Title
Tool Use in Measuring: Second and Fourth Graders' Mediation of Their Linear Estimates Using Rulers and Paper Strips on Number Lines

Author
Kang, Bona

Publication Date
2017

Peer reviewed|Thesis/dissertation

Tool Use in Measuring:
Second and Fourth Graders' Mediation of Their Linear Estimates
Using Rulers and Paper Strips on Number Lines

By

Bona Kang

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Education

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Geoffrey Saxe, Chair
Professor Michelle Wilkerson
Professor Silvia Bunge

Fall 2017

Tool Use in Measuring: Second and Fourth Graders' Mediation of Their Linear Estimates

Using Rulers and Paper Strips on Number Lines

by

Bona Kang

Abstract

Tool Use in Measuring:
Second and Fourth Graders' Mediation of Their Linear Estimates
Using Rulers and Paper Strips on Number Lines

by

Bona Kang

Doctor of Philosophy in Education

University of California, Berkeley

Professor Geoffrey Saxe, Chair

This dissertation investigates the interplay between second and fourth graders' use of rulers and paper strips to solve number line measurement problems. Drawing upon and extending Vygotsky's (1978, 1986) approach to mediation in problem solving, and Saxe's (2012) treatment of microgenesis of representational activity, I analyzed 36 second and 39 fourth graders' measurement activity as they participated in videotaped semi-structured interviews. At each grade, students were randomly assigned to ruler, short paper strip, or long paper strip conditions, and solved five number line problems first without the tool available and then five similar problems with the tool available. For each problem, two numbers were labeled (e.g., 4 and 6) and students were asked to place a third number appropriately (e.g., 9). Problems and tools were designed to elicit unitizing strategies (splitting, iterating, counting) as students coordinated tool with target to produce a linear measure. I report findings related to the *precision* of students' estimates contrasting tool absent and present conditions, and findings related to students' *adaptation of tools to serve measurement functions*. Analyses of students' precision revealed that with or without the support of tools, fourth graders were more precise than second graders, though the utility of tool varied with tool and problem type. The long paper strip was particularly difficult for second graders, and at both grades, many students rejected using the long paper strip. Students generally became precise with tools when a solving a problem type that displayed a linear unit of 1 instead of a larger unit, though this varied with tool type. On other problem types, tools interfered with precision, and on still other problem types, tools interfered with the precision of second, but supported the precision of fourth graders. Analyses of strategies revealed three levels in students' ability to coordinate a unitization of tool and target as they tried to adapt tools to serve measurement functions. I argue that these findings can productively inform our understanding of the interplay between tool use in students' developing measurement activity as well as the design of instructional problem environments to support students' tool using in linear measurement.

Table of Contents

Acknowledgements

This dissertation was completed with the help and support of several individuals.

I am grateful to my dissertation committee—Geoff Saxe, Michelle Wilkerson, and Silvia Bunge—for the feedback and support they provided to advance the conceptualization, design, analysis, and writing of the study.

I would like to thank Geoff, my advisor, who invested so much into my ideas throughout my years in graduate school. He was incredibly generous with providing feedback on my written work and ideas at all hours of the day and night and enriched my incipient understandings of theory, methods, research design, and writing. Throughout my time in graduate school, Geoff exposed me to productive models for collaboration, co-writing, research study design, apprenticeship, and how to think deeply about my and others' research. These are invaluable lessons that I will carry forward.

I would like to thank Michelle and Silvia for their perspectives on my dissertation. Michelle's advice has brought new insights to my work and helped me think more critically about the directions to which I can extend my work. Silvia's outside perspective has been invaluable in influencing and improving the design components of my research.

I would not have survived the first half of my time in graduate school without the endless support of Maryl Gearhart. Maryl was an exceptional role model for mentoring emergent scholars and effectively managing a robust research project. I am thankful for her faith in my potential since my undergraduate years. She has provided invaluable intellectual feedback and emotional support that was key to my early development in graduate studies.

I would like to send many thanks to Aki Murata, who was not just an amazing colleague and mentor, but also a dear friend who helped me through my toughest years in graduate school. She helped me realize what kind of academic identity I could create for myself, connected me with teachers and classrooms in which I piloted my incipient dissertation ideas, drank copious amounts of coffee and wrote with me in cafes for many hours, and provided endless encouragement throughout my journey.

I am truly grateful to my friends Jing Li, Eddie Europa, and Jennifer Lum. The physical writing and completion of this dissertation would have been impossible without the regular check-ins and writing sessions with them during the last two years. Thank you for pushing me, believing in me, making me laugh, writing with me, eating with me, and helping me feel like a real human being throughout this process.

I have also benefitted from all former and current members of the LMR research group and Saxe advisee group who have provided invaluable intellectual support to my work: In particular, Anna Casey, who somehow even after moving to Seattle channeled all of her positive energy to Berkeley so that I could "FSITG"; Kathryn Lanouette and Anna Zarkh, who truly helped me believe in the value of my own ideas and contributed greatly to the foundational design of the tasks in my dissertation; Heather Fink and Nickie Leveille Buchanan for being instrumental in helping me secure data collection sites after I continuously failed to do so on my own for 2 years; Marie Le for pushing my thinking, helping me vent when I needed to, and being a genuine friend both in and out of school during such an isolated experience; Darrell Earnest for repeatedly checking in with me and providing moral support and advice even when I was terrible at responding to his kind messages in a timely manner.

I am also thankful for the Summer and Missy writing groups. Hee-jeong Kim, Elissa Sato, Dan Reinholz, Chooza Moon, Serena Rhee, and many others showed me productive working models of peer support and writing feedback, models that I aim to continue beyond Berkeley. Thank you for helping me realize that successful support groups are possible for both sustaining productivity and fostering a sense of community.

I would also like to thank other individuals from the GSE and broader Berkeley community. My undergraduate research assistants Jazmin Marie Garcia and Monica Edith Silva provided incredible help not only in analyzing my data, but also contributed intellectually to the design of my coding schemes and inspired me to improve as a research mentor. Min-jeong Jung never hesitated to listen to my worries, and gave endless advice on overcoming graduate school life challenges and looking ahead. Dr. Frank Worrell and Lisa Kala provided the resources and opportunities for piloting my dissertation ideas; without Frank and Lisa, the start of this dissertation would not have been possible. Additionally, special thanks to the staff, especially Caron Williams, Kate Capps, Ilka Williams, and Rosa Garcia, who answered all of my questions and helped me navigate the intricacies of the Berkeley bureaucracy. I also benefitted immensely from the Research in Cognition and Mathematics Education (RCME) community. Additional thanks to Emilie Dandan, who has been my longtime outside-department moral support in Tolman.

I am indebted to Paul Li and my undergraduate students, who helped me (re)discover my love for teaching and mentoring, which in turn, helped propel my academic research. They pushed me to move beyond the status quo and inspired me with their curiosity and stories. I am especially thankful to Paul; he has been a true cheerleader of my capacity as a colleague, instructor, researcher, and mentor, and this helped me push through the last year.

Lency Olsen welcomed me into her elementary school classrooms for several years. I have gained many insights working with her students that informed the design of my study, and I have very much appreciated her interest and efforts in engaging her students in the world of mathematics.

I am grateful to my long-time friends Viv Choi and Maelin Wu, who believed and reminded me endlessly that I would finish, and supported me through the ups and downs of graduate school. I am thankful to Dex Liu who has repeatedly played a crucial role in ensuring that I could stay in Berkeley and finish my degree.

I would like to thank my parents for believing in me despite the many years and immensely huge Pacific Ocean that separated us, and my younger sister Sophia for keeping me grounded and sometimes having to reverse roles because her older sister still lived in a bubble protected from the larger society and workplace.

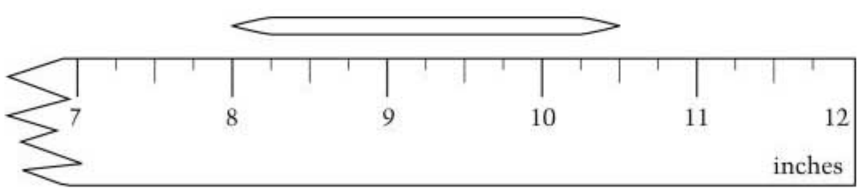
My deepest gratitude goes to the teachers, parents, and students who contributed to this study. The teachers showed interest in my research and opened their classroom doors for me, the parents provided me an opportunity to interview their children, and the students agreed to be interviewed by a complete stranger (me). This dissertation is dedicated to the students: Thank you for openly sharing with me so many insights to your mathematical thinking, for genuinely enjoying my number line racecourses and asking if you can play again, and also for telling me fun facts about cats, "pizza burritos," and why zero is a special number. You are the true motivation for why I do what I do.

Chapter 1: Introduction

This dissertation investigated the *interplay* between children's emerging understandings of linear measurement and their uptake of different types of tools to support their problem solving activity in linear measurement.

Measuring length is ubiquitous in everyday activity and serves as a mathematically important topic in the classroom in of itself as well as groundwork for developing ideas in other mathematical topics, such as number and operations, geometry, functions, approximation, and proportional reasoning (Lehrer & Schauble, 2000). The Common Core State Standards for Mathematical Content (National Governors Association Center for Best Practices & Council of Chief State School Officers [NGA Center & CCSSO], 2010) support the learning of linear measurement ideas in grade 2, and support reasoning about fractions and operations on linear representations by grade 3.

Despite its significance, linear measurement is still a topic that elementary school children find difficult to learn as evidenced by their national assessment records (Drake 2014; Kloosterman, Rutledge & Kenney, 2009; US Department of Education, 2003, 2006, 2012). Figure 1 illustrates the kind of difficulties that many students display with a task used in the 2003 National Assessment of Educational Progress (NAEP). On this particular task, 80% of fourth grade students and 42% of eighth grade students provided incorrect answers, typical solutions including 10.5 inches (resulting from focusing only on the right end of the toothpick) or 3.5 inches (resulting from starting a count of 1 at the 8-inch mark) (US Department of Education, 2003). Thus, educators and researchers are concerned about how we can support children in not only learning the skills to measure, but also what it means to measure (National Council of Teachers of Mathematics [NCTM], 1989, 2000).



What is the length of the toothpick in the figure above?

Figure 1. Sample linear measurement problem from 2003 NAEP (Correct answer: 2.5 inches)

Linear measurement tools play a central role in supporting or constraining linear measurement understandings (Barrett & Clements, 1996; Boulton-Lewis, Wilss, & Mutch, 1996; Nunes, Light, & Mason, 1993; Saxe, Earnest, Sitabkhan, Haldar, Lewis, & Zheng, 2010; Stephan & Clements, 2003), and we need to better understand how children adapt tools to support their measurement activity. For example, young children may prefer to use rulers to non-standard tools such as sticks or body parts to measure object lengths, but actually measure more correctly with the non-standard tool before becoming more precise with the ruler (Kotsopoulos, Makosz, Zambrzycka, & McCarthy, 2015). However, because rulers are more interesting to children, they may engage children in more meaningful measurement activities and be less demanding in teaching children how to measure length compared to tools with arbitrary unit sizes (Boulton-Lewis et al., 1996; Nunes et al., 1993).

Empirical research on how children shift in how they make sense of and use different types of tools in linear measurement activity has typically focused on children's use of tools to produce a single linear measure or to indirectly compare lengths of objects as the final problem solving goal (Clements, 1999; Clements & Stephan, 2004; Kamii & Clark, 1997). These types of activities involve students directly aligning the tool against an object and accumulating a count of a particular unit (e.g., 10 blocks or 15 inches) to measure the length or using the length of the tool to indirectly compare the lengths of two different objects. However, tools are rarely co-quantified with the target object to generate multiple linear measures; few systematic empirical studies have focused on how children transform a tool and coordinate a linear measure produced by the transformed tool as a means for solving a superordinate goal in problem solving activity. In a typical length measurement task, the object's length is unknown, and children do not need to transform their tool multiple times in order to produce two different measures on the same object. The whole tool either serves as the unit size (e.g., the width of a wooden block) to be tiled or iterated to the length of an object of interest, or the tool is already unitized to help determine the length of an object (e.g., measuring the full length of a toothpick with a ruler like in Figure 1). Tools are rarely used multiple times in the same measuring act and the object does not need to be unitized more than once. A more in-depth investigation of children's developing tool use understandings can be accomplished by illuminating how children are able to quantify and coordinate their adaptation of different tools generate multiple linear measures in the same problem space.

Tool using on number line problems may be productive contexts for understanding children's developing linear unit understandings. Children's estimates on number lines have been used as resources for exploring children's numerical cognition (Booth & Siegler, 2006, 2008; Dehaene, Izard, Pica, & Spelke, 2008; Ebersbach, Luwel, Frick, Onghena, & Verschaffel, 2008; Moeller, Pixner, Kaufmann, & Nuerk, 2009; Reinert, Huber, Nuerk, & Moeller, 2015; Siegler & Booth, 2004; Siegler & Opfer, 2003), examining how children coordinate numerical values with linear units (Saxe et al., 2010, 2013), modeling children's informal strategies for reasoning about relationships between numbers (e.g., arithmetic problems) (Gravemeijer & Stephan, 2002; Fosnot & Dolk, 2001; Murata, 2008; Murata & Kattubadi, 2012; NCTM 2000, 2003), and revealing children's measurement skills (Barth & Paladino, 2011; Cohen & Sarnecka, 2014; Petitto, 1990), but the effect of using measurement tools on placing linear estimates on number lines has rarely been the focus (see Saxe et al., 2013 and Petitto, 1990 for examples of measurement tool use on number lines). Children's interpretations and constructions of number lines display how experienced they are in coordinating numerical values and linear units, and the goal of estimating number locations on a line mediate their thinking process (Saxe et al., 2010, 2013). Some children will place numbers accurately on the line, with correct orders and correct placements that coordinate the lengths of intervals with the values of tick marks. Others may display partially developed understandings, such as placing numbers in order from left to right, but not preserving the same interval length to represent the same interval value on the line. The use of measurement tools on number lines may provide resources for children to reason about the relationship between linear unit and numerical value; children's interpretations about the relationship between number and linear units can be overtly displayed in their placements of number on number lines and children's mediation of their linear measurement activity may interact the tools they use to organize their linear estimation activity. Adequate linear estimates will involve co-quantifying the tool with values on the number line. If students use the tools in an unexpected manner, such as not coordinating a consistent unit length between the tool and number line or leaving gaps for each iteration when placing numbers, this may help reveal how they reason about linear distances in relation to numerical values. Using number lines as a measurement context would allow for examining the interplay in which children's linear unit understandings can advance through the construction of measurement tools in problem solving, and their tool use is guided by the progress of their linear unit understandings.

Children's Unitizing in Measurement Problem Contexts

To better identify the shifting conceptual challenges in linear measurement for children, we need a way to examine the relationship between what kinds of linear unit understandings children bring to bear in a measurement problem, and the process of how children may mediate their unitization of target objects to produce linear measures. Consider the number line problem shown in Figure 2 to examine what kinds of linear unit understandings may be coordinated to solve a linear measurement problem without tools. On this problem, a line segment has the numbers 1/3 and 1 pre-labeled but is otherwise blank, and the goal is to mark the location for the number 2 on the line. For a successful solution, this problem requires an understanding of how linear units and numbers are represented on a linear model such as the number line, and how to transform and operate on linear units (e.g., subdividing units or iterating units) to generate unlabeled numbers on the line. The linear distance between 1/3 and 1 on the number line represents a linear unit of 2/3 (Figure 3a), which can be subdivided into shorter linear units and iterated across the line. A possible solution for this problem then involves two steps: (1) Subdivide the given length between 1/3 and 1 into half to generate shorter lengths that each represent a linear unit of 1/3 (Figure 3b), and (2) iterate a linear unit of 1/3 three times end-to-end from 1 to place 2 on the line (Figure 3c).

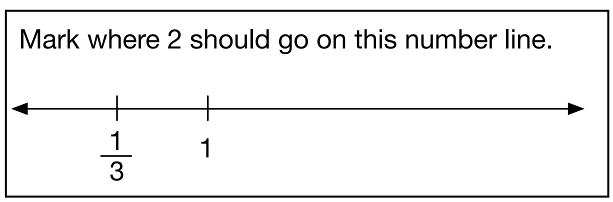


Figure 2. Number line problem with two numbers, 1/3 and 1, marked with a target value of 2

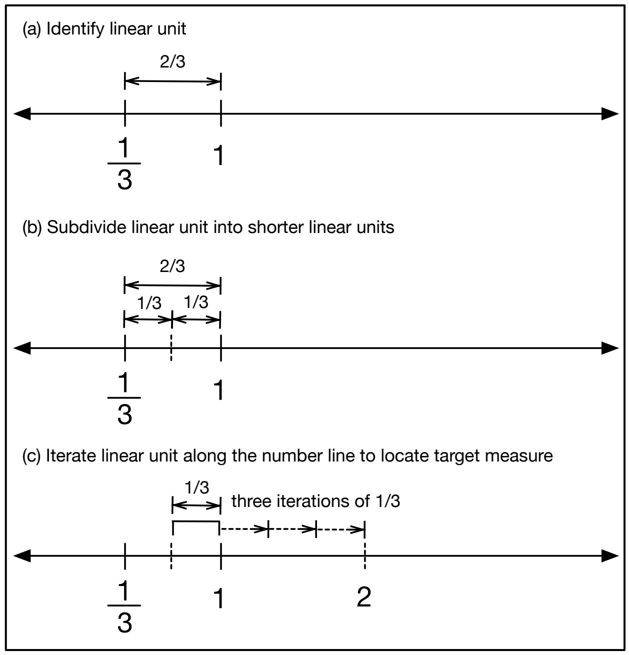


Figure 3. Subdividing and iterating linear units to produce an adequate position for 2 on the number line

We may not be able to illuminate what kinds of conceptual supports children may need or use in linear measurement activity if we only focus on the precision of children's final estimates. Based upon pilot data and children's documented developing unit understandings, I expect that children will encounter various kinds of unitization difficulties in producing an adequate solution to problems such as the one presented in Figure 2. Younger children may simply understand 2 to follow 1 and just place 2 towards the right of the line without any regard for linear unit relationships. Older children, on the other hand, may appreciate that the linear distance between 1/3 and 1 is key to solving the problem, but may be unclear on how to make use of that distance to position 2 adequately, still resulting in an imprecise measure. Only focusing on the accuracy of the final estimate of a linear measurement problem, such as for 2 on the number line problem above, would not readily reveal what kinds of unitization actions students coordinate or may have the potential to generate in linear measurement activity.

Children's Construction of Tools in Measurement Problem Contexts

Analyzing children's measurement tool constructions to support their problem solving may help illuminate what kinds of conceptual coordinations children produce in linear measurement activity. Children's conceptual interpretations of the problems at hand may influence their ability to adapt tools to support their solutions as well as the character of their tool using approaches. I expect that younger children who have yet to appreciate the meaning of linear units may find tools unnecessary or an added challenge to constructing a solution. Older children may attempt to use tools in different ways to help them measure on the line, though their success may be affected by the material affordances of the tool and how they try to perform operations with linear units. For example, imagine children are presented with the problem in Figure 1 and are asked to use a ruler or a paper strip to help them place a pointer for the location of 2 on the line segment. For younger children, it may not be obvious to create a measure between 1/3 and 1 using either tool, and there may be no differences in their problem solutions with the presence of tools. They may even find the presence of tools distracting.

For older children, however, each tool may present different challenges and affordances. An older child might consider that the distance between 1/3 and 1 is important. If the paper strip is a rectangular piece of paper that fits between 1/3 and 1, it may seem more obvious to fit the paper strip (rather than the ruler) inside the interval for this reason. However, without a metric understanding of linear units, the child may find it much less obvious to then use the paper strip to represent and preserve the established linear unit of 2/3 on the line, or to fold or cut the paper strip to represent shorter linear units on the line, or to iterate the paper strip without gaps along the line to mark 2 on the number line; the child may fail to coordinate the length of the paper strip with linear units on the number line or may iterate the piece with gaps or overlaps, resulting in a non-metric location for 2 (Figure 4).

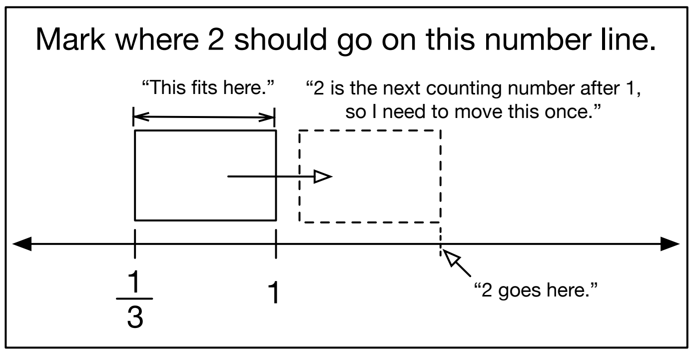


Figure 4. Example of how a child may use a paper strip and place 2 imprecisely on the number line

In contrast, I expect that the ruler will present different challenges than the paper strip. The ruler already has numeric units marked throughout, allowing for better precision and preservation of unit lengths. However, the child now needs to coordinate between defined numerical units on the ruler and on the number line. The child not only has to consider how to measure the length of the linear unit between 1 and 1/3 as being a certain amount of inches using the ruler, but the child also needs to consider how three more 1/3s beyond 1 represents the location for 2, and how those shorter linear units should be represented on the ruler in inches (Figure 5).

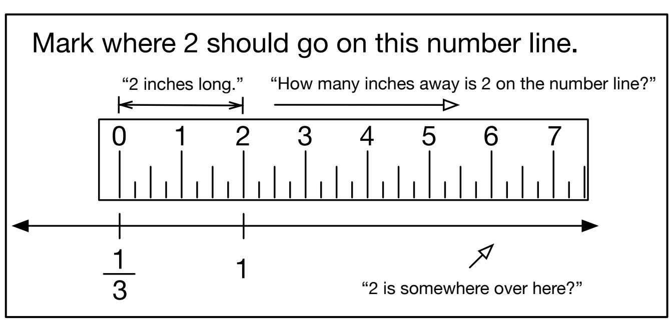


Figure 5. Example of how a child may use a ruler to place 2 on the number line

In this dissertation, I describe a developmental investigation of the interplay between children's linear unit understandings and measurement tool constructions in linear measurement problem solving activity; I report findings on second and fourth grade children's understandings of linear units in the domain of whole numbers, and children's appropriation of different types of artifacts (paper strips, ruler) as tools to solve measurement problems in the same numerical domain. I argue that in solving length estimation problems, (1) children who are early in developing linear unit understandings have difficulty treating artifacts as tools, and their estimates of length are more accurate when they rely on intuitions and not tools; (2) children who are later in developing linear unit understandings show growing efforts to quantify artifacts (paper strips, rulers) as they quantify the target of their measurement activity, but their performance is uneven—different artifacts create specialized challenges in producing estimates of length; and (3) the most advanced children display sophisticated strategies in generating tools with the artifacts, creating adaptations of different kinds of materials to solve a range of linear measurement problems, and their precision is similar with or without the tools.

To support my arguments, I report analyses of semi-structured interviews with second and fourth graders who solved number line estimation tasks using only perceptual strategies (no physical aids) and then with one of three types of tools (ruler, short paper strip and long paper strip). The analyses focused on students' shifts in precision as well as the differences in students' increasing abilities to coordinate their tool units with the number line units by grade, tool type, and problem type.

Review of the Literature

Measurement is an everyday act and mathematical activity that deepens children's knowledge of space (Lehrer, 2003; Piaget, Inhelder, & Szeminska, 1960) and knowledge of number (Saxe et al., 2010). Mathematics educators and researchers have repeatedly emphasized the importance of children's need to not just learn the skills of measuring, but to develop an understanding of fundamental measurement concepts to be mathematically proficient (National Council of Teachers of Mathematics [NCTM], 1989, 2000, 2003; National Governors Association Center for Best Practices & Council of Chief State School Officers [NGA Center & CCSSO], 2010). In other words, children should learn what it *means* to measure in addition to focusing on the *procedures* of how to measure. In order to support children's learning of linear measurement ideas, there is a need to better understand the conceptual coordinations children encounter when they draw in measurement tools to support multiple linear measures in a measurement problem context.

In linear measurement problem solving, children's conceptualizations of linear unit understandings emerge from their coordination of two types of unitizing activities: unitizing objects' lengths and unitizing linear measurement tools (Clements, Battista, & Sarama, 1998; Nunes et al., 1993; Saxe and Moylan, 1982; Stephan & Clements, 2003). These are two processes that must be coordinated by putting linear units represented on the target object (e.g., number line) to correspondence with linear units represented on the tool (e.g., paper strip or ruler) to make a meaningful linear measure. Children who are transitioning in their linear measurement understandings face several challenges in using tools to produce a linear measure of a target object. They may experience difficulty in quantifying the target or tool into units and in constructing a correspondence between the target and tool such that the tool represents a length of measure. Different kinds of tools, however, can provide differential support in mediating linear measurement activity and understandings (Barrett & Clements, 1996; Boulton-Lewis et al., 1996; Clements et al., 1998; McClain, Cobb, Gravemeijer, & Estes, 1999; Nunes et al., 1993; Saxe et al., 2010; Solomon, Vasilyeva, Huttenlocher, & Levine, 2015; Stephan & Clements, 2003). Varying the target estimates (e.g., varying the unit length or positions of values to be placed on the number line) can also help reveal what kinds of linear unit understandings children use to support their unitization activity and the challenges they may face in coordinating the target units with their tool (Saxe et al., 2010, 2013).

In this section, I review literature that considers the ways in which this coordination presents challenges to children as they make an effort to generate linear measures using two types of tools: Non-standard tools and standard tools. Next, I present the types of unitizing actions that need to be coordinated to produce linear measures. Then I consider a productive problem context that may illuminate the interplay between children's shifts in tool use and their linear unit understandings.

