Minding the Gap: Children's Difficulty Conceptualizing Spatial Intervals as Linear Measurement Units

Article *in* Developmental Psychology · September 2015

DOI: 10.1037/a0039707

Tracy Solomon
SickKids

Susan Cohen Levine
University of Chicago

Developmental Psychology

Minding the Gap: Children's Difficulty Conceptualizing Spatial Intervals as Linear Measurement Units

Tracy L. Solomon, Marina Vasilyeva, Janellen Huttenlocher, and Susan C. Levine

Online First Publication, September 21, 2015. http://dx.doi.org/10.1037/a0039707

Developmental Psychology

0012-1649/15/$12.00 http://dx.doi.org/10.1037/a0039707

Minding the Gap: Children's Difficulty Conceptualizing Spatial Intervals as Linear Measurement Units

Tracy L. Solomon
Hospital for Sick Children (CYMHRU), Toronto, Ontario, Canada

Marina Vasilyeva
Boston College

Janellen Huttenlocher and Susan C. Levine
University of Chicago

Understanding measurement units is critical to mathematics and science learning, but it is a topic that American students find difficult. In 3 studies, we investigated the challenges underlying this difficulty in kindergarten and second grade by comparing performance on different versions of a linear measurement task. Children measured crayons that were either aligned or shifted relative to the left edge of either a continuous ruler or a row of discrete units. The alignment (aligned, shifted) and the measuring tool (ruler, discrete units) were crossed to form 4 types of problems. Study 1 showed good performance in both grades on both types of aligned problems as well as on the shifted problems with discrete units. In contrast, performance was at chance on the shifted ruler problems. Study 2 showed that performance on shifted discrete unit problems declined when numbers were placed on the units, particularly for kindergarteners, suggesting that on the shifted ruler problems, the presence of numbers may have contributed to children's difficulty. However, Study 3 showed that the difficulty on the shifted ruler problems persisted even when the numbers were removed from the ruler. Taken together, these findings suggest that there are multiple challenges to understanding measurement, but that a key challenge is conceptualizing the ruler as a set of countable spatial interval units.

Keywords: measurement, unit, spatial reasoning, mathematics, quantitative development

Measurement is a key aspect of mathematics with widespread application in the science, technology, engineering, and mathematics (STEM) disciplines as well as in everyday life (Lehrer, 2003; Sophian, 2007; Wilson & Rowland, 1993). A critical concept in understanding measurement is the notion of a unit. Units bring precision to quantifying continuous quantities—without units, one would only be able to quantify approximately. The present investigation focused on children's understanding of units in the context of measuring length with a ruler, one of the earliest measurement activities children encounter in school. Learning to use a ruler introduces children to the notion of units of continuous quantity, including the critical ideas of partitioning into equal units and the inverse relationship between the number of units and unit size. These ideas are central not only to understanding measurement but to understanding a variety of mathematical topics, including division, fractions, place value, and later, advanced topics such as integrals. Achieving a strong conceptual grasp of the notion of a unit in the context of ruler measurement may thus have positive, reverberating effects throughout the mathematics curriculum.

Previous research suggests that students find measurement a challenging topic in elementary school and beyond (Clements & Bright, 2003; Mullis, Martin, Gonzalez, & Chrostowski, 2004; National Center for Educational Statistics, 2009). One task that exemplifies this difficulty is determining the length of an object when the object is not aligned with the zero point on a ruler—a *shifted ruler problem*. For example, children find it highly challenging to measure the length of a stick correctly when the left edge of the stick is aligned with the 2-in. mark on the ruler. There is a tendency to make one of two kinds of errors. Children either read off the number on the ruler corresponding to the right edge of the object, suggesting procedural knowledge with little understanding of what the answer represents (Clements, Battista & Sarama, 1998), or they count the number of hash marks on the ruler that correspond to the extent of the object, indicating a lack of understanding of the ruler units (Kamii,

Tracy L. Solomon, Child and Youth Mental Health Research Unit, Department of Psychiatry, Hospital for Sick Children (CYMHRU), Toronto, Ontario, Canada; Marina Vasilyeva, Lynch School of Education, Boston College; Janellen Huttenlocher, Department of Psychology, University of Chicago; Susan C. Levine, Departments of Psychology and Comparative Human Development, University of Chicago.

This research was supported in part by a National Science Foundation (NSF) Science of Learning Center Grant to Susan C. Levine (SBE-1041707, SBE-0541957), the Spatial Intelligence and Learning Center (SILC), and in part by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A120184 to the Hospital for Sick Children. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

Correspondence concerning this article should be addressed to Tracy L. Solomon, Child and Youth Mental Health Research Unit, Department of Psychiatry, Hospital for Sick Children (CYMHRU), 555 University Avenue, Toronto, Ontario, M5G 1X8 Canada. E-mail: tracy.solomon@sickkids.ca

1995, 2006; Lehrer, Jenkins, & Osana, 1998; Lindquist & Kouba, 1989; Martin & Structhens, 2000; Petitto, 1990). The hash-mark response is incorrect because counting hash marks yields a number that is one greater than the number of spatial interval units that correspond with the length of the object. Similar findings have been reported when children are asked to measure objects with a broken ruler, such as a ruler that is missing the first few inches (Nunes, Light, & Mason, 1993; see also Kellman & Massey, 2013). Although it is well established that children find shifted ruler problems challenging, few experimental studies have probed the nature of their difficulty.

Solving shifted ruler problems requires a conceptual understanding of measurement. The crucial concept to grasp is that measurement involves decomposing a continuous quantity into discrete, uniform units and then counting those units. With a ruler, it means dividing the ruler into equal spatial interval units and understanding how the ruler markings—the hash marks and numbers—demarcate those units, and thus can be used to enumerate them. To understand the difficulty children experience acquiring this notion, it is instructive to consider what they might bring to the task of learning to measure and how it might influence their uptake from early measurement instruction.

Previous research points to at least two cognitive biases that might influence children's thinking about measurement. One bias is the tendency to count only bounded, individuated objects (see, e.g., Sophian, 2007), sometimes referred to as "default" units (Brooks, Pogue & Barner, 2011) or "countable entities" (Shipley & Shepperson, 1990). For example, preschool and even young elementary schoolchildren have difficulty regarding contiguous parts or arrays of objects as countable units (Brooks et al., 2011; Shipley & Shepperson, 1990; Sophian, 2007; Sophian & Kailihiwa, 1998). Moreover, when children are asked to enumerate whole objects, they often include separated parts of a single object in their count, such as counting a single fork broken into two pieces as two forks (Brooks et al., 2011). Interestingly, this error decreases when the object parts have a familiar label, such as a wheel that is separated from the rest of a bicycle (Brooks et al., 2011; Giralt & Bloom, 2000), or in other words, when the object parts themselves are more likely to be viewed as individuated objects.

