distinguishing conceptual responses from mechanical responses. Responses coded as conceptual skewed towards higher LT level ratings and featured a larger range of behaviors within the response. This suggests that the mimic, mechanical (i.e., procedural only), and conceptual codes are useful indicators of relative sophistication.

5.4.4. LT level

The final category was the LT level (Table 1) assigned to the child. We coded LT levels in two ways. We assigned an LT level to performance on each area task (i.e., each intervention trial as well as each pretest and posttest item had such a level assignment, for a total of 15 per child). If the child showed evidence of behaviors or thinking from two levels on one specific task, the higher (almost always the latter) one was coded. This was based on the behaviors we observed and the child’s verbal explanation(s) that were indicative of thinking at a particular level as defined in the LT (Table 1). We used these individual codes to help identify shifts in thinking for each child during the study. Separately, for each assessment (pretest and posttest), we made a determination about the child’s dominant level of thinking at the beginning and the end of the study by the highest level (test items were designed to elicit different levels of thinking).

Table 5

Codes for general approach to determining area.

Code (abbreviation)	Observed Behavior/Strategy
Mimic	Applies a procedure that closely mirrors the behaviors previously observed in intervention videos.
Mechanical	Applies a procedure to yield an answer without evincing an understanding of “why” the procedure measures area.
Conceptual	Applies a procedure to yield an answer and indicates an understanding of “why” the procedures used yield an accurate measurement of area.

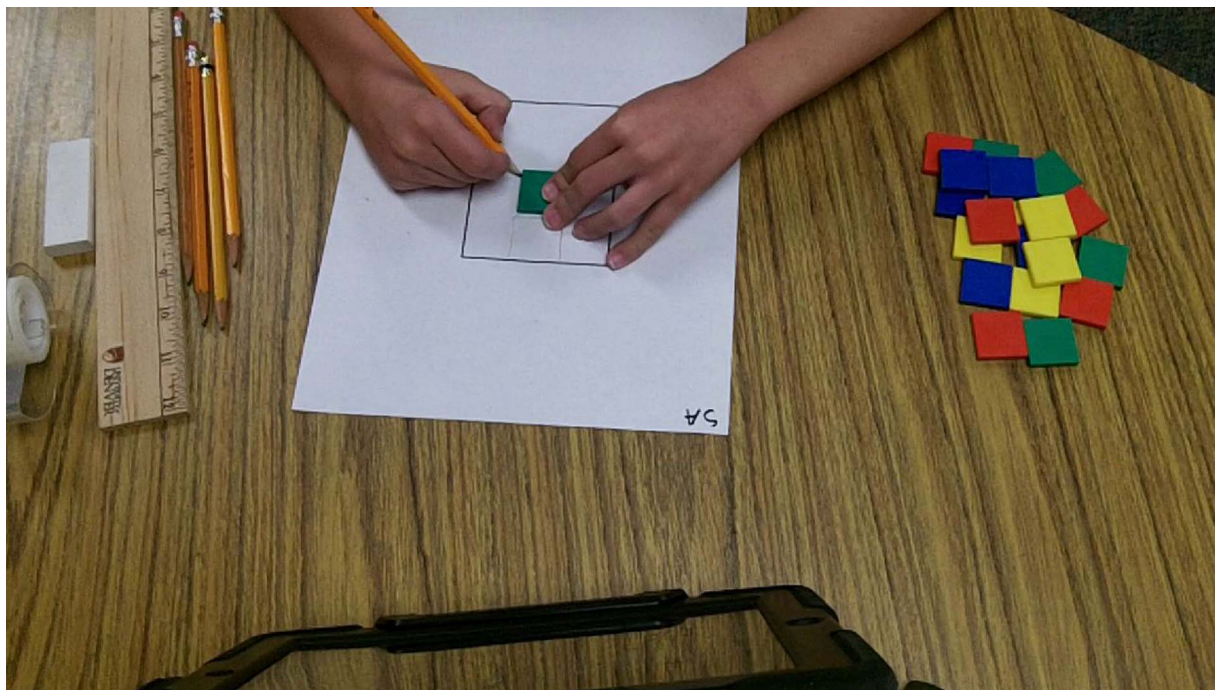


Fig. 1. Illustration of child S2-4's strategies on the first trial.

5.4.5. A coding example

Our coding of the first two trials on day two illustrate how we used the codes to reveal change. On the first trial (see Fig. 1), child S2-4 (S group, grade 2, student #4) drew individual squares to cover by tracing one tile, anchoring each new square to one previously drawn. They counted each drawn square before reporting a numeric answer. The student did not provide a unit with their numeric answer until they were prompted to do so, but it was incorrect. Hence, on this trial the codes for broad observable behaviors were Traces tile, Counts, and Gives unit (Table 3), and the codes for specific, observable behaviors Anchoring, Attempts to maintain equal size unit, and Gives unit prompted/incorrect (Table 4) were applied. Because these behaviors were different from the intervention's procedure and seemed to be informed by conceptual principles, the general approach was coded as Conceptual (Table 5) and the LT level (Table 1) was coded as AURR. On the following trial, the child used the inch markings on the ruler to draw row and column line segments. Thus, the broad observable behavior codes for this trial were Demarcates (lines), Traces tile, Counts, Ruler, and Gives unit (Table 3); the specific, observable behavior codes were Anchoring, Attempts to maintain equal size unit, Recognizes row or column, Gives unit unprompted/correct (Table 4); the general approach was coded as Conceptual (Table 5); and the LT level (Table 1) was coded as ICS.

On the second trial (Fig. 2) the new strategies of demarcating and using a ruler were employed. These new strategies were accompanied by the recognition of a row or column, and giving the correct unit without prompting from the facilitator. This represents a change from tracing single squares until the rectangle was covered, to demarcating inch-wide rows and columns. Although both approaches yield a correct answer, the second is more sophisticated. The child also stated the unit without being prompted, indicating an understanding of the nature of measurement. In terms of analysis, examining which codes are changed or added and which codes remained the same reveals what changed from one trial to the next, indicating that a change point has occurred, as well as indicating the form of those changes.

6. Results

6.1. Correctly determining the area of a rectangle

We first examined if children improved in their ability to correctly determine the area of a rectangle, a necessary condition for examining relative efficacy of the interventions. For this, we used a score of 1 to indicate a child correctly determined the area, 0.5 for correctly determining the area after a pre-determined prompt, and 0 for an incorrect response. In Fig. 3 and looking at all children, we show the percentage of correct scores (y-axis) by intervention group (C, S, and I) in determining the area of each rectangle in Trials 1 through 9 (x-axis).