Tool types in linear measurement activity.

There are principally two different types of tools that individuals can specialize to serve length-measuring functions. On the one hand, *non-standard tools* make use of unconventional units such as a blank strip of paper or a wooden block. These types of tools may be more familiar to children who are learning to measure, but they are typically not quantified into units, especially with units that are widely used, such as inches or centimeters. On the other hand, *standard tools* make use of conventional units for which there are international as well as national standards, such as rulers pre-labeled with centimeter or inch units. In schooling contexts, children typically learn how to use standard tools, which are already quantified with equally spaced units on the tools for the purpose of calibrating the precision of measures.

Non-standard and standard tools each present different kinds of challenges to children as they draw the tools into linear measurement activity, and there is no general consensus on which type of tool should be introduced to students first in school. Some researchers assert that children should be introduced to measurement by reasoning about lengths of everyday objects (e.g., width of classroom) using non-standard tools (e.g., a child's foot) before being introduced to formal systems of standardized measure (Haylock & Cockburn, 1989; Kastner, 1989; McDonough & Sullivan, 2011), and this view is still reflected in many classroom curricula (Clements, 1999; Kamii & Clark, 1997; Smith III, Males, Dietiker, Lee, & Mosier, 2013). Non-standard tools provide opportunities for students to discuss important measurement ideas, such as the need for equal-size units and not leaving gaps between displacements of a non-standard tool along an object's length (McClain, Cobb, Gravemeijer, & Estes, 1999). However, others assert that young children do not necessarily find standardized tools more difficult to use than non-standard tools (Boulton-Lewis et al., 1996; Clements, 1999; Nunes et al., 1993). Standard tools allow for individuals to publically communicate about target measures, which may help facilitate a more precise solution for the measurement activity (e.g., "it starts from the line above number 1 and ends at the line above number 3 on the ruler") and still provide opportunities to learn ideas about linear units and how to measure them. Students as young as kindergarten-age also prefer to use standard tools like rulers over non-standard tools (Kotsopoulos et al., 2015), and this preference and interest in standard tools may be more effective in advancing their linear unit understandings (Boulton-Lewis et al., 1996; Kamii & Clark, 1997; Kotsopoulos et al., 2015).

In the subsequent sections, I describe specific challenges to unitization that children may encounter using standard tools and non-standard tools in linear measurement activity.

Standard tools.

Children find standard tools that are pre-quantified with units such as inches or centimeters challenging to use in measurement activity. Learning how to use pre-labeled units such as standard units is an important part of the elementary school curriculum as children are expected to understand and use standardized linear measures as early as 2nd grade (NGA Center & CCSSO, 2010). However, both young and old children display persistent difficulty in understanding how to conceptualize standard tools like rulers as a collection of linear units, and how to coordinate features on rulers with target objects to produce a linear measure.

A common problem that is used to assess children's understanding of linear measurement is the broken ruler task, such as the 2003 NAEP task presented earlier in Figure 1, in which an object is depicted above a representation of a ruler such that it is not aligned at the zero-point on the ruler and the goal is to figure out the length of the object. Two common incorrect responses for this task were 10.5 inches (14% of fourth graders and 7% of eighth graders) and 3.5 inches (23% of fourth graders and 20% of eighth graders), and these reflect particular conceptual difficulties children face in trying to unitize rulers and coordinate those actions on object lengths.

Two key challenges in ruler use in linear measurement activity are highlighted in both the students' NAEP responses and the literature. First, one type of challenge is relating the *whole* length of a target to the corresponding part of the ruler (Kamii, 2006). This may result in focusing on only one end point of the target object and reading the number on the ruler that corresponds to the edge of the target (Clements et al., 1998; Drake, 2014; Kamii, 2006; Nunes et al., 1993). Students may have measured the toothpick in Figure 1 as 10.5 inches long because the right edge falls above the 10.5-inch mark on the ruler. This reveals a (1) lack of unitizing the ruler as a collection of units, and (2) lack of unitizing the target object (the toothpick) as a collection of units, leading to (3) the tool (ruler) and target (toothpick) not being put to correspondence to produce a numerical length that represents the span of the object.

Second, another challenge is conceptualizing a ruler as a set of countable continuous *spatial* units. This may result in starting a count of measure from 1 instead of 0 or counting the inch tick marks on the ruler between the two end points of the target object, both which result in an answer that is one inch longer than the actual answer (Drake, 2014; Kamii, 1995, 2006; Lehrer, 2003; Lehrer, Jenkins, & Osana, 1998; Lindquist & Kouba, 1989; Martin & Structhens, 2000; Petitto, 1990). This may explain the erroneous response for 3.5 inches in the 2003 NAEP problem presented earlier. Bragg and Outhred (2004) also report this as a key ruler challenge for children. When they asked fifth and sixth grade children what features on the ruler are counted when measuring length, many children responded that the tick marks on the ruler rather than the interval lengths between tick marks should be counted. This type of response reveals a lack of coordinating what the units of the ruler represent, which are continuous lengths, and not simply discrete counts of numbers or tick marks. More recently, Solomon et al. (2015) found that younger (kindergarten and second grade) students also persisted in having difficulty coordinating the length between interval marks on rulers to measure lengths of crayons for this reason. When crayons were off-set from the zero-point on rulers, both age groups heavily relied on tick mark counting strategies to report a measure for crayon lengths. Even when rulers with no numbers were used in this way, the second grade children used tick mark counting strategies 75% of the time they produced erroneous measures. Thus, we might anticipate that children in the elementary grades will display challenges in using features on rulers to record linear measures of targets, particularly when the target's edge is not already aligned at the end of the ruler where 0 is marked.

Non-standard tools.

Non-standard tools are not pre-quantified, and can provide flexibility in children's unitizing activities because they do not have a predetermined enumerated unit size. However, these tools still present challenges to children who are learning how to measure. Non-standard tools come in a variety of sizes often with no clear markers or numerals demarcated, leading to children's difficulties in quantifying them into linear units as well as corresponding them to targets so that they represent object lengths. Here, I identify two challenges for children when using non-standard tools in linear measurement activity.

First, treating a non-enumerated tool such as paper strip or a stick or a cutout of a shoe (all non-standard tools) as a consistent, spatial unit of measure is challenging. When asked to directly measure a target using such non-standard tools, children display tendencies of leaving gaps or generating overlaps when iterating the tools as a single unit, treating linear measures as a discrete count of a non-standard tool's units instead of a span of linear space that is represented by the end-to-end collection of the tool's unit lengths (Hiebert 1981, 1984; Lehrer, 2003; Lehrer, Jacobson, Kemeny, & Strom, 1999, Nunes et al., 1993). This also results in difficulties in splitting the tool into shorter units. For example, Lehrer et al. (1999) found that children in the second grade did not initially see problems with leaving gaps between concatenations of textbooks when measuring the width of a bookcase in the classroom. They were also unsure how to count a measure when the distance of interest could not be covered with a whole number amount of units of textbooks (i.e., a partial unit had to be created and added), resulting in linear measures that were shorter or longer than the actual lengths of objects. Only after careful guidance from the teacher and much classroom discussion did they gradually come to accept that the units should be equal in size, have no gaps between iterations, and a part of a unit should fill in any partial distances. Similarly, Hiebert (1981, 1984) found that when first grade children were shown a crooked path composed of five medium-sized Cuisenaire rods (C-rod length = 7 cm) and asked to reproduce the same length using shorter C-rods (5cm), children typically chose the same number of rods instead of using more than five rods to reproduce the same length. In this case, the children were corresponding discrete counts of units, but did not treat the rods as representing a partitioned length of the path. Nunes et al. (1993) discovered that 6- and 8-year-old children successfully accomplished indirect measuring activities with partners over the telephone using tick marks and numbers already represented on rulers compared to pieces of string because the strings had to be iterated along the span of the target (a line drawn on a piece of paper) or split into shorter lengths to produce a more precise count of length measure. In other words, one reason non-standard tools may be challenging for students is because they have to reinvent iteration or subdivision of spatial, consistent units with the tool to produce a quantified measure.

Second, even when children show emergent understandings that a non-standard tool can represent a length of an object, variations in non-standard tool lengths affect transitive reasoning between object lengths. For example, Kamii and Clark (1997) reported that, when presented a task goal to prove that one line is longer than another, children were able to use a 12-inch long tagboard strip, which was similar to the length of the line of interest to make judgments by aligning the strip against each line by second grade, but faltered in accomplishing the same task using a 1.75-inch long small block as a single unit until fourth grade. With a tool that was similar to the length of the target object, children more readily represented target lengths in comparison activity, but with a tool that was grossly different in size from the target object (and in this case, needed to be iterated), younger children experienced difficulty in representing targets lengths to compare lengths. Thus, we might anticipate that children in the middle elementary grades and younger will display challenges in transforming non-standard tools to function as devices for recording linear measures depending on the discrepancy between the length of the tool and the length of the target.

Linear measurement: A coordination of unitizing actions.

Measuring length involves identifying a linear unit and displacing it along the extent of an object (Piaget & Inhelder, 1948/1956; Piaget et al., 1960), which involves particular unitizing actions.[1] As illustrated in Figure 3 earlier, children may use various actions to generate, conserve, and measure linear units on line segments. They can generate new linear units by subdividing (*splitting action*) a line segment into shorter segments. This length then is conserved as it is displaced (*iterating action*) a particular number of counts (*counting action*) along the extent of a target object to produce a final linear measure. In the same way, tools can also be unitized using splitting, counting, and iterating actions. For example, a paper strip can be folded into shorter pieces (splitting). A folded piece then can be displaced along the length of an object (iterating) and the number of iterations (counting) would produce the final measure of the object's length. The coordination of splitting, iterating, and counting actions are key in understanding children's linear unit understandings that emerge in their problem solving activity, and my dissertation utilizes these actions in analyzing children's measurement strategies.

Number lines: A productive problem context for examining linear measurement activity.

In this dissertation, I regard unbounded number line problems[2] as a productive context for investigating the interplay between children's linear unit understandings and their measurement tool transformations. Here, I articulate the utility of this problem space in illuminating children's emergent linear unit understandings.

Recent research studies that used unbounded number line problems have documented how particular measurement strategies help reveal an individual's sense of linear unit ideas on line segments (Cohen & Blanc-Goldhammer, 2011; Ebersbach, Luwel, & Verschaffel, 2015; Reinert et al., 2015; Saxe, Shaughnessy, Gearhart, & Haldar, 2013). For example, both children and adults may use skip-counting strategies to measure a position for a number because they visualize multiple iterations of the pre-labeled interval length across the line (Cohen & Blanc-Goldhammer, 2011; Ebersbach et al., 2015; Reinert et al., 2015). Or children may treat the pre-labeled interval as a unit of 1 regardless of what the pre-labeled numbers are, such as iterating

[1] In Saxe, de Kirby, Kang, Le, & Schneider's (2015) framework, the authors describe how counting, displacing, and subdividing actions may serve as resources for learning formal concepts of measurement on the number line. In this dissertation, these sensorimotor actions are introduced as unitizing actions that can be operated both on targets (e.g., number line) as well as tools (e.g., rulers, paper strips) in measurement problem solving activity.

[2] An unbounded number line has an interval pre-labeled with two numbers such as 0 and 1, but is otherwise empty. An unbounded number line problem asks the individual to position a number outside of those two pre-labeled numbers on the number line. In other words, given only two numbers marked on an otherwise empty number line, the task is to position a third number outside of the interval.

the interval between 7 and 9 (a linear unit of 2) two counts instead of one count to the right to place 11 on the line (Saxe et al., 2013).

Unbounded number line problems also provide a context in which to investigate how individuals may transform a tool multiple times to represent a linear measure. Unbounded number line problems inherently include multiple linear measurement acts in the same problem space that would require the problem solver to transform a tool multiple times to represent multiple linear units across the number line. To solve an unbounded number line problem, one has to produce a linear measure for the pre-labeled interval and then use that as a means for producing a new shorter or longer linear measure to locate a number that is not shown on the number line. Thus, a tool needs to be adapted such that the length of the pre-labeled interval on the number line is represented on the tool and then transformed again to represent a different length on the number line to locate the missing number. Consequently, children will display different challenges in tool transformation depending on which numbers are positioned on the line, the tool's material properties, and their own linear unit understandings.

For example, consider two different unbounded number line problems for which a child has to use a ruler to locate a number on the number line. The first problem has the numbers 3 and 7 marked 4cm apart and the task is to place 10 on the line. The second problem has the numbers 3 and 5 marked 4cm apart and the task is to place 7 on the number line (Figure 6).

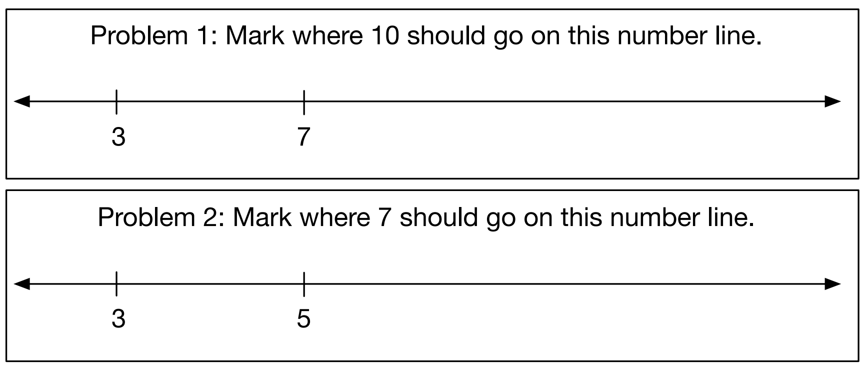


Figure 6. Two unbounded number line problems that vary in their pre-labeled values and target values

Even though the physical distance between the two pre-labeled numbers is the same between both problems, the particular values presented in each problem require the child to use different unit coordinations between the linear units on the line and units on the ruler. The approach for coordinating numeric units on the ruler and on the number line in the first problem is relatively less challenging compared to the second problem. For the first problem, each centimeter unit on the ruler can represent a linear unit of 1 on the line. For the second problem, however, the child has to coordinate the fact that 4cm on the ruler represents a linear unit of 2 on the line (Figure 7).

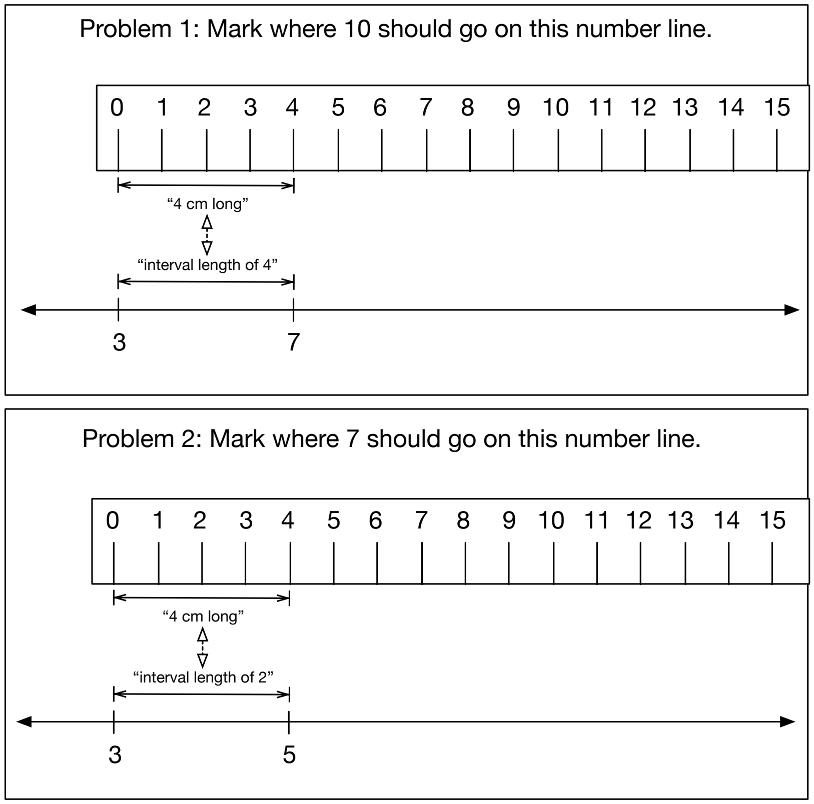


Figure 7. Coordinating units on a ruler with units on unbounded number line problems

In contrast, if the child is asked to use a rectangular paper strip to solve these problems, different challenges arise. A paper strip that is 4cm wide can be fitted onto the line and then folded, cut, or iterated without necessarily having to think about specific numerical relationships already marked on the tool such as the case of rulers. However, parts of the paper strip still have to be coordinated with linear units on the number line. For problem 1, the paper strip would need to be folded or cut so that a quarter of the paper strip represents a linear unit of 1 on the line. Then this shortened strip can be iterated 3 counts to locate 10 on the line. For problem 2, the whole length of the paper strip would represent a linear unit of 2 on the line, and since 7 is two units away from 5 on the line, the paper strip can be iterated without being folded to generate a linear measure of 1. In this case, the second problem may present fewer unit coordination challenges for the child. In comparison to the shorter paper strip, a paper strip much longer than the pre-labeled intervals on number line problems would need an additional folding action to measure the interval before further unitizing the tool and number line and coordinating between the two.

Thus, unbounded number line problems can present challenges between different kinds of numbers and different tool types that may illuminate how children coordination between unitizing actions. On an unbounded number line problem, one has to coordinate actions with linear units on the line as well as the measuring tool. Existing linear units may need to be split into shorter linear units (e.g., halving an interval of 2 to find the length of an interval of 1 or folding a paper strip into half) or iterating linear units without gaps or overlaps across the line to produce longer linear measures (e.g., iterating a linear unit of 1 three counts to find the length of an interval of 3). The problem features provide opportunities to elicit various kinds of measurement strategies, which in turn may provide insights into the challenges that children display in transforming tools with different material properties to function as linear measurement devices on the number line.

Conceptual Framework

To investigate children's conceptual coordinations in linear measurement activity, I focus on children's unitizing actions on tools and targets. I both extend Vygotsky's (1978, 1986) approach to analyzing tool use[3] in problem solving activity, and utilize Saxe's (2012) treatment of microgenesis of representational activity. The approach of this investigation treats children's use of different kinds of tools to support their linear measurement problem solving activity as being regulated by two processes: (1) children's ability to conceptualize the target as a quantifiable object that can be measured in linear units, and (2) children's efforts to unitize artifacts such that they become linear measurement tools[4] adapted to the constraints of the measurement problem.

Vygotsky's approach to analyzing tool use in problem solving activity.

In Vygotsky's (1978, 1986) approach to development, he identified two distinct lines of development that are interwoven in the child's behavior: On the one hand, there is a "natural" line of development in elementary processes that have biological roots and a quality of immediacy, such as the visual perception of independent elements in a visual field or sensory memory. On the other hand, there is a higher psychological "cultural" line of development characterized by the child's degree of mastery in the use of speech and tools to mediate these elementary processes to serve new functions voluntarily. For example, children may use speech to describe actions and relationships between elements in a picture or use objects in the environment to serve as mnemonic devices. Vygotsky viewed the latter kind of development—voluntarily controlling elementary psychological functions like memory—as being key to understanding cognitive development in children; qualitative transformation in the child's thinking is enabled by the mediational process of drawing in artifacts or language to function as "tools" to organize problem solutions.

To illustrate Vygotsky's approach to studying psychological development, I turn to his study of mediated memory. In this study, individuals of different ages played a memory game in which they were supposed to answer a set of questions such as "What is your favorite color?" and "What color is your shirt?" while following two rules: to not say particular color names (e.g., cannot say blue and red) and to not say any color name twice. In the initial condition, the

---

[3] Vygotsky (1978) distinguishes between *sign operations* and *tool use* in his approach. Though both signs and tools serve a mediating function in activity, they differ in the way that they orient human behavior. Tools are externally oriented, and are used to make changes in the activity itself. Signs are internally oriented, and are used to create and alter psychological operations. In this particular study, I refer to sign operations as "tool use" for ease of communication. Physical artifacts such as rulers and paper strips will be analyzed for their internally-oriented mediating function in problem solving; the artifacts will be presented to participants for the purpose of observing how they are drawn into measurement activity to aid conceptual coordinations of linear units.

[4] For the purposes of the remainder of this manuscript, physical artifacts presented to the children as measurement aides are termed "tools" regardless of whether the artifacts are productively drawn into problem solving activity or not. However, it is important to note that tools serve particular functions, meaning that artifacts are *transformed into* tools. An artifact that is not drawn into activity to serve a particular problem-solving function is yet a "tool."

participants were not given any tools, and in subsequent conditions they were given tools in the form of nine color cards to help them answer the series of questions successfully.

Vygotsky found that with age, individuals shifted in the way that they used the color cards to help mediate their memory and be successful at the game. Preschool-age children were not very successful in the memory game, and did not perform better when the color cards were available for use because they were not able to use them as tools to help them play the game better. School-age children progressively used the color cards productively as tools to help them play the game, and performed considerably better with the color cards than without them. For example, without the color cards, they made several mistakes by repeating color names, but with the cards, they were able to turn over cards of colors they named to help them remember that they cannot repeat those color names. Adults performed significantly better than children without physically relying on the color cards even when they were available for use.

Thus, the character of mediated activity and character of successful solutions to the game activity shifted across age groups; the color cards were first external to problem solving and not helpful for the memory game, then they were incorporated into problem solving as external tools to organize successful solutions to the game, and later they became covert in mastering the memory game. The shifting ways in which individuals used tools in their problem solving activity as a function of age also reflected the shifts in mastery of memory processes.

This developmental approach to tool use for mediating problem solving activity has potential to illuminate cognitive shifts in linear measurement problem solving, where individuals are engaged in organizing solutions for complex problems of quantifying, comparing, and manipulating magnitudes of length. Vygotsky did not focus on practices of linear measurement in particular when investigating developmental shifts in individuals, but the domain can also be distinguished between a natural line of development and the development of higher psychological functions. On the one hand, even infants are able to discriminate changes in magnitudes of length in their visual field (de Hevia & Spelke, 2010),[5] and young children are able to use perceptual strategies to compare objects in the same visual field in non-metric ways (e.g., "this object is longer than the one next to it") (Haylock & Cockburn, 1989; Steffe & Hirstein, 1976).

On the other hand, however, if a task requires metric reasoning or indirect use of measures, such as to measure and manipulate linear units to generate a new length in the problem featured in Figure 2, the activity becomes difficult to organize with perceptual strategies and requires means to regulate unit understandings and operations to construct adequate solutions, such as using tools to function as measuring devices. Younger children may falter at this kind of task, and not know how to use tools to effectively organize their problem solving. As children grow older, they may construct more effective strategies for reasoning about length in metric ways, drawing in tools to function as measuring devices. Thus, a shift towards a higher psychological function of linear measurement activity can be observed when children cannot rely on their natural perceptual strategies and draw in tools to organize their problem solving.

---

[5] I posit that the divide between natural and higher measurement understanding is analogous to psychologists' findings of children's concepts of number. Infants show abilities to nonverbally discriminate small numbers of items as individual items and large numbers of items as distinct sets without the aid of counting operations (e.g., Feigensen, Dehaenes, & Spelke, 2004; Xu & Spelke, 2000) and as children grow older, they learn number words that help build concrete numerical concepts (e.g., Fuson, 1988).

In this study, I extend Vygotsky's approach in two ways. First, I argue that shifting conceptual understandings of the problem regulate the mediated problem solving activity. Implicit in Vygotsky's approach to studying development is that changes in conceptual development are also occurring with age and may regulate mediated activity. For example, in the memory game, Vygotsky explicitly addresses age as a variable, but presumably he does not consider this a causal variable of the change in mediated activity. Rather, part of what regulated how different age groups used the color cards was the ability to identify different colors as well as understanding of the goals and relationships between components of the memory game. Thus, understandings of linear measurement could be analyzed by how children—who are at various stages in learning about how to measure length—differentially appropriate tools to function as measuring devices in problem solving activity.

Second, I introduce variations in tool types to study how children differentially appropriate different types of tools into their problem solving in linear measurement activity. Vygotsky did not explicitly explore how variations in tools would interact with children's conceptual understandings and affect problem solving activity. Yet, different types of tools have different affordances and constraints that may present different challenges to children in problem-solving activity. Children's shifting conceptual understandings would govern how they perceive the utility of different types of tools and may lead to different tools serving different kinds of functions in length measuring activity.

For example, reconsider again the number line example from Figure 2. One way to study the development of children's linear measurement activity would be to analyze the varying interactions between children's shifting understandings of linear units and their use of one tool, such as a paper strip. Another way to study the development of linear measurement activity would be to look at how children's shifting understandings of linear units would interact with how they use more than one tool—such as a paper strip and a ruler—differentially to measure and perform operations with linear units. This latter approach would better reveal the different kinds of conceptual challenges children face in using linear units, because different tools will present different kinds of obstacles in solving the problem. The paper strip is flexible, but has no inherent numerical property, and the ruler is already an apparent measuring device, marked with units and numbers, but these markings would have to be coordinated with the unit and number relationships constructed on the object of measure.

Thus, building on Vygotsky's developmental approach, this study focuses on the interplay between the shifts in tool use in measurement activity across age groups and children's linear unit conceptualizations, and variations in tools are introduced to study how children appropriate different types of tools into measurement devices to mediate their linear measurement problem solving.

Microgenesis in linear measurement activity.

Mathematical problem solving takes form in microgenetic activity (Saxe, 2012). Generating a mathematically coherent solution is a *process* in which individuals construct problem-solving goals based on their conceptualization of the problem and coordinate their cognitive as well as available material and social resources to organize their approach to the problem (Saxe, 2012; Vygotsky, 1978). This study adapts this approach and treats linear measurement problem solving as a microgenetic process. To understand individuals' linear unit understandings, their purposeful efforts at constructing mathematically coherent solutions are analyzed.