A second relevant bias is the tendency to estimate continuous quantities based on spatial cues, without counting units. For example, Huntley-Fenner (2001) found that preschool children have difficulty applying discrete units to continuous quantities. In one study, preschoolers were presented with pairs of transparent boxes containing glasses filled with sand. One box contained two glasses, whereas the other box contained three glasses, of sand. Presented with the same pair of boxes, children performed well when asked, "Does one box contain more glasses?" but significantly worse when asked, "Does one box contain more sand?" Critically, their performance on the latter question did not differ from performance when the boxes contained piles of sand (without glasses) that differed in amount by the same ratio as in the glasses condition. Huntley-Fenner suggested that although children could reason in terms of the discrete glasses, they perceived the sand as a continuous nonindividuated substance, and thus did not apply the highly salient, discrete units (the glasses) to quantify the amount of sand more precisely.

So how might the early biases to count discrete objects and to use spatial cues to quantify continuous quantities influence children's interpretation of early measurement activities that are commonly used in classroom instruction? Early instruction in measurement typically includes measuring length with nonconventional units, such as using paper clips to measure the length of a book. Although the real goal of this activity is to understand that by counting the paper clips, children are actually enumerating the spatial intervals underlying them, children could interpret the task differently. Specifically, the bias to count discrete objects and yet not to view them as spatial interval units that are helpful in quantifying continuous extent may lead children to view the task merely as an object counting exercise. Indeed, their tendency to leave gaps between units and to overlap units in executing the iteration procedure supports the notion that they do not understand the discrete object/spatial-interval relation (Bragg & Outhred, 2004; Lehrer, 2003). The point is that even when children are successful at measuring length with nonconventional units, that is, when they are able to iterate the unit correctly and to report the correct number of nonconventional units along the object's length, their good performance does not necessarily reflect a conceptual understanding of measurement.

Another common measurement activity is measuring the length of objects with the conventional ruler. Early ruler activities typically involve learning to align the left edge of the ruler and the to-be-measured object and to read off the number on the ruler that corresponds to the object's right edge. If children grasp the discrete-unit/spatial-interval relation from their experience measuring with nonconventional units, then they might be able to apply that knowledge to discern the underlying structure of the ruler and to recognize that the hash marks and the numbers demarcate the spatial interval units. However, if children do not grasp this relation, they may learn the procedure of ruler measurement without achieving any understanding of the ruler structure and markings. When presented with shifted ruler problems, their lack of a conceptual understanding of a unit of continuous amount may become apparent as they rely on surface markings in accordance with their early biases and measurement experiences, and thus arrive at an incorrect answer.

Specifically, the tendency to count discrete objects may lead children to count the hash marks—discrete markings that are more "object-like" than the actual ruler units, which are a series of adjacent spatial intervals that share a common boundary. Alternatively, children may construe the ruler task simply in terms of carrying out the read-off procedure whether or not the object and the ruler are aligned. In other words, they could execute this procedure on aligned ruler problems without any awareness that it works because the number on the ruler aligned with the right edge of the object indicates the number of spatial intervals that correspond with the object's length. It is therefore not surprising that on shifted ruler problems, the two most common errors are using the hash-mark-counting and read-off strategies (Kamii, 1995; Lindquist & Kouba, 1989; Martin & Structhens, 2000; Petitto, 1990).

Present Research

The present research probed the nature of children's difficulty with ruler measurement. In particular, we examined whether young children have difficulty conceptualizing a ruler in terms of as a set of countable, spatial interval units (see also Kellman & Massey, 2013). In three experiments, we compared children's performance on aligned and shifted linear measurement problems when measuring with discrete, adjacent units (coins or circles) with their performance on the identical problems when measuring with a continuous ruler. To our knowledge, no previous work has directly compared children's ability to measure length on aligned and shifted problems with discrete units versus with a continuous ruler. All three experiments involved children in kindergarten and second grade to illuminate children's thinking about measurement after relatively little instruction, and to gain a sense of the development of their thinking after additional measurement instruction and experience. To elucidate the source of children's difficulty with measurement units, we examined their level of performance as well as the nature of their errors on the different types of linear measurement problems.

In Experiment 1, children measured the length of crayons that were either aligned or shifted relative to a row of discrete units or to a ruler. We reasoned that contrasting performance with discrete and continuous measuring tools would allow us to disentangle the challenges associated with the lack of alignment between the object and the measuring tool from those associated with understanding the spatial interval units on a ruler. If the challenge on the shifted ruler problems is that children do not know what, if anything, to make of the lack of alignment between the object and the measuring tool, then they should perform well on the aligned problems and poorly on the shifted problems with both types of measuring tools. However, if the challenge is related to difficulty understanding the spatial interval units on a ruler, a different pattern should obtain. Children should perform well on both types of problems with the discrete units, as they could simply count the coins that correspond to the object length. They should also perform well on the aligned ruler problems in which they could apply their procedural knowledge about rulers and obtain the correct answer, even without a conceptual understanding of the ruler units. But they should perform poorly on the shifted ruler problems, as success would necessitate treating the spatial intervals on the ruler as countable units.

In Experiments 2 and 3, we investigated children's understanding of how the ruler numbers demarcate the spatial interval units. In Experiment 2, we kept the numbers on the ruler and also placed numbers on the discrete units. We predicted that performance on the shifted discrete unit problems would decline because, without a clear understanding of how the numbers demarcate the units, the numbers might trigger the read-off strategy. We expected a greater, negative impact of the numbers on performance with the discrete units in kindergarten compared with second grade because younger children have less well-developed inhibitory control (Davidson, Amso, Anderson, & Diamond, 2006; Williams, Ponesse, Schachar, Logan, & Tannock, 1999). Even so, based on our hypothesis that children's main difficulty is conceptualizing the spatial interval units on the ruler, we predicted that they would continue to find shifted ruler problems more challenging than shifted discrete unit problems.