We see in the positive slopes of the regression lines that each group showed improvement in correctly determining the area through the trials. Children in the S and I groups, with slopes of 0.05 and 0.03, respectively, showed more improvement than children in the C group, with a slope of 0.01. On average, over the nine trials, the percentage of children in the S group answering correctly

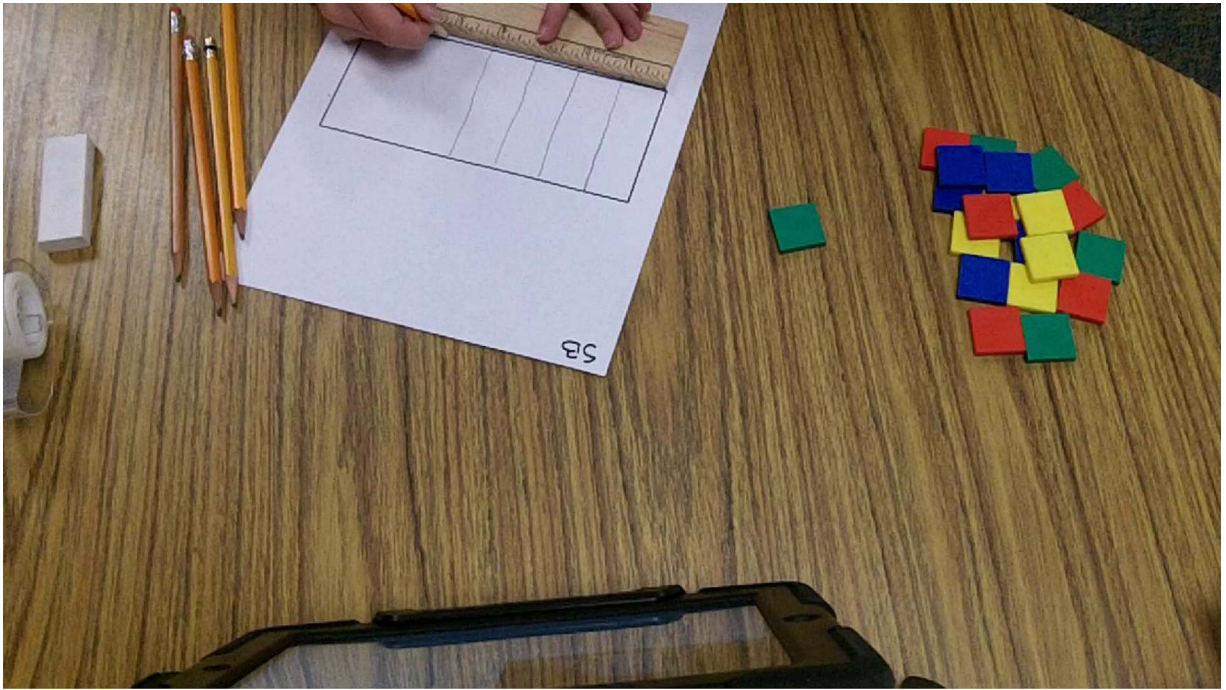


Fig. 2. Illustration of child S2-4's strategies on the 2nd trial.

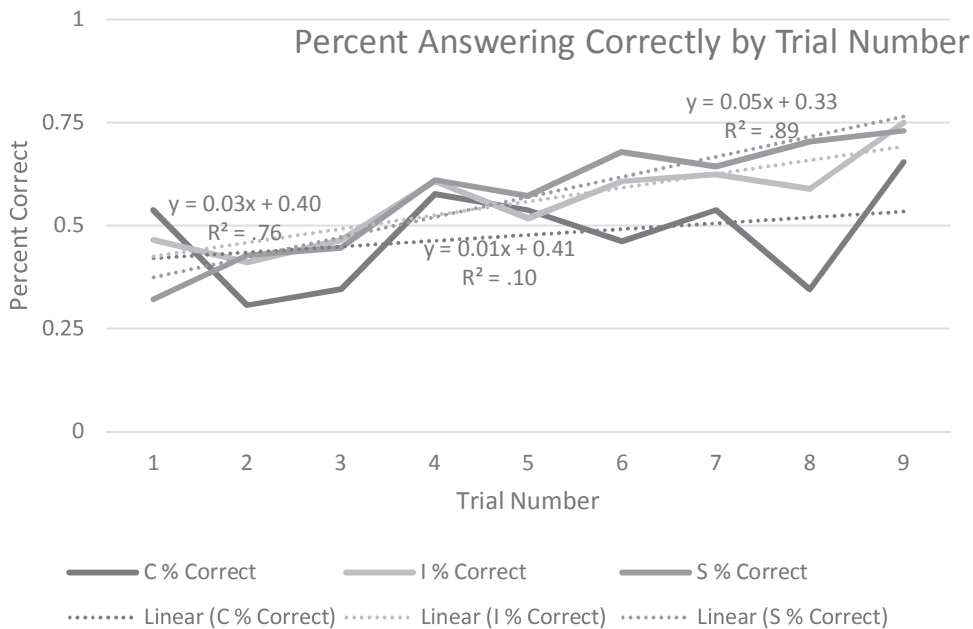


Fig. 3. Percentage of children determining the area of each rectangle by intervention group.

increased by about 39%, in the I group by about 27%, and in the C group by about 11%. The R^2 for the C group indicates that only about 10% of the variability is explained by the model. The larger R^2 values of 89% and 76% for the S and I groups, respectively, indicate that the analysis explains more of the variance for the two experimental interventions than the C intervention. Taken together, the larger R^2 values along with the larger positive slopes for the C and I groups suggest that the two experimental interventions were more effective in helping children learn to correctly determine the areas of rectangles than the C intervention.

Recall that one goal of our study was to replicate the previous study with slightly younger children. To examine whether grade level moderated the results, we compare the percentage of children correctly determining the area by grade level in Figs. 4–6.

In all three intervention groups we see children in Grade 3 more successful in correctly determining the area of the rectangles than

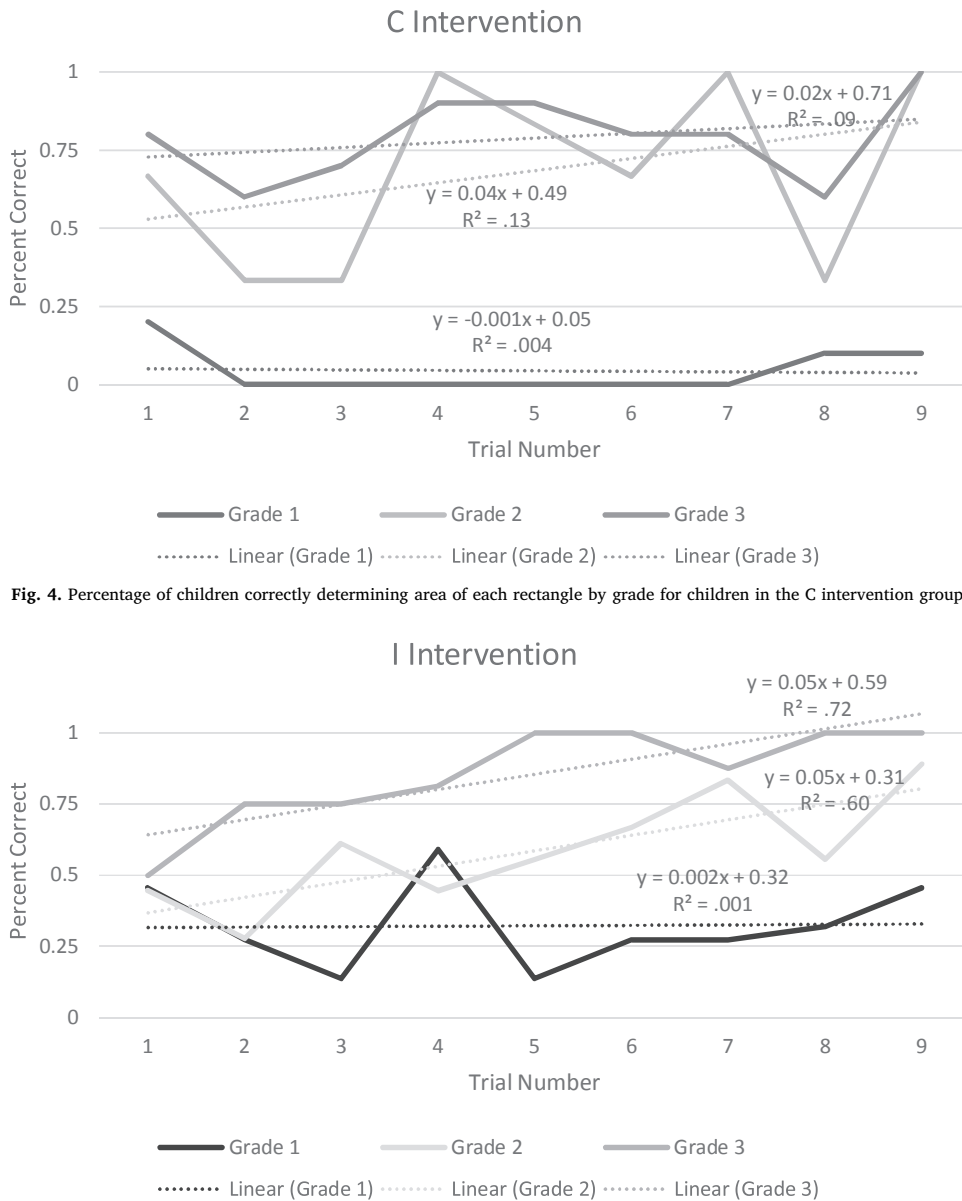


Fig. 4. Percentage of children correctly determining area of each rectangle by grade for children in the C intervention group.