Using measurement tools to measure length is a microgenetic activity. In the process of determining an object's length—or the filled space between two endpoints of an object (Piaget et al., 1960)—one has to quantify a target of measurement (e.g., the location for a number on the number line) and artifact that becomes a tool for measurement, and construct a correspondence between the target and tool such that the tool represents a measure of the target. This coordinated effort between the quantifying of the target and tool and constructing a correspondence between the two are bootstrapped; the activities do not take on meaning on their own or follow a particular sequence of production. Rather, these unitization activities take on meaning in an individual's active efforts to represent one (e.g., the quantified target) as the other (e.g., the quantified tool) (Saxe, 2012; Werner & Kaplan, 1963). To illustrate, consider a more mature child's microgenetic process of solving the number problem presented earlier using a paper strip as a measuring tool (Figure 8).

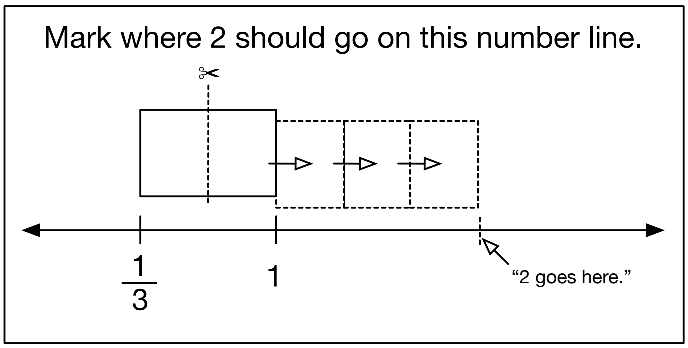


Figure 8. Example of child using a paper strip to locate 2 on the number line

In this number line problem, the child supports her measurement activity to locate a number on the number line by generating a correspondence between two quantification activities such that her tool represents a linear measure. To approach this problem, she treats the target—the number line—as a quantifiable object, and at the same time, she treats the tool—a paper strip—as quantifiable into units. In her efforts to position 2 adequately on the number line, she coordinates her quantification of the target and tool such that number line interval lengths correspond to lengths of the paper strip, and lengths of the paper strip correspond to interval lengths on the line. Through this process, she generates a numerical meaning for the paper strip such that it represents a linear unit on the number line (Figure 9).

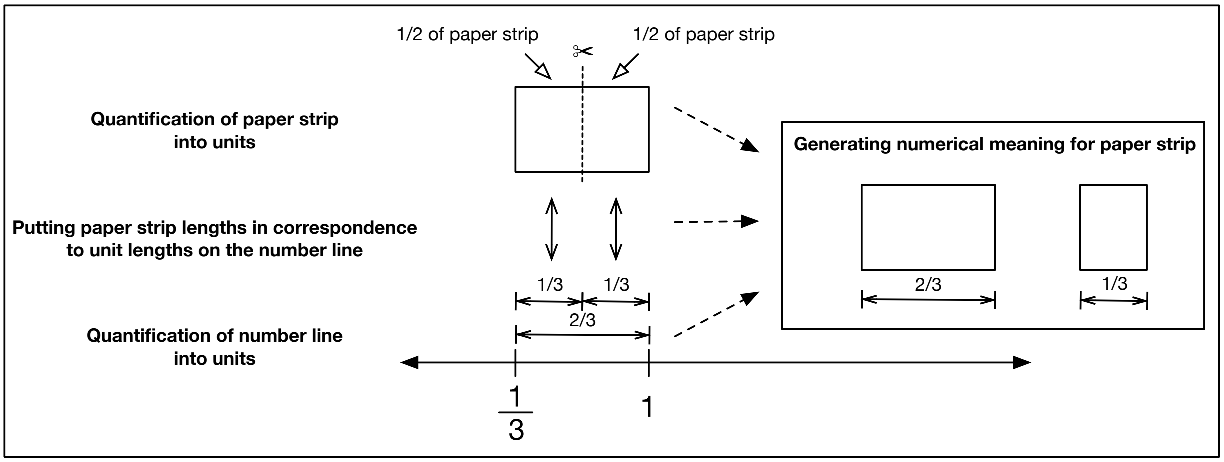


Figure 9. Microgenesis of metric meaning for the length of paper strip

In this study, I posit that two processes of quantification regulate one another when a child makes an effort to construct a metric solution to a linear measurement problem using tools, such as in the example above. These processes constrain and enable the construction of qualitatively different kinds of solutions in linear measurement activity. On the one hand, individuals unitize the target object being measured. The target itself can be conceptualized as being a length magnitude composed of a certain amount of linear units by splitting the target's length into equal-sized units. For example, in the number line problem, the interval between 1/3 and 1 on the number line can be conceptualized as being composed of two linear units of 1/3, and the location of 2 can be conceptualized as being three linear units of 1/3 to the right of 1 on the line. On the other hand, individuals unitize the tool they are using to serve a measuring function. A linear measurement tool can also be conceptualized to be splittable into equal-sized linear units. For example, a paper strip can be folded once to represent two half-unit lengths, or iterated in one direction three counts to represent a length of three units. These two unitizing activities—on the target and on the tool—must be coordinated to produce a coherent metric solution to a measurement problem. Each linear unit conceptualized on the object, such as an interval length on the number line, must correspond with a linear unit conceptualized on the measurement tool, such as the length of a short paper strip. An interval length of 1/3 on the number line can be represented by half of the length of a paper strip, and reciprocally, half of the paper strip can represent a magnitude of 1/3 on the line.

These unitizing activities and how they regulate one another will vary for children who are transitioning in their linear unit understandings. Children will be challenged in their linear measurement activity by their conceptualizations of the problem, and the various material properties of physical tools whose affordances or constraints in representing lengths may or may not be apparent to children. For example, a child who does not appreciate that an established unit length should remain consistent in a measuring act would not be able to adequately unitize objects or tools or put to correspondence the units on an object to the appropriate units on a tool to represent the same length. Likewise, a child who does not take into consideration the tool's material property may find difficulty in coordinating the two unitizing activities. For instance, if the paper strip from the above example problem is not folded or cut or iterated across the line, there would not be an adequate way to generate a metric representation for the length between 1 and 2 on the line. As children learn to recognize the need for equal-sized linear units in a count of measure and how to unitize an object by splitting lengths or iterating them across a targeted span, they also learn to discern the material affordances and constraints of tools as they unitize on them.

Research Questions.

Two research questions guided the present study of the development of students' measurement tool use.

First, what is the character of students' tool-using approaches on number line measurement problems? In other words, how do students' abilities to coordinate between their tool unitization and target (number line) unitization activities advance by grade, problem type, and tool type?

Second, how do students mediate the precision of their number line estimation activity using tools? Specifically, how are students' precision of their target value estimates on number line problems influenced with the introduction of tools, and how does this activity interact with grade level, tool type, or problem type?

Overview of the Dissertation

In this dissertation, I present a mixed design interview study of second and fourth grade students and their measurement tool using activity on number line estimation problems. I examine the interplay between children's linear unit understandings and their adaptation of different tools to support their measurement activity.

Chapter 2 presents the methods used in this dissertation. I describe the features of the research design and how I made use of both between-subjects and repeated measures comparisons, and the methods I used to conduct the semi-structured interviews. General analytic approaches are also presented here, but are presented in more detail with the results in the subsequent chapter.

Chapter 3 reports more in-depth analytic approaches and findings of students' measurement tool using activity on number line problems. I first describe how I used the conceptual framework to develop an analysis of developmental shifts in students' tool to target coordinations. Students' tool-target coordination approaches were compared across tool using conditions, grade, and problem type. Then I describe the precision of students' estimates and report findings on how tool using affected students' linear estimates on different problem types.

Chapter 4 returns to my discussion of conceptualizing children's developing linear unit understandings and tool using. I reflect on the results of the interview findings and utility of the conceptual framework, and how this approach may help illuminate children's emergent linear unit understandings. I then discuss the limitations of the study design that may have obscured students' actual transitional understandings. I consider future directions that may extend and enrich this work.

Chapter 2: Research Design and Methodology

This chapter describes the methods used in the current study. I describe the participants, how the interview study features were designed to reveal the development of students' measurement tool using understandings in the context of number line estimation tasks, the procedures of how the study was conducted, and the general approach to analysis.

Participants

Participants were one hundred and six students in second grade (29 girls, 19 boys) and fourth grade (36 girls, 22 boys) from two public elementary schools in the San Francisco Bay Area. At the beginning of the fall 2016 term, 74 students were recruited from School A (35 second graders (22 girls, 13 boys); 39 fourth graders (23 girls, 16 boys)), and 32 students were recruited from School B (14 second graders (7 girls, 6 boys); 18 fourth graders (13 girls, 6 boys)). In School A, approximately 36.5% of the student body in the 2016-2017 school year was Asian, 3% African American, 2.9% Filipino, 13% Hispanic or Latino, 0.5% Pacific Islander, 27.9% White, 15.7% mixed, and 0.3% other, and in School B, approximately 7.3% of students were Asian, 10.7% Black or African American, 8.9% Hispanic or Latino, 60.4% White, 11.3% mixed, and 1.3% other. In School A, 6.3% of students were eligible for free or reduced-price meals. and in School B, about 8.4% of the students were eligible for free or reduced-price meals (Ed-Data, 2017). Only students who received parent permission were interviewed for the research study.

I selected students in second and fourth grade at the beginning of the school year for two reasons. First, Common Core State Standards for Mathematics Content (CCSSM), adopted widely across the country by 42 states, District of Columbia, and four US territories, recommends that students start to indirectly measure two objects using a third object in the first grade, and learn how to use standard measurement tools in the second grade (NGA Center & CCSSO, 2010). Given the CCSSM's recommendations, I expected that second grade students in the fall term would have had little formal instruction in the use of standard tools, and some formal instruction with non-standard tools. The study of second graders' tool using strategies would thus allow for analyses of children's incipient knowledge of measurement and the use of tools to support their measurement activity. In contrast, I expected that fourth grade students in the fall term would have had significantly more experience with using both standard and non-standard tools and be able to adapt some tools into their measuring activity. However, their tool-using strategies would illuminate the persisting challenges different kinds of tools may impose on students as they try to unitize them and measure different target lengths.

The second reason for selecting these two grade levels was that summer pilot interviews with children who had completed grades 1 through 5 by June 2016 revealed that those who *completed* second grade, but not those who completed first grade, were showing differentiated ways in using different types of tools in similar ways to what I had expected for older children, and those who *completed* fourth grade, but not those who completed third grade, were showing little to no variation in their use of tools during problem solving. In fact, students who had completed fourth and fifth grades were placing numbers adequately regardless of tool type or number line target values.

Participant recruitment process.

After the study protocol was approved by the university's Internal Review Board in August 2016, I approached the school sites' principals to obtain informal approval. Then I proceeded with the corresponding school districts' formal approval process. Once authorized by the school districts, I obtained formal approval letters from principals and contacted the schools' second and fourth grade teachers. Students were recruited from classrooms whose teachers who were interested in the research study.

Parent permission forms and student assent forms were color-coded, one color for parents and another for students, and distributed by second grade and fourth grade teachers to their students during third week of September for School A and during the second week of October for School B. I collected completed forms the following week. Students who received parent permission and assented to the study were interviewed in the months of October and November in 2016.

Interview Study Design

Overview.

I used a mixed design to analyze students' developing understanding of linear measurement. As shown in Figure 10, 106 students (48 second and 58 fourth graders) were randomly assigned to one of three *tool types* (ruler, short paper strip, and long paper strip). Within each treatment group, I engaged students with semi-structured interviews in which I first asked them to estimate the position of target values—whole numbers—on five line segments that had two whole numbers pre-labeled (*given interval*) *without* using any physical aids (*tool absent* condition). After placing their estimates, participants, depending upon condition assignment, were presented with a tool and again asked to estimate numerical positions on five similar line segments *using* the assigned tool (*tool present* condition). The linear measurement problems varied with respect to whether the given interval between the two pre-labeled whole numbers was a *unit interval* (a unit value of 1) or a *multiunit interval* (a unit value that was a whole number greater than 1). A target value's distance from the given interval (*measured interval*) varied for multiunit interval problems in two ways: (1) at a *whole multiple* distance of the given interval length, or (2) at a *fractional multiple* distance of the given interval length. Thus, this study used a 2 (grade: second grade, fourth grade) × 3 (tool type: ruler, short paper strip, long paper strip) × 2 (tool presence: absent, present) × 2 (given interval: unit, multiunit) × 2 (measured interval: whole multiple, fractional multiple (for only multiunit interval problems)) design. The following subsections describe the study features in more detail.

Tool Type	Ruler				Short Paper Strip				Long Paper Strip			
Tool Presence	Tool Absent		Tool Present		Tool Absent		Tool Present		Tool Absent		Tool Present	
Given Interval	Unit	Multiunit	Unit	Multiunit	Unit	Multiunit	Unit	Multiunit	Unit	Multiunit	Unit	Multiunit
Measured Interval	WM	WM, FM	WM	WM, FM	WM	WM, FM	WM	WM, FM	WM	WM, FM	WM	WM, FM
Grade 2	S01-S16 (n=16)				S17-S32 (n=16)				S33-S48 (n=16)			
Grade 4	S49-S68 (n=20)				S69-S86 (n=18)				S87-S106 (n=20)			

Figure 10. Research design of interview study (WM = whole multiple, FM = fractional multiple)

Assignment of participants to study conditions.

Students in each grade were randomly assigned to one of three tool treatment groups (ruler, short paper strip (SPS), long paper strip (LPS)), balanced by gender and classroom.

Within each tool group, two problem sets were counterbalanced; half of each group received Set A for the tool absent condition, and Set B for the tool present condition, and the other half received the opposite for tool presence conditions. The order of problems within each problem set was randomized for each student using an online list randomizer (random.org).

To randomly assign students in each grade to the different treatment conditions, I first obtained each classroom's roster of participating students and divided each into two lists of girl and boy groups alphabetized by last name. Within each group, students' names were randomized using an online list randomizer (random.org). After appending each classroom's boy lists into one combined classroom boy list and then girl lists into one combined classroom girl list, these two randomized lists of participants were each assigned a sub-treatment condition in this order: Ruler, problem set A for tool absent condition and problem set B for tool present condition (RAB); Ruler, problem set B for tool absent condition and problem set A for tool present condition (RBA); Short Paper Strip, problem set A for tool absent condition and problem set B for tool present condition (SAB); Short Paper Strip, problem set B for tool absent condition and problem set A for tool present condition (SBA); Long Paper Strip, problem set A for tool absent condition and problem set B for tool present condition (LAB); Long Paper Strip, problem set B for tool absent condition and problem set A for tool present condition (LBA).

Materials.

The following materials were used in the interviews (see Figure 11): Three types of tools (ruler, short paper strip, long paper strip), 13 individual number line problem worksheets, and three animal pointers (fox, rabbit, turtle).

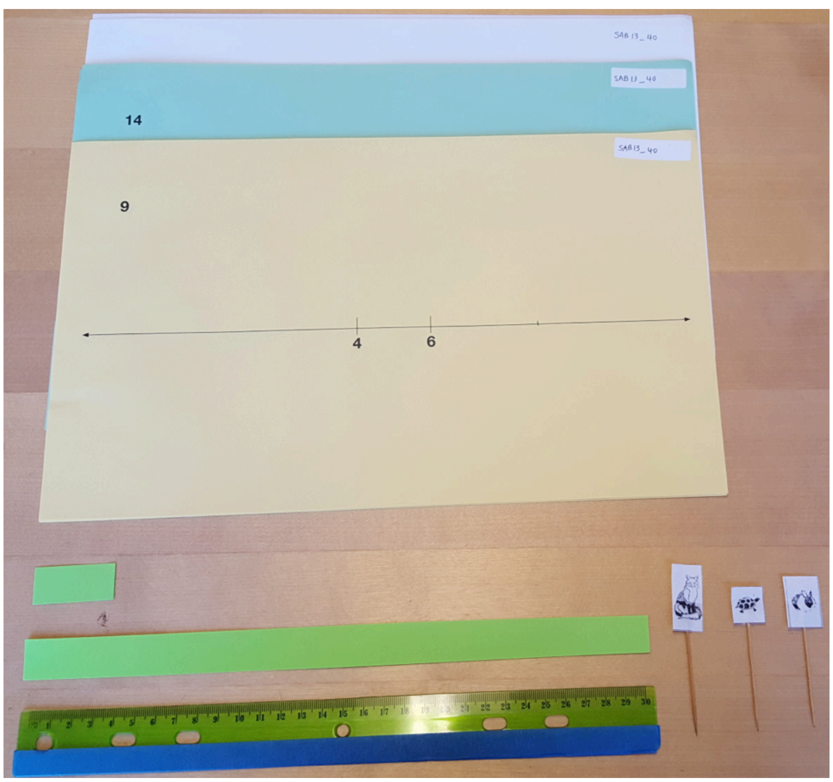


Figure 11. Interview materials: Problem sheets (top), tools (bottom left), animal pointers (bottom right)

The ruler was a neon green plastic ruler with 12 inches (calibrated to 16th of an inch) marked on one edge and 30 centimeters (calibrated to millimeters) marked on the other edge. The inch side was covered with blue painter's tape to prevent participants from using inch units. The short and long paper strips were cut from neon green 20-pound printer paper. The short strip's dimensions were 1.9 cm by 4 cm (the length of the given interval on each problem), and the long strip's dimensions were 1.9 cm by 30.9 cm (the full length of the ruler).

The number line problems included three orientation problems (printed on white paper), five tool absent problems (pastel green paper), and five tool present problems (pastel yellow paper). On each printed problem sheet, a whole number (target value) was printed in the upper left corner, and two tick marks, set 4cm apart, were pre-labeled with whole numbers (given interval) and centered on a line segment 33cm long.[6] To fit the length of the line segment on problem sheets, the problems were printed on legal size paper (11 inches (27.94cm) by 14 inches (35.56 cm)).

Each animal pointer had a monochrome picture of the animal taped onto one end of a toothpick and was between 75 mm and 80 mm tall.

Tool types.

I included three tool treatment groups to study variation in children's uptake of particular standard and non-standard tools[7] in their problem solving on number line problems. Tool absent and tool present conditions were included across standard and non-standard tool conditions to compare and contrast how students' efforts to unitize tools affected the character of their measurement activity.

Rulers and paper strips have different affordances and constraints for being transformed into linear units. On the one hand, rulers are standard tools that are marked with enumerated units, but these numbered units may not readily align with numerical relationships constructed on number lines. The ruler was chosen to observe how children might support this coordination in measuring linear segments on the number line. On the other hand, paper strips are non-standard tools that are flexible and have no inherent numerical property. In other words, paper strips have to be further transformed into units (e.g., folded into a particular interval length) to measure linear segments on the number line. The two different paper strip lengths were chosen to observe whether a particular length would support measurement activity better by resembling a potential unit for iteration or by being much longer than the given interval and resembling the shape of a ruler. The short paper strip was designed to fit the provided interval in each given problem exactly. The long paper strip was designed to look like a non-calibrated ruler (same length as ruler), but needed to be first transformed into the length of the short paper strip to be used effectively in the given linear measurement problems.

---

[6] This length of the given interval was to allow for sufficient variation in how students might approach the given interval type and measured interval type combinations within each problem set. This length allowed for a possible student solution in which the given interval could be iterated three counts to the right without running out of space on the sheet, a strategy that was common in my informal classroom and pilot study observations when the target value was three units or more than one of the given interval values.

[7] Recall that non-standard tools have no inherent numerical properties and can be unitized to serve as measurement devices, such as a shoe being iterated and counted along an object to measure it. Standard tools are already unitized with units for the purpose of serving as measurement devices, such as rulers marked in inches or centimeter cubes.

Number line problems: Given intervals and measured intervals.

The goal of the number line problems was to elicit three types of actions that could support students' unitization strategies. These actions consisted of splitting (subdividing linear distances), iterating (displacing a preserved linear distance), and counting (accumulating a final count of measure of a particular linear distance).

Structurally, the number line problems included a given interval of (*a, b*) with target value *c* to be positioned at a measured interval (*b, c*). A given interval referred to the linear distance between pre-labeled values *a* and *b*, and a measured interval represented the linear distance between the given interval and target value *c* (i.e., distance between pre-labeled value *b* and target value *c*) Positioning target value *c* on a number line entailed using the given interval between *a* and *b* to produce a measured interval between *b* and *c*. (see Figure 12).

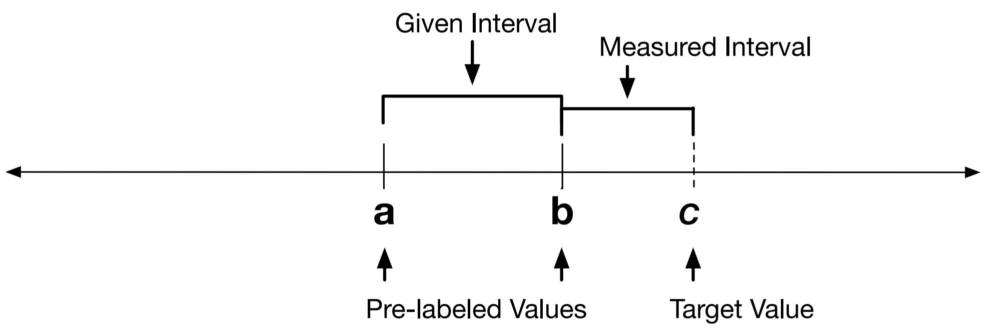


Figure 12. Structure of number line problem showing given interval of (*a, b*), target value *c*, and measured interval (*b, c*)

In the interviews, each number line problem was presented as a racecourse on which three animal figures were running. I chose this thematic context because this measurement context might better encourage students to think about linear distances between nonconsecutive numbers on number lines (Saxe et al., 2013). The two pre-labeled whole numbers for the given interval on the number line represented two of the animal figures' positions in the race in miles, and the target value (also a whole number), printed in the upper left corner, represented the third animal's position in the race. Students were tasked with measuring the target value location for an animal after being given two animals' distances on the racecourse. For example, consider a problem where the turtle has run 3 miles, the rabbit has run 7 miles, and the fox—to be positioned by the child—will run 10 miles. Here, the given interval is the linear distance between 3 (the turtle) and 7 (the rabbit), which is 4 units (or miles), and the measured interval is the linear distance between 7 (the rabbit) and 10 (the fox), which is 3 units (or miles). Thus, an adequate solution is to position the fox 3 units to the right of the rabbit on the number line (Figure 13). However, the correct length for 3 units on this problem can only be derived from the given interval value and length. To measure 3 units, one would need to first generate the length of a linear unit (unit of 1) by splitting the given interval into 4 equal length segments. Then this measure could be displaced to the right three times end-to-end to position the fox on the racecourse.

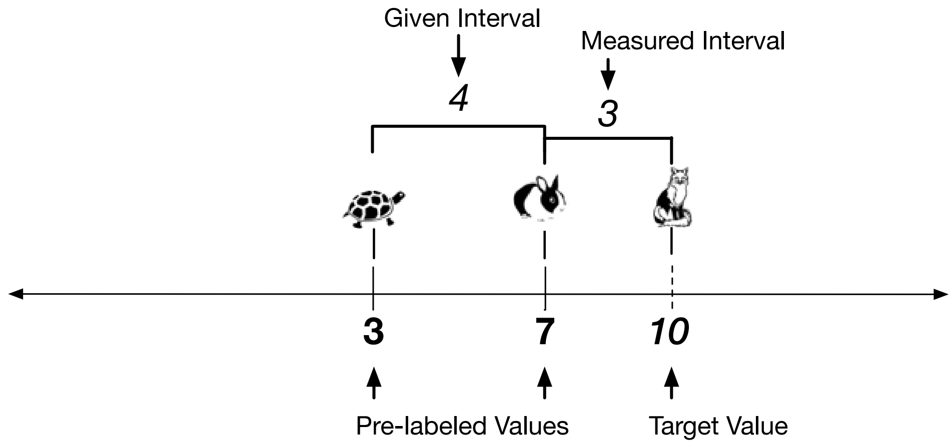


Figure 13. Placing the fox at a measured interval of 3 units given an interval of 4 units between the turtle and the rabbit

Number line problems: Problem type.

I generated five different types of number line problems by varying the given interval and measured interval of number line problems. For example, to produce a measured interval on a number line problem adequately, problems varied in whether the given interval should be split into shorter segments before being iterated a number of counts to the target value's position on the number line or the given interval could be iterated a number of counts to the target value's position on the number line without any splitting actions to generate shorter segments. The possible combination of actions varied depending on what values were used for the problems.

I utilized two types of given intervals to encourage certain unitization actions on the number line problems: *unit intervals* and *multiunit intervals*. A unit interval consisted of two consecutive whole numbers (e.g., 7 and 8 or 5 and 6), and a multiunit interval consisted of two nonconsecutive whole numbers (e.g., 6 and 8 or 3 and 7). To position the target value on a unit interval problem adequately, children needed to iterate the given interval (a linear distance of 1) and accumulate a count of their linear measure. To position the target value on a multiunit interval problem adequately, children needed to split the given interval (a linear distance greater than 1) into shorter linear units and iterate and count a linear measure.