In Experiment 3, we removed the numbers from the ruler and also from the discrete units, effectively eliminating the possibility of using the read-off strategy. We predicted that if the ruler numbers lured children toward the read-off strategy, effectively obscuring a deeper understanding of the ruler units, then children would show better performance on the shifted ruler problems with the numbers removed, and perhaps perform as well as on the shifted discrete unit items. However, if the challenge for children extends beyond simply inhibiting number information, to difficulty conceptualizing the ruler intervals as countable units, then removing the numbers from the ruler should not lead to improved performance and children should continue to show an advantage on shifted discrete unit problems compared with shifted ruler problems. In this case, children might enumerate the most salient discrete features on the ruler—the hash marks demarcating the ruler intervals.

Experiment 1

Method

Participants. The participants were 16 kindergarteners (mean age = 6.1 years) and 35 second graders (mean age = 8.2 years). The children were recruited from an urban, private school and were predominantly White and middle to upper-middle class; boys and girls were roughly equally represented. Children were randomly assigned to receive either the aligned or the shifted block of test problems first, and one of two forms of the test battery (see Materials). Testing took place in the spring term of the school year.

Materials. Children received 16 test problems and eight filler problems. Each problem was printed on a sheet of paper and placed into a transparent plastic cover. Figure 1 shows examples of the test problems. Each test problem depicted a different color crayon positioned horizontally above a gray-scale measuring tool. For half of the problems, the measuring tool was a ruler, and for the remaining half it was a row of coins. One-inch increments were clearly marked on the ruler and numbered. The coins were 1 in. in diameter and there were no spaces between them. On each problem, the ruler or row of coins was either 7 or 8 in. long, and four response choices were presented below the measuring tool.

For the eight aligned problems, the left edge of the crayon was aligned with the left edge of the ruler (four problems) or row of coins (four problems). For the eight shifted problems, the left edge of the crayon was shifted to the right of the left edge of the ruler (four problems) or row of coins (four problems). The crayon was shifted either by 1 in. or one coin (two ruler problems and two coin problems), 2 in. or two coins (1 ruler problem and 1 coin problem), or 3 in. or three coins (one ruler problem and one coin problem). The crayons were 3, 4, 7, and 8 in. long for the aligned problems, and 3, 4, and 7 in. long for the shifted problems. The problems on the ruler task and the coin task were equivalent, except for the measuring tool. For example, in one aligned ruler problem, the left edge of a 3-in. crayon was flush with the left edge of an 8-in. ruler (Figure 1, top left), and in the corresponding coin problem, the same length crayon was flush with the left edge of a row of eight coins (Figure 1, bottom left). Similarly, for one of the shifted ruler problems, the left edge of a 3-in. crayon was 2 in. to the right of the left edge of the ruler (Figure 1, top right), and in the corresponding coin problem, the same length crayon was two coins to the right of the left edge of the row of coins (Figure 1, bottom right).

The 16 test problems included four types of questions: four aligned ruler problems, four corresponding aligned coins prob-

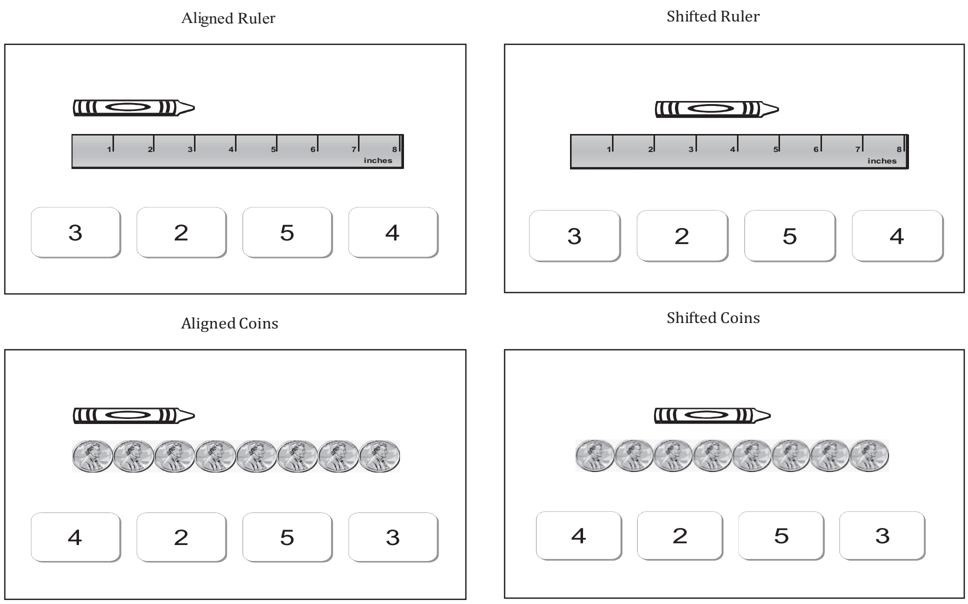

Figure 1. Examples of the four types of problems used in the experiments (in Experiments 2 and 3, the coins were replaced with circles).

lems, four shifted ruler problems, and four corresponding shifted coins problems. The eight filler problems required children to reason about area or length; however, these problems did not involve using a measuring tool (e.g., in one problem, children saw four different-sized sticks and were asked to point to the longest one). The eight aligned problems plus four of the filler problems were assembled in random order to form one block of problems, and the eight shifted problems plus the four remaining filler problems were similarly arranged to form a second block of problems. Children received either the aligned block or the shifted block of problems first.

The four response options were single digit numbers consisting of the correct response and three foils. On the shifted problems, the foils included a read-off response, the number on the ruler that was above the right end of the crayon, and a hash-mark response, the number of hash marks that corresponded to the length of the crayon. When the crayon was shifted by 2 or 3 in., the read-off and hash-mark responses led to different answers (two ruler and two coins problems). In contrast, when the crayon was shifted by 1 in., the read-off and hash-mark responses led to the same answer (two ruler and two coins problems). The "other" responses—the remaining one or two foils for the shifted problems, and all three foils on the aligned problems—were randomly selected from the Numerals 1 to 8. For the aligned problems, all strategies—correct unit counting strategy, read-off strategy, and hash-marks counting strategy—led to the correct answer (there was no hash mark at the zero point on the ruler; see Figure 1). For each of the two test orders (aligned block first or shifted block first), there were two test forms. The test problems in the two forms were identical except for the spatial position of the four response options, which differed across the two forms.

Ten plastic checkers and a Styrofoam cup were used for the counting task. Index cards (3 × 5 in.) bearing the numbers 1 to 10 (2 in. tall) were used for the numeral reading task (see Procedure).

