EXPLORING ATTENTION TO NUMERICAL FEATURES IN PROPORTIONAL REASONING: THE ROLE OF REPRESENTATIONS, CONTEXT, AND INDIVIDUAL DIFFERENCES

Michelle Ann Roddy Hurst

A dissertation

submitted to the Faculty of

the department of Psychology

in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Boston College
Morrissey College of Arts and Sciences
Graduate School

August, 2017

EXPLORING ATTENTION TO NUMERICAL FEATURES IN PROPORTIONAL REASONING: THE ROLE OF REPRESENTATIONS, CONTEXT, AND INDIVIDUAL DIFFERENCES

Michelle Ann Roddy Hurst

Advisor: Sara Cordes, PhD

Human infants show relatively sophisticated abilities to track and use proportional information. However, by the age of 6, children tend to make predictable errors in their proportional reasoning and later encounter significant challenges in many aspects of formal fraction learning. Thus, one of the central questions motivating this research is to identify the factors leading to these difficulties, in light of evidence of early intuitions about these concepts. In the current dissertation, I address this question by investigating the tradeoff between attending to proportional magnitude information and discrete numerical information about the components (termed “numerical interference”) across both spatial (i.e., area models, number lines) and symbolic (fractions, decimals) representations of proportion information. These explorations focus on young children (5-7 year olds) who have yet to receive formal fraction instruction, older children (9-12 year olds) who are in the process of learning these concepts, and adults who have already learned formal fractions. In Project 1, I investigated how older children and adults map between symbolic and spatial representations, particularly focusing on their strategies in highlighting componential information versus magnitude information when solving these mapping tasks. In Projects 2 and 3, I explore the malleability of individual differences in this numerical interference in 4- to 7-year-old children. Across the three projects, I suggest that although numerical interference does impact proportional reasoning, this

over-attention to number can be reduced through modifying early experiences with proportional information. These findings have implications for education and the way we conceptualize numerical interference more generally.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENTS

There are so many people and organizations I am eternally grateful for, without whose support this dissertation, as well as the rest of my graduate career, would not have been possible. First, of course – thank you to my committee. I have been extremely appreciative of your feedback and your mentorship at each stage of this process.

My advisor, Sara Cordes: for the many drafts and late night emails, and engaging with my often long-winded discussions about research and science—thank you. I look forward to continuing to work on our on-going projects and collaborations.

I am also indebted to the rest of the Boston College Infant and Child Cognition Lab, including the lab managers, grad students, and post-docs whose support, both in the lab and out, kept me sane and organized. Alison and Raychel: for being the best lab managers and allowing me to really rely on others when I needed to. Karina and Sophie: for having an amazingly friendly and collaborative approach to maintaining the lab (and having the occasional much needed happy hour). Nadia: for being the best office mate and having an informal professional development discussion almost every Tuesday. And, in particular, my many, many, amazing and dedicated research assistants. I would need an entire extra page to list you all—but please know how much I appreciate all the hours you spent recruiting, collecting data, and coding. This truly would have been impossible without you.

In addition to the ICCL, the faculty and grad students at Boston College in both the Psychology Department and the Lynch School of Education have played a large role in my developing abilities as a scientist and a researcher—from thoughtful questions at

Grad Research Day, to small meetings and discussions, to teaching me about design and statistical principles, I very much appreciate all of the interactions I've had while at BC.

Of course, I am particularly thankful to all of the families who participated in our studies, as well as our off-site testing partners who helped us reach even more children and parents than we possibly could have on our own: Living Laboratory at the Museum of Science, Boston, MA; Boston Children's Museum, Boston, MA; Acton Discovery Museum, Acton, MA; many schools, after school programs, day cares, and summer camps from the greater Boston, MA area.

Thank you to the Natural Science and Engineering Research Council of Canada (NSERC) and the Boston College Department of Psychology for funding my research and me over the years.

Thank you to my family—for encouraging me from a young age to ask questions, and when you don't know the answer, to figure it out.

Last, but not least, James. I cannot express how much I have appreciated your support throughout this journey. Your partnership has allowed me to do more than I ever thought possible and I so look forward to the rest of our adventures.