To systematically vary the unitization actions on number line problems with unit and multiunit given intervals, I made use of two types of measured intervals: *whole multiple intervals* and *fractional multiple intervals*. The relationship between given intervals and measured intervals could be expressed as $n \times$ *given interval* = *measured interval*. Here, a measured interval was considered a whole multiple interval if n, the factor of the expression, was a whole number. If n was a fraction, then the measured interval was considered a fractional multiple interval. Since whole multiple intervals were whole multiples of the given interval length, the given interval could be iterated and counted to generate a measured interval on the number line. Splitting was not a necessary action for placing the target value. Fractional multiple intervals were fractional multiples of the given interval length, so a splitting action was necessary on the given interval (into quarters or halves) to generate shorter linear measures. Then those shorter units could be iterated and counted to produce a measured interval for the target value on the number line problem. Figure 14 shows the particular combinations that were used for this study, and each problem type is described below.

Problem	Given Interval	Measured Interval
1	Unit interval (1 unit)	Whole multiple (2 × 1 unit = 2 units)
2	Multiunit interval (2 units)	Whole multiple (1 × 2 units = 2 units)
3	Multiunit interval (2 units)	Fractional multiple (1.5 × 2 units = 3 units)
4	Multiunit interval (4 units)	Fractional multiple (0.5 × 4 units = 2 units)
5	Multiunit interval (4 units)	Fractional multiple (0.75 × 4 units = 3 units)

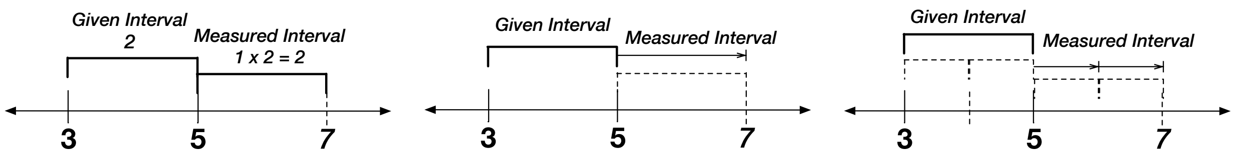
Figure 14. Five problem types designed for the study

The *Unit Interval Problem* (problem type 1) had a target value to be placed at a whole multiple measured interval by a factor of 2. For this problem type, the given interval was one unit, and a target value was to be placed two units away (2 × 1 unit = 2 units) from the given interval (e.g., Given (5, 6) place 8 on the number line). A special feature of this problem type was that the given interval did not need to be split into shorter linear segments to generate other whole numbers on the number line. A numerically coherent solution preserved the given interval length with a tool, and iterated that preserved length towards the right two times without gaps to position the target value.

The *Multiunit Interval Problem with Whole Multiple Measured Interval* (problem type 2) was a multiunit interval problem of 2 units with a target value to be placed at whole multiple measured interval by a factor of 1. For this problem type, the target value was to be placed two units away (1 × 2 units = 2 units), which also happened to be equivalent to the given interval length by a factor of 1 (i.e., whole multiple) (e.g., Given (7, 9) place 11 on the number line). To produce an estimate, the given interval could be split into shorter linear units to generate other whole numbers on the number line. However, the splitting step was not absolutely necessary for solving the two-unit interval problems; since the given interval was 2 units and the measured interval was also 2 units, it was possible to simply measure the given interval and iterate that length to produce an estimate for the target value.

Two types of numerically coherent solutions were possible for this type of multiunit interval problem. One involved measuring the given interval length with the presented tool and splitting that measure on the tool into half to represent a unit interval. The halved linear measure could then be iterated two counts to the right. The other involved using the full measure of the given interval on the tool to represent 2 units, and iterating that preserved length once towards the right to position the target value.

To illustrate the possible unitization actions to produce an adequate solution, consider the following number line problem that had a two-unit interval (3, 5) as the given interval and target value 7 (Figure 15a). The measured interval was 2 (linear distance between 5 and 7), which was a whole multiple of 2 (the given interval) by a factor of 1 (1 × 2 = 2). To position 7 on the number line, the given interval could be iterated one count to the right (Figure 15b) *or* the given interval could be split once to create two linear units of 1. Then that split linear measure could be iterated two counts to the right (Figure 15c). Both approaches involved iterating and counting actions, but a splitting action was necessary only for the latter approach, and not the former.



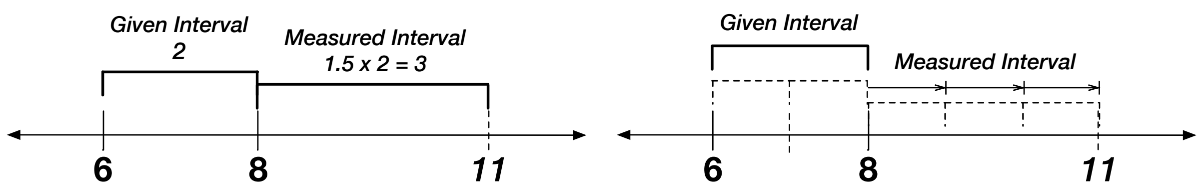
a. Given interval (3, 5) and whole multiple measured interval (5, 7)

b. Iterating the given interval one count to generate the measured Interval

c. Splitting the given interval once and iterating that unit two counts to generate the measured interval

Figure 15. Example of iterating, counting, and splitting actions to produce a whole multiple measured interval from the given interval (3, 5) and place 7 on the number line

The *Multiunit Interval Problem with 1.5-Fractional Multiple Measured Interval* (problem type 3) was a multiunit interval problem of 2 units with a target value to be placed at a fractional multiple measured interval by a factor of 1.5. For this problem type, the target value was to be placed 3 units away (1.5 × 2 units = 3 units) (see Figure 16a for an example problem). To produce an estimate for the target value's position on the number line, the given interval (or an equivalent length) needed to be split into shorter linear units to generate other whole numbers on the number line. An example of a numerically coherent solution to this problem would be to split the given interval into quarters to generate a unit interval length, and then to iterate that unit length three times to the right to position the target value. For example, on a problem with given interval (6, 8) and target value 11, the given interval had to be split into half to create a unit of 1, so that the shorter unit could be iterated three counts to the right to place the target value of 11 (Figure 16b). For a problem with a multiunit given interval and a fractional multiple measured interval, all three unitization actions were needed. Unlike problem types 1 and 2, a splitting action was necessary for this type of problem (and problem types 4 and 5 outlined below).



a. Given interval (6, 8) and fractional multiple measured interval (8, 11)

b. Splitting the given interval once and iterating that unit three counts to generate the measured interval

Figure 16. Example of iterating, counting, and splitting actions to produce a fractional measured interval from the given interval (6, 8) and place 11 on the number line

The *Multiunit Interval Problem with 0.5-Fractional Multiple Measured Interval* (problem type 4) was a multiunit interval problem of 4 units with a target value to be placed at a fractional multiple measured interval by a factor of 0.5. For this problem type, the target value was to be placed 2 units away (0.5 × 4 units = 2 units) (e.g., Given (10, 14), place 16 on the number line). To produce an estimate for the target value, at least one splitting action was necessary: the given interval could be split into four unit interval lengths and iterated twice to the right, or the given interval could be split into two two-unit lengths and iterated once to the right.

The *Multiunit Interval Problem with 0.75-fractional Multiple Measured Interval* (problem type 5) was a multiunit interval problem of 4 units with a target value to be placed at a fractional multiple measured interval by a factor of 0.75. For this problem type, the target value was to be placed 3 units away (0.75 × 4 units = 3 units) (e.g., Given (5, 9), place 12 on the number line). To produce an estimate for the target value, at least one splitting action was necessary. The given interval could be split into quarters to produce a linear unit which could be iterated three times to the right, or the given interval could be split into half to produce a two-unit length, which could be iterated once, and then split again into a unit interval, which would be iterated once more.

For the interviews, I created two sets of problems with each set comprising one problem for each problem type. In other words, each problem type had two forms that were structurally the same but used different values. Figure 17 below depicts these number line problems organized by problem type and problem set.

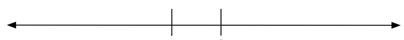
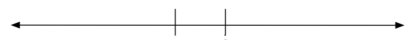
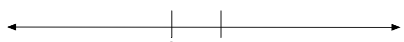
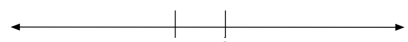
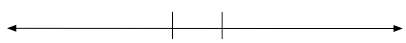
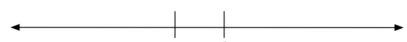
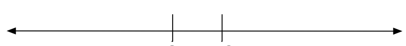
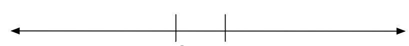
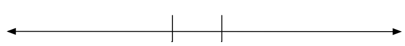
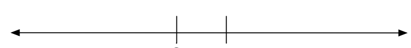
Problem Type	Set A	Set B
1: Unit given interval with whole multiple (2) measured interval	10  7 8	8  5 6
2: Multiunit (2) given interval with whole multiple (1) measured interval	7  3 5	11  7 9
3: Multiunit (2) given interval with fractional multiple (1.5) measured interval	11  6 8	9  4 6
4: Multiunit (4) given interval with fractional multiple (0.5) measured interval	14  8 12	16  10 14
5: Multiunit (4) given interval with fractional multiple (0.75) measured interval	12  5 9	10  3 7

Figure 17. Illustration of two problem sets designed for the interview study

Interview Procedures

Students were individually interviewed from each class in alphabetical order by last name during the school day in a quiet hallway or office at the school site. On average, the interviews took about 20 minutes for second grade participants and 15 minutes for fourth grade participants to complete. The video camera was positioned such that the recording would capture audio, the students' worksheets, hand gestures, and tool use, but not any other identifying features, such as their faces. At the beginning of each interview, I introduced myself and the general purpose of the study. Students were told that they could choose to stop the interview at any time. I then turned on the camera to begin the videotaped portion of the interview. Each problem was presented one at a time, and completed problems were turned over and placed away from the students' view before the subsequent problem was presented. At the end of the interview, the students were thanked for their participation and the recording was stopped. Afterwards, students were invited to ask any questions they had about the interview and escorted back to their classrooms.

For each participant, the videotaped portion of the interview progressed through four phases in chronological order: number line orientation, tool absent condition, tool orientation, tool present condition.

During the first phase, *Number Line Orientation*, students were familiarized with three number line tasks using materials that resembled those to be used in the subsequent phases of the interview. My goal was to determine whether students: (1) understood the task at hand (placing animals on a racecourse or numbers on the number line); (2) articulated that numbers increase in value from left to right on the number line; and (3) comprehended the interviewer's questions and were able to explain their thinking to the interviewer (me). The first task was a blank number line presented to students as a racecourse onto which they were asked to place three animals (a turtle, a rabbit, and a fox), such that the turtle ran the least, the fox ran the most, and the rabbit ran more than the turtle but less than the fox. The second and third tasks had two numbers pre-labeled, and students was asked to place the third number on the line and explain their solution. Afterwards, they were asked follow up questions about the order of numbers on the line (e.g., "Can 5 go here (to the left of 2)? Why or why not?").

During the second phase, *Tool Absent Condition*, students solved a set of five number line problems from set A or set B depending on condition assignment, without the use of any physical aids. I recorded and observed the precision of participants' measures. Participants were asked to sit on their hands so that they could only use their eyes to estimate positions of target values. For each problem, I presented the racecourse story while placing two animals on the pre-labeled interval values on the number line, and then moved the third animal pointer from left to right across the line until students said "stop!" for their target value estimation. At this point, I stopped the animal pointer and marked its location with a pen on the line.

In the third phase, *Tool Orientation*, students were familiarized with their assigned tool. The purpose of this phase was to illuminate whether students had any challenges in transforming the tool to represent different lengths. To this end, students were asked specific unitization questions about their assigned tool during the third phase of the interview. For the ruler condition, participants were asked to identify a length of 4 centimeters. Then they were asked if it was possible to identify that length elsewhere on the ruler. The same questions were asked for the length of 2 centimeters. For the short paper strip and long paper strip conditions, participants were asked to fold the strip to show two pieces of the same size. Subsequently, participants were asked to fold the strip to show four pieces of the same size.

In the final phase of the interview, *Tool Present Condition*, students solved a set of five number line problems (different set from the tool absent condition) and were asked to use their assigned tool to help measure and position the animal at a target distance on the racecourse. The purpose of this phase was to illuminate students' unitizing strategies on their measurement tools and the target number lines, and how they coordinated between tool and target to produce linear estimates. The racecourse context was used again. I placed two animals on the given interval and asked them to place the third animal at the target value's location. After participants indicated the location of the animal for a problem, I marked the location with a pen and then asked participants how they used the tool to support their solution. If participants did not use the tool to support their estimation activity even after being asked at least twice, I followed up with questions such as, "Is it possible to use [the tool] on this number line? Why? Why not?" Figure 18 provides an illustration of the interview protocol for the tool absent and tool present conditions (See Appendix A for the full interview protocol).

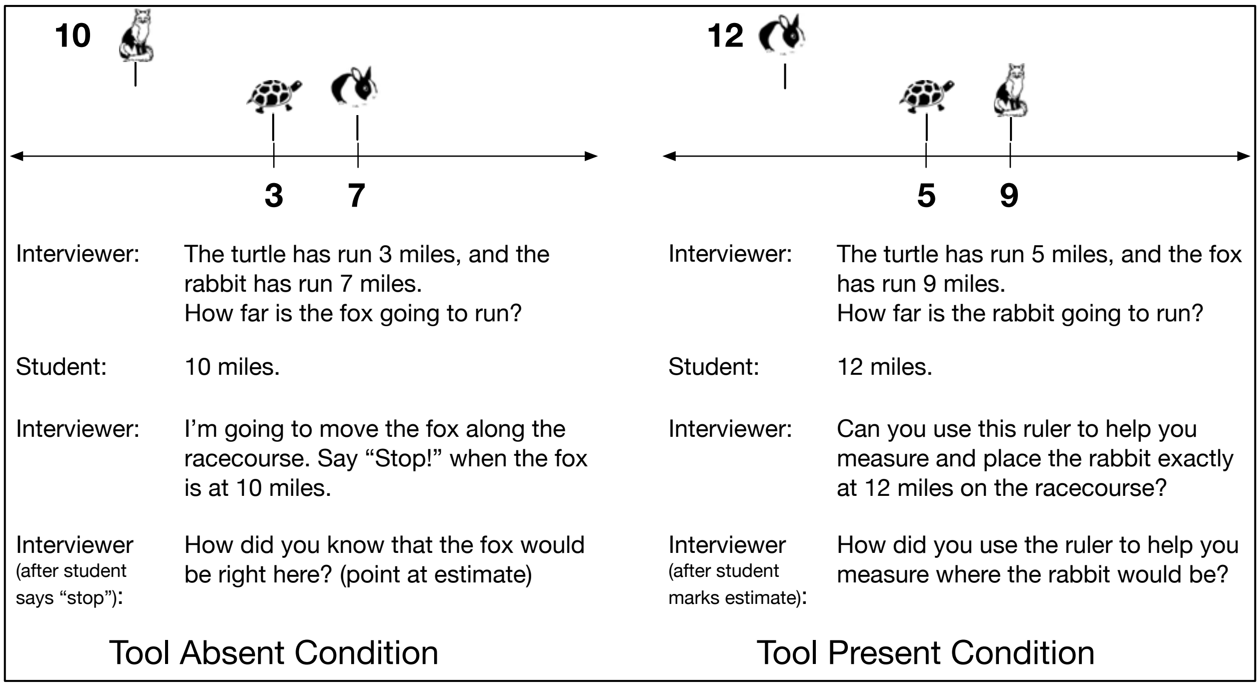


Figure 18. Illustration of interview protocol for tool absent and tool present conditions of one problem type for a participant

Data storage procedures.

At the end of each interview day, the interview video files were compressed into smaller files and transferred to an encrypted hard drive in a locked office. Students' worksheets were filed in a locked office separate from the location of the encrypted drive with their video data. All files (videos and worksheets) were labeled with corresponding participant IDs, and the participant ID key was stored in a locked office separate from the location of the encrypted drive and worksheets.

General data analysis procedures.

Out of the data corpus of 106 interviews, I analyzed 75 students' tool using approaches in linear measurement activity using two data sources: the video recordings of tool present conditions, and the students' linear estimates on their problem worksheets. Using a random number generator function in Microsoft Excel, I randomly selected 36 students in second grade, and 39 students in fourth grade from the tool treatment groups. The students included in the analyses were ones that demonstrated understanding that numbers increased in value from left to right on the number line in the number line orientation phase of the interview. Additionally, all of the 75 students were able to unitize their assigned tool (e.g., identify 4cm on the ruler, fold a paper strip into half) in the tool orientation phase of the interview. Table 1 displays the distribution of students in each condition for the 75 students.

Table 1
Number of Students Whose Tool Using Approaches Were Analyzed (by Grade and Tool Type)

	Ruler	Short Paper Strip	Long Paper Strip	Total
Second Grade	12	12	12	36
Fourth Grade	13	14	12	39
Total	25	26	24	75

As a first step to analyze students' tool using approaches, I developed a Tool-Target coordination coding scheme (presented in Chapter 3) and coded video records of students' problem solving. The aim of this analysis was to examine and document whether there were qualitatively different ways in which students unitized tools and targets and coordinated between the two, and to see if students' approaches differed by grade, tool condition or problem type. I used the video coding software Studiocode™ and identified instances of iterating, splitting, and counting actions students displayed on their tools and on the number line problems they solved. These instances were then labeled to identify if the actions were appropriate to achieve an adequate solution to the given problem. Using the conceptual framework, I categorized different coordinations of actions into three levels of student coordinations between tool unitization efforts and target (number line) unitization efforts. These codes were applied to each problem solution (in the tool present condition), and recorded in a Microsoft Excel database for each participant. To analyze students' precision of linear estimates, I used a ruler to measure the linear distance between a student's estimate and the given interval for all tool absent and tool present problems. These were recorded in the Microsoft Excel database. I compared students' estimates to the ideal linear distances to calculate precision scores for tool absent and tool present conditions. This process is presented in more detail in Chapter 3 with the corresponding analysis. The focus of this analysis was to understand if and how students' precision of estimates shifted as a result of the introduction of a tool (i.e., changes in precision from the tool absent to tool present conditions). These measures were compared and contrasted between grades and tool treatment groups across problem types.

Chapter 3: The Development of Measurement Tool Use on Linear Estimation Problems

Introduction to the Analysis

The measurement framework presented in chapter 1 indicates that students' abilities to produce linear measurements should vary as both a function of type of tool used and the properties of the problems they solve; further, the affordances of the tools should vary over children's development. Accordingly, I presented students with tasks in which I systematically varied features of problems and features of tools.

For problems, recall that I presented number lines that varied in two principal ways: I varied the *given interval size* (the distance, in number line units, between the two pre-labeled values) and I varied the to be *measured interval size* (the distance, in number line units, of the target value from the rightmost endpoint of the given interval). The variation in the given interval on a problem was either a *unit interval* (a linear unit of 1) or a *multiunit interval* (a linear unit greater than 1: either 2 units or 4 units). The variation in the to be measured intervals was either *whole multiples* of the given interval size (the same size or twice the size) or *fractional multiples* of the given interval size (0.5, 0.75 or 1.5 times the size) (Figure 19). I included these variations in order to motivate different unitization actions (e.g., iterating, splitting). For instance, a unit interval like (7, 8) could be iterated across the number line to measure a whole multiple measured interval (2 times one unit) and place 10 on the line (Figure 20). In contrast, a multiunit interval like (6, 8) would need to be split into two linear units in order to measure a fractional multiple measured interval (1.5 times 2 units or a total of 3 units) and produce an estimate for 11 on the number line (Figure 16).

Unit Interval Problem with Whole Multiple Measured Interval	Multiunit Interval Problem with Whole Multiple Measured Interval	Multiunit Interval Problems with Fractional Multiple Measured Intervals		
2 multiple (Problem 1)	1 multiple (Problem 2)	1.5 multiple (Problem 3)	0.5 multiple (Problem 4)	0.75 multiple (Problem 5)
Given (7, 8) place 10; Given (5, 6) place 8	Given (3, 5) place 7; Given (7, 9) place 11	Given (6, 8) place 11; Given (4, 6) place 9	Given (8, 12) place 14; Given (10, 14) place 16	Given (5, 9) place 12; Given (3, 7) place 10

Figure 19. Problem types and their features

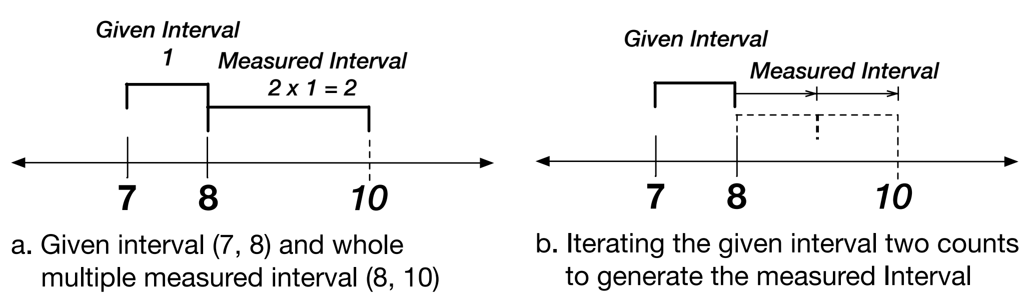


Figure 20. Example of number line unitization on problem with given interval of (7, 8) and target value 10

For tools, I presented three types—a ruler, a short paper strip, or a long paper strip—that differed in their material properties by either being flexible with no pre-labeled units (long paper strip and short paper strip) or being rigid and having pre-labeled units (ruler). Like problem features, tool features were varied to afford different kinds of unitization actions. For example, a long paper strip could be aligned with a given interval and folded to represent the length of that interval on the number line. A short paper strip could be fitted into a given interval without any initial folding or folded in half to represent a shorter length and iterated across the line to generate other numbers on the number line. The ruler could be used to measure the length of the given interval in centimeters, and that measure could be reproduced, split, or marked elsewhere on the number line using the enumerated tick marks on the ruler.

General expectations about students' use of tools by problem type and grade level. At second grade (earlier in development), I expected students to struggle in producing numerical estimates on targeted number line problems due to difficulties in co-quantifying targets (number lines) and tools (paper strips, ruler) with the use of a consistent unit, which would not support or perhaps hinder the precision of estimates. Tool-target coordinations would be more tractable with problems that allowed for fewer splitting or iterating actions, such as the unit interval problem (two iterations of given interval, Figure 20) and the multiunit interval problem with whole multiple measured interval (one iteration of given interval if given interval is not split to generate linear unit), leading to better precision of estimates. In contrast, tool-target coordinations would be more challenging for problems that presented a multiunit interval with a fractional multiple measured interval, such as the multiunit interval problem with a 1.5-fractional multiple measured interval (splitting of given interval and three iterations of new unit, Figure 16), possibly leading to no improvement or even a decrease in precision as compared to when solving problems without tools. Different tools would also vary in their affordances for unitizing as a function of problem type. For example, compared to a long paper strip that has to be folded first to measure a given interval, second grade students may more readily fit the short paper strip into the given interval on the number line, and then iterate this as the unit length across the number line to position a target value. This would be an adequate solution for positioning other whole numbers on a number line with a given interval of (7, 8). However, this approach—treating their short paper strip length as representing a linear unit on the target—would *not* be an adequate solution for a problem that has a multiunit interval such as (6, 8) as the given interval, leading to an overestimation for the location of the target value.

At fourth grade (later in development), I expected students to produce numerical estimates on targeted number line problems by co-quantifying targets and tools with the use of a consistent unit, leading to better precision in linear estimates with the use of tools. Similar but to a lesser degree compared to second grade students, I still anticipated that fourth grade students' tool-target coordinations would be more successful (and lead to better precision) on number line problems with unit intervals or whole multiple measured intervals as contrasted to multiunit interval problems with fractional multiple measured intervals, and that different tools would vary in their affordances for unitizing as a function of problem type. For example, I expected that the majority of fourth graders would be able to co-quantify units on tools and targets with the use of any tool on unit interval problems or multiunit interval problems with whole multiple measured intervals without much difficulty. However, coordinating numerical units on the ruler with numerical units on the number line might be more challenging for multiunit interval problems with fractional measured intervals. Students would need to produce more complex conceptual coordinations when using a ruler because of the two numerical scales they have to keep track of during their measuring activity. For instance, if a given interval of 2 linear units (e.g., (6, 8)) on the number line is 4cm long on the ruler, then a linear unit is equivalent to 2cm, and thus, a measured interval that is 3 linear units long would be 6cm long (i.e., 11 is positioned 3 linear units greater than given interval (6, 8)) (Figure 21).