Procedure. Participants were tested individually in a quiet area of their school. The experimenter sat beside the child at a table with the appropriate binder of problems centered in front of the child. Children were told that they were going to look at some pictures and that the experimenter would ask them questions about the size of the things in the pictures. The experimenter then turned to the first problem and asked "How many inches (or coins) long is the crayon?" If children responded inappropriately—for example, if they said "inches"—the experimenter repeated the question beginning with "Listen carefully. . . ." If a response was not one of the four choices, the experimenter encouraged children to select their answer from the four rectangular boxes. In cases for which children spontaneously changed their initial response, only the last answer was used for scoring. For the filler problems, the instructions were tailored to the nature of the problem (e.g., "Which stick is the longest?").

A counting task and a numeral reading task were administered immediately after the measurement problems to rule out the possibility that any difficulties on the measurement problems occurred because of poor counting skills or poor numeral recognition. For the counting task, children were asked to count the 10 plastic disks and to place them into the cup as they counted.

They were instructed to count the disks aloud, one at a time. For the numeral reading task, the index cards were presented individually, in random order, and children were required to report the number depicted. Each child completed each task twice so that we could examine performance consistency. The experimenter recorded children's responses and noted any errors.

Results

For each of the four types of questions (aligned ruler, aligned coins, shifted ruler, and shifted coins), performance on the four test problems was converted to percent correct. Preliminary analyses revealed that performance was not affected by whether or not children received the aligned or shifted block of test problems first. We therefore analyzed the data collapsed across this variable. The least significant difference test was used for all post hoc pairwise comparisons, and one-sample t tests were used for comparisons with chance-level performance. Chance was set at 25%, as children selected the correct response from four response options.

Panel A in Figure 2 shows the results for Experiment 1. The only significant effect in a 4 (Question Type) × 2 (Age Group) analysis of variance (ANOVA), with question type as a within-subjects variable and age group as a between-subjects variable, was a main effect of question type, $F(3, 147) = 88.77$, $p < .0001$ ($\eta_p^2 = .70$). Post hoc analyses showed that performance on the shifted ruler problems was significantly worse than performance on each of the other types of problems, all ps $< .0001$. No other pairwise comparisons were significant. The means for the aligned and shifted coins problems and the aligned ruler problems were significantly above chance for both age groups (all ps $< .0001$), whereas the mean for the shifted ruler problems was not significantly different from chance for either age group.

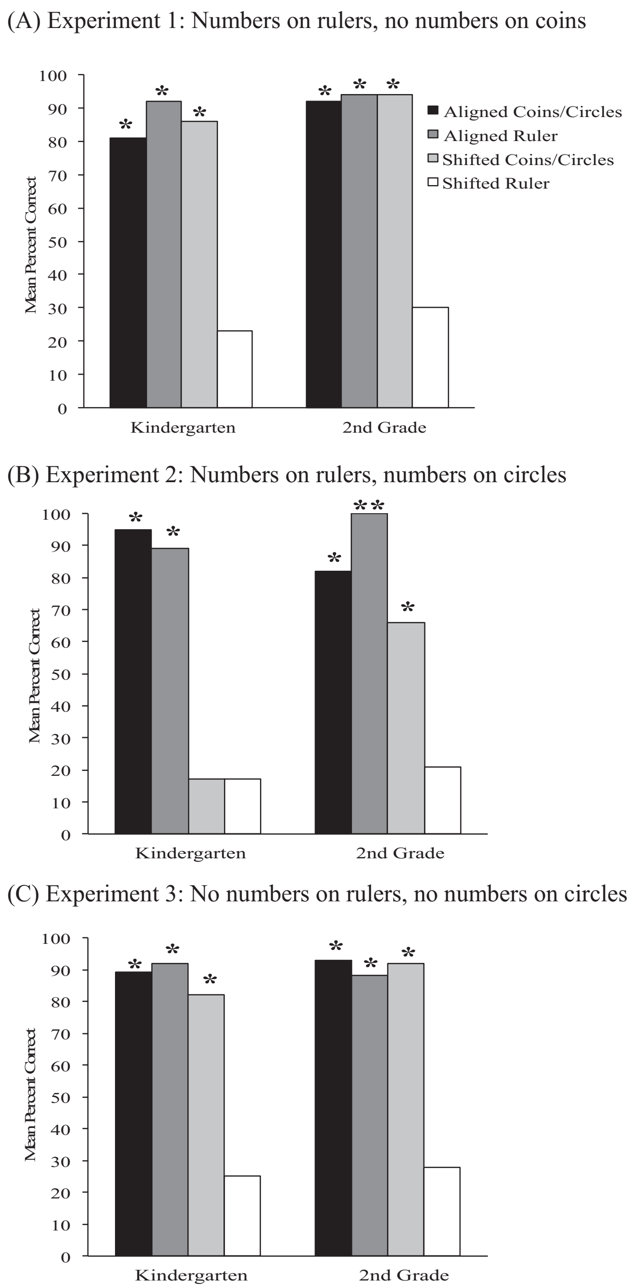

Figure 2. Mean percent of correct responses for the four types of problems in Experiments 1 to 3. * denotes means significantly greater than chance, $p < .05$; ** indicates a mean of 100% correct.

To elucidate the nature of children's difficulty with the shifted ruler problems, we computed the frequency of the different types of responses for one of the shifted ruler problems, in which the read-off and hash-mark strategies led to different answers, and its corresponding shifted coins problem.[1] Panel A in Figure 3 shows the breakdown of the four types of responses ("correct," "read-off," "counting hash marks," and "other") for each grade in Experiment 1. As is clear in the figure, performance was excellent on the shifted coins problem. In contrast, on the corresponding shifted ruler problem, all of the errors in both age groups were either read-off or hash-mark-counting errors. However, whereas the kindergarten children selected the read-off response 56% of the time and the hash-mark response 19% of the time, the second graders selected the read-off response 17% of the time and the hash-mark response 52% of the time. A Fisher's exact test crossing age group with the distribution of error types (read-off vs. hash-mark) revealed a significant effect of age group, $\chi^2(1) = 8.23$, $p < .004$.

Performance on both the counting and numeral reading tasks was errorless on each of the two administrations, indicating that children's errors on the measurement task could not be attributed simply to insufficient counting skills or to a lack of familiarity with Arabic numerals.