CHAPTER 1: GENERAL INTRODUCTION

Human infants and children (McCrink & Wynn, 2006; Denison & Xu, 2010), as well as non-human animals (Rakoczy, Cluver, Saucke, Stoffregen, Grabener, Migura, & Call, 2014; Rugani, McCrink, de Hevia, Vallortigara, & Regolin, 2016; Tecwyn, Denison, Messer, & Buchsbaum, 2016) show substantial abilities to attend to the relation between two quantities. That is, not only can infants, children, and adults represent discrete numerical information about a single quantity (e.g., Cordes & Brannon, 2008; Barth, Kanwisher, & Spelke, 2003; Barth, Lamont, Lipton, & Spelke, 2005; Halberda & Feigenson, 2008), they can also represent the relation between two quantities, in the form of a ratio or proportion. For example, infants as young as 6 months old can track the ratio of two quantities and use proportional information to make inferences about the probability of certain outcomes (McCrink & Wynn, 2006; Denison, Reed, & Xu, 2013). Despite these early intuitions, children encounter substantial difficulty when learning about formal fractions and decimals in elementary and middle school (National Mathematics Advisory Panel, 2008; Ni & Zhou, 2005; Lortie-Forgues, Tian, & Siegler, 2015; Resnick, Nesher, Leonard, Magone, Omanson, & Peled, 1989). Thus, a major question that motivates my work is: *Why is there such a disconnect between children's early intuitions about probability and the formal fractions (and related concepts) they learn in school*?

One potential factor impacting the difficulty of understanding formal proportional information is the incredible variability in how it can be represented. Proportion can be represented via two distinct symbolic notations (fractions and decimals). Moreover, proportion can also be represented spatially using both area models (e.g., pie charts) and linear models (e.g., a number line). Thus, learning formal proportion requires a flexible ability to perceive proportional information across a variety of contexts and representations.

However, the variability in these representations is not random: although all proportional representations provide information about proportional magnitude, only some also provide discrete numerical information about the parts. Thus, we can group these representations into two categories: those that preserve discrete information about the components and those that represent proportion in a continuous manner. For example, spatial models can either be discrete (e.g., a rectangle, divided into equal-sized pieces) or continuous (e.g., an undivided rectangle, where the number of parts cannot easily be counted). In the same vein, fraction notation provides information about the number of parts out of the whole, whereas decimal notation simply denotes proportional magnitude, without information about the number of parts. Importantly, substantial research has shown that when discrete numerical information about the components is available, children and adults have substantial difficulty attending to the proportional magnitude (e.g., Atagi, DeWolf, Stigler, & Johnson, 2016; Boyer, Huttenlocher, & Levine, 2008; DeWolf, Bassok, & Holyoak, 2014; Hurst & Cordes, 2016; Jeong, Levine, & Huttenlocher, 2007; Ni & Zhou, 2005). As such, they may be more inclined to indicate that 4/9 is larger than 2/3 (when comparing fraction notation or spatial representations), simply because the former has a greater number of salient parts (4>2). This tendency to focus on discrete numerical information, at the expense of proportional magnitude, can be described as *numerical interference.* The central topic of the current dissertation is whether the tendency for whole number information to interfere or compete with proportional magnitude is malleable. In other words, when using representations that preserve both magnitude and discrete information about the parts, can people's attention be turned toward the magnitude information and away from the parts?

Spatial Representations

Recent guidelines for fraction instruction suggest including multiple visual representations for proportional information, including area models (e.g., a pie chart or rectangle) as well as number line models (National Governors Association Center for Best Practices, 2010). Research with whole numbers has suggested that linear representations may have substantial benefits for learning whole numbers (Ramani & Siegler, 2008; Siegler & Ramani, 2009). However, given that proportional information involves both holistic magnitude and component parts, it is less clear whether number lines would also be the best spatial representation for learning proportion, or whether other representations (e.g., area models) may also be particularly useful.

In line with findings from whole numbers that number lines may be particularly useful, some evidence indicates that number lines are better for teaching children about fraction magnitudes than pie charts (Hamden & Gunderson, 2017; Wang & Siegler, 2013), possibly because number lines allow for a more transparent way to compare the numerical magnitudes of fractions, decimals, and whole numbers within the same visual representation (Cramer, Post, & DelMas, 2002; Saxe, Diakow, & Gearhart, 2013). For example, in one study teachers implemented an experimental curriculum that included emphasizing relative order and unit intervals on a number line with both fractions and integers (Saxe et al., 2013). Students who received the experimental curriculum outperformed their business-as-usual peers on a variety of problems involving integers and fractions, as well as problems with and without number lines (Saxe et al., 2013).However, the experimental curriculum differed from the business-as-usual curriculum in other ways, beyond just the spatial representation of a number line, including the content of the fraction lessons, which had an emphasis on core definitions involving units and order (Saxe et al., 2013). Thus, it remains an open question how critical the number line representation is per se, as opposed to the more general benefits of the concepts the number line representation was used to teach.