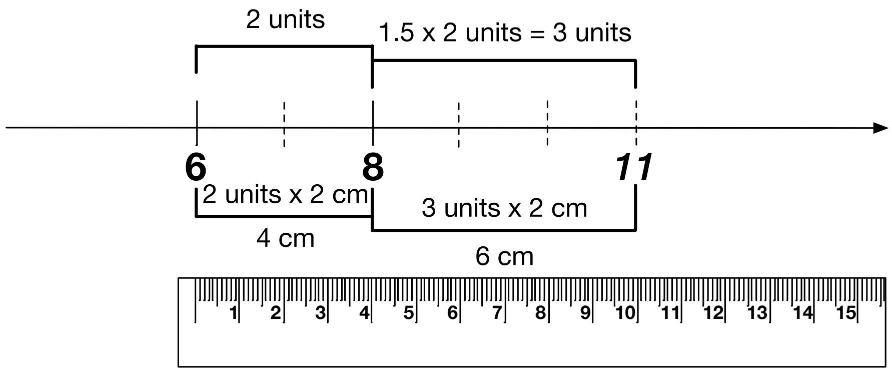


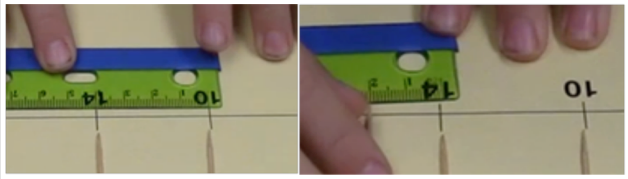
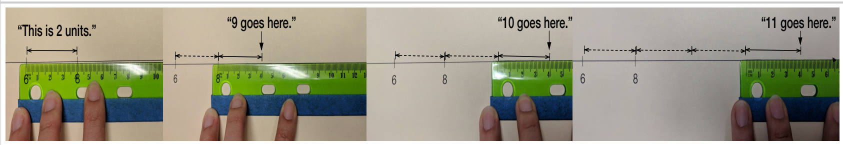
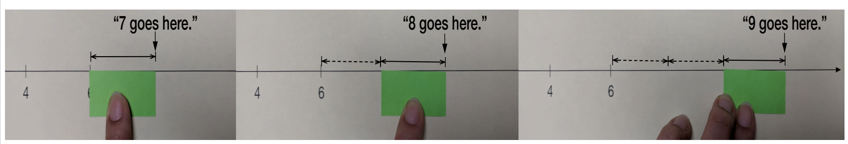
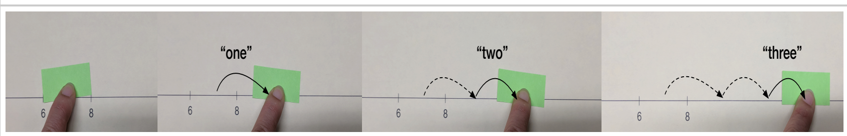
Figure 21. Coordinating ruler units with number line units on multiunit interval problem with 1.5-fractional multiple measured interval

Thus, I expected to find developmental levels in tool-target coordinations to vary by students' abilities to unitize tools and targets and their coordination between the two in measuring activity. These levels were conceptualized into a coding scheme, which is introduced in the following section along with observations of students' coordination levels.

Analysis of Students' Tool-Target Coordination Levels: A Focus on Unitizing

Recall from the conceptual framework presented in Chapter 1 that students must generate tool-target coordinations via co-quantification of units to adequately estimate the position of numbers on the presented number lines. In this section, I describe how I utilized the framework to examine students' tool-target coordination levels in coordinating between unitization of tools and unitization of targets in measurement activities. Here, I introduce my ordinal coding scheme and present observations of students' levels that illustrate different unitizing coordinations between tools and targets. As I argued earlier, a focus on coordinated unitizing activities is central to the conceptual development of measuring: a linear measurement tool can serve as a useful measure for a target only when individuals conceptualize a correspondence between a unit on a tool and a unit on the target—that is, the individual implicitly quantifies tool and target with corresponding units.

To characterize students' tool-target coordination levels into conceptually different categories, I identified each problem-solving observation as a display of one of three different levels. The three-level tool-target coordination coding scheme is presented in Figure 22. The coordination levels were ordered in sophistication of unit coordination between tool and target object. At the lowest level (Level 1), students did not coordinate units on the target number line with units on the presented tool to produce estimates, at the middle level (Level 2), students partially coordinated units on the target number line and their tool, and at the highest level (Level 3), students adequately unitized and coordinated the target number line and their assigned tool to produce numerically coherent estimates. Though students' coordination levels varied somewhat by tool type and problem type, their coordination levels for each problem were still consistently categorized into these three levels with high reliability[8].

LEVEL	DESCRIPTION OF COORDINATION	EXAMPLE
Level 3	Target and tool are each unitized, with units on the tool corresponding to units on the target.	 *Given 10 and 14 marked on the number line, student places his estimate for 16 on the number line. He measures the given interval as 4cm long, correctly stating that 1cm on the ruler is equivalent to 1 unit on the number line. He measures 2cm to the right of 14 to place his estimate correctly on the number line (since 16 is 2 more units than 14).*
Level 2	Tool is unitized, and target may or may not be unitized. Units on the tool are not always coordinated to units on the target. On the number line, the *given interval* is left unmeasured or not explicitly stated as the source of target unitization; the estimate is produced using units afforded by the tool, not coordinated with a unitization of the given interval on the number line.	 *Student unitizes a given interval of 6 to 8 as being 2 units long as he fits his ruler to measure 4cm for the interval. Then he states he will measure 3 more units from 8 to place 11, but uses 4cm for each linear unit instead of 2cm.*  *Given 4 and 6 marked on the number line, student places her estimate for 9. She does not fit her tool into the given interval, and instead treats her paper strip to represent a linear unit of 1 on the number line. She iterates her paper strip 3 times from 6 to place 9 on the number line.*
Level 1	Tool is rejected, or the tool is not used to produce an estimate on the target, or the estimate is produced from displacing the tool on the target like a counter, but not through coordinating between tool units and target units.	 *Short paper strip is moved along the number line, but with inconsistently sized displacements for each linear unit on the number line.*

[8] An overall percent agreement of 94.72% was established by three coders (myself and two undergraduate assistants) for 24 videos (8 videos were randomly chosen from each tool condition, balanced by grade and gender), which was 32% of the coded video data corpus.

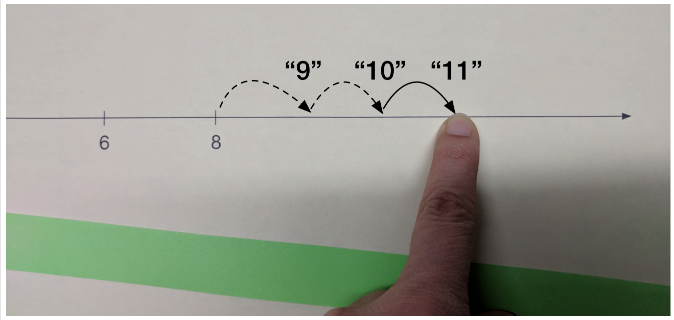
		 *Long paper strip is just placed at the bottom of the problem sheet and with no attempt to use it to produce a measure of the target. Student uses her index finger to hop on the number line as she counts up to her target.*

Figure 22. Tool-target coordination scheme

At Level 1, students did not coordinate tool units with target units to produce a linear measure; the tool and target were not co-quantified to corresponding units. Level 1 coordination occurred in all three tool conditions and problem types in two distinct ways.

First, some students did not attempt to use the assigned tool in any way to produce a linear measure of the target. For example, students placed the tool nearby the number line, but did not use it to physically or verbally coordinate with linear units. Instead, students used their fingers to hop and count numbers directly on the number line or eyeballed a location to position their target value for the problem (Figure 23). Their explanations included statements such as, “I just counted and found my answer.” At this level of performance, some students also stated that they could not use the provided tool to solve the problem because they did not know how to use it.

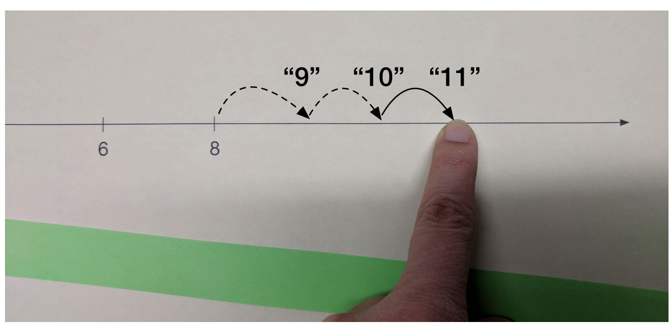


Figure 23. Level 1 example: Student placed long paper strip to the side and solved the number line problem with their finger

Second, some students used the tool on the number line to produce an estimate, but the tool only supported counting activity, not an activity that made use of a consistent linear unit; tool units were not coordinated to target units to represent linear measures. In the case of paper strip conditions, the estimate was produced from inconsistently-sized displacements of the tool as the student counted numbers, resulting in linear units of different lengths (i.e., unequal distances between consecutive numbers). For example, students fitted a short paper strip in the given interval between 6 and 8 on the number line, and moved it three times with inconsistent displacements to the right as they counted up three consecutive numbers from the given interval to position 11 on the line (Figure 24). For ruler conditions, students arbitrarily placed the ruler underneath the number line (sometimes aligned to the left edge of the number line) and using their finger to keep track, counted each numbered centimeter mark from 1 to the target value (Figure 25). In this type of Level 1 coordination, students used the tool to track their count of a number sequence on the number line instead of representing and preserving a co-quantified and consistent linear measure in coordination with the number line.

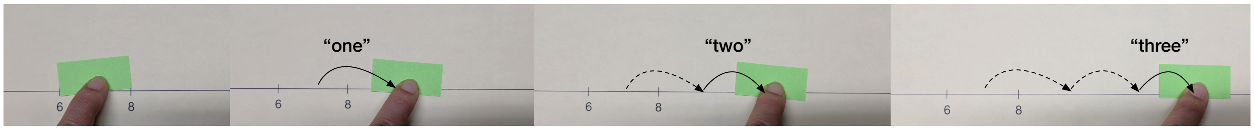


Figure 24. Level 1 example: Student used a short paper strip to support counting activity on the number line

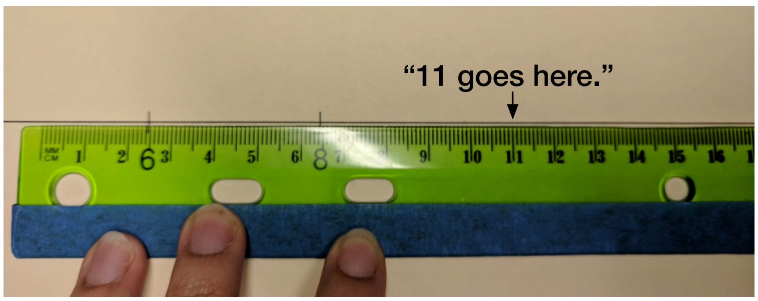


Figure 25. Level 1 example: Student placed the ruler on the problem sheet, but focused only on counting the centimeter marks on the ruler to provide an estimate on the number line

At Level 2, students only partially coordinated their construction of tool and target units in two distinct ways.

First, in some cases, the given interval was left unmeasured (not visibly unitized at all); students appeared to produce estimates using units afforded by the tool without respecting units constructed on the target by the given interval. For instance, without measuring the given interval or acknowledging the interval values, a student arbitrarily decided that the length of the short paper strip represented 1 linear unit and iterated the paper strip to estimate the position of the target value. This meant that though students were constructing units on the target number line using their assigned tool, their unitization of the tool was uncoordinated with linear units of the target. As a result, students' solutions were generally incorrect. When students used this method to produce estimates, short paper strips were always assumed to be the length of a unit of 1, and 1cm was most often assumed to be the length of a unit for rulers (see Figure 26 and Figure 27). For long paper strips, students most often placed their finger on the paper strip about a finger pinch away from the left edge, and assumed that length as a unit. For example, on a number line with 4 and 6 marked for the given interval, a student first arbitrarily decided that their short paper strip represented a linear measure of 1 without measuring the given interval. Then, starting from 6 on the number line (outside of the given interval), they used the short paper strip to position 7, 8, and then 9—the target value—three iterations to the right of 6 on the number line. This solution was incorrect, since the paper strip was actually equivalent to the length of 2 linear units on this number line problem.

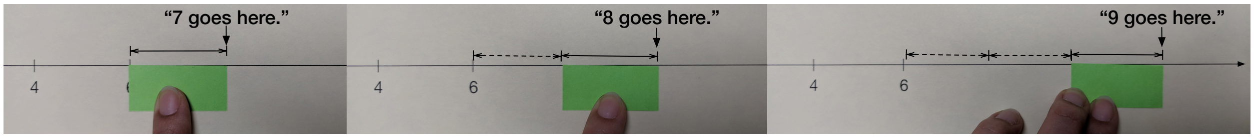


Figure 26. Level 2 example: Student ignored the given interval and used the short paper strip as a unit to measure the target value's position

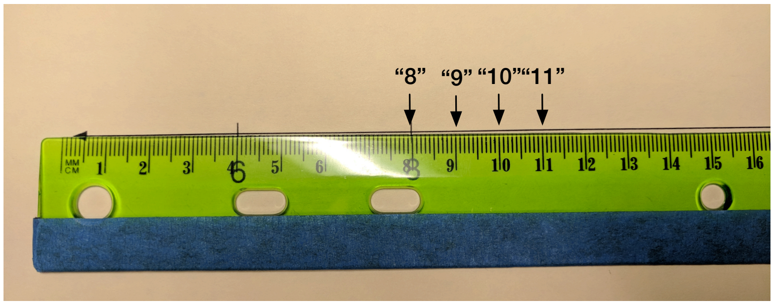


Figure 27. Level 2 example: Student aligned an enumerated centimeter mark (8cm) to the multiunit interval of two units on the number line (to tick mark for 8 on the line) and then counted each centimeter thereafter for estimates of 9, 10, and 11

Students who unitized their tool without respect to units on the target and assumed an ad hoc unit for their tool did happen to produce a correct estimate for the location of a target number in two cases, one for the unit interval problem (problem 1) in the short paper strip condition and the other for the four-unit interval problems (problem 4 and problem 5) in the ruler condition. When students assumed a short paper strip represented 1 linear unit on a unit interval problem (e.g., 7 and 8 marked on the line) without measuring the given interval, students' estimates were still correct, because the length of the short paper strip was actually equivalent to the given interval. When students assumed 1cm on the ruler represented 1 linear unit on a four-unit interval problem (e.g., 10 and 14 marked on the line) without measuring the given interval, students also happened to be correct in their estimates because 1cm did equal the length of a unit on the number line. However, students' coordination levels in these two cases were still coded as Level 2 if they did not visibly measure the given intervals and they used level 2 coordinations consistently on other problems.

Second, in other cases of level 2 coordinations, students displayed difficulty in coordinating scales of the tool and target units as the same size. Students unitized both the tool and the target, but the size of the linear unit constructed on the tool differed from the size of the linear unit constructed on the target, leading to incorrect estimates for target values. For example, some students stated that the given interval between 6 and 8 on the number line was 2 linear units, and fit the ruler on the target to measure the given interval as 4cm long. Then they proceeded to explain that 11 should be measured 3 linear units to the right of 8 on the number line (which was correct). However, instead of corresponding 4cm on the ruler to 2 linear units on the number line, they treated 4cm to represent 1 linear unit on the number line, and iterated 4cm three times to the right of the given interval. This resulted in the estimate of 11 being placed 12cm—instead of 6cm—to the right of 8 on the number line (Figure 28). This type of level 2 coordination in paper strip conditions were similar in principle; students unitized the given interval and fit their paper strips within it, but then represented a differently-sized unit on their tool to measure their estimate for the target values.

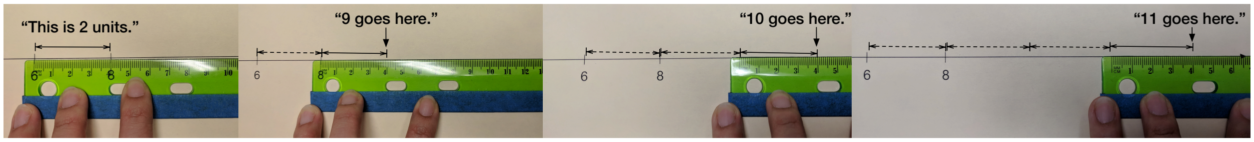


Figure 28. Level 2 example: Student measured a two-unit interval as 4cm long, but then used 4cm to represent each additional unit from the given interval

At Level 3, target and tool were each unitized with units on the tool corresponding to units on the target; students consistently generated, represented, and preserved linear measures on the target number line using the provided tool, achieving more adequate coordinations than those produced at levels 1 and 2. For example, some students measured the given interval between 10 and 14 on the number line as being 4cm long on their ruler, and then stated that one linear unit on the number line was 1cm long; the initial challenge of measuring the given interval with the tool was resolved. Then they continued to represent one linear unit as 1cm on their ruler to unitize the rest of the number line and positioned the target value of 16 two centimeters to the right of 14 on the number line (since 16 was two 1cm-long linear units greater than 14) (Figure 29); the tool and target were unitized on the same scale throughout the measuring activity. Another example for the same problem was students who first folded the edge of their long paper strip to fit the length of the given interval. Then they tore off the folded edge, and folded the smaller piece into fourths to represent 4 linear units. This accordion-like piece was then placed to the right of the given interval, and students pointed to the second fold (the halfway mark of the paper strip) for their estimate for the target value (Figure 30). Students who used this approach also continued to consistently coordinate a linear unit on the target number line with a unit on their tool on the same scale. Students who used a level 3 coordination with the short paper strip on this problem also represented a linear unit on a quarter of their paper strips.

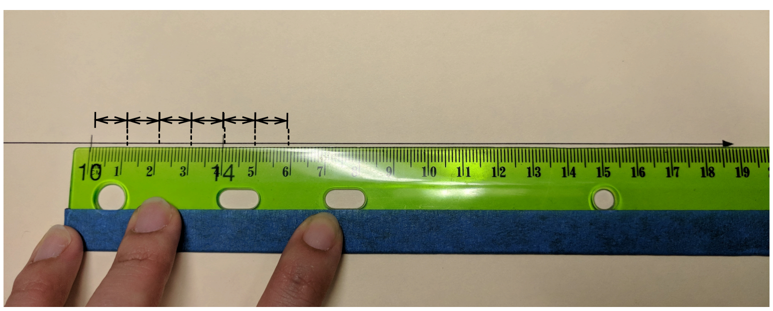


Figure 29. Level 3 example: Student coordinated 1cm on the ruler with 1 linear unit on a number line with four-unit given interval to position 16 on the number line

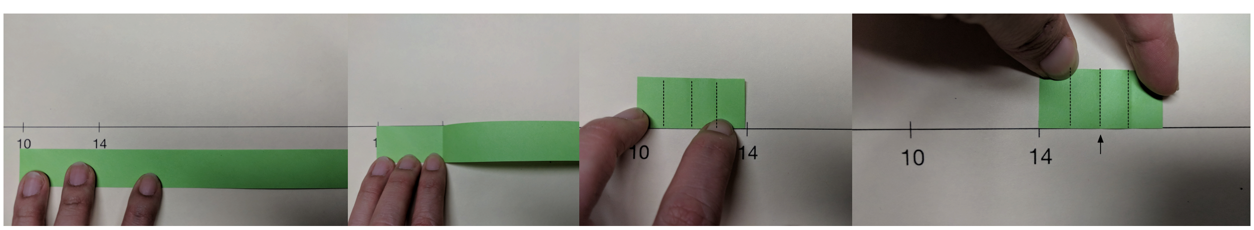


Figure 30. Level 3 example: Student folded the long paper strip to fit the given interval, and then tore off that segment. Then the shorter segment was folded into quarters to coordinate 1/4 of the strip to 1 linear unit on the number line to position 16 on the line

Classifying my observations into these three coordination levels helped demonstrate the distinct unitizing challenges students encountered when adapting tools into their measuring activity. Students in both grades displayed all levels in all tool conditions, suggesting that both younger and older children experienced similar tool-target coordination challenges. However, I noted that some levels were more prominent than others for particular grades and tool conditions, and that for many students, the levels that they manifested varied with problem type. To understand students' developing tool-using behavior as a function of tool condition and problem types, I analyzed systematically students' level of tool-target coordinations as an interaction between tool type and problem type for both second and fourth graders, and I present these analyses in the next section.

Analysis of Unitizing in Tool-target Coordinations: A Focus on Development

In this section, I describe the analyses of students' levels in coordinating between unitization of tools and unitization of targets in measurement activities. The goal of these analyses is to better illuminate when and how particular tools or problem features interacted with students' efforts to generate linear estimates with the use of tools. The considerations presented in Chapter 1 led me to expect that individuals' abilities to create tool-target unitizing coordinations are dependent upon the material properties of the tool on the one hand (e.g., can the tool be folded into half?), and the properties of the target on the other (e.g., can the marked interval on the number line be split into half to generate another whole number?), and that students' coordination levels would vary by age. To evaluate the merit of these expectations, I analyzed students' tool-target coordination levels as a function of tool conditions, problem types, and grade levels. My analyses focus on the affordances that particular tools offer students and the way these affordances may vary by problem types and grade levels. More specifically, I analyzed whether particular tools and problem types led students to support more sophisticated tool-target coordinations and the challenges that particular tools and problem types posed for students' tool-target coordinations.

To study whether there were developmental shifts in how children used tools in their problem solving activity, I built upon frameworks articulated by Vygotsky (1978, 1986) on the development of tool use and Saxe (2012) on the microgenesis of tool using activity. The frameworks informed my study design and analyses of my results.

Following Vygotsky, I expected the presented tool's function to vary over students' development. Early in development (e.g., most second graders and perhaps some fourth graders), students should not adapt tools to serve measurement functions (and perhaps even interfere with perceptual approaches to estimating linear distance). Later in development (e.g., some second graders and some fourth graders), students should adapt tools in overt activity to begin to serve measurement functions as students begin to use a tool to organize linear estimates in measuring activities. Still later in development (e.g., most fourth graders), students' tool using activity may become fluid and even used in mental acts of measurement.

Following Saxe (2012), I expected to find, complementing Vygotsky's treatment of tool using, students to vary in their microgenetic quantifications in their tool-target coordinations, though these coordinations would vary with the affordances of tools and problems. For students early in their development, I expected students to be unable to coordinate between tool units and target units, and thus presented tools would often not be used to serve measurement functions in students' task performances, with variations expected by problem type and tool type. For intermediate students, I expected students to partially coordinate between tool units and target units; in other words, students should begin to use tools to serve a functional value in their measurement activity on a broader range of problems, and at least partially quantify the tool or target in tool-target coordination efforts. For the most advanced students, I expected students to no longer experience any difficulty in coordinating between tool units and target units, and that this coordination would be consistent across problems and tool types.

Distribution of tool-target coordination levels across conditions.

In the first step of my analysis of students' coordination levels, I examined broad patterns in how coordination levels were distributed across conditions. In other words, how did students' tool-target coordinations vary for each grade, tool type, and problem type? The approach entailed examining which level of tool-target coordination older students used more (or less) compared to younger students, and whether this pattern was the same or different between tool groups and problem types. For instance, I wondered, did a particular tool afford higher level coordinations compared to other tools? Or did certain problem features elicit higher level coordinations than others? Was this different for fourth graders versus second graders? Figure 31 displays bar charts that represent the percent distribution of students' coordination levels for each problem type as a function of their grade and tool condition.

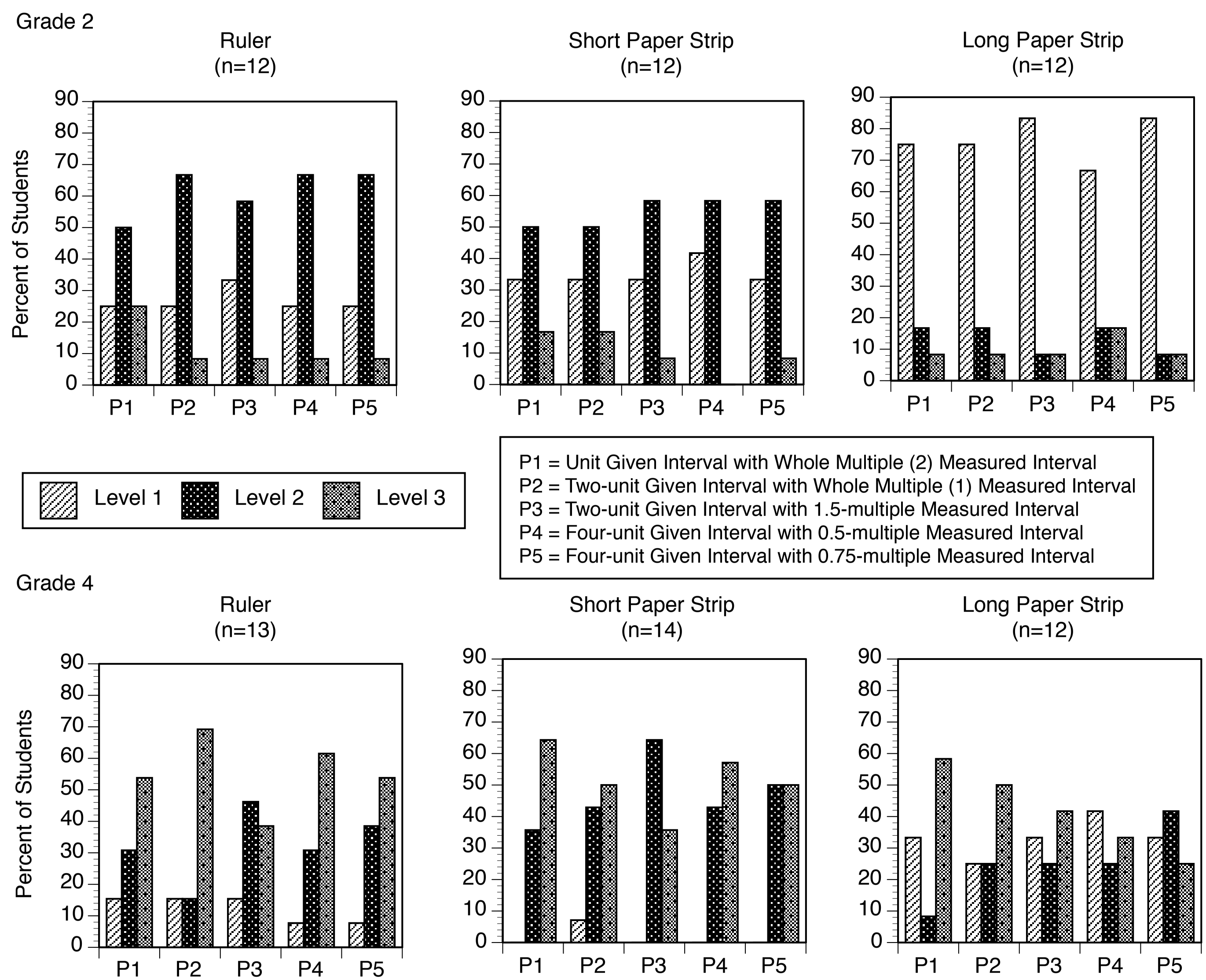


Figure 31. Percent of students' tool-target coordination levels by grade, tool type, and problem type

To determine whether grade level, tool condition, and/or problem type influenced the levels students displayed, I conducted a three-way repeated measures analysis of variance (ANOVA) with students' coordination levels (a score of 1, 2, or 3) on each problem as the dependent variable: 2 (grade: second, fourth) × 3 (tool type: ruler, short paper strip, long paper strip) × 5 (problem type, repeated).[9] The ANOVA revealed a significant difference in tool-target coordination levels based on tool type ($F (2, 69) = 4.51$, $p < 0.05$) and grade level ($F(1, 69) =27.63$, $p < 0.01$). There was no problem effect within groups ($p=0.082$), and there was no interaction effect between tool types and grade levels on tool-target coordination levels ($p=0.833$). Figure 32 displays means of tool target coordination levels by grade and tool type. At fourth grade, means for each tool type were above level 2 (Ruler: 2.43, Short paper strip: 2.50, Long paper strip: 2.08), and at second grade, means for each tool condition were below level 2 (Ruler: 1.85, Short paper strip: 1.75, Long paper strip: 1.33). Scheffe's post hoc tests on tool conditions revealed that there was a significant difference in means between the long paper strip condition and ruler condition (Scheffe $p = 0.030$), and a significant difference between the long paper strip condition and the short paper strip condition (Scheffe $p = 0.027$). Means for the ruler condition and short paper strip condition were higher than the long paper strip condition.