Fig. 5. Percentage of children correctly determining area of each rectangle by grade for children in the I intervention group.

children in either Grade 2 or Grade 1, and children in Grade 2 more successful than children in Grade 1 (albeit with little difference in the C group between Grades 2 and 3).

In Grade 1, we see from the near-zero slopes of the regression lines that both the C and I interventions appear ineffective at helping children correctly determine the area of a rectangle. In the S group, however, a slope of 0.04 indicates that children showed an average overall increase of about 30% in their correctness scores across the nine trials. Additionally, the R^2 value for this group indicates that about 65% of the variability in the group can be explained by this analysis. This is an indication that the increase can be attributed, in large part, to the intervention. Grade 2 children in the C group appeared to show a larger increase than children in either Grade 1 or Grade 3. However, the R^2 for this group also indicates that only about 13% of the variation can be attributed to the model. Additionally, several Grade 2 children spontaneously talked about working with arrays in their math classes during the study, potentially confounding our interventions (we did not ask for elaborations to avoid further confounding).

Children in Grades 2 and 3 showed similar increases in the I and S groups. Regression line slopes for Grade 2 were 0.055 and 0.051 for the I and S groups, with corresponding R^2 values of 0.60 and 0.63, respectively; slopes for Grade 3 were 0.053 and 0.064, with corresponding R^2 values of 0.72 and 0.73, respectively. Thus, children in both the I and S groups improved in correctly determining the area of the rectangles in those grades.

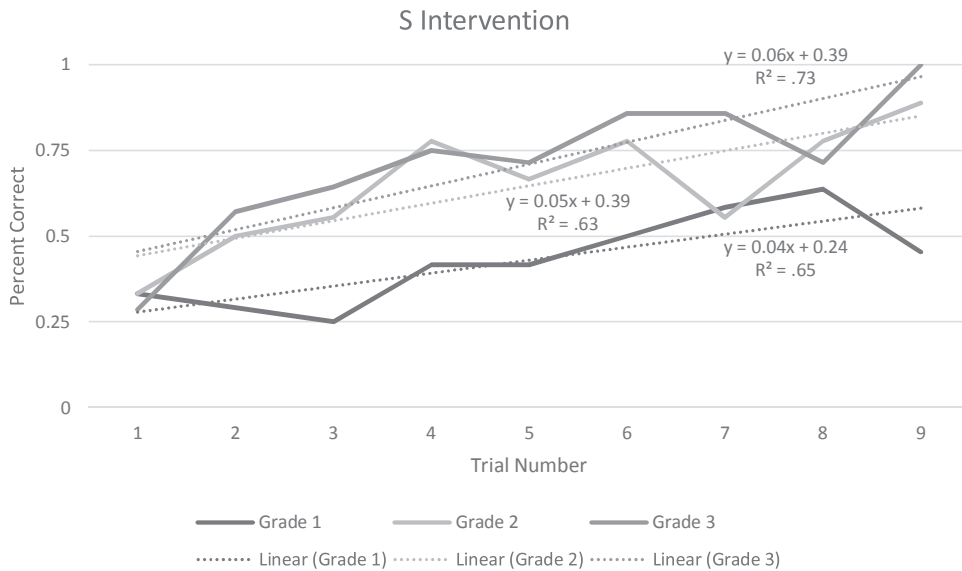


Fig. 6. Percentage correctly determining area of each rectangle by grade for children in the S intervention group.

6.2. Patterns of change in levels of thinking

We also analyzed children's movement through LT levels, which produced several consistent patterns of changes, discussed in the following sections. After identifying levels on the pre- and posttest (one level for each test as a whole) and for each trial, we identified those children demonstrating clear movement between contiguous levels during the intervention (which we termed "growth points"). We explored the behavioral and strategic changes that we observed between two trials. Overall, 24 of the 70 children showed evidence of at least a full level change during the study. Of those 24, four children were coded as No Claim (NC) or Area Quantity Recognizer (AQR) on most trials in their intervention. NC was coded when we are unable to assign an LT level based on a child's observed behaviors during a trial, usually because the child did not provide an answer or we did not see not sufficient behavioral evidence to assign one; AQR was used to indicate we only saw evidence of recognizing area but no behavior indicative of higher LT level approaches. Although these children eventually revealed higher LT strategies, there were multiple intermediate trials with either no response or with exactly the same response, so we could not determine specific growth points relevant to our research questions, and therefore, they were included in the pre-post comparisons but not in the analyses of specific growth points. The LT levels (see Table 1) most commonly observed included Physical Coverer and Counter (PCC), Complete Coverer and Counter (CCC), Area Unit Relater and Repeater (AURR), and Initial Composite Structurer (ICS).

6.2.1. Level change in the C group

In Table 6, we show the levels coded on each of the tests and trials in the intervention for the two children in the C group who demonstrated growth. One child moved from PCC at the pretest to CCC and remained at one of those two levels. The other child moved from ICS to ARCS, seeming to use the modeled procedure.

Table 6
LT level codes for children demonstrating growth in the C group.

	C1-1	C3-2
Pre	PCC	ICS
T1	CCC	ICS
T2	CCC	ICS
T3	CCC	ICS
T4	CCC	ARCS
T5	PCC	ICS
T6	CCC	ICS
T7	CCC	ARCS
T8	CCC	ARCS
T9	CCC	ARCS
Post	CCC	ARCS

Note: C1-1 indicates C group, grade 1, student #1; T = Trial; PCC = Physical Coverer and Counter; CCC = Complete Coverer and Counter; ICS = Initial Composite Structurer; ARCS = Area Row and Column Structurer.

Table 7

LT level codings for children demonstrating growth in the I group.

	I1-1	I1-2	I2-3	I2-4	I1-5	I2-6	I3-7	I2-8	I2-9	I3-10
Pre	PCC	PCC	PCC	PCC	CCC	CCC	CCC	AURR	AURR	AURR
T1	CCC	PCC	NC	AURR	CCC	CCC	PCC	NC	NC	AQR
T2	CCC	PCC	NC	AURR	CCC	CCC	PCC	NC	NC	ICS
T3	CCC	PCC	ICS	ICS	AURR	AURR	PCC	NC	NC	ICS
T4	CCC	CCC	NC	ICS	CCC	CCC	AURR	NC	ICS	ICS
T5	CCC	CCC	ICS	ICS	AURR	ICS	ICS	NC	ICS	ICS
T6	CCC	NC	ICS	ICS	ICS	ICS	AURR	NC	ICS	ICS
T7	CCC	ICS	ICS	AURR	CCC	ICS	AURR	ICS	ICS	ICS
T8	CCC	ICS	ICS	AURR	ICS	ICS	AURR	ICS	ICS	ICS
T9	CCC	ICS	ICS	ICS	ICS	ICS	AURR	ICS	ICS	ICS
Post	CCC	AURR	CCC	AURR	ICS	AURR	ARCS	ICS	ICS	ICS

Note: I1-1 indicates I group, grade 1, student #1; T = Trial; NC = No Claim; PCC = Physical Coverer and Counter; CCC = Complete Coverer and Counter; AURR = Area Unit Relater and Repeater; ICS = Initial Composite Structurer; ARCS = Area Row and Column Structurer.