On the other hand, when learning about the part-whole components of proportional information, pie charts and other area models may actually be a more useful representation (Cramer, Wyberg, & Leavitt, 2008). In particular, the Rational Number Project (RNP), one of the largest fraction curriculum reform projects, has found the use of area models in emphasizing fraction concepts about part-whole to be effective (e.g., Cramer, Post, & delMas, 2002; Cramer & Wyberg, 2009). Furthermore, pie charts may be particularly useful for representing relative amounts in graphical representations (Shah & Hoeffner, 2002). Thus, pie charts may be more useful for representing information about the components, rather than the proportional magnitude.

However, there are still several open questions about how specific this preference might be. That is, although pie charts are an area model, with obvious information about the part-whole components, can they be interpreted in terms of both proportional magnitude and the specific components, depending on the context? Similarly, are number lines approached using a part-whole perspective that allows individuals to think about the component parts in addition to the proportional magnitude?

Symbolic Representations

Children are taught two formal symbolic representations for proportion, fractions and decimals, both of which pose substantial difficulty for them. For example, the bipartite structure of fraction notation (i.e., a/b) often leads children to treat fraction symbols as two distinct whole numbers, rather than a coherent unit (e.g., Ni & Zhou, 2005). Similarly, children often incorrectly assume that "longer decimals = larger values" (e.g., 0.312 > 0.43) and misunderstand the role of zero (e.g., in 0.05 vs. 0.5 vs. 0.50; Desmet, Gregoire, & Mussolin, 2010; Durkin & Rittle-Johnson, 2015). Importantly, these two symbolic notations are not only distinct in terms of the challenges they pose to young learners, but they also convey different amounts of information. Decimal notation only represents magnitude and does not retain any information about the specific components (e.g., 2 out of 4 and 4 out of 8 are both represented as 0.5). On the other hand, fraction notation represents specific information about the components as well as magnitude information (e.g., 2/4 and 4/8 represent equivalent magnitudes, but have different numerical components).

Given these differences, it is unsurprising that people tend to favor some symbols over others, dependent upon the context, as each notation has distinct advantages. In particular, decimal notation has been found to be the most transparent means for representing magnitude information (DeWolf, Grounds, Bassok, & Holyoak, 2014; Hurst & Cordes, 2016), particularly in continuous (as opposed to discrete) contexts (Rapp, Bassock, DeWolf, & Holyoak, 2015). Fraction notation, on the other hand, tends to be preferred when discrete information about the components is relevant (DeWolf et al., 2014; Rapp et al., 2015) and when adults are performing rational number arithmetic (Hurst & Cordes, 2016).

Thus, the distinct formal symbolic notations taught in the classroom have both implicit and explicit differences in their relative affordances. Critically, the retention of part-whole information in fraction notation lends itself to greater numerical interference of the component parts when judging holistic magnitude (e.g., Bonato, Fabbri, Umilta, & Zorzi, 2007). However, it is not impossible for people to access magnitude information from fractions. Rather, children and adults show evidence of being able to represent the magnitudes of fractions, decimals, and whole numbers along an integrated mental continuum when the use of componential information from fractions is prevented (e.g., Bonato et al., 2007; Schneider & Siegler, 2010; Ganor-Stern, 2013, 2012; Hurst & Cordes, 2016; Hurst & Cordes, under review - B).

Yet, there are still several open questions about how well people can think about the magnitudes of fractions in particular, given that they can preserve both magnitude and part-whole information. Given that number lines are particularly oriented toward magnitude information, one way to emphasize the magnitude of fractions may be to map them to number lines (e.g., Hamden & Gunderson, 2017). Investigating how numbers are translated onto number lines has been a fruitful area of research for whole numbers (Barth & Paladino, 2011; Berteletti, Lucangeli, & Zorzi, 2012; Booth & Siegler, 2006; Cohen & Blanc-Goldhammer, 2011; Siegler & Opfer, 2003). For proportional information, this becomes particularly relevant because of the extensive use of spatial representations, both in terms of area models and linear models, for representing proportion in the classroom.