Grade	Tool Type	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
Second Grade	Ruler	1.850	.165	1.522	2.178
	Short Paper Strip	1.750	.165	1.422	2.078
	Long Paper Strip	1.333	.165	1.005	1.662
Fourth Grade	Ruler	2.431	.158	2.115	2.746
	Short Paper Strip	2.500	.152	2.196	2.804
	Long Paper Strip	2.083	.165	1.755	2.412

Figure 32. Descriptive statistics for tool-target coordination levels by grade and tool type

I now consider more carefully the frequency distributions contained in Figure 31 in order to gain an understanding of the patterns of tool-target level shifts that were the source of grade level and tool type effects. Figure 31 reveals that second graders primarily used level 1 and level 2 tool-target coordinations with level 2 being the mode for two out of the three tool types (ruler and short paper strip). In contrast, fourth graders primarily used level 3 coordinations as their modal level across the tool types. The exception to students' modal level patterns was the long paper strip condition. At second grade, the modal level for the long paper strip condition was level 1, not level 2. At fourth grade, though the modal level for three out of five problem types was level 3, many students used level 1 coordinations in the long paper strip condition and rarely used level 1 coordinations in the other tool conditions.

At second grade, the mode of coordination levels for each problem was level 2 for ruler and short paper strip conditions while the mode of coordination levels for each problem was level 1 for the long paper strip condition. Fifty percent or more of the second grade students in the ruler and short paper strip conditions used a level 2 coordination for every problem type. In the long paper strip condition, however, a much larger majority (a range of 66.7% to 83.3%) of

[9] Two assumptions were violated in conducting this ANOVA: My dependent variable (coordination level) was an ordinal measure and not a continuous one, and my sample did not display a normal distribution. However, ANOVAs are still robust with respect to the violations of these assumptions, which is why this method was used.

second graders used a level 1 coordination compared to a much smaller percentage of students who used a level 2 coordination (a range of 8.3% to 16.7%) for every problem type. Second graders were frequently able to partially coordinate between tool units and target units for the ruler and short paper strip conditions across problems, but experienced more difficulty in adapting long paper strips into their measurement activity.

At fourth grade, though the overall mode for each tool group was level 3, there were still problems for which the mode was not level 3 in each tool condition, and the next most frequently used tool-target coordination level was not always level 2. For the ruler and short paper conditions, the mode was level 2, not level 3, for the two-unit (multiunit) interval problem with 1.5-fractional multiple measured interval (problem 3), and for the four-unit interval problem with 0.75-fractional multiple measured interval (problem 5). While 46.2% ruler using students and 64.3% short paper strip using students used a level 2 tool-target coordination on problem 3, only 38.5% of ruler using students and 35.7% of short paper strip using students used a level 3 coordination. For the long paper strip condition, the modal level for problem 5 was level 2, and the modal level for the four-unit interval problem with 0.5-fractional multiple measured interval (problem 4) was level 1, which was the second most frequently used coordination level for the rest of the problem types (the frequency was the same for level 1 and level 2 for the two-unit interval problem with whole multiple measured interval (problem 2)). Like second graders, fourth graders also had more difficulty adapting the long paper strip into their measurement activity compared to the two other tools as evidenced by the three different modes of coordination levels across the problem types. Additionally, problem 3 challenged fourth grade students' coordination efforts more than other problem types for the ruler and short paper strip conditions, while problem 4 and problem 5 challenged coordination efforts of fourth graders the most in the long paper strip condition.

The ANOVA and frequency distribution findings show that with age, students coordinated tool units and target units better across conditions; fourth graders were better able to adapt tools to serve as measuring tools in their measurement activity compared to second graders by coordinating their tool units with target units. However, fourth graders still continued to have problems with tool-target coordinations, as evidenced by the presence of level 1 and level 2 coordinations across tool conditions, and coordination level modes that were not always level 3 for every problem. Additionally, long paper strips were less accessible for students in both grades to use effectively as measurement tools to support their problem solving compared to rulers and short paper strips.

Interrelationships between tool-target coordination levels by problem types.

To better understand to what extent students' tool-target coordinations developed in concert across problems within each of the three tool conditions, I generated a correlation matrix of coordination levels between problems for tool conditions using Spearman's rho (r_s) (Figure 33). There were significantly positive correlations in coordination levels between problems for all tool conditions (α=0.05) except for between problem 1 (P1: unit given interval with whole multiple measured interval) and problem 3 (P3: multiunit given interval with 1.5-fractional multiple measured interval) in the ruler condition. With the exception of this one non-significant correlation, a student's coordination level on one problem was generally similar to their coordination level on another problem, regardless of grade level or tool condition. More than half of the inter-problem relationships in the ruler condition were moderate correlations ($0.39 < r_s < 0.60$), while the rest were stronger. In contrast, all of the short paper strip correlations were strong ($0.59 < r_s < 0.80$) or very strong ($0.79 < r_s$), and this was the same for the long paper strip students with the exception of one moderate relationship.

Ruler

	P1	P2	P3	P4	P5
P1		.433*	0.387	.708**	.581**
P2			.538**	.708**	.581**
P3				.577**	.428*
P4					.806**
P5					

Short Paper Strip

	P1	P2	P3	P4	P5
P1		.843**	.737**	.803**	.715**
P2			.837**	.807**	.735**
P3				.773**	.799**
P4					.777**
P5					

Long Paper Strip

	P1	P2	P3	P4	P5
P1		.741**	.604**	.764**	.562**
P2			.861**	.729**	.835**
P3				.650**	.822**
P4					.620**
P5					

Note. *p<.05, **p<.01

Figure 33. Correlations of tool-target coordination levels between problems by tool type

Students' tool-target coordination performance on the multiunit interval problem with 1.5-fractional multiple measured interval (problem 3) was less consistent with their performance on other problems in the ruler condition, especially with the unit interval problem (problem 1). Recall that unitization coordinations differed between problem 1 and problem 3, and the ruler may have posed additional challenges. For problem 3, students needed to use their tool and split a two-unit given interval into half to generate a linear unit and measure three units to place the target value (e.g., Given (6, 8) place 11). With the ruler, this involved measuring the given interval as 4cm long, and recognizing that since 2 linear units corresponded to a length of 4cm, 1 linear unit corresponded to a length of 2cm. This relationship could then be used to measure 3 more linear units as 6cm to place the target value. In contrast, for problem 1, the given interval was already a linear unit, so no splitting actions were necessary for students to measure 2 more units for the target value (e.g., Given (5, 6) place 8). Only one tool-target relationship needed to be kept consistent between the ruler and number line, that one linear unit corresponded to 4cm, which could be used to coordinate a measure of 2 units as 8cm to place the target value. To better understand the nonsignificant correlation between problem 1 and problem 3 for ruler condition students, I created a contingency table to analyze the crosstabulation of coordination levels for the ruler condition (Figure 34). Out of 25 students in the ruler condition, six students performed at higher coordination levels on problem 1 than problem 3 while two students performed at higher coordination levels on problem 3 than problem 1. Thus, coordinating tool-target relations on problem 3 using the ruler was more challenging for more students than coordinating tool-target relations on problem 1 using the ruler.

	Problem 3 Tool-Target Coordination Levels			
Problem 1		Level 1	Level 2	Level 3
Tool-Target	Level 1	3	1	1
Coordination	Level 2	1	9	0
Levels	Level 3	2	3	5

Figure 34. Count of ruler students' tool-target coordination levels on problem 1 and problem 3 (N=25)

The presence of more moderate correlations between problems in the ruler condition suggests that relatively more students vacillated in their coordination levels when using the ruler but not the paper strips. One possible source of this vacillation was that it was not uncommon for an individual student to mix two different methods for measuring a linear unit with their ruler on some, but not all of the problems. Even when students initially generated a correct tool to target unit coordination on their ruler (e.g., measuring a two-unit interval as 4cm, and splitting it in half to generate 1 linear unit as 2cm), sometimes they switched to an incorrect tool-target relationship afterwards to estimate the target value. Instead of using the zero-point (0cm) of the ruler as the beginning of their iterated measure, they started with the left edge of the ruler. The zero-point of the ruler was off-set from the left edge of the ruler by 3mm, so this method reflected an inconsistent coordination of a linear measure during problem solving activity (Figure 35). Students who used a paper strip to estimate linear measures did not use this type of mixed approach.

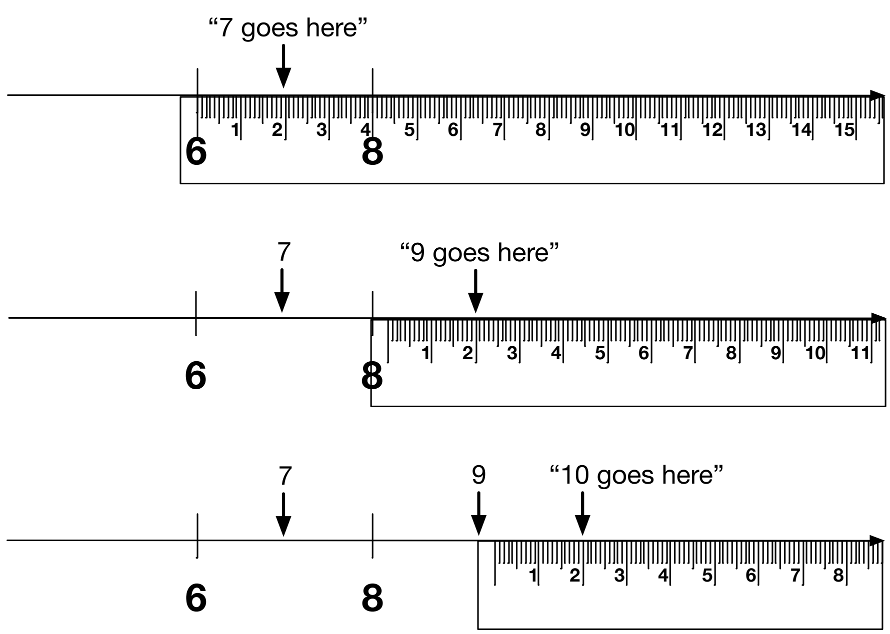


Figure 35. Example of student initially using zero-point of ruler to measure unit and then the edge of ruler as the start of the iterated measure

Another source of the ruler using students’ relatively inconsistent coordination levels across problems was that some students used a higher-level coordination (level 2 or level 3) on some problems, but then only used perceptual features of the ruler to count a number sequence on other problems (level 1). This method involved placing the ruler anywhere on the line, and counting the enumerated centimeter marks from 1 to the target value to place their estimate.

Thus, though the sample sizes were very small to make any strong claims, it is possible that some features of the ruler (enumerated tick marks or off-set origin, which the paper strips did not have) and problem 3 (coordinating splitting and iterating actions on a tool with such features) presented unique unit coordination challenges for students who used rulers. This may be evidence for a transitional period in which students were still learning to establish a consistent relationship between their tool and target with the use of a ruler.

In sum, the ANOVA results, frequency distributions, and the inter-problem correlations show that fourth graders used more sophisticated tool-target coordination levels than second graders; that stated, not all tools were consistently accessible to students regardless of grade level. The long paper strip was much more difficult for both grade levels to support linear measurement activity compared to rulers and short paper strips. Across problems, students were generally consistent in the level of tool-target coordination they used, but the ruler condition presented relatively more coordination challenges than the paper strip conditions.

In the following section, I focus on relations between solutions with tools and without tools. To this end, I analyze how students' precisions of estimates were influenced by the introduction of tools.

Analysis of Tool Efficacy on Precision of Estimates: A Focus on Development

The prior analyses on developmental shifts in tool-target coordination levels illuminated how children qualitatively transformed their use of tools to mediate their linear measurement activity. I now shift to a focus on whether children's use of the introduced tools influenced the precision of their measurements as contrasted with their problem solving when they did not have access to those tools.

Following Vygotsky's (1978) treatment of a "natural line" of development (unmediated) and a "cultural line" of development (mediated), I expected that the shift of students' precision of estimates would be different from the tool absent to the tool present conditions as a function of age. For students early in the development of the use of tools to mediate estimating distances (e.g., most second graders), I expected the introduced tools to either interfere with or not serve a functional value in students' measurement precision; rather, estimates would be rooted in unmediated perceptual reactions to measurement questions (Vygotsky's natural line), resulting in imprecise measurements; further, for these students, the introduction of tools would either lead to greater imprecision (the tools serving to confuse the situation) or no change in precision. In terms of Saxe's (2012) treatment of microgenesis in tool using, I did not expect to see students engage with a unitization of tool or target in a tool-target coordination. For intermediate students, I expected tools to begin to serve a value and help improve students' measurement precision; these students would show early efforts to appropriate tools to advance their measurement activity, leading to greater precision of linear estimates. For the most advanced students, I expected precision scores to improve with the use of tools, and further, for students to no longer be dependent upon tools for precision when tools were not available, using internal meditational forms rooted in prior tool using activity; these students would be able to mediate their linear measurement activity both with introduced tools and mentally (without tools).

Based upon my prior analyses on tool-target coordination levels (building on Saxe's (2012) microgenetic framework), I expected that the particular tool conditions would also interact with students' precision. Specific tool features might interfere or support students' mediation of estimation activity in different ways. I expected that the short paper strip would better support students to mediate their unitization activity on the number line because its size (which was equivalent to the given intervals on the problems) afforded the iteration of a measure without an additional operation on the tool itself (splitting it or treating it as a multiple). However, students who were only able to partially coordinate tool-target relationships would fail to recognize that the paper strip could be folded to generate shorter measures to afford better precision. The ruler would also be more readily adapted into students' measurement activity because students are familiar with its use and purpose from classroom instruction, but its pre-labeled enumerated units and off-set zero-point might interfere with students' efforts to use the ruler to mediate their estimates. The long paper strip would serve the least functional value to students because it would be unclear to students on how to transform a tool that looks similar to a ruler (it is the same length) but has no demarcations to guide measurement. This confusion would lead to more unmediated and imprecise estimation activity in the long paper strip condition compared to the other tools. Thus, I expected that there would be fewer advanced students and more unmediated, imprecise estimations in the long paper strip condition compared to the short paper strip or rulers, and relatively more intermediate and advanced students in the ruler and short paper conditions, and hence, greater precision in estimates compared to the long paper strip condition.

Measurement of precision scores.

I adapted Siegler and Booth's (2004) method[10] for calculating the precision of students' linear estimates on the number line problems. This captured the difference between students' measures and the ideal measure scaled by the ideal measure. The closer a precision score was to 0, the more precise the student's estimate was for the problem. The absolute value was taken because I was concerned with the magnitude of students' imprecision, and to avoid using negative values when comparing means of students' precision scores across conditions.

$$Precision\ Score = \left|\frac{Student's\ Estimate - Ideal\ Estimate}{Ideal\ Estimate}\right|$$

To measure students' estimates, I used the same ruler students used in their interviews to measure the physical distance between the student's estimation mark and the rightmost endpoint[11] of the given interval to the tenth decimal place in centimeters. These measures were recorded in a database, categorized by each individual student and their solutions for each problem they solved. The ideal estimate was the physical distance at which the target value of a problem should be placed to the right of the rightmost endpoint. For example, for a problem with

[10] Siegler and Booth's (2004) Percent Absolute Error (PAE) formula was used for bounded number line problems for which students were asked to estimate the location of a target value *between* two pre-labeled values. Thus, their formula utilized the actual values presented in the problem (pre-labeled values, target value, value of student's estimate). Since this study's problems involved unbounded number line problems for which students were asked to estimate the location of a target value *outside* of two pre-labeled values, I used the physical distances of numbers on the number line to produce precision scores.

[11] The rightmost endpoint is the larger value of the given interval in the number line problem. For example, 6 is the rightmost endpoint, or larger value of the given interval (5, 6), and 8 is the rightmost endpoint of the given interval (6, 8).

a 4-centimeter given interval of (5, 6), the ideal estimate for target value 8 was 8 centimeters because 8 should be placed two 4-centimeter units to the right of 6—the rightmost endpoint of the given interval—on the number line. Thus, if a student's estimate for this problem was 8cm, then their precision score would be 0 (precise). However, if the student's estimate was longer (overestimate) or shorter (underestimate) than 8cm, then their precision score would be greater than 0 and considered imprecise.

Unexpected tool rejections.

An unanticipated finding that needed to be considered before analyzing precision scores was that a number of students claimed that they could not or did not know how to use the presented tool even when encouraged to do so. Some students rejected the tool for every problem while others rejected the presented tool for a subset of the problems (Figure 36 and Figure 37). At second grade, at least some students rejected the presented tool in every tool type; most rejection instances occurred in the long paper strip condition (66.67% rejecting the tool on every problem and 8.33% rejecting the tool on some of the problems). At fourth grade, only some of the students in the long paper strip condition rejected the tool (16.67% rejecting the tool on every problem and 33.33% rejecting the tool on some of the problems). No tool rejections occurred in the ruler or short paper strip conditions at fourth grade. Students' rejection of the tools itself is a phenomenon consistent with Vygotsky's treatment of tool using. Early in development according to Vygotsky, students are expected to rely on intuitive or perceptual processes regardless of the availability of or encouragement to use a tool. Thus, the finding of the prevalence of tool rejection among second but not fourth graders is a phenomenon that would be consistent with Vygotsky's expectations. Further, the rejection by fourth graders of the long paper strip points to the lack of its affordances for mediating problem solving for these participants.

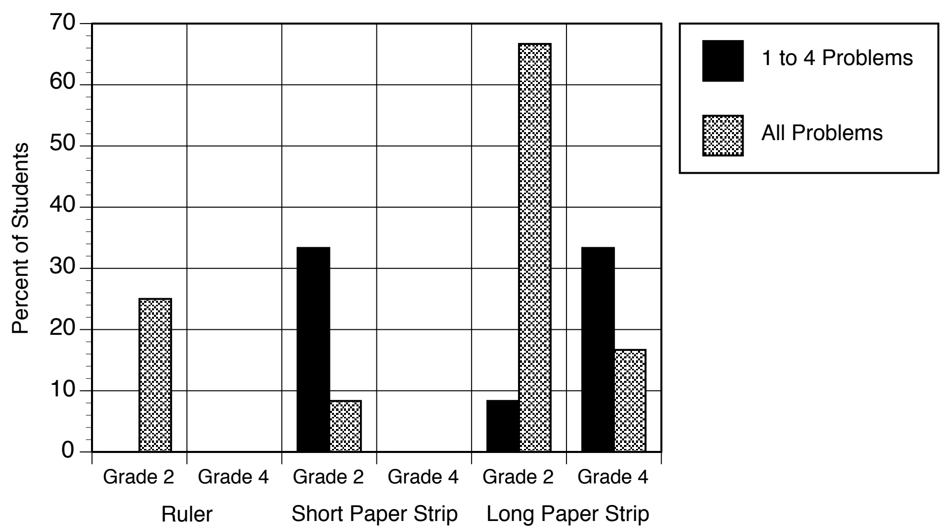


Figure 36. Percent of students who rejected the presented tool for some problems (1 to 4) and all problems (5) as a function of tool type and grade level

		Rejected Tool for Some Problems	Rejected Tool for All Problems
Ruler	Second Grade (n = 12)	0 (0%)	3 (25%)
	Fourth Grade (n = 13)	0 (0%)	0 (0%)
Short Paper Strip	Second Grade (n = 12)	4 (33.33%)	1 (8.33%)
	Fourth Grade (n = 14)	0 (0%)	0 (0%)
Long Paper Strip	Second Grade (n = 12)	1 (8.33%)	8 (66.67%)
	Fourth Grade (n = 12)	4 (33.33%)	2 (16.67%)

Figure 37. Number of students who rejected the presented tool for some problems (1 to 4) and all problems (5) as a function of tool type and grade level

Whether or not students rejected the use of the presented tool, students produced estimates for each problem in the tool using conditions. When students rejected the tool for the tool present condition problems, their solutions could not be construed in any direct fashion as mediated by the tool. That said, the mere presence of the tool could have conceivably influenced students' performance even if students rejected the tool. I chose to proceed with my analyses of students' precision scores as a function of tool absent versus tool present conditions in three ways that allowed some accommodation for students' unanticipated tool rejections, treating the interpretation of each analysis cautiously and in light of the others. Each method carries with it different assumptions about students' precision scores in the tool present condition when they had rejected the use of the tool.

Treating tool rejecting students' precision scores as those that were influenced by the presence of the tool.

For this analysis, I proceed as initially planned in my experimental design. I assumed that students' precision scores in the tool present condition were valid data whether students did or did not reject the tool. This analysis assumes that the mere presence of the tool influenced tool-rejecting students' precision estimates in the tool present condition.

To determine whether grade level, tool type, or tool presence influenced the precision of students' estimates between tool absent and tool present conditions, I conducted a three-way repeated measures analysis of variance (ANOVA) with students' sum of precision scores across problems as the dependent variable: 2 (grade: second, fourth) × 3 (tool type: ruler, short paper strip, long paper strip) × 2 (tool presence: tool absent, tool present, repeated). Figure 38 contains student mean precision scores that were revealed in this analysis. The ANOVA revealed main effects for grade, tool type, tool presence, and an interaction between grade and tool presence.

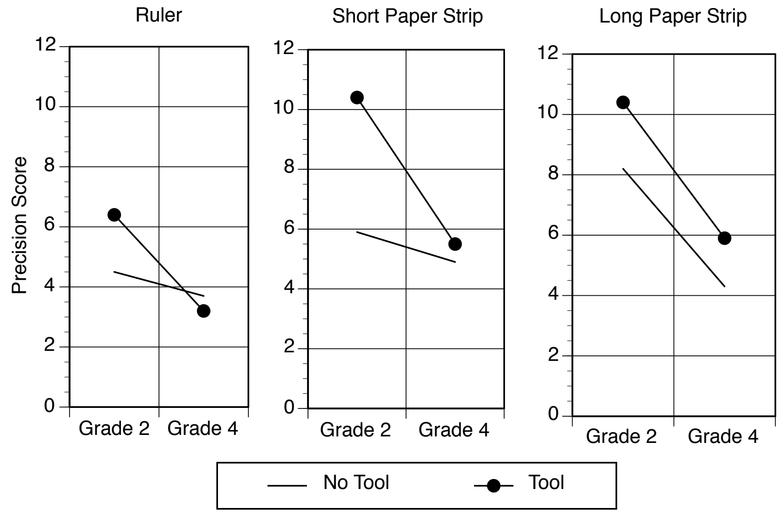


Figure 38. Shifts of precision score sum means (0 = precise) for no tool (tool absent) condition and tool (tool present) condition by grade level and tool type

The main effect for grade revealed that overall second graders were less precise than fourth graders (F (1, 69) =11.77, $p < 0.05$). Second graders' means were higher (less precise) than those of fourth graders (Second grade: 7.65, 95% CI [6.37, 8.94]; Fourth grade: 4.58, 95% CI[3.35, 5.82]).

The main effect for tool type revealed that students were less precise on at least one tool type compared to another (F (2, 69) = 3.53, $p < 0.05$). A post hoc Scheffe's test showed that mean for students in the ruler condition differed significantly from the mean for students in the long paper strip condition (Scheffe $p = 0.048$). Students in the ruler condition displayed lower means (more precise) than long paper strip condition students; the mean difference in precision score sums between the ruler group and long paper strip group was -2.78 (95% CI [-5.55, -0.017]). Compared to rulers, long paper strips were less effective in supporting students to mediate their measurement activity to become more precise. There were no interaction effects between groups.

The main effect for tool presence revealed that on average, students became significantly impaired in their precision scores with the introduction of tools (F (1, 69) = 10.46, $p < 0.01$). From tool absent to tool present conditions, the mean of students' precision score sums increased (i.e., they became more imprecise) by 1.745 (Tool absent: 5.25, 95% CI [4.33, 6.16]; tool present: 6.99, 95% CI [5.83, 8.15]).

There was one interaction effect of grade within students' shifts in precision score sums from tool absent to tool present conditions (F (1, 69) = 4.44, $p < 0.05$). To investigate the source of the interaction effect, I used a paired sample t-test to compare precision score sums between tool absent and tool present conditions for each grade level. There was a significant difference in means between tool presence conditions for second graders ($t_{35} = -2.75$, $p < 0.01$), but not fourth graders. On average, second graders became more imprecise by a precision score increase of 1.76 (95% CI [-3.064, -0.46]).