Discussion

Experiment 1 revealed a stark difference between performance on the shifted ruler problems and on the three other types of problems in both age groups. Analysis of children's errors suggested that their poor performance on the shifted ruler problems was not the result of guessing. Children tended to rely on two incorrect strategies: reading off the number on the ruler that corresponded to the right edge of the crayon and counting the hash marks on the ruler that corresponded to the length of the crayon. These two strategies accounted for all of the errors on the shifted ruler

[1] For one of the two shifted ruler problems in which the read-off and hash-mark-counting strategies led to different answers, the read-off option was omitted from the answer choices presented to some of the children in Experiment 1 because of an oversight. We therefore did not include the problem (and its matching coins counterpart) in the analyses comparing the mean percentage of correct responses with chance. This had a negligible impact on the results; the means for the shifted coins and ruler problems were 74% and 5% averaged across all four problems, and 71% and 3% based on the three problems remaining after the problems in question were removed. Accordingly, the error analyses presented in Experiment 1 are based on one shifted ruler problem and one matching shifted coins problem. However, as we corrected the problematic items before conducting Experiments 2 and 3, the error analyses for those experiments were based on two shifted ruler problems and their corresponding two shifted circles counterparts.

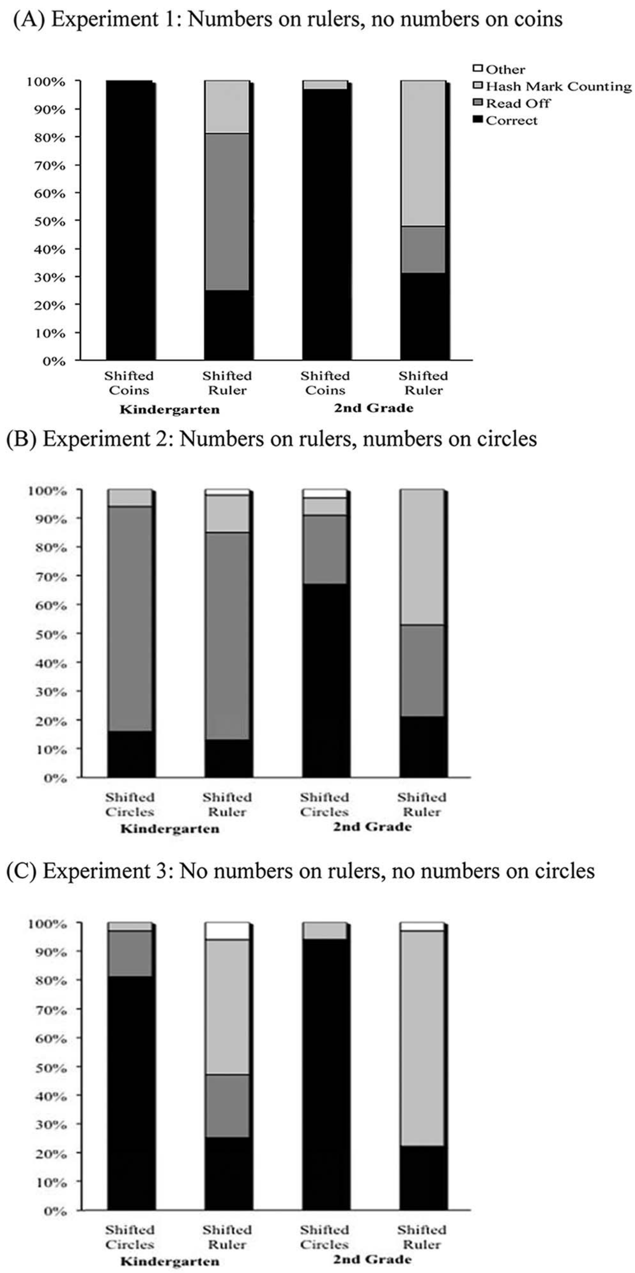

Figure 3. Children's responses on the shifted problems in Experiments 1 to 3.

problems in both age groups. Interestingly, kindergarteners relied more on the read-off strategy and second graders relied more on the hash-mark-counting strategy, suggesting developmental changes in children's (albeit still erroneous) thinking about the shifted ruler task (a point we return to in the General Discussion).

Children's much poorer performance on the shifted ruler problems compared with the shifted coins problems also could not be attributed to the misalignment per se and lends support to the idea that their difficulty may be related to a poor understanding of the ruler units. However, discrete measuring tools such as coins and rulers differ in more than just the nature of the units. On standard rulers, the units are also numbered and this may be distracting, especially for younger children. To explore the possibility that the ruler numbers may have lured children toward the read-off strategy, we manipulated the presence or absence of the numbers on the measuring tools, in two follow-up experiments. In Experiment 2, we added numbers to the discrete unit problems and kept the numbers on the rulers. We predicted that this might lead to an increase in read-off responses, and a corresponding decrease in correct responses, on the shifted discrete unit items. Given that kindergarteners relied on the read-off strategy on the shifted ruler problems more than second graders in Experiment 1, we expected a more pronounced effect of adding numbers to the coins in the younger group. We administered the same ruler items (aligned and shifted) in Experiment 2 that we used in Experiment 1, in order to compare directly performance on the shifted ruler problems and shifted discrete unit problems with numbers, and also to replicate our findings on shifted ruler problems from Experiment 1.

Experiment 2

Method

Participants. The participants in Experiment 2 were 16 new kindergarteners and 17 new second graders (mean ages = 5.9 and 8.3 years, respectively). The children were recruited from the same urban, private school that participated in Experiment 1. There were roughly equal numbers of boys and girls in each age group.

Materials and procedure. The test battery used in Experiment 2 was identical to the test battery in Experiment 1, with two exceptions: The coins were replaced with circles measuring 1 in. in diameter, and the row of circles was numbered from left to right beginning with the number "1." Circles were used instead of coins so that the numerals, which were placed in the center, were the only potentially distracting information on the circles. Thus, in Experiment 2, the units were numbered on both measuring tools (numbers condition). We did not include the counting and number reading tasks from Experiment 1 because of the very high level of performance on both tasks in both age groups.

Results

Panel B in Figure 2 shows the results for Experiment 2. A 4 (Question Type) × 2 (Age Group) ANOVA revealed a main effect of question type, $F(3, 93) = 75.08$, $p < .0001$, $\eta_p^2 = .71$, and a Question Type × Age Group interaction, $F(3, 93) = 9.58$, $p < .0001$, $\eta_p^2 = .24$. For the kindergarteners, there was an overall effect of alignment, with worse performance on the shifted problems than on the aligned problems—for both the circles and the ruler, $p\text{s} < .0001$. There were no significant differences in performance between the aligned circles and aligned ruler problems, or between the shifted circles and shifted ruler problems. For the second graders, performance was worse on the shifted ruler problems than on the aligned ruler problems ($p < .0001$), but there was no difference between the means for shifted and aligned circles problems, as was the case in Experiment 1. Furthermore, in contrast to the younger children, performance on the shifted circles problems was better than on the shifted ruler problems ($p < .05$).