We now look at the children in the experimental groups who showed evidence of growth during the intervention. We start with the I group (see Table 7, ordered left to right from lower to higher starting LT levels, and by grade within levels).

6.2.2. Level change PCC into CCC in the I group

Two children in the I group (I1-1 and I1-2) exhibited a level change from PCC (Physical Coverer and Counter) to CCC (Complete Coverer and Counter). This growth was characterized by the elimination of errors (especially of alignment), a movement from simple drawing of small regions within the larger region to tiling and tracing, and increased attention of filling regions completely. Although they did not build and physically iterate composite units, which was demonstrated in the video, they appeared to learn how square units can be used to constitute a complete covering of the rectangle; that is, although they did not grasp the creation of a composite unit or physical iteration, they attended more to covering the region with square units, eliminating gaps to fill the space.

6.2.3. Level change from CCC to AURR in the I group

No consistent behavioral patterns could be gleaned from the two children (I1-5 and I2-6) who moved from CCC to AURR (Area Unit Relater and Repeater), but both children evinced this change point between the second and third rectangles presented on Day 1.

6.2.4. Level change from CCC to ICS in the I group

Only one child (I2-3) made the jump to ICS (Initial Composite Structurer), but because the change appeared to skip a level, it may be relevant to research question 3. This child demonstrated CCC behaviors on the pretest, but could not be reliably coded on the first two trials of the first day, and then showed ICS behaviors on the third trial. Therefore, it is not possible to say if the child operated at the AURR level for a period of time. (This child scored lower than ICS on the posttest. Hierarchic Interactionism posits that children will often not operate at only one level but may evince an earlier level depending on internal factors—e.g., fatigue, as there were three items rather than each session's single task—or external factors such as distraction.)

6.2.5. Level change from AURR to ICS in the I group

Even though two (I1-5 and I2-4) of the four children (including I3-7 and I3-10) exhibiting this growth did not create a manipulative composite unit as a strategy, demonstration in the videos seemed to catalyze an understanding of rows and columns. That is, children moved from tiling and tracing with anchoring to iteration, as shown in the video, as well as multiplicative reasoning; for example, they would iterate a unit the entire length of the row or column drawing tick marks instead of tracing the entire square, or create a composite unit and iterate this unit-of-units throughout the region. Some would iterate both a row and then a column before using multiplicative reasoning to find the area.

Finally, we look at the children in the S group who showed evidence of growth during the intervention. In Table 8, we provide a summary of the levels coded for each of these children on each trial during the intervention.

6.2.6. Level change from PCC into CCC in the S intervention

Children in the S intervention more explicitly adopted the subdividing strategy than did children in the other interventions adopt the strategy they viewed. However, this appeared to represent initial experimentation with the strategy because all four children (S1-1, S1-2, S1-3, and S2-4) starting at the PCC level were coded as using mechanical strategies with errors (e.g., tiling, tracing the tiles, and counting); three of the four children (all but S1-3) then mimicked the subdividing strategy, decreasing the tracing of tiles and use of anchoring and increasing ruler use, demarcation, and drawing full line segments.

6.2.7. Level change from CCC into AURR in the S intervention

On the first trial of the second day, two children (S2-4 and S2-5) demonstrated level change from the CCC level to the AURR level, changing from mimicking behavior to behavior indicative of greater conceptual understanding. For example, in prior tasks S2-5 employed a demarcation strategy using a ruler, but did not line up the ruler correctly, and drew lines with relative haste. This resulted

Table 8

LT level codings for children demonstrating growth in the S group.

	S1-1	S1-2	S1-3	S2-4	S2-5	S1-6	S2-7	S2-8	S1-9	S2-10	S3-11	S3-12
Pre	PCC	PCC	PCC	PCC	PCC	CCC	CCC	CCC	AURR	AURR	AURR	ICS
T1	PCC	NC	CCC	PCC	CCC	PCC	CCC	ICS	AURR	PCC	AURR	ICS
T2	PCC	PCC	CCC	CCC	CCC	CCC	CCC	AURR	AURR	ICS	ICS	ARCS
T3	PCC	CCC	CCC	CCC	CCC	CCC	ICS	ICS	ICS	ICS	ICS	ARCS
T4	PCC	CCC	CCC	AURR	AURR	CCC	ICS	AURR	AURR	ICS	ICS	ICS
T5	CCC	CCC	AURR	ICS	AURR	ICS	ICS	AURR	ICS	ICS	ICS	ICS
T6	CCC	CCC	AURR	ICS	AURR	ICS	ICS	AURR	CCC	ICS	ICS	ARCS
T7	CCC	NC	AURR	CCC	AURR	ICS	ARCS	AURR	ICS	ICS	ICS	ICS
T8	CCC	CCC	AURR	ICS	AURR	ICS	ARCS	AURR	ICS	ICS	ICS	ARCS
T9	CCC	CCC	AURR	ICS	AURR	ICS	ARCS	ICS	ICS	ICS	ICS	ARCS
Post	CCC	CCC	AURR	CCC	AURR	ICS	ICS	ICS	ICS	ICS	ICS	ARCS

Note: S1-1 indicates S group, grade 1, student #1; T = Trial; NC = No Claim; PCC = Physical Coverer and Counter; CCC = Complete Coverer and Counter; AURR = Area Unit Relater and Repeater; ICS = Initial Composite Structurer; ARCS = Area Row and Column Structurer.

in inconsistent unit sizes. On this task, she lined up the ruler correctly, and spent more time making sure lines were drawn in a straight manner across the rectangle. This was interpreted as evidence of the child understanding the need for equal size units.

6.2.8. Level change from CCC into ICS in the S intervention

Development from CCC into ICS by three children (S1-6, S2-7, and S2-8) showed a robust growth pattern, indicating progression from mimicking to conceptual understanding of spatial structure and area. As with the CCC to AURR change (i.e., the two children who evinced AURR behaviors during the transition), these children initially mimicked, but with errors in ruler use and alignment while demarcating, which lead to unequal-size subdivisions and then progressed to careful and error-free demarcation. Further, at the ICS level, these children constructed composite units, skip counted those units, and reported the total number along with the correct measurement unit. Regarding research question 3, although the table seems to show a skipped level, because two of the three showed signs of AURR during the transition, there is little evidence of a departure from the developmental progression.

6.2.9. Level change from AURR into ICS in the S intervention

Similarly, these three children (S1-9, S2-4, and S3-11) moved from tracing or tiling to subdivision, always using the ruler for careful measurement then sometimes using a ruler to draw and other times drawing freehand to create rows and columns.

6.2.10. Comparing interventions on LT level growth

The I and S interventions were similar in their effectiveness at promoting level change from one level to another: 10 children (36%) in the I group demonstrated observed LT level growth compared to 12 (43%) in the S group, and 2 (15%) in the C group.

6.3. Procedural knowledge, conceptual understanding, and transfer

6.3.1. Comparing interventions' effects on children's movement from mechanical solutions (procedural only) to conceptual understanding

There was a difference between the interventions in the pace children moved to a conceptual approach; that is, the number of tasks we presented to a child before observing evidence of growth into the next LT level. We presented each child with a total of 15 tasks—three pretest items, nine intervention trials, and three posttest items—and examined the point at which the child first demonstrated a conceptual approach (see definitions in Table 5). We included pretest items in this analysis because we considered the possibility of observed growth due to exposure to these items. For this comparison, we only considered children who demonstrated growth into a higher LT level. We excluded the two children from the C group due to the small sample size, as well as two children from the S group because they consistently demonstrated conceptual approaches beginning with the second item of the pretest, suggesting that their conceptual understanding was consolidated prior to the intervention. The remaining 10 children from the I group and 10 children from the S group all demonstrated growth into conceptual approaches during the intervention.