In sum, this analysis revealed four findings. First, students' precision of estimates on average became impaired with the presence of a tool. Second, fourth graders were on average more precise than second graders. Third, with the introduction of a tool, second graders' shifts in precision scores became significantly imprecise compared to fourth graders' shifts in precision scores. Fourth, between tool conditions, the long paper strip was significantly less accessible for students to mediate the precision of their estimates compared to the ruler.

Treating students who rejected tools as missing data.

For this analysis, I assumed those estimates in the tool conditions for which students rejected the tool as missing data since students did not use their presented tool to estimate the target. I took as "real data" only those estimates for which students attempted to use the tool. As a result of the assumption, my sample size was reduced, especially for the long paper strip condition. Consequently, I eliminated the long paper strip condition from this analysis, treating tool-rejecting participants as missing data in these analyses. The students who remained in the analysis were as follows: At second grade, 9 out of 12 students never rejected the ruler and 7 out of 12 students never rejected the short paper strip. No fourth graders rejected the ruler (13 students) or short paper strip (14 students) (Figure 39).

Ruler	Grade 2	9
	Grade 4	13
Short Paper Strip	Grade 2	7
	Grade 4	14
Total		43

Figure 39. Number of students who never rejected the ruler and short paper strip tools

The 3-way ANOVA on student precision scores (2 (grade: second, fourth) × 2 (tool type: ruler, short paper strip) × 2 (tool presence: tool absent, tool present, repeated)) revealed similar between-subjects and within-subjects effects as the earlier approach of treating the tool rejecting students' estimates as valid data. First, a main effect for tool presence revealed that students' precision scores became significantly imprecise when using the tool ($F(1, 39)=11.58$, $p < 0.01$). Second, a main effect for grade level ($F(1,39)=10.07$, $p < 0.01$) revealed that on average, second grade students were significantly more imprecise than fourth grade students (second grade mean: 7.30, 95% CI [5.80, 8.80]; fourth grade mean: 4.33, 95% CI [3.19, 5.48]). Third, a main effect for tool type ($F(1,39)=6.75$, $p < 0.05$) revealed that the mean for students in the short paper strip group was significantly higher (more imprecise) than of those in the ruler group; the short paper strip was less accessible than the ruler for students to mediate the precision of their estimates. Fourth, there was an interaction effect of precision score sum means within grade levels ($F(1, 39)=10.65$, $p < 0.01$). A paired sample t-test revealed that the source of this interaction was that second graders, but not fourth graders, became significantly imprecise from tool absent to tool present conditions ($t_{15}= -3.50$, $p < 0.01$, mean difference = -4.17, 95% CI [-6.72, -1.63]). Figure 40 illustrates the difference of average precision score sums of each tool presence condition as a function of grade and tool type.

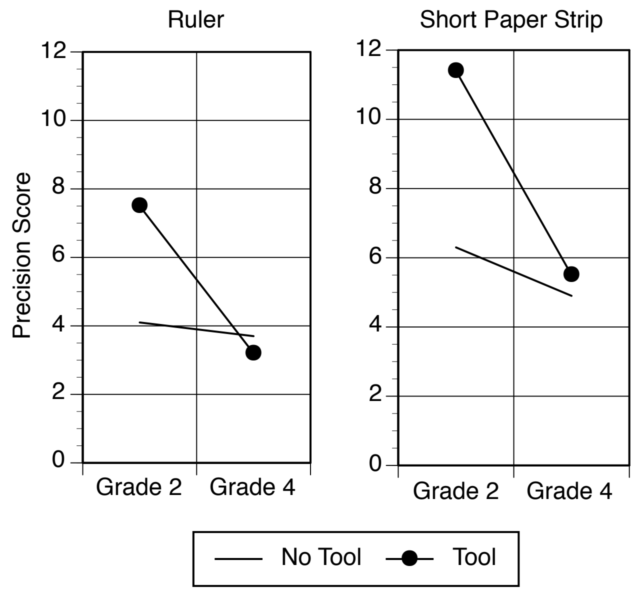


Figure 40. Shifts of precision score sum means (0 = precise) for no tool (tool absent) condition and tool (tool present) condition by grade for each tool type

Treating tool-rejecting students' precision scores for the tool present condition as identical to those for the tool absent condition.

Like the second analysis, in my third analysis, I regarded tool-rejecting participants' estimates in the tool present condition as problematic. These students' estimates in the tool present condition would be suspect because they might have misinterpreted the intended purpose of the second request for an estimate, inferring that their estimate was wrong and they were being given a second chance to improve (in other words, treating the second question as a Gricean conversational implicature (see, for example, Siegal, 1991)). Such an interpretation might have prompted students to produce two different estimates for a problem. Thus, I reassigned tool rejecters the same precision score in the tool present condition that they were assigned in the tool absent condition. To confirm the merit of this approach, I analyzed tool-rejecting participants' relative precision scores between the tool absent and present conditions. If my expectations about the unreliability of students' precision scores in the tool present condition are well founded, I would expect some students to show less reliability across conditions and some to show greater reliability. Such a finding would support the merit of this analytic approach, and I found that students did vary in their precision of estimates by grade and tool presence, and tool type for second graders' shifts from tool absent to tool present conditions.

The ANOVA that included substitution scores revealed three findings, two of which mirrored the previous two analytic approaches (treating tool-rejecting students' estimates as valid, and treating tool-rejecting students as missing data). First, the ANOVA revealed that on average, students significantly became imprecise in precision scores from tool absent to tool present conditions ($F(1, 69) = 8.52$, $p < 0.01$). Second, second graders were significantly more imprecise compared to fourth graders ($F(1, 69) = 6.53$, $p < 0.05$; second grade mean: 7.09, 95% CI [5.70, 8.49]; fourth grade mean: 4.61, 95% CI [3.27, 5.96]). These two findings were similar to previous analyses. However, unlike the previous two analytic approaches, this analysis did not reveal a main effect of tool type or an interaction effect of tool presence and grade. Instead, the third result was that there was a significant interaction effect of grade level, tool type, and tool presence ($F(2, 69) = 3.64$, $p < 0.05$). A paired-sample t-test of students' precision scores from tool absent to tool present conditions over grade and tool revealed that second graders in the short paper strip condition, but not others, became significantly more impaired in precision ($t_{11} = -3.128$, $p < 0.05$, mean difference=-3.02, 95% CI [-5.14, -0.89]). Figure 41 illustrates the difference of average precision score sums of each tool presence condition as a function of grade and tool type.

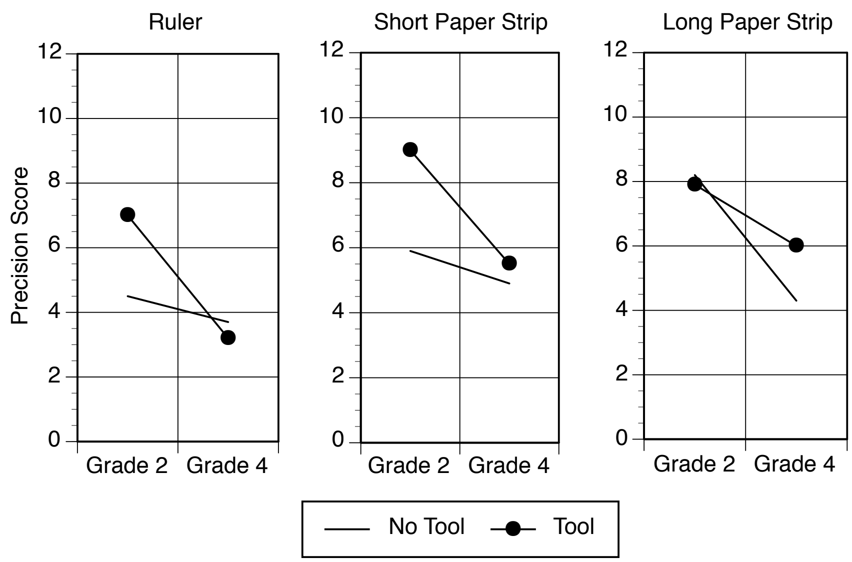


Figure 41. Shifts of precision score sum means (0 = precise) for no tool (tool absent) condition and tool (tool present) condition by grade for each tool type

Summary of the three tool-rejection accommodation analyses.

The large number of unanticipated tool-rejections by students on the long paper strip condition compared to the other conditions supports my expectation that the long paper strip presents more difficulty for students that then other tool types. In the subsequent analyses, I accommodated the rejections in three ways and revealed results that matched some, but not all, of my earlier expectations of students' tool using and their precision scores. I had expected an interaction of students' precision of estimates with tool type as they shifted from unmediated estimations without tools to mediated estimations using tools. For age, I had expected more students in fourth grade than second grade would have been able to mediate their estimates using their assigned tools, leading to an improvement in precision from tool absent to tool present conditions. Though all three analytic approaches revealed that fourth grade students were more precise than second grade students on means of precision scores, students in both grades generally became more imprecise. In Vygotsky's treatment of development and tool using, this might mean that students in both grades were generally early in development and unable to use tools to mediate the precision of their estimates. However, more fourth graders than second graders might have begun to use their tools to serve a measurement function.

For tools, as I noted earlier, I had expected that the ruler and the long paper strip would be the more difficult for students to use to support their linear measurement activity compared to the short paper strip and improve their precision. A larger proportion of students rejected the use of the long paper strip compared to other tools, which supports the expectation that the long paper strip was most challenging for students to adapt into their linear measurement activity. For the first analytic approach (treating tool-rejecting students as valid data), I found that the long paper strip students were significantly more imprecise (higher precision score means) than the ruler condition students, but not the short paper strip condition students. For the second analytic approach (treating tool-rejecting students as invalid data), the long paper strip condition along with all tool-rejecting students were dropped, and results showed that short paper strip students were significantly more imprecise (higher precision score means) than the ruler condition students. Following Vygotsky for the two analytic approaches, it is possible that there were fewer advanced students and more unmediated imprecise estimations in one or both of the paper strip conditions compared to the ruler condition. For the third analytic approach (treating tool-rejecting students' tool absent and present estimates as identical), findings indicated that second grade students became significantly impaired in their precision with the introduction of the short paper strip. This suggests that for students early in their development of tool using, the short paper strip particularly interfered with their estimation activity.

The next analysis focused on uncovering a possible intermediate group of students who did improve their precision on problems. For this analysis, I assumed that the mere presence of the tool influenced students' performance even if they rejected the tool. The above analyses all used precision score sums in the tool absent and tool present conditions as the dependent variable, which may have obscured individual students who did improve in precision when using a tool to mediate their estimates. An overall impairment in precision may have resulted from two other possible sources. First, there were five problem types that students encountered in each tool presence (tool absent and tool present) condition, and it is possible that students were able to mediate their precision with a tool on some problems but not others. Second, it might be that more students became impaired in precision with the introduction of tools than those who improved in precision with the use of a tool. If either of these cases occurred, this would have resulted in an overall increase in precision score sums, leading to a conclusion that students became more imprecise with tools present. An intermediate group of students who were able to mediate their precision of estimates with their assigned tool may have been overlooked because the number of students were fewer or many but not all problems were interfering with students efforts to use their tool.

Students' precision score shifts.

To determine whether there may have been an intermediate group of students who did improve their precision of estimates with the use of a presented tool, I tabulated how many students improved their precision on the majority of problems, or impaired their precision on the majority of problems, or had an equal number of improved and impaired precision instances (including the case of no instances of either). I found that there were three students at fourth grade (2 ruler, 2 short paper strip) who improved on all problems with the use of their tool, and two students at second grade (1 ruler, 1 long paper strip) whose precision became impaired on all problems with the use of their tool. There were no individual students who showed no difference in precision scores for the entire problem set as a whole. The rest of the individual students each improved precision on some problems and did not improve precision on others with the use of their tool. Six of these students showed no difference in precision on only one problem. Figure 42 displays how many students in each tool condition and grade level more often improved their precision with the tool, or more often impaired their precision with the tool, or evenly improved and impaired their precision across problems with the tool.

		Less precise more often with tool use	More precise more often with tool use	Evenly less and more precise with tool use
Ruler	Second Grade (n = 12)	5 (41.67%)	7 (58.33%)	0 (0%)
	Fourth Grade (n = 13)	3 (23.08%)	9 (69.23%)	1 (0.08%)
Short Paper Strip	Second Grade (n = 12)	9 (75%)	3 (25%)	0 (0%)
	Fourth Grade (n = 14)	5 (35.71%)	9 (64.29%)	0 (0%)
Long Paper Strip	Second Grade (n = 12)	6 (50%)	6 (50%)	0 (0%)
	Fourth Grade (n = 12)	6 (50%)	6 (50%)	0 (0%)

Figure 42. Tendency of students' precision shifts within problem sets from tool absent to tool present conditions

I found that in any given tool condition at each grade, there were students who became more precise more often with the use of a tool. Further, in the ruler condition at both grades and in the short paper strip condition at fourth grade, this population outnumbered the number of students who became less precise more often with the use of a tool. This confirmed that though the earlier ANOVA results suggested that most student groups on average became impaired in their precision with the introduction of tools, there was a subgroup of students who were able to mediate their precision of estimates with the use of a tool. This subgroup's performance may have been obscured because the average of their precision score decreases was less than the precision score increases of those who became less precise with a tool.

It is possible that students were exhibiting more precision or less precision as a function of problem type. A large percentage of students became more precise with a tool on more problems than more imprecise with a tool. More second graders became more precise than less precise on three or more problems with the use of a ruler. The majority of fourth grade students were more precise on three or more problems with the use of a ruler or short paper strip. Half of the long paper strip users displayed this behavior. This suggests that students' overall mean of precision became worse not because they were unable to use their tool to mediate their estimation activity, but because some problems were particularly difficult for students to mediate their linear estimates using their tool; coordinations between tool units and target number line units might have been more difficult for some problems than others, resulting in different frequencies of precision score decline or improvement by problem. Thus, the above analysis was further decomposed into students' improvements or no improvements (by collapsing impairment and no difference) in precision for each problem as a function of tool type and grade (Figure 43).

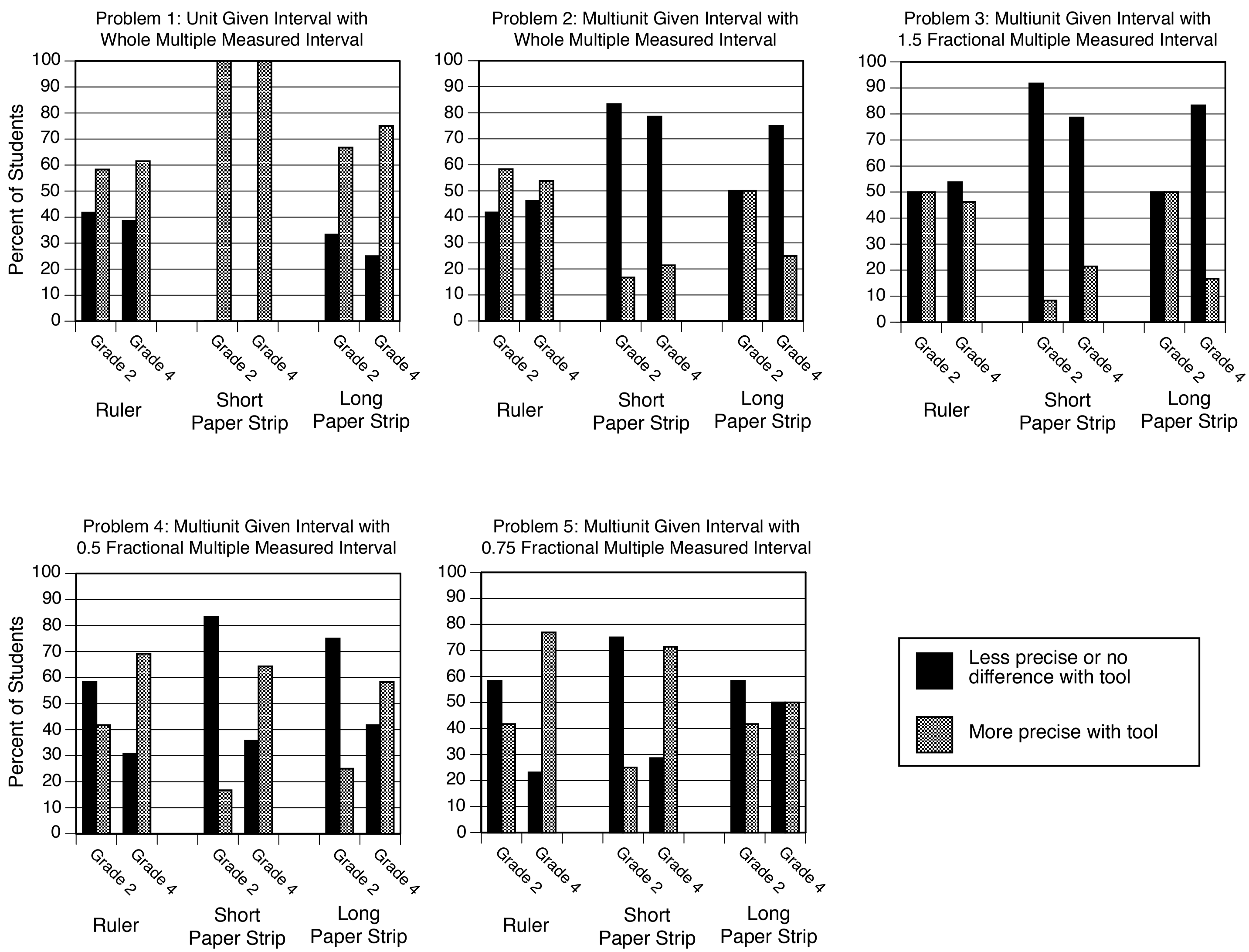


Figure 43. Percent of students shifting to lesser or greater precision for each problem with tool as a function of tool type and grade

The frequency analysis revealed that in all but one condition, there were students who improved precision when tools were introduced into their measurement activity, but the relative percentage of students who improved precision or impaired precision differed as a function of problem type. For the unit interval problem (problem 1), more than 50% of students increased in precision with the use of their tool in each grade and tool type (second grade: 58.3% ruler, 100% short paper strip, 66.7% long paper strip; fourth grade: 61.5% ruler, 100% short paper strip, 75% long paper strip). Notably, all students in the short paper strip condition increased in precision on problem 1. For the multiunit interval problem with whole multiple measured interval (problem 2), however, only students in the ruler condition had a slightly higher percentage of students who improved in precision with the use of a tool (58.3% at second grade and 53.8% at fourth grade) compared to those who did not improve in precision. For the multiunit interval problem with 1.5-fractional multiple measured interval (problem 3), there were no instances at each grade or tool type in which more students improved with the use of a tool than students who did not. For the multiunit interval problems with 0.5-fractional multiple interval (problem 4) and 0.75-fractional multiple measured interval (problem 5), only fourth grade students exhibited a higher percentage of those who improved their estimate precision with their tool (problem 4: 69.2% ruler, 64.3% short paper strip, 58.3% long paper strip; problem 5: 76.9% ruler, 71.4% short paper strip).

At both grades, students who were able to mediate their precision of estimates with the use of a tool were most successful on the unit interval problem (problem 1), especially in the short paper strip condition. This indicated that students were typically able to coordinate the unit interval on the number line with their tool. This makes sense given that the unit interval did not need to be further split into shorter units to generate a linear estimate on the number line, and students only needed to coordinate one measure with their tool on the number line. With the short paper strip, all students in each grade improved their precision on this problem. The short paper strip did not need to be further split into shorter units to represent a linear unit on the target number line. Compared to the ruler or long paper strip, fewer coordinations between units were needed compared to the ruler. For the ruler, a magnitude of 4cm needed to be coordinated with 1 linear unit on the number line, and for the long paper strip, it first needed to be folded to fit and represent a linear unit).

For the multiunit interval problems (problems 2-5), second and fourth graders displayed visibly different directions of precision shifts with the presence of the ruler and the short paper strip. At both grades for problem 2 and problem 3, which had given intervals of 2 units, a large percentage—more than 70%—of students did not improve in precision with the short paper strip. For problem 4 and problem 5, which had given intervals of 4 units, the short paper strip continued to be largely inaccessible for precision improvement for second graders, but the inverse was true for fourth graders. Similarly, the ruler was less accessible for more second graders in improving their precision while it was more accessible for more fourth graders in improving their precision.

On multiunit interval problems, second graders and fourth graders behaved differently in their use of the ruler and short paper strip. For most second graders, the ruler and short paper strips interfered with students' production of linear estimates of target values (as revealed by the decline in precision when using the tool). In contrast, at fourth grade, these tools supported students' estimates on four-unit interval problems (as revealed by improved precision from tool absent to tool present conditions). The developmental advance reveals challenges that many fourth graders have overcome in tool-target coordinations on this problem type. Indeed, most second graders treated the short paper strip as a linear unit on target number lines; this type of coordination on multiunit interval problems led to gross overestimations of target values. Many second graders also treated 4cm (the length of the given interval) on the ruler as a linear unit; coordinating 4cm as a linear unit also led to gross overestimations of target values. In contrast, the majority of fourth graders were able to overcome these coordination challenges and improve their precision on four-unit multiunit interval problems. For more fourth graders than second graders, tools began to serve a functional value and students became more precise in their estimates of target values with the introduction of a tool.

Chapter 4: Discussion

In the present study, I argued that we need more systematic empirical studies that investigate the conceptual challenges children encounter in using tools for length estimation, particularly when children have to coordinate multiple linear measures in problem solving activity. Children's use of measurement tools has been studied by documenting their successes and challenges using non-standard tools (e.g., paper strips) and standard tools (e.g., rulers) (Clements, 1999; Haylock & Cockburn, 1989; Kastner, 1989; McDonough & Sullivan, 2011). These have included different kinds of studies, such as depicting children's advances and challenges in generating a linear measure or comparing the length of two objects using non-standard tools or rulers (Boulton-Lewis et al. 1996; Bragg & Outhred, 2004; Clements, 1999; Hiebert, 1981, 1984; Kamii & Clark, 1997; Kotsopoulos et al., 2015; Lehrer et al., 1999; Nunes et al., 1993; Solomon et al., 2015) and developing or evaluating instructional sequences or pedagogical approaches for teaching linear measurement (Barrett et al., 2011; Kostopoulos et al., 2015; McClain et al., 1999; Szilagyi, Clements, & Sarama, 2013). Though some linear measurement studies point to possible trajectories through which children increase their abilities to operate on linear measures (without a focus on tool using) (e.g., Barrett & Clements, 2003), few have focused directly on the mediational strategies children use with tools to support their linear estimates. Number line problems help illuminate children's understanding of linear units (Fosnot & Dolk, 2001; Petitto, 1990; Saxe et al. 2010, 2013), but how children use different types of tools to operate on linear units on number lines has not been studied extensively. A focus on how children use tools to mediate their linear estimates in number line contexts may provide a more comprehensive understanding of how children advance from perceptual to internalized measurement strategies.

To support my inquiry, I adapted Vygotsky's (1978) framework on mediation and Saxe's (2012) treatment of microgenesis of tool using, focusing on children's developing use of tools to produce numerical estimates on number line. In terms of development, I examined how students mediated their linear measurement activity using tools and how tool using influenced their estimates compared to when no tools were available. These analyses focused on whether there were students early in developing measurement tool use who did not find tools to serve a measurement function, students later in development who began to mediate and improve their measurements using tools, and advanced students who were able to mediate their measurements with and without tools. To analyze how students produced measurements using tools, I focused on students' microgenetic constructions of the coordination between tools (three types with varying properties) and target number lines (five types with varying properties). These analyses identified the manner in which students were able to generate a relationship between two quantification activities: (1) transforming tools to quantify them into units, and (2) using interval values to identify linear units on the number line.

In this chapter, I review the key contributions of this dissertation. Then I discuss the limitations of the study, implications of the findings, and next steps to extend the work.

Review and Discussion of Key Findings

To study the interplay between children's linear unit understandings and their adaptation of different tools to measure linear targets, I interviewed second grade and fourth grade students as they solved linear measurement problems (placing numbers on unbounded number lines) and presented three different types of tools—a ruler and paper strips of two different lengths—to support their problem solving. All problems were designed to elicit coordinations between particular unitizing actions (splitting, iterating, counting) on the targets (number line problems) and tools. I expected that younger children would encounter more difficulty generating and preserving units on number lines using rulers compared to paper strips. However, as more unitizing actions were required to solve a problem, children would find the short paper strip, which was the length of the pre-labeled given intervals on the number lines, more accessible than the long paper strip, which was as long as the ruler. I also anticipated that even though older children might be more precise in generating linear estimates on number line problems than younger children, their tool using strategies would reveal persisting challenges in coordinating between tool units and target units, especially for rulers. Overall, students' mediational approaches using measurement tools showed context specificity. The character of students' tool-target coordinations varied by problem type, and there was marked variation in levels of performance across tool treatment groups at a given grade level.

Many fourth graders were able to overcome challenges in coordinating the co-quantification of tools and target number lines with a consistent unit to generate more precise estimates than most of the second graders, but at both grades, the majority of students were more imprecise with the use of a tool compared to a perceptual strategy. The majority of fourth graders mediated their linear estimates by coordinating their tool units with number line units while the majority of second graders were only able to partially coordinate between tool units and target units, leading to greater imprecision compared to fourth graders. However, some second grade students were able to use their tool to mediate their measurements adequately, and many fourth grade students still experienced some difficulty as evidenced by the presence of lower performance across tool conditions and problem types. Thus, at both grades, there were students early in their ability to adapt tools to serve a measurement function for their number line estimates, and those who were later in development and able to transform the tools to serve a measurement function.