Mapping Between Symbolic and Spatial

Given that the common symbolic (fractions, decimals) and spatial (number line, area model) representations of proportion show differences in how transparently they convey magnitude information, it may be that the mapping between these symbols and spatial representations also show critical differences in how well they align with each other. Evidence reveals that adults may be particularly error-prone when mapping between a number line and a fraction (Hurst, Relander, & Cordes, 2016) and prefer to align decimals with number lines and fractions with pie charts when mapping between symbols and space (Hurst, Santry, Relander, & Cordes, in prep). This is likely due to differences in the type of information best conveyed by each representation: both decimals and number lines are thought to be preferred for conveying magnitude, but not part-whole information (which is conveyed by fraction notation; DeWolf et al., 2014; Hurst & Cordes, 2016; Hurst et al., in prep). Thus, although evidence suggests that each of these representations may be particularly beneficial for distinct contexts, how symbolic and spatial representations may be used in conjunction with each other and, in particular, how spatial representations may impact people's understanding of symbolic representations are relatively open questions. In Project 1, I explore how fraction notation, specifically, aligns with distinct spatial representations. In particular, I investigate how adults spontaneously think about fraction magnitudes and whether children and adults can be encouraged to think about fraction magnitude when primed with a number line, rather than a pie chart.

Discrete and Continuous Representations

Unlike whole-number numerical quantity, proportional magnitudes represent the *relation* between two quantities. Importantly, the two component quantities can be either continuous (e.g., representing the size of parts) or discrete (e.g., representing the number of parts). For example, proportion in a continuous area model may be judged as the relative amount of red area compared to the amount of blue area in a red and blue rectangle. Alternatively, this same area model can be discretized by partitioning (divided into equal-sized pieces), so that the relation becomes the number of red pieces versus the number of blue pieces. Although this perceptual distinction is minor, the continuous or discrete nature of a proportional representation can substantially change the way that children attend to proportional information. In particular, when asked to judge the relative proportion of two area models, children tend to perform significantly worse when the proportional amounts are presented with discrete, divided up shapes than when they are presented continuously, without broken up pieces (e.g., Boyer et al., 2008; Jeong et al., 2007). Furthermore, this numerical interference is only evident when all the area models involved are discrete, allowing for a comparison of number as well as proportion (Boyer et al., 2008). This suggests that this difficulty does not just stem from the presence of discrete information, but is specifically about the reliance on a numerical equivalence strategy instead of a relational strategy when number matching is made possible.

Although this numerical interference in the presence of discrete numerical information is robust, emerging as early as age 6, the mechanisms responsible for numerical interference are relatively unexplored. Why is it that children and adults preferentially attend to numerical information at the expense of proportional magnitude? In particular, what factors may impact numerical interference and is it malleable or preventable? On one hand, whole numbers are a central component of early math learning. Early number knowledge is entirely focused on counting and mapping number words and symbols to discrete quantity (LeCorre & Carey, 2007; Condry & Spelke, 2008; Hughes, 1986; Hurst, Anderson, & Cordes, 2017; Sinclair, Siegrist, & Sinclair, 1983). Furthermore, whole number information is so important that recent studies have found that preschoolers' tendency to spontaneously attend to number is predictive of their math ability (Gray & Reeves, 2016). Thus, it may be that this tendency to focus on the numerical features (i.e., numerical interference) is a necessary component of learning early math concepts, which has a necessary emphasis on whole numbers, and that it is not until after formal instruction and deep conceptual learning that children are able to overcome this numerical interference and attend to proportional magnitude.

On the other hand, it may not be that whole numbers are *over*-emphasized, but rather that proportional information is *under*-emphasized. Attending to relative amounts may also be a particularly important early developing skill (Denison, Reed, & Xu, 2013; Duffy, Huttenlocher, & Levine, 2005). Furthermore, number may not always be extremely salient, but instead depends on the saliency of the other relevant features (Chan & Mazzocco, 2017), which may suggest that emphasizing proportion may make number less salient in these contexts. Thus, numerical interference may not arise because of an overt focus on number, but rather numerical interference may be evident because proportional information is not salient or available to the child in that context (Alibali & Sidney, 2015). If this is the case, then we may be able to encourage children to attend to proportional magnitude by increasing the saliency of proportional information within a specific context.