In the I group, the interval at which children demonstrated a change to a conceptual approach was usually after five or six trials and eight or nine tasks (recall this includes the pretest items). In the S group, it was typically after three trials and six tasks. Across all tasks, 29% of the I group responses were rated as conceptual, 53% as mechanical, 10% as mimic, and 7% as mixed mimic and mechanical. In the S group, 47% were rated as conceptual, 36% as mechanical, 12% as mimic; 3% as mixed mechanical and mimic, and 1% as mixed procedural and conceptual. These results suggest that elements of the S intervention were more effective at eliciting a conceptual approach relative to the I and C interventions.

6.3.2. Transfer of learning in the I and S interventions

We explored questions of transfer of learning by examining performance on Items 2 and 3 of the posttest. In Item 2, a near transfer item (resembling trials in the sessions), we presented children with a rectangle and told them the area was 10 square inches, then we asked them to “draw how the 10 square inches fit.” In Item 3, a far transfer item (beyond competencies taught in the trials), we presented only a blank piece of paper and asked children to “draw a rectangle that has an area of 8 square inches,” as well as to “show

how the 8 square inches fit.”

In the I group, 10 children demonstrated LT level growth. On the two posttest items, only one of these children attempted to construct and iterate composite units (haltingly lining up and moving tiles, but then putting them aside and using a different strategy) despite this being a defining characteristic of the intervention. Similarly, only three of six children coded as demonstrating a conceptual approach on Items 2 and 3 of the posttest (the transfer items) utilized another characteristic of the I intervention, the drawing of tick marks along one side of a rectangle. On Item 3, one of five children demonstrated a mechanical approach and three out of five children demonstrated a conceptual approach to this item.

The most common strategies employed across Items 2 and 3 of the posttest by children in the I group involved using 1-inch square tiles for purposes other than covering an area with the tiles, such as tracing or drawing of tick marks (three did not use tiles or used them only once). Children may have generated these strategies from features of the videos, although they were not explicitly demonstrated in the videos.

In the S group, 12 children demonstrated LT level growth. Children in the S group used the demarcation strategy more than the I group (121 vs. 34 uses overall; 20 vs. 6 uses on Items 2 and 3), which is a defining characteristic of this intervention, on Items 2 and 3 of the posttest. For example, 11 children used demarcation on at least one of the items and nine used demarcation on both items; seven used a ruler to guide demarcation (a characteristic of the video), whereas four used a tile to guide demarcation (one used both).

Other characteristic elements of the S intervention that children applied in these two items were the use of a ruler to measure sides of the rectangle (six children) and the labeling of the dimensions of the sides of the rectangle (four children). In contrast, a notable element of the S intervention that was seldom used on the posttest was skip counting; only three of 12 children demonstrated skip counting, only two verbalized their counting as in the video, and only one gestured across a row as in the video.

Of the 12 children, 10 demonstrated an approach to Item 2 of the posttest that was coded as conceptual. Of these children, all 10 employed a demarcation strategy, which none had used on the pretest. Seven used a ruler and three used a tile, the latter was never shown in the videos. On Item 3 of the posttest, two were coded as mechanical, one was coded as a mixture of mimic and mechanical, and nine were coded as conceptual (including one who was remarkable for going beyond the need for careful demarcation by abstracting what it represented, demarcating in a hasty way, unconcerned with keeping the lines straight and even, but rapidly drawing the correct number of squares). Of the three children who labeled the dimensions of the rectangle, all were coded as ICS or ARCS.

After watching the videos, some children demonstrated behaviors that they did not demonstrate previously. Although these behaviors were not copies of what they saw in the video, we conclude that the new behaviors were inspired by the videos because they featured aspects of the strategies demonstrated in the videos, and there was commonality across children in the form of these new behaviors.

Thus, the I and S interventions were similar in promoting overall LT level growth; however, the subdividing approach of the S intervention appeared easier for the children to incorporate and transfer to novel contexts relative to the iterating approach of the I intervention. Although children in both groups demonstrated a greater breadth of strategy use on the posttest, the S intervention was more effective at teaching children specific procedures and simultaneously promoting conceptual and transferable learning.

7. Discussion and implications

Measurement of area is important in and of itself. If learned well, it also builds a foundation for other STEM topics including multiplication of integers and rational numbers, fractions and operations with fractions, and composition of geometric figures. Unfortunately, in the United States, the teaching and learning of area measurement is not adequate. To help ameliorate this situation, we replicated and extended a study we conducted previously with intermediate-grade students (Cullen et al., *in press*) to evaluate the effects on younger children of three interventions, two of which were designed to support children's growth in measuring rectangular regions but in different ways. Here we investigated how these strategic activity settings would support or not support younger children's grasp of important concepts that enable them to structure planar space, enabling them to improve their ways of measuring the area of rectangular figures.

7.1. Correctly determining the area of a rectangle

Our first research question addressed how the three instructional conditions affect children's production of an accurate numerical measurement and their strategy use. We also asked whether effects differed for children of different grades, especially for those earlier than the typical introduction of instruction on area.

As hypothesized, children in both the I (Iteration of a composite unit) and S (Subdividing into rows and columns) groups increased their correctness scores more than children in the C (Control) group. Examining the effects by grade level, Grade 2 children in the C group increased their correctness scores more than the C-group children in either Grade 1 or Grade 3. The authors of the *Common Core State Standards—Mathematics* (National Governor's Association Center for Best Practices & Council of Chief State School Officers, 2010) introduce the foundation for multiplication in Grade 2. Additionally, children in our study talked about working with arrays in their math classes during the study, which may explain the higher correctness scores among children in Grade 2 than those in Grade 1 and Grade 3.

Both the C and I interventions appeared generally ineffective with Grade 1 children. Although the S intervention was more effective, these first graders learned less than older children in that intervention. Grades 2 and 3 children showed similar increase in correctness scores in the I and S interventions. This finding is also important because teachers in Grades 2 and 3 may find either

model for intervention leading to improved performance if they measure that by correctness alone. However, we found that the S intervention supported greater transfer across tasks and supported conceptual growth more often, an issue to which we return.

7.2. Patterns of change in levels of thinking and movement through levels

The second research question was: Are the patterns in children's observable behaviors before and after the shift from measuring area by operating on collections of individual units to coordinating rows and columns, including how they relate linear and area units consistent with the developmental progression?

Examining children's learning by LT level, children in the I group who progressed from the Physical Coverer and Counter (PCC) to the Complete Coverer and Counter (CCC) appeared to gain a greater understanding and procedural competence in filling regions with units. They did not use composite units nor physical iteration but may have established a goal sketch (Siegler, 1996) of such complete fillings that directed their own tiling and tracing. I-group children who progressed from CCC to Area Unit Relater and Repeater (AURR) did so rapidly, after only two trials on the first day. Children who progressed from AURR to Initial Composite Structurer (ICS) also did not use composite units as they were demonstrated in the videos, but the I intervention appeared to catalyze their perception and replication of the rows and columns structure. Children who had tiled and traced using anchoring now formed rows and columns. In summary, the I intervention, with an emphasis on composite units and iteration, featured a visual representation of a single row. However, children did not use the video strategy of creating and iterating a composite unit when solving the trials themselves. Instead, the I intervention appeared to provide a model of array creation from which children abstracted the row and/or column structure that allowed them to use other strategies (such as skip counting) that depends on such conceptions and builds proficiencies leading to the ICS level. That is, they learned to use a unit of units to move beyond counting single square units as well as to apply that unit of units repeatedly to determine the area.