In kindergarten, the means for the aligned ruler and aligned circles problems were significantly above chance ($p\text{s} < .0001$), but the means for both the shifted ruler and shifted circles problems were not significantly different from chance. In contrast, in second grade, the only type of problem on which children's performance was not significantly higher than chance was the shifted ruler problems.

To further test the hypothesis that adding numbers to the circles was deleterious to children's performance, we compared the data from Experiments 1 and 2. A 2 (Experiment) × 2 (Age Group) × 4 (Question Type) ANOVA revealed main effects of experiment, $F(1, 80) = 14.82$, $p < .0001$, $\eta_p^2 = .16$, grade, $F(1, 80) = 8.06$, $p < .01$, $\eta_p^2 = .09$, and question type, $F(3, 240) = 146.06$, $p < .0001$, $\eta_p^2 = .65$. The key pattern of findings was captured by a three-way interaction of Question Type × Experiment × Grade, $F(3, 240) = 6.59$, $p < .001$, $\eta_p^2 = .08$. Follow-up analyses showed that for both age groups, the only significant difference between experiments occurred on the shifted coins/circles problems, with better performance in Experiment 1 when the coins did not have numbers, $p < .02$. However, the magnitude of the difference between the shifted coins/circles conditions in the two experiments was greater in kindergarten (69%) than in second grade (28%).

Panel B in Figure 3 shows the error analysis for Experiment 2. The breakdown of responses is reported for the two shifted ruler problems in which the read-off and hash-mark-counting responses led to different answers as well as for the two corresponding shifted circles problems. As in Experiment 1, virtually all of the children's errors could be accounted for by these two types of errors. For the shifted ruler problems, a chi-square analysis of the distribution of the error responses by age group revealed a significant effect of grade, $\chi^2(2) = 22.02$, $p < .00001$. Kindergarteners selected the read-off response 72% of the time and the hash-mark response 13% of the time, whereas second graders selected the read-off response 32% of the time and the hash-mark response 47% of the time. For the shifted coins problems, the chi-square also revealed a significant effect of grade, $\chi^2(2) = 10.75$, $p < .005$. As Figure 3b shows, kindergarteners committed far more errors than second graders, selecting the read-off response 78% of the time and the answer corresponding to the hash response 6% of the time. In contrast, the corresponding percentages for second graders were 24% and 6%.

Discussion

The results from Experiment 2 for the ruler problems replicate the results for the same problems in Experiment 1. In both experiments, performance in kindergarten and in second grade was very good on the aligned ruler problems and significantly better than on the shifted ruler problems. The results from Experiment 2 also indicate that adding numbers to the circles was deleterious to performance on the shifted circles problems, with a more marked effect in the kindergarteners. Indeed, kindergarteners' performance on the aligned and shifted circles problems in Experiment 2 now paralleled their performance on the analogous aligned and shifted ruler problems in Experiments 1 and 2. Second graders' performance on the shifted ruler problems was significantly worse than on the other problems in both experiments; however, their performance on the shifted circles problems in Experiment 2 was significantly worse than on the corresponding problems in Experiment 1.

The results for the error analysis support our prediction that adding numbers to the circles would lead to an increase in the use of the read-off strategy. Comparing the error patterns on the shifted coin problems in Experiment 1 (Figure 3, Panel A) with the shifted circles problems in Experiment 2 makes this clear (Figure 3, Panel B). Children in both age groups committed more read-off errors in Experiment 2, but the size of the effect was much larger in the younger children. This supports our notion that when the object being measured is not aligned with the zero point on the ruler, the numbers on the measuring tool that correspond to the right edge of the object may pose a distraction, especially to the younger children.

The results for Experiment 2 raise the question of whether or not removing the numbers from the ruler would lead to improved performance on the shifted ruler items. We addressed this possibility in Experiment 3. We predicted a significant reduction in the responses that correspond to using the read-off strategy. The critical question, however, was whether or not the expected decrease in the use of the read-off strategy would be offset by an increase in correct responding or by an increase in hash-mark-counting errors. If removing the numbers from the ruler leads to improved performance, the results would point to a greater understanding of units of linear measurement than is suggested by children's low performance on the shifted ruler problems in Experiments 1 and 2. If, however, the reduction in read-off choices is offset by an increase in hash-mark responding, the results would support the hypothesis that children have difficulty understanding that the spatial intervals on a ruler constitute the units of measurement.

Experiment 3

Method

Participants. The participants in Experiment 3 were 16 new kindergarteners and 18 new second graders (mean ages = 5.7 and 8.3 years, respectively). The older children were recruited from the same urban private school that participated in Experiments 1 and 2. The younger children were recruited from an urban, private school in a different city. However, the two schools served comparable populations—children from predominantly middle- to upper-middle-class families. There were roughly equal numbers of boys and girls in both grades.

Materials and procedure. The stimuli and procedure used in Experiment 3 were the same as those used in Experiment 2, with one exception: There were no numbers on either the rulers or the circles (no-numbers condition).

Results

Panel C in Figure 2 shows the results for Experiment 3, which were highly similar to the results from Experiment 1. A 4 (Question Type) × 2 (Age Group) ANOVA revealed a main effect of question type, $F(3, 96) = 61.10$, $p < .0001$, $\eta_p^2 = 0.7$, and no other significant effects. As in Experiment 1, children in both age groups performed well and significantly better on the aligned circles, aligned ruler, and shifted circles problems compared with the shifted ruler problems, all ps $<.0001$. No other pairwise comparisons were significant. In both groups, the means for the aligned circles, aligned ruler, and shifted circles problems were significantly above chance (all *p*s $<.0001$), whereas the mean for the shifted ruler problems was not significantly different from chance.

To further examine the impact of removing the numbers from the ruler on children's performance, we compared the data from Experiment 3 with the data from Experiment 1. The only significant effect from a 2 (Experiment) $\times$ 4 (Question Type) $\times$ 2 (Age Group) ANOVA was a main effect of question type, $F(3, 243) = 147.10$, $p < .0001$, $\eta_p^2 = .65$, consistent with the patterns that obtained when the data for the two experiments were analyzed separately. Notably, there were no significant main effects or interactions involving the experiment variable.