Compared to rulers and short paper strips, long paper strips were the most challenging for students in both grades to adapt into their number line measurement activity. Students found it challenging to unitize the long paper strip by corresponding a segment of it onto the number line, even though they were able to provide estimates using eyeball strategies or their fingers. Long paper strips resembled rulers in shape, but a segment first had to be cut, folded, or visualized as the length of the given interval on the number line to serve a measurement function. The long paper strip in particular was rejected by a large number of students, especially at second grade. Amongst the students in the long paper strip condition, a large majority of second graders and half of the fourth grade students rejected the tool on at least one problem. When encouraged to use the long paper strip, most of the tool-rejecting students claimed that they did not know how to use the tool and some reasoned that it would be possible to measure estimates if there were enumerated marks (like a ruler) on the paper strip. In contrast, only a small number of second grade students in the ruler and short paper strip conditions rejected the tool and no fourth grade students rejected the ruler or short paper strip. The length of the short paper strip fit in the given interval, providing a visual cue that it the tool and number line interval could be coordinated in some fashion. Even though not all students used the ruler correctly, students were familiar with the tool and knew it could be used to measure, and readily tried to use the enumerated tick marks or the edge of the ruler to generate a linear estimate. Students may have been more likely to adapt the long paper strip if an intermediate step was shown, such as fitting the edge of the paper strip to the interval for the student or folding the edge of the strip to the length of the given interval.

On average, fourth grade students were more precise than second grade students, but for students at both grades some tools either interfered or did not serve a functional value. Students who used long paper strips became significantly more imprecise than students who used rulers. This means that even though many students (especially at second grade) rejected the use of the long paper strip, they still became more imprecise by the mere presence of the tool. Second graders, but not fourth graders, became significantly imprecise from the use of short paper strips. In contrast, ruler-using fourth graders, as a group, improved the precision of their estimates. Students who were later in development in tool using likely recognized that they could use the ruler's equally spaced enumerated units as a resource for generating new lengths, which may be why relatively more fourth grade students were able to better improve the precision of their estimates compared to students with the short paper strip.

There were students who improved their precision with the use of their tool as a function of problem type in each grade and tool type. The most instances occurred for the unit interval problem (e.g., Given (7, 8) place 10): all of the short paper strip using students improved their precision with the use of their tool, and for other tools, more students improved than impaired in precision with the use of their tool. As expected, some students did find that mediating linear measures with their tool was relatively easier when a splitting action was unnecessary. The ruler and short paper strip benefitted more fourth graders than second graders on the four-unit interval problems (e.g., Given (10, 14) place 16 or given (5, 9) place 12). The source of the problem and grade interaction on precision shifts was likely due to second graders' persistent difficulty with using the short paper strip to mediate their linear estimates on multiunit interval problems. The two-unit interval problem with 1.5-fractional multiple measured interval (e.g., Given (6, 8) place 11) was the only problem for which the relative proportion of students who improved in precision with their tool was not greater than those who did not improve in precision. The features of this problem may have presented more challenges to students' efforts to mediate their measurement activity.

Thus, many second grade students and some fourth grade students were unable to use the presented tools to mediate their unitizing coordinations and improve their precision. Rather, the tools were often a hindrance and did not serve a functional value for coordinating quantified units. This was most evident in the case of the long paper strip, and second graders who were presented with the short paper strip. The presentation of the tool was detrimental to students' efforts in tool-target coordination and improving precision of estimates. Yet, in both grades, there were intermediate students who became more precise with the use of the tool. The tools began to serve a functional value for this group, and students improved in their performance with tools relative to their performance without tools, especially on the unit interval problem and for fourth graders who used the ruler. Some of these students may have been advanced students who displayed little difference in their precision scores from tool absent to tool present conditions, which may mean that they were already covertly mediating unit coordinations and did not necessarily need the provided tools to improve their precision of estimates. It is possible that the advanced students were fourth graders who were presented with the ruler or the short paper strip and improved in their precision of estimates using their tool. These students exhibited a small amount of shift in precision from tool absent to tool present conditions, which suggests that these students may have not depended on the ruler or short paper strip to place their estimates.

Limitations

In this section, I discuss the limitations of this study design, focusing on six aspects that have implications for future research: (1) adding scaffolding tasks, (2) varying location of given intervals, (3) additional interviews, (4) providing tool preference, (5) expanding age groups, and (6) additional numerical domains.

Adding scaffolding tasks.

A limitation of the tasks presented to the students was that students were not given any suggestions for unitization even when they rejected the use of the tool. Recall that all students were able to unitize their assigned tool before solving problems with their tools, such as folding the paper strip into “two equal pieces” or showing the length of 2cm on the ruler. However, many students, especially at second grade, claimed they could not use their tool to measure lengths on the number line. The interview protocol could have included follow-up suggestions for students to use a unitization strategy they demonstrated earlier with the tool. This would have better illuminated their developing use of measurement tools.

Varying location of given intervals.

For all problems, given intervals were centered on the line segments. This was originally done in order to avoid providing students with a zero-point on the number line. However, students may have still assumed that 0 was at the left end of the lines, and they may have been covertly unitizing the number line with this in mind, either leading to overestimations or assumptions that units did not have to be equally spaced. If more time or multiple interviews were incorporated into the interviews, I could have added more number line tasks that also varied in the location of the given intervals (e.g., at the left end, at the right end, etc).

Additional interviews.

Repeated questioning may have prompted students to solve the problems in tool absent and tool present conditions differently, because they viewed the interviewer as violating conversational rule and requesting for a different estimate. Adding a time delay, such as a week or two, between no tool and tool conditions may have helped mitigate this. Adding more interviews with time delays in between may have also allowed for opportunities to ask students to evaluate hypothetical other students’ solutions or measurement strategies as a means to understand students’ abilities to coordinate between tools and targets.

Providing tool preference.

The time constraints of the school sites and the number of participants was not large enough to allow for systematic comparisons in which each student could try more than one type of tool. Having each student try each tool type or choose a preferred tool from all types may have helped reveal better which tools they can better adapt to measurement activity.

Expanding age groups.

The majority of second and fourth graders became more imprecise in their linear estimates with tool use. Additionally, by the time data collection started, students in second grade at the selected school sites became quite familiar with rulers. Interviewing even younger students as well as much older students may have shown a better spectrum of tool using development.

Additional numerical domains.

This interview design only included whole numbers, and many of the older students performed very well on the problems as a result. Coordinating splitting, counting, and iterating actions with fractional values would have been more challenging for students, and the addition of number line tasks with fractions may have allowed better insight into how students coordinate between their tool and targets.

Implications and Next Steps

Tool-target coordination and precision findings showed that students' performance differed as a function of age, tool type, and problem type. Adapting Vygotsky's (1978) framework on mediation and tool using and Saxe's (2012) treatment of microgenesis allowed for an integrated conceptualization of students' measurement tool using activity in a novel way. Students' developing use of tools was analyzed in terms of how the use of tools influenced their measurements, and their ability to co-quantify tools and targets.

Unbounded number lines afford a context in which a given interval on the line can serve as a means to generate a new length. Interval and target values can be modified to require multiple linear unit operations, such as splitting an interval twice before iterating and counting. By examining how children adapt different types of tools on unbounded number line problems to support their linear estimates, one may better illuminate children's emergent linear unit understandings in terms of the varied unit coordinations they produce between tools and targets (number lines). Depending on the given values of a number line problem, and the properties of the provided tools, children might encounter different unitizing challenges that might also influence the precision of their estimates.

Assessing students' understandings of measurement tool using should include students' explanations on how they would unitize the tool, the target, and how those two would be corresponded to one another. Precision shifts from no tool use to tool use on number line problems did not adequately represent students' abilities to coordinate between tool units and target units. Many fourth graders and several second graders clearly demonstrated conceptual understanding in quantifying and coordinating tools and targets, but many students in both grades did not necessarily improve in precision. This may have been due to motor coordination abilities than lack of understanding how to use their tools.

The affordances or challenges in students' use of measurement tools partly arise from an interaction between the tool's physical features and problem features. The difficulty of tool and problem combinations will differ for age groups, so a variety should be presented both in the earlier grades as well as later grades instead of focusing on presenting one type of tool before the other (e.g., non-quantified non-standard tools before pre-quantified standard tools) or not having problems that include splitting, iterating, and counting actions for younger children.

In next steps, I intend to address some of the study's limitations to build a more comprehensive understanding of students' use of measurement tools. First, the current tasks will be piloted with more students in additional grades to determine whether adding additional age groups would help illuminate more variety in unit coordination strategies. Second, in addition to generating multiple linear measures on number lines, students will also be given the opportunity to evaluate the accuracy of number line targets placed by other (hypothetical) students using their provided tool. Third, students will be able to choose from different types of tools to support their measurement activity, and then be asked to use a less preferred tool to solve problems as well. Fourth, if students are unable to unitize their tool in any fashion, then they will be provided with suggestions on transforming the tool (e.g., What if you folded the paper strip like this?). Fifth, students will be interviewed more than once to minimize asking students to solve one type of problem more than once in a session. Starting with these five revisions may provide better insight in students' generative understandings in measurement tool using.

Concluding Remarks

This dissertation presented a systematic study of students' developing use of different kinds of measurement tools in number line problem contexts. The findings supported the utility of focusing on mediation and conceptualizing measurement tool using in terms of co-quantifying the given tool with a number line target. There was also evidence that though students could independently unitize number lines and tools, many are unable to coordinate between the two. Older children were generally more precise and better able to coordinate between their tool units and number line units, but measurement tool difficulties still persisted. The longer non-standard tool (long paper strip), which did not appear to have an obvious measurement function, was particularly difficult for students to perceive as a useful tool to support their linear estimates. The variation in tool types and problem types helped reveal these findings. More research needs to explore and understand how students advance their tool to target unit coordinations, and the findings of this research study present a productive design and context to pursue this goal.

References

Barrett, J. E., & Clements, D. H. (1996, October). *Representing, connecting and restructuring knowledge: A microgenetic analysis of a child' s learning in an open-ended task involving perimeter, paths, and polygons.* Paper presented at the 18th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education, Panama City, FL.

Barrett, J. E., & Clements, D. H. (2003). "Quantifying Length: Fourth-Grade children's developing abstractions for measures of linear quantity." *Cognition and Instruction*, *21*(4), 475-520.

Barrett, J., Cullen, C., Sarama, J., Clements, D., Klanderman, D., Miller, A., & Witkowski Rumsey, C. (2011). Children's unit concepts in measurement: A teaching experiment spanning grades 2 through 5. *ZDM: The International Journal on Mathematics Education, 43*(5), 637-650.

Barth, H. C., & Paladino, A. M. (2010). The development of numerical estimation: Evidence against a representational shift. *Developmental Science, 14*, 125- 135.

Booth, J. L., & Siegler, R. S. (2006). Developmental and individual differences in pure numerical estimation. *Developmental Psychology, 41*, 189–201.

Booth, J. L., & Siegler, R. S. (2008). Numerical magnitude estimations influence children's learning. *Child Development, 79*(4), 1016-1031.

Boulton-Lewis, G. M., Wilss, L. A., & Mutch, S. L. (1996). An analysis of young children's strategies and use of devices for length measurement. *The Journal of Mathematical Behavior*, *15*(3), 329–347.

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.

Clements, D. H. (1999). Teaching length measurement: Research challenges. *School Science and Mathematics, 99*(1), 5-11.

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* (1st ed., Vol. 1, pp. 201–226). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Clements, D. H., & Stephan, M. (2004). Measurement in preK-2 mathematics. In D. H. Clements, J. Sarama & A.-M. DiBiase (Eds.), *Engaging young children in mathematics: Standards for early childhood mathematics education* (pp. 299-317). Mahwah, NJ: Lawrence Erlbaum Associates.

Cohen, D. J., & Blanc-Goldhammer, D. (2011). Numerical bias in bounded and unbounded number line tasks. *Psychonomic Bulletin and Review,18*, 331–338.

Cohen, D. J., & Sarnecka, B. W. (2014). Children's number-line estimation shows development of measurement skills (not number representations). *Developmental psychology*, *50*(6), 1640.

de Hevia, M. D., & Spelke, E. S. (2010). Number-space mapping in human infants. *Psychological Science*, *21*(5), 653–660. doi: http://doi.org/10.1177/0956797610366091

Dehaene, S., Izard, V., Spelke, E., & Pica, P. (2008). Log or linear? Distinct intuitions of the number scale in western and Amazonian indigene cultures. *Science, 320,* 1217-1220. doi: 10.1126/science.1156540

Drake, M. (2014). Learning to measure length: The problem with the school ruler. *Australian Primary Mathematics Classroom*, *19*(3), 27–32.

Ebersbach, M., Luwel, K., Frick, A., Onghena, P., & Verschaffel, L. (2008). The relationship between the shape of the mental number line and familiarity with numbers in 5- to 9-year old children: Evidence for a segmented line. *Journal of Experimental Child Psychology*, *99*, 1-17.

Ebersbach, M., Luwel, K., & Verschaffel, L. (2015). The relationship between children's familiarity with numbers and their performance in bounded and unbounded number line estimations. *Mathematical Thinking and Learning*, *17*(2-3), 136–154. doi: http://doi.org/10.1080/10986065.2015.1016813

Education Data Partnership (Ed-Data) (2017). *School profile*. Retrieved from www.ed-data.org

Feigenson, L., Dehaene, S., & Spelke E. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*, 307-314.

Fosnot, C. T., & Dolk, M. L. A. M. (2001). *Young mathematicians at work*. Portsmouth, NH: Heinemann.

Fuson, K. C. (1988). *Children's counting and concepts of number.* New York: Springer-Verlag Publishing.

Gravemeijer, K., & Stephan, M. (2002). Emergent models as an instructional design heuristic. *Symbolizing, modeling and tool use in mathematics education*, 145-169.

Haylock, D. W., & Cockburn, A. D. (1989). *Understanding early years mathematics*. London, England: Paul Chapman Publishing.

Hiebert, J. (1981). Cognitive development and learning linear measurement. *Journal for Research in Mathematics Education*, *12*(3), 197. doi: http://doi.org/10.2307/748928

Hiebert, J. (1984). Why do some children have trouble learning measurement concepts? *The Arithmetic Teacher*, 19–24.

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*(3), 154–158.

Kamii, C., & Clark, F. B. (1997). Measurement of length: The need for a better approach to teaching. *School Science and Mathematics*, *97*(3), 116–121.

Kastner, B. (1989). Number sense: The role of measurement applications. *The Arithmetic Teacher*, *36*(6), 40–46.

Kloosterman, P., Rutledge, Z., & Kenney, P. A. (2009). Exploring Results of the NAEP: 1980s to the Present. *Mathematics Teaching in the Middle School*, *14*(6), 357.

Kotsopoulos, D., Makosz, S., Zambrzycka, J., & McCarthy, K. (2015). The Effects of Different Pedagogical Approaches on the Learning of Length Measurement in Kindergarten. *Early Childhood Education Journal*, *43*(6), 531-539.

Lehrer, R. (2003). Developing understanding of measurement. In J. Kilpatrick, W.G. Martin, & D. Schifter (Eds.), *A Research Companion to Principles and Standards for School Mathematics* (pp. 179–192). Reston, VA: National Council of Teachers of Mathematics.

Lehrer, R., Jacobson, C., Kemeny, V., & Strom, D. (1999). Building on children's intuitions to develop mathematical understanding of space. In E. Fennema & T. Romberg (Eds.), *Mathematics classrooms that promote understanding* (pp. 63-87). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

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

Lehrer, R., & Schauble, L. (2000). Modeling in mathematics and science. In R. Glaser (Ed.), *Educational design and cognitive science* (Vol. 5, pp. 101–159). Mahwah, NJ: Lawrence Erlbaum Associates.

Lindquist, M. M., & Kouba, V. L. (1989). Measurement. In *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 *Results from the seventh mathematics assessment of the NAEP* (pp. 193–234). Reston, VA: National Council of Teachers of Mathematics.

McClain, K., Cobb, P., Gravemeijer, K., & Estes, B. (1999). Developing mathematical reasoning within the context of measurement. In L. Stiff (Ed.), *Developing mathematical reasoning (1999 Yearbook of the Nation Council of Teachers of Mathematics).* Reston, VA: National Council of Teachers of Mathematics.

McDonough, A., & Sullivan, P. (2011). Learning to measure length in the first three years of school. *Australasian Journal of Early Childhood*, *36*(3), 27-35.

Moeller, K., Pixner, S., Kaufmann, L., Nuerk, H.-C. (2009). Children's early mental number line: Logarithmic or decomposed linear? *Journal of Experimental Child Psychology, 103,* 503-515. doi: 10.1016/j.jecp.2009.02.006

Murata, A. (2008). Mathematics teaching and learning as a mediating process: The case of tape diagrams. *Mathematical Thinking and Learning, 10*(4), 374-406. doi: 10.1080/10986060802291642

Murata, A., & Kattubadi, S. (2012). Grade 3 students' mathematization through modeling: Situation models and solution models with multi-digit subtraction problem solving. *Journal of Mathematical Behavior, 31,* 15-28. doi: 10.1016/j.jmathb.2011.07.004

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: Author.

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

National Council of Teachers of Mathematics. (2003). *Learning and teaching measurement*. Reston, VA: National Council of Teachers of Mathematics.

National Governors Association Center for Best Practices and Council of Chief State School Officers. (2010). *Common Core State Standards for Mathematics*. Washington, DC: Author.

Nunes, T., Light, P., & Mason, J. (1993). Tools for thought: The measurement of length and area. *Learning and Instruction*, *3*(1), 39–54.

Petitto, A. L. (1990). Development of Numberline and Measurement Concepts. *Cognition and Instruction*, *7*(1), 55–78.

Piaget, J., & Inhelder, B. (1948/1956). *The child's conception of space.* London: Routledge & Kegan Paul.

Piaget, J., Inhelder, B., & Szeminska, A. (1960). *The child's conception of geometry.* New York: Basic Books.

Reinert, R. M., Huber, S., Nuerk, H., & Moeller, K. (2015). Multiplication facts and the mental number line: evidence from unbounded number line estimation. *Psychological Research,* 1-9. doi: 10.1007/s00426-013-0538-0

Saxe, G. B. (2012). *Cultural development of mathematical ideas: Papua New Guinea studies*. New York, NY: Cambridge University Press.

Saxe, G. B., de Kirby, K., Kang, B., Le, M., & Schneider, A. (2015). Studying cognition through time in a classroom community: The interplay between "everyday" and "scientific" concepts. *Human Development*, *58*(1), 5–44. doi: http://doi.org/10.1159/000371560

Saxe, G. B., Earnest, D., Sitabkhan, Y., Haldar, L. C., Lewis, K. E., & Zheng, Y. (2010). Supporting generative thinking about the integer number line in elementary mathematics. *Cognition and Instruction, 28*(4), 433-474. doi: 10.1080/07370008.2010.511569

Saxe, G. B., & Moylan, T. (1982). The development of measurement operations among the Oksapmin of Papua New Guinea. *Child Development*, *53*(5), 1242–1248.

Saxe, G. B., Shaughnessy, M. M., Gearhart, M., & Haldar, L. C. (2013). Coordinating numeric and linear units: Elementary students' strategies for locating whole numbers on the number line. *Mathematical Thinking and Learning, 15*(4), 235-258. doi: 10.1080/10986065.2013.812510

Siegler, R. S., & Booth, J. L. (2004). Development of numerical estimation in young children. *Child Development, 75*(2), 428–444. doi: 10.1111/j.1467-8624.2004.00684.x

Siegler, R. S., & Opfer, J. E. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological Science, 14,* 237-243.

Smith III, J. P., Males, L. M., Dietiker, L. C., Lee, K., & Mosier, A. (2013). Curricular treatments of length measurement in the United States: Do they address known learning challenges?. *Cognition and Instruction*, *31*(4), 388-433.

Solomon, T., Vasilyeva, M., Huttenlocher, J., & Levine, S. C. (2015). The long and the short of it: Children's difficulty conceptualizing spatial intervals on a ruler as linear units. *Developmental Psychology, 51*, 1564-1573.

Steffe, L. P., & Hirstein, J. J. (1976). *Children's thinking in measurement situations (National Council of Teachers of Mathematics Yearbook)*. Reston, VA: National Council of Teachers of Mathematics.

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

Szilágyi, J., Clements, D. H., & Sarama, J. (2013). Young children's understandings of length measurement: Evaluating a learning trajectory. *ZDM-The International Journal on Mathematics Education, 44*, 581-620. Retrieved from www.jstor.org/stable/10.5951/jresematheduc.44.3.0581

U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP). (2003). *2003 Mathematics Assessment*. Retrieved from http://nces.ed.gov/nationsreportcard/itmrlsx/detail.aspx?subject=mathematics

U.S. Department of Education, Institute of Education Sciences, National Center for Education

Statistics, National Assessment of Educational Progress (NAEP). (2006). *2006 Mathematics Assessment*. Retrieved from http://nces.ed.gov/nationsreportcard/itmrlsx/detail.aspx?subject=mathematics

U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP). (2012). *2012 Mathematics Assessment*. Retrieved from http://nces.ed.gov/nationsreportcard/itmrlsx/detail.aspx?subject=mathematics

Vygotsky, L.S. (1978). *Mind in Society. The Development of Higher Psychological Processes* (M. Cole, V. John-Steiner, S. Scribner, E. Souberman (Eds.)). Cambridge, MA: Harvard University Press.

Vygotsky, L.S. (1986). *Thought and language.* (Revised Edition, A. Kozulin [Ed.]), Cambridge, MA: MIT Press.

Werner, H., & Kaplan, B. (1963). *Symbol formation: An organismic developmental approach to language and the expression of thought*. New York: Wiley.

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition, 74*, B1-B11.

Appendix A: Interview Protocol and Script

Name: _____________ Grade: ____ Date & Time:_________________ Treatment: ________

Order of Set A:
Order of Set B:

Phase	Script
Intro	"As you already know, my name is Ms. Bona (or Ms. K) and I'm from UC Berkeley. I'm trying to understand how students like you think about math. Today, I'm going to ask you to solve some problems about numbers and number lines. You may have never seen some of these problems before, and that's okay. This isn't a test, and there isn't a right or wrong answer. Sometimes I'll ask you to explain what you wrote and how you thought about the problem to help *me* understand how people learn. Okay?" [Turn on camera.]
Orientation (empty line)	"This line shows a racecourse, and the fox, rabbit, and turtle are running from this side (left) all the way over to this side (right)." "The turtle has run just a small distance, and fox has run a large distance. Can you place the turtle and fox on the racecourse?" "The rabbit has run more than the turtle but less than the fox. Can you place the rabbit on the racecourse?" "Who has run the most distance in the race?" "Who has run the smallest distance in the race?" "Did the rabbit run more or less than the fox?" "Did the rabbit run more or less than the turtle?"
Orientation to number lines (2 problems)	"Okay, now we're going to look at some number line problems. For these problems I'm going to ask you to mark the place for one number on each problem." **TARGET VALUE:** "What is this number?" **[Point to target value on upper left corner and wait for child to state the number. Repeat the number child stated.]** **"**Can you point with your finger where you think [target value] should go on this number line?" **[Use pen to mark line where child points.]** "Why did you put [target value] there?" Ask follow-up questions to determine level of comprehension of activity.
Main: Tool absent condition (5 problems)	"Okay, now we're going to place animals on a racecourse. BUT this time, I'm going to ask you to sit on your hands or keep your hands on your lap so that you can't use them to help you. I'm going to move the animals across the racecourse like this [demonstrate], and you say 'STOP!' when you think the animal gets to the correct

	spot. OK?" "On this problem, [animal 1] has run x miles, and [animal 2] has run y miles. **What's this number? [point to target value]** [Animal 3] has run z miles! I'm now going to move the [animal 3]. Tell me 'STOP!' when [animal 3] gets to z miles. OK?"
Orientation to tool use	<u>Paper strips:</u> "Can you fold this piece of paper so that there are **TWO** pieces of the same size?" **"How did you do that? Can you show me?"** "Can you fold this piece of paper so that there are **FOUR** pieces of the same size?" **"How did you do that? Can you show me?"** <u>Ruler:</u> "Can you show me how **LONG** 4 centimeters is on the ruler?" "**Is there any other place** to show how LONG 4 centimeters is on the ruler?" **"How did you do that? Can you show me?"** "Can you show me how **LONG** 2 centimeters is on the ruler?" "**Is there any other place** to show how LONG 2 centimeters is on the ruler?" **"How did you do that? Can you show me?"** Take notes on what the child does to the tool, along with a quick yes/no evaluation of whether they were able to unitize the tool as asked.
Main: Tool present condition (5 problems)	"Okay, now we're going to place animals on a racecourse. BUT this time, I'm going to ask you to use this paper strip/ruler to help you measure. Place the animal on the racecourse once you think you know where the animal should be. After you place the animal, I will mark the animal's location on the racecourse with my pen." "On this problem, [animal 1] has run x miles, and [animal 2] has run y miles. **What's this number? (point to target value)** [Animal 3] has run z miles! Use the paper strip/ruler to help you measure and **find exactly where** [animal 3] is on the racecourse." "Can you explain how you used the paper strip/ruler to help you measure where the fox is on the racecourse?" Follow-up questions: "When you said you __________, why did you do that?" "Can you show me?" "Can you talk out loud while you show me?" [Turn off camera.]
Closing	"Thank you, that's the end of the interview. Your answers were really helpful! Do you have any questions you want to ask me?" "You may return to your classroom now. Thank you!"
Transitions	

	CHECK: - **Bell Schedule:** How much time left? Can we move to next student or wait? - **SD card:** How much time left? Need to switch to a new one? - **Camera power:** Is it still running on power adapter or battery? - **WRITE DOWN TIME on protocol (and any other info that's missing).** - **Label and paper clip student's work, stow away, get next one out.**