Unlike those in the I group, children in the S group more directly and completely adopted that intervention's subdividing strategy. This appeared initially to be mimicry of the demonstrated strategy, especially for children who progressed from PCC to CCC. For these children, tiling and drawing were replaced by demarcation as the strategy of choice, as a result of mimicking the intervention video. S-group children who progressed from the CCC to AURR replaced mimicking the demarcation strategy to evince explicit signs of conceptual understanding, especially by recognizing the need for equal-size units. However, they did not appear to use rows and columns, and even abandoned their mimicry of ruler use, and instead traced tiles to draw these equal-size units. Compared to those in the I intervention, children in the S intervention needed more video viewings (i.e., three to four) to engender this change. Children who progressed from CCC to ICS showed clear signs of moving from mimicry to conceptually-driven strategies based on children's concept of rows and/or columns as composite units, evidenced by their careful (and generally error-free) demarcation and skip counting of rows or columns. Additionally, although few children at the CCC level identified the unit label (i.e., "square units") with their answer, all provided them once they reached ICS. This may also be a result of seeing the S intervention video, where units are identified at the end of the video, and is further evidence of the S-intervention videos' utility in engendering specific behaviors.

Finally, children in the S intervention who progressed from the AURR to the ICS or Area Row and Column Structurer (ARCS) levels during the study also showed a recognition of rows and columns and used rulers and freehand drawing, seemingly from adopting the demarcation and line-drawing strategy from the video. In summary, the S intervention, with an emphasis on subdivision and its visual presentation of a completely subdivided array, was more successful than the I intervention in demonstrating a strategy that these young children could reproduce and use to build conceptual understanding. The visual residue—the drawn images left showing the partitioning structure—may have contributed to the S-video's impact compared to the I-video which left no such visual record. Nearly all S-group children eventually utilized the demonstrated strategy to measure area. Thus, children's behaviors were consistent with those predicted by the LT's developmental progression (Table 1), answering the second research question.

The third, related research question was: Do children operating at level $n - 2$ who experience the modeling of advanced-level strategies from level n learn these directly, or do they evince behaviors indicating that they develop through the sequence of levels ($n - 2$, $n - 1$, then n) as hypothesized in our theoretical framework (Sarama & Clements, 2009)? The finding that almost all (20 of 24 with two additional children who vacillated between contiguous levels) children moved through the LT levels (e.g., n to $n + 1$) even when the intervention was at a higher level (e.g., $n + 2$, $n + 3$, or $n + 4$ given n is the pretest level of various children) provides empirical support for our LT for area measurement. Most children were observed progressing a level (or two) at a time (sometimes during a single session or trial) rather than directly adopting behaviors of the target level, even as they repeatedly viewed a strategy that was from that more advanced (often considerably higher) level.

This is not to say that children advanced lock step through the levels nor that every intermediate level was or could be observed. According to the theory of Hierarchic Interactionism (Sarama & Clements, 2009), this would not be expected. The theory posits that various levels of thinking grow in tandem to a degree, and "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" and that "under conditions of increased task complexity, stress, or failure this probability level decreases and an earlier level serves as a fallback position" (Sarama & Clements, 2009, p. 21). However, it is important that even when children did not evince an intermediate level (e.g., AURR) for a full trial (e.g., as they moved CCC into ICS) most children showed behaviors indicative of the intermediate level during their transitions. Along with the recognition that certain behaviors were not emphasized by the study's tasks (e.g., there were no different size units or questions about how the number would change if the unit size changed as would engender AURR behaviors), and that children do not always evince each level of learning (Steffe & Cobb, 1988) especially with intense instruction (Clements & Sarama, 2014), there is little evidence of departure from the developmental progression. Again, the most cogent evidence is that even though the instruction targeted a single higher level, no child directly evinced the target level that was modeled. Instead, all children

developed through the intermediate levels, despite the absence of any instruction addressing levels between the child's initial level of thinking and the target level. This provides support for the hypothesized sequence of the developmental progression, although this study does not address the relative efficacy of instruction based on that developmental progression.

7.3. Procedural knowledge, conceptual understanding, and transfer

The fourth research question, especially relevant given the emphasis of the interventions on procedures, was: How do the three treatment conditions affect children's use of strategies and their development of procedural and conceptual knowledge, including transfer of learning? We asked specifically whether children who advance (a) increase in procedural knowledge only, or (b) progress from procedural to conceptual knowledge states, or (c) develop intertwined procedural and conceptual knowledge in all phases of learning. The aforementioned progress through each level of the hypothesized developmental progression suggests that children were building more than just procedural knowledge, which was intended by the foundation of both interventions in the construct of spatial structuring. Although what we observed suggested that knowledge and application of procedures was observed before signs of explicit understanding (cf. Rittle-Johnson & Siegler, 1998), we believe that (at least) intuitive learning of concepts was occurring, especially because of the spatial structuring that was embedded in the two experimental interventions, and that the overall results support the position that the development of procedures and concepts is mutually interactive and reinforcing (as theorized in Hierarchical Interactionism, Sarama & Clements, 2009). Because the videos' explicit content consisted of trials that children were asked to complete and demonstrations of procedures (albeit with a clear grounding in spatial structuring in both S and I, an issue to which we return), the question arises of what engendered conceptual learning? Several theories of learning, although they differ, posit an internal, largely unconscious process involving the mental creation of records of experience, a comparison and identification of patterns within these records in an implicit effort to construct more sophisticated processes and conceptions (Sarama & Clements, 2009; Siegler, 1996; Simon, Placa, & Avitur, 2016; Simon, Tzur, Heinz, & Kinzel, 2004). Consider the child in the I group who jumped from CCC to ICS. The child was coded as CCC on the pretest items, using only the procedural tools of tiling and tracing. However, by the end of the first instructional session, this child was able to create a composite unit and physically iterate that unit, an ICS level approach. During the intervening trials, the child's approaches were largely devoid of any focused behaviors until the final trial of the day. We interpret this change as a rapid cognitive reorganization (in our theory, "new experiences can engender rapid change to a new level, Sarama & Clements, 2009, p. 21).

Similarly, other children appeared to engage in mental activity based on their actions on regions, units, and, eventually, composite units. Across several trials, with increasing adoption of demonstrated procedures embodying spatial structuring, they appeared to abstract a regularity, such as the repetition of a tile to form a row, and such rows can be formed with intersecting line segments or each row has the same number of units in it (cf. Simon & Tzur, 2004). This leads to the development of an anticipatory structure in which they foresee that drawing intersecting lines will create the array of squares they had more tediously traced—and more accurately and efficiently. This has been described as psychological curtailment, a tacit encapsulation process in which a mental activity gradually stands in for another physical or mental activity. Curtailment is an abbreviation of a meaningful process, yielding an efficient method that, unlike one merely copied from a demonstration, is conceptually grounded (Clements & Burns, 2000).

The mental system then seeks to re-process, control, and incorporate—to understand—this new process (representational re-description in Karmiloff-Smith, 1986; Sarama & Clements, 2009). A new pattern of thinking and new concepts (i.e., a new LT level) is thereby constructed that can guide and generalize the new conceptual-procedural mental structure. One experience is usually insufficient. Rather, 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 (Sarama & Clements, 2009). Successful application leads to the increasing use of a particular level.

Although concepts were not explicitly verbalized in the three interventions, it is critical to note that the mechanical procedures of the control (C) intervention were ineffective. The clear, repeated, embodiment of spatial structuring in the S and I interventions differentiated them from the C intervention and thus provided the experiential, educative experiences that promoted children's learning.