Panel C in Figure 3 shows the breakdown of responses for the two shifted ruler problems in which the read-off and hash-mark-counting responses were different, and for their shifted circle counterparts. For the shifted ruler problems, a chi-square analysis of the distribution of the error responses by age group revealed a significant effect of age group, $\chi^2(2) = 10.53$, $p < .005$. As is clear in the figure, removing the numbers from the ruler completely eliminated the read-off response in second graders. They selected the hash-mark-counting response 75% of the time, virtually every time they responded incorrectly (they selected the "other" response 3% of the time). Kindergarteners selected the hash-mark response 47% of the time, the choice corresponding to the read-off response 22% of the time, and the "other error" choice 6% of the time. Both age groups made relatively few errors on the shifted circles problems, as was the case in Experiment 1.

Discussion

The critical finding in Experiment 3 was that removing the numbers from the ruler did not lead to improved performance in the shifted ruler condition. In both age groups, performance in Experiment 3 (no numbers on the ruler or the circles) did not differ significantly from performance in Experiment 1 (numbers on the ruler but not on the coins). Children in both age groups performed significantly better on aligned ruler problems and on aligned and shifted circles problems compared with the shifted ruler problems. They also performed, significantly above chance on all types of problems, except the shifted ruler problems.

As expected, removing the numbers from the ruler significantly reduced (and for the second graders, eliminated) the read-off response. However, the decrease in the use of this incorrect strategy was offset by a tendency to rely more on another erroneous strategy—hash-mark counting—rather than by an increase in correct responding. The replacement of read-off errors with hash-mark-counting errors rather than with correct responses suggests that having numbers on the ruler did not simply mask a deeper understanding of the ruler units, but rather that children found the ruler units themselves difficult to conceptualize.

General Discussion

In three experiments, we found that kindergarteners and second graders performed well when measuring an object that was aligned with the zero point on a ruler, but very poorly when measuring an object that was shifted to the right of the zero point on the ruler, replicating a well-established performance pattern (Kamii, 1995, 2006; Lindquist & Kouba, 1989; Martin & Structhens, 2000; Petitto, 1990). The unique contribution of the present work stems from comparing performance on aligned and shifted problems with continuous rulers to the identical problems with discrete units, with and without numbers. The results illuminate the challenges children face in coping with the shifted ruler scenario.

One possibility we considered—that children do not know what to make of the misalignment of the object and the measuring tool—was not supported by our findings. Although performance in both age groups was poor on the shifted ruler problems, performance on the identical problems with discrete units (provided they were not numbered) was very good (see Experiments 1 and 3). As success on the shifted unit problems necessitates counting only those units that extend along the length of the crayon (and not the ones to the left of the crayon), the high level of performance in both age groups on these problems indicates that even kindergartners take the misalignment of the object and the measuring tool into account when measuring the length of the object.

Another possibility, taken up in Experiments 2 and 3, was that children's poor performance on shifted ruler problems reflects difficulty inhibiting misleading numerical information. We found some support for this notion in Experiment 2, when we placed numbers on the discrete units. Both age groups, but especially the kindergarteners, performed poorly on the shifted unit problems. Indeed, kindergartners now did as poorly on these problems as they did on shifted ruler problems. The greater tendency in the younger children to use the read-off strategy on the discrete units may be explained by their less well-developed inhibitory control (e.g., Williams et al., 1999), and greater tendency to rely on surface features, in general, compared with older children (e.g., Gentner & Rattermann, 1991).

A particularly important finding was that although second graders' performance on shifted discrete unit problems also declined when there were numbers on the units, they still performed significantly better on those problems than on shifted ruler problems. Moreover, when we removed the numbers from the ruler in Experiment 3, their performance on the shifted ruler problems without numbers remained lower than their performance on the shifted discrete unit problems without numbers, and, critically, was no higher than their performance on the shifted ruler problems with numbers in Experiments 1 and 2. Taken together, these results, show that while the presence of numbers on the ruler contributes to children's difficulty on shifted ruler problems, it does not fully account for this difficulty.

We therefore posit that children's difficulty on shifted ruler problems reflects a lack of understanding of spatial intervals as countable units. As Kellman and Massey (2013) have argued, the structure of the ruler is not automatically perceived by looking at it, but rather it has to be learned. Lacking a conceptual understanding of ruler units, children either resort to counting the wrong unit—the hash marks on the ruler—or to using a procedure that is successful on aligned ruler problems—reading off the number that corresponds to the object's right edge. In contrast, on shifted discrete unit problems, they are able to count the number of units that correspond with the length of the object perhaps without understanding that what they are actually counting are the spatial intervals that underlie these objects.

It is important to bear in mind that the present work involved samples of middle- to upper-middle-class children and that our findings may thus be limited in their generalizability. Students of lower socioeconomic status (SES) tend to be behind their higher SES peers in math achievement (Jordan & Levine, 2009; National Mathematics Advisory Panel, 2008), and thus might be more inclined to rely on the read-off strategy, and more susceptible to misleading surface information on the shifted problems. However, as all children are likely to gain experience with measurement in the classroom, if not at home, lower SES students' performance may resemble that of their middle and upper SES peers by the second grade. Further research elucidating the contribution of SES to children's developing understanding of measurement could provide important insights regarding measurement instruction with diverse students.

Future Directions and Conclusions

Children's persistent difficulty understanding the nature of linear units raises questions about instructional strategies that might help them recognize spatial intervals on a ruler as countable units. Although the present study did not directly address the role of instruction, our findings are consistent with the possibility that children's errors stem, at least in part, from the shallow instruction they receive.

But what constitutes adequate instructional support? One possibility is to make the intervals on a ruler more salient to children, such as by using different or perhaps alternating colors. Another possibility is to introduce young children to measurement with rulers that do not have numbers, to help avoid the pitfalls of the read-off strategy. If children are presented with measurement activities involving discrete units, it is important to structure these activities in ways that make the mapping between discrete units and spatial intervals more explicit. For example, children can be guided through an activity in which they assemble rulers out of discrete units, number those units in order, and then use the ruler to measure the length of various objects. Another activity may involve measuring with different objects of identical length to emphasize the fact that such discrete units of measurement occupy the same amount of space. For example, children might gain an understanding that units are spatial intervals by learning that measuring with alternating inch-long chips and the inch-long blocks yields the same answer as measuring with either one of these units.

A study by Levine and colleagues (Levine, Kwon, Huttenlocher, Ratliff & Dietz, 2009) tested another approach to make the mapping between discrete units and spatial intervals more explicit. This study compared second grade children's performance on shifted ruler problems after they received one of two different kinds of training. One kind of training involved measuring with aligned rulers and with transparent inch-long chips in separate activities, as is frequently done in early mathematics instruction. The other kind of training involved having children place transparent inch-long chips on top of the ruler units. Performance on shifted ruler problems was better following the second, but not the first, kind of training, suggesting that explicitly linking discrete units with a continuous ruler can facilitate thinking about the spatial intervals on the ruler as measurement units.