Thus, both the I and S interventions increased children's ability to determine area measurements correctly in these contexts and raise children's level of thinking and functioning, as delineated by our LT (Table 1), to approximately the same degree. The success of both interventions, despite their differences, supports the notion that spatial structuring is an important component of area measurement proficiency and a worthwhile educational goal (Battista et al., 1998; Outhred & Mitchelmore, 2000; Sarama & Clements, 2009). Also supported is the notion that a combination of animated gestures (Ayres, Marcus, Chan, & Qian, 2009; Cook et al., 2013; Cook & Goldin-Meadow, 2006; Höffler & Leutner, 2007; Hu et al., 2015) and linked verbalizations in a dual-modality presentation technique (Mousavi et al., 1995; Tindall-Ford et al., 1997) is an effective way of developing spatial structuring competencies.

Further, the S intervention was more efficient at facilitating a conceptual understanding of area relative to the I intervention; that is, children in the S group evinced conceptual approaches sooner and more frequently than those in the I group. The S-group children also applied learned strategies in near and far transfer items more often and showed greater signs of conceptual learning in these novel contexts. Thus, the subdivision process, more directly related to spatial structuring, was more effective, providing additional evidence for the importance of this construct.

7.4. Comparisons to the first study and implications

Comparing our results to the first study with older students (Cullen et al., *in press*), we find both similarities and some differences.

The first study found that the I intervention was more effective than the other two interventions at eliciting ICS-level strategies and the S intervention was more effective than the other two interventions at eliciting the postulated, more sophisticated ARCS level. Both interventions, then, supported LT transitions, just as we found with a younger range of children, but the first study found greater increase in correctness and higher-level strategies. In this study, we found approximately the same effect on correctness, but also that the S intervention more effectively promoted conceptual approaches. Thus, both studies agree that the S intervention led to particularly improved learning and that seeing a complete record of the array may have been a key feature supporting that learning. It supported the development of new mental objects and actions, including the conceptual notion of units of units (rows and columns).

Thus, professional development in area measurement may be enhanced by the incorporation of the developmental progression (Table 1) as a tool to promote understanding of children's thinking and learning of area measurement. Although this study did not include an intervention that follows the levels of the LT, instantiations of the specific interventions, either the S or a combination of the S and I interventions, may be useful, especially if enhanced with explicitly conceptual approaches such as class dialogue. Research should evaluate how to best instantiate them in classrooms. For example, young children might play with creating arrays with blocks and teachers should be aware of how to support and discuss such activity to build a foundation for shape composition and spatial structuring and bringing preschool and kindergarten children to the PCC and CCC levels in their construction of arrays with square manipulatives. Then, primary-grade teaching can include specific instructional activities and strategies to support children's movement through the hypothesized developmental progression—this is why LTs include three components: mathematical goal, developmental progression, and instructional activities and strategies correlated with each level (Clements & Sarama, 2014). Making composite units as in this study's I intervention helps children develop the competencies of iteration and composition of units and units. Then, structuring rectangular regions as in the S intervention, intentionally building spatial structuring of two-dimensional space, may help children progress to the goal of thinking at the ARCS level. In these efforts, there may be a place for video demonstrations that correspond to a child's specific work on a trial-by-trial basis. This kind of feedback and strategic modeling may not typically be supplied in direct correspondence to the specific case a student is working to solve, as we did in this study, which may have been important. More work is needed to design and check whether such classroom activities are productive.

References

- Ayres, P., Marcus, N., Chan, C., & Qian, N. (2009). Learning hand manipulative tasks: When instructional animations are superior to equivalent static representations. *Computers in Human Behavior*, 25(2), 348–353. <http://dx.doi.org/10.1016/j.chb.2008.12.013>.
- Baroody, A. J., Lai, M.-L., & Mix, K. S. (2005, December). *Changing views of young children's numerical and arithmetic competencies*. Paper presented at the National Association for the Education of Young Children, Washington, DC.
- Barrett, J. E., Clements, D. H., & Sarama, J. (2017a). 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*, 16 Reston, VA: National Council of Teachers of Mathematics.
- Barrett, J. E., Clements, D. H., Sarama, J., Miller, A. L., Cullen, C., Van Dine, D. W., Klanderma, D., ... (2017b). Chapter 6 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* Reston, VA: National Council of Teachers of Mathematics Journal for Research in Mathematics Education monograph series (Vol. 16, pp. 129–148).
- Battista, M. T., & Clements, D. H. (1996). Students' understanding of three-dimensional rectangular arrays of cubes. *Journal for Research in Mathematics Education*, 27, 258–292. <http://dx.doi.org/10.2307/749365>.
- Battista, M. T., Clements, D. H., Arnoff, J., Battista, K., & Borrow, C. V. A. (1998). Students' spatial structuring of 2D arrays of squares. *Journal for Research in Mathematics Education*, 29, 503–532. <http://dx.doi.org/10.2307/749731>.
- Bokosmaty, S., Sweller, J., & Kalyuga, S. (2014). Learning geometry problem solving by studying worked examples: Effects of learner guidance and expertise. *American Educational Research Journal*, 52(2), 307–333. <http://dx.doi.org/10.3102/0002831214549450>.
- Carpenter, T. P., Coburn, T., Reys, R. E., & Wilson, J. (1975). Notes from National Assessment: Basic concepts of area and volume. *Arithmetic Teacher*, 22, 501–507.
- Cavanagh, M. (2008). Reflections on measurement and geometry: Area measurement in year 7. *Reflections*, 33(1), 55–58.
- Clements, D. H., & Battista, M. T. (2001). Length, perimeter, area, and volume. In L. S. Grinstein, & S. I. Lipsey (Eds.). *Encyclopedia of mathematics education* (pp. 406–410). New York, NY: RoutledgeFalmer.
- Clements, D. H., & Burns, B. A. (2000). Students' development of strategies for turn and angle measure. *Educational Studies in Mathematics*, 41, 31–45.
- Hypothetical learning trajectories [Special issue]. D. H. Clements, & J. Sarama (Eds.). *Mathematical Thinking and Learning*, 6(2).
- Clements, D. H., & Sarama, J. (2007). Early childhood mathematics learning. In F. K. Lester (Vol. Ed.), *Second handbook of research on mathematics teaching and learning: Vol. 1*, (pp. 461–555). New York, NY: Information Age Publishing.
- Clements, D. H., & Sarama, J. (2009). *Learning and teaching early math: The learning trajectories approach*. New York, NY: Routledge.
- 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). New York, NY: Information Age Publishing.
- Clements, D. H., Battista, M. T., & Sarama, J. (2001). Logo and geometry. *Journal for Research in Mathematics Education Monograph Series*, 10. <http://dx.doi.org/10.2307/749924> Reston, VA: National Council of Teachers of Mathematics.
- Clements, D. H., Sarama, J., & Liu, X. (2008). Development of a measure of early mathematics achievement using the Rasch model: The research-based early maths assessment. *Educational Psychology*, 28(4), 457–482. <http://dx.doi.org/10.1080/01443410701777272>.
- Clements, D. H. (1992). Elaboraciones sobre los niveles de pensamiento geométrico [Elaborations on the levels of geometric thinking]. In A. Gutiérrez (Ed.). *Memorias del Tercer Simposio Internacional Sobre Investigación en Educación Matemática* (pp. 16–43). València, Spain: Universitat De València.
- Confrey, J., Maloney, A. P., Nguyen, K. H., & Rupp, A. A. (2014). Equipartitioning: A foundation for rational number reasoning. Elucidation of a learning trajectory. In A. P. Maloney, J. Confrey, & K. H. Nguyen (Eds.). *Learning over time: Learning trajectories in mathematics education* (pp. 61–96). New York, NY: Information Age Publishing.
- Cook, S. W., & Goldin-Meadow, S. (2006). The role of gesture in learning: Do children use their hands to change their minds? *Journal of Cognition and Development*, 7(2), 211–232. http://dx.doi.org/10.1207/s15327647jcd0702_4.
- Cook, S. W., Duffy, R. G., & Fenn, K. M. (2013). Consolidation and transfer of learning after observing hand gesture. *Child Development*, 84(6), 1863–1871. <http://dx.doi.org/10.1111/cdev.12097>.
- Cullen, A. L., Eames, C. L., Cullen, C. J., Barrett, J. E., Sarama, J., Clements, D. H., & Van Dine, D. W. (in press). Effects of three interventions on children's spatial structuring and coordination of area units. *Journal for Research in Mathematics Education*.
- Cuneo, D. (1980). A general strategy for quantity judgments: The height + width rule. *Child Development*, 51, 299–301.
- Greeno, J. G., Riley, M. S., & Gelman, R. (1984). Conceptual competence and children's counting. *Cognitive Psychology*, 16, 94–143.
- Höfler, T. N., & Leutner, D. (2007). Instructional animation versus static pictures: A meta-analysis. *Learning and Instruction*, 17(6), 722–738. <http://dx.doi.org/10.1016/j.learninstruc.2007.09.013>.

- Hiebert, J. C., & Wearne, D. (1993). Instructional tasks, classroom discourse, and students' learning in second-grade classrooms. *American Educational Research Journal*, 30, 393–425.
- Hirstein, J. J. (1981). The second national assessment in mathematics: Area and volume. *Mathematics Teacher*, 74, 704–708.
- Hu, F.-T., Ginns, P., & Bobis, J. (2015). Getting the point: Tracing worked examples enhances learning. *Learning and Instruction*, 35, 85–93. <http://dx.doi.org/10.1016/j.learninstruc.2014.10.002>.
- Karmiloff-Smith, A. (1986). From meta-processes to conscious access: Evidence from children's metalinguistic and repair data. *Cognition*, 23, 95–147.
- Lehrer, R., Jacobson, C., Thoyre, G., Kemeny, V., Strom, D., Horvath, J., ... Koehler, M. (1998a). 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: Erlbaum.
- Lehrer, R., Jenkins, M., & Osana, H. (1998b). 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: Erlbaum.
- 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.
- Lindquist, M. M., & Kouba, V. L. (1989). Geometry. In M. M. Lindquist (Ed.). *Results from the fourth mathematics assessment of the National Assessment of Educational Progress* (pp. 44–54). Reston, VA: National Council of Teaching of Mathematics.
- Mousavi, S. Y., Low, R., & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of Educational Psychology*, 87, 319–334.
- Mullet, E., & Paques, P. (1991). The height + width = area of a rectangle rule in five-year-olds: Effects of stimulus distribution and graduation of the response scale. *Journal of Experimental Child Psychology*, 52(3), 336–343.
- National Assessment of Educational Progress (2013). *NAEP data explorer*. Retrieved from website: <http://nces.ed.gov/nationsreportcard/naepdata/report.aspx>.
- National Governor's Association Center for Best Practices, & Council of Chief State School Officers (2010). *Common core state standards*. Washington, DC: Council of Chief State School Officers and the National Governors Association Center for Best Practices.
- Outhred, L. N., & Mitchelmore, M. C. (1992). Representation of area: A pictorial perspective. In W. Geeslin, & K. Graham (Vol. Eds.), *Proceedings of the sixteenth psychology in mathematics education conference: Vol. 2*, (pp. 194–201). Durham, NH: Program Committee of the sixteenth Psychology in Mathematics Education Conference.
- 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.
- Piaget, J., & Inhelder, B. (1967). *The child's conception of space*. New York, NY: W. W Norton Langdon, F. J. & Lunzer, J. L., Trans.
- Piaget, J., Inhelder, B., & Szeminska, A. (1948/1960). *The child's conception of geometry*. New York, NY: Basic Books.
- Polanin, J. R., Tanner-Smith, E. E., & Hennessy, E. A. (2015). Estimating the difference between published and unpublished effect sizes: A meta-review. *Review of Educational Research*, 86(1), 207–236. <http://dx.doi.org/10.3102/0034654315582067>.
- Rittle-Johnson, B., & Siegler, R. S. (1998). The relation between conceptual and procedural knowledge in learning mathematics: A review. In C. Donlan (Ed.). *The development of mathematical skills* (pp. 75–110). East Sussex, United Kingdom: Psychology Press.
- Rulence-Paques, P., & Mullet, E. (1998). Area judgment from width and height information: The case of the rectangle. *Journal of Experimental Child Psychology*, 69(1), 22–48.
- Sarama, J., & Clements, D. H. (2009). *Early childhood mathematics education research: Learning trajectories for young children*. New York, NY: Routledge.
- Siegler, R. S. (1996). *Emerging minds: The process of change in children's thinking*. New York, NY: Oxford University Press.
- Siegler, R. S. (2006). Microgenetic analyses of learning. In W. Damon & R. M. Lerner (Series Ed.) & D. Kuhn & R. S. Siegler (Vol. Eds.), *Handbook of child psychology: Vol. 2. Cognition, perception, and language* (6th ed., pp. 464–510). Hoboken, NJ: John Wiley & Sons.
- Simon, M. A. (1995). Reconstructing mathematics pedagogy from a constructivist perspective. *Journal for Research in Mathematics Education*, 26(2), 114–145. <http://dx.doi.org/10.2307/749205>.
- Simon, M. A., & Tzur, R. (2004). Explicating the role of mathematical tasks in conceptual learning: An elaboration of the hypothetical learning trajectory. *Mathematical Thinking and Learning*, 6(2), 91–104. http://dx.doi.org/10.1207/s15327833mtl0602_2.
- Simon, M. A., Placa, N., & Avitzur, A. (2016). Participatory and anticipatory stages of mathematical concept learning: Further empirical and theoretical development. *Journal for Research in Mathematics Education*, 47(1), 63–93.
- Simon, M. A., Tzur, R., Heinz, K., & Kinzel, M. (2004). Explicating a mechanism for conceptual learning: Elaborating the construct of reflective abstraction. *Journal for Research in Mathematics Education*, 35, 305–329.
- Smith, J. P., Males, L. M., Dietiker, L. C., Lee, K., & Mosier, A. (2013). Curricular treatment of length measurement in the United States: Do they address known learning challenges? *Cognition and Instruction*, 31(4), 388–433. <http://dx.doi.org/10.1080/07370008.2013.828728>.
- Smith, J. P., 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. <http://dx.doi.org/10.1080/10986065.2016.1219930>.
- Steffe, L. P., & Cobb, P. (1988). *Construction of arithmetical meanings and strategies*. New York, NY: Springer-Verlag.
- Thomassen, A., & Tibosch, H. (1991). A quantitative model of graphic production. In G. E. Stelmach, & J. Requin (Eds.). *Tutorials in motor science* (pp. 269–281). Dordrecht, the Netherlands: Kluwer Academic Publishers.
- Tindall-Ford, S., Chandler, P., & Sweller, J. (1997). When two sensory modes are better than one. *Journal of Experimental Psychology: Applied*, 3(4), 257–287. <http://dx.doi.org/10.1037/1076-898X.3.4.257>.
- Wilson, P. H. (2014). Learning trajectories and professional development. In A. P. Maloney, J. Confrey, & K. H. Nguyen (Eds.). *Learning over time: Learning trajectories in mathematics education* (pp. 227–242). New York, NY: Information Age Publishing.