The present research comparing children's performance measuring objects with discrete units to measuring objects with a continuous ruler suggests that a key challenge to children's understanding of linear measurement is conceptualizing the spatial intervals on a ruler as the relevant units of measure. Other work suggests that this challenge extends to other aspects of measurement. For example, Vasilyeva, Casey, Dearing, and Ganley (2009) have shown that elementary school students also have difficulty imposing two- and three-dimensional units onto continuous expanses of area and volume. Indeed, the notion of a unit of continuous quantity may be fundamental to understanding many aspects of mathematics. For example, learning about fractions requires understanding that [1/4] of a continuous quantity such as a pie involves dividing the pie into four equivalent pieces, and that the fraction refers to one of those pieces. In later years, learning about integrals in calculus requires comprehending how a continuous quantity can be divided into increasingly finer units. Identifying and addressing the challenges children face in treating the spatial intervals on a ruler as countable units may have implications not only for understanding measurement but also for understanding other areas of mathematics in which the relation between continuous and discrete quantities is a key concept.

References

Bragg, P., & Outhred, L. (2004). A measure of rulers—The importance of units in a measure. In M. J. Hoines & A. B. Fuglestad (Eds.), *Proceedings of the 28th Conference of the International Group for the Psychology of Mathematics Education* (pp. 159–166). Bergen, Norway: Bergen University College.

Brooks, N., Pogue, A., & Barner, D. (2011). Piecing together numerical language: Children's use of default units in early counting and quantification. *Developmental Science, 14,* 44–57. http://dx.doi.org/10.1111/j.1467-7687.2010.00954.x

Clements, D. H., Battista, M. T., & Sarama, J. (1998). Development of geometric and measurement ideas. In R. Lehrer & D. Chazan (Eds.), *Designing learning environments for developing understanding of geometry and space* (pp. 201–225). Mahwah, NJ: Erlbaum.

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

Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia, 44,* 2037–2078. http://dx.doi.org/10.1016/j.neuropsychologia.2006.02.006

Gentner, D., & Rattermann, M. J. (1991). Language and the career of similarity. In S. Gelman & J. Byrnes (Eds.), *Perspectives on language and thought: Interrelations in development* (pp. 225–277). New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511983689.008

Giralt, N., & Bloom, P. (2000). How special are objects? Children's reasoning about objects, parts, and holes. *Psychological Science, 11,* 497–501. http://dx.doi.org/10.1111/1467-9280.00295

Huntley-Fenner, G. (2001). Why count stuff? Young preschoolers do not use number for measurement in continuous dimensions. *Developmental Science, 4,* 456–462. http://dx.doi.org/10.1111/1467-7687.00187

Jordan, N. C., & Levine, S. C. (2009). Socioeconomic variation, number competence, and mathematics learning difficulties in young children. *Developmental Disabilities Research Reviews, 15,* 60–68. http://dx.doi.org/10.1002/ddrr.46

Kamii, C. (1995, October). *Why is the use of a ruler so hard?* Paper presented at the Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education, Columbus, OH.

Kamii, C. (2006). Measurement of length: How can we teach it better? *Teaching Children Mathematics, 13,* 154–158.

Kellman, P. J., & Massey, C. M. (2013). Perceptual learning, cognition, and expertise. *Psychology of Learning and Motivation, 58,* 117–159. http://dx.doi.org/10.1016/B978-0-12-407237-4.00004-9

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

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

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

Levine, S. C., Kwon, M., Huttenlocher, J., Ratliff, K. R., & Dietz, K. (2009). Children's understanding of ruler measurement and units of measure: A training study. In N. A. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 2391–2395). Austin, TX: Cognitive Science Society.

Lindquist, M. M., & Kouba, V. L. (1989). Measurement. In M. M. Lindquist (Ed.), *Results from the fourth mathematics assessment of the National Assessment of Educational Progress* (pp. 35–43). New York, NY: Teachers College Press.

Martin, W. G., & Structhens, M. E. (2000). Geometry and measurement. In E. Silver & P. A. Kenney (Eds.), *Results from the Seventh Mathematics Assessment of the NAEP* (pp. 193–234). Reston, VA: National Council of Teachers of Mathematics.

Mullis, I., Martin, M., Gonzalez, E., & Chrostowski, S. (2004). *TIMSS 2003 international mathematics report: Findings from IEA's trends in international mathematics and science study at the fourth and eighth grades*. Chestnut Hill, MA: Boston College.

National Center for Educational Statistics. (2009). *NAEP Questions*. Retrieved on from http://nces.ed.gov/nationsreportcard/itmrls/startsearch.asp

National Mathematics Advisory Panel. (2008). *Foundations for success: The final report of the National Mathematics Advisory Panel*. Washington, DC: U. S. Department of Education.

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

Petitto, A. L. (1990). Development of numberline and measurement concepts. *Cognition and Instruction, 7,* 55–78. http://dx.doi.org/10.1207/s1532690xci0701_3

Shipley, E. F., & Shepperson, B. (1990). Countable entities: Developmental changes. *Cognition, 34,* 109–136. http://dx.doi.org/10.1016/0010-0277(90)90041-H

Sophian, C. (2007). *The origins of mathematical knowledge in childhood*. New York, NY: Erlbaum.

Sophian, C., & Kailihiwa, C. (1998). Units of counting: Developmental changes. *Cognitive Development, 13,* 561–585. http://dx.doi.org/10.1016/S0885-2014(98)90007-7

Vasilyeva, M., Casey, B., Dearing, E., & Ganley, C. (2009). Measurement skills in low-income elementary school students: Exploring the nature of gender differences. *Cognition and Instruction, 27,* 401–428. http://dx.doi.org/10.1080/07370000903221809

Williams, B. R., Ponesse, J. S., Schachar, R. J., Logan, G. D., & Tannock, R. (1999). Development of inhibitory control across the life span. *Developmental Psychology, 35,* 205–213. http://dx.doi.org/10.1037/0012-1649.35.1.205

Wilson, P. S., & Rowland, R. (1993). Teaching measurement. In P. Wilson (Ed.), *Research Ideas for the Classroom, National Council of Teachers of Mathematics Research Interpretation Project* (pp. 171–194). New York, NY: McMillan.

Received May 6, 2014
Revision received June 1, 2015
Accepted July 27, 2015 ■