Lastly, recent evidence suggests that children's spontaneous focusing on relations between quantities has been shown to predict formal fraction understanding (McMullen, Hannula-Sormunen, Laakkonen, & Lehtinen, 2016). As such, there may be important individual differences in children's tendency to attend to proportional information. These ideas will be explored in Projects 2 and 3.

The Current Dissertation

The presence of numerical interference in the context of representations that preserve both magnitude and part-whole information about the components is well established. However, there are still several open questions about how easily people can attend to magnitude information, even when componential information is available.

In this dissertation, I present three projects, with a total of six experiments, investigating adults' and children's understanding of proportional magnitude across various representations (e.g., pie charts and number lines; fractions and decimals; discrete and continuous area models; verbal labels), representations known to convey proportional magnitudes and information about the components. In so doing, I am able to shed light on the specific kinds of representations that may be more beneficial for communicating information about the holistic magnitude than information about the components, as well as provide some insight into the underlying structure of proportional representations more generally.

Thus, in all three projects, I investigate how children and/or adults think about proportional magnitude using distinct external representations. In Project 1, I focus on how formal representations of magnitude, in the form of fractions, decimals, number lines, and pie charts, are used to represent magnitude. To do so, I investigate the spatial representations that people rely upon when judging the proportional magnitude represented by fraction symbols (Project 1, Experiment 1), as well as whether priming children and adults to think about fractions using number lines (which emphasize relative magnitude), as opposed to pie charts (which retain part-whole information) impacts their ability to judge the relative magnitude of fraction symbols (Project 1, Experiments 2 and 3).

In Projects 2 and 3, I turn toward investigating the factors that impact children's numerical interference much earlier in development, well before children have learned formal fractions. In Project 2, I investigate individual differences in children's proportional reasoning and whether children's experiences with continuous proportional information can decrease their focus on discrete components and toward proportion magnitude. Building on this work, in Project 3 I investigate how formal fraction labels (e.g., "three-fourths") may be implicated even in children's early processing of non-symbolic visual proportional information, well before they have formally learned fractions. Although the tasks in all three projects are focused on the context of proportion *magnitude*, the availability of discrete numerical information about the parts is manipulated or compared across distinct representations in order to investigate how the availability of information about the parts may be at odds with information about the magnitude.

Across all three projects, I will argue that the perceptual and semantic make up of the representations (e.g., discrete vs. continuous; linear vs. area; numerical labels vs. categorical labels) impact the way these representations are used, precisely because they signal whether the relevant information to be processed is holistic proportional magnitude or the numerical information from the components. However, other aspects of the context or other available representations (e.g., verbal labels) may be able to direct attention toward the relevant information.

CHAPTER 2: FRACTION MAGNITUDE: MAPPING BETWEEN SYMBOLIC AND SPATIAL REPRESENTATIONS OF RATIO

Abstract

Recent research in children's understanding of rational numbers has focused on children's understanding of fraction magnitude. This research has led to several recent studies investigating the usefulness of using a number line for teaching fractions. However, fraction notation is not exclusively used for magnitude information and in fact, may be particularly poorly suited for magnitude information. We report three experiments investigating how adults and children think about symbolic fraction magnitudes in relation to number lines and pie charts. Our data reveal that adults do not tend to spontaneously rely on number line representations, and that practice with mapping rational numbers to number lines did not improve performance on a subsequent symbolic magnitude comparison task relative to practice with mapping the same magnitudes to pie charts. However, adults and children showed evidence of distinct external strategies for working with number lines compared to pie charts. For both fractions and decimals, people were more likely to use a magnitude estimation strategy with number lines and a part-whole partitioning strategy with pie charts. We discuss the limitations and implications of these findings for future research and education.

Formal fractions pose a substantial challenge for many students, who make both procedural and conceptual errors throughout fraction education and beyond (e.g., Christou & Vosniadou, 2012; National Mathematics Advisory Panel, 2008; Ni & Zhou, 2005; Lorti-Forges, Tian, & Siegler, 2015; Vamvakoussi & Vosniadou, 2010). However, in addition to this difficulty, learning and understanding fraction concepts is particularly critical for later math skills, including Algebra (Booth & Newton, 2012; Siegler et al., 2012). Thus, it is essential to understand the specific difficulties children face and investigate ways to help children overcome them. One particularly critical aspect of fractions is that they represent a relation between two quantities and this relation has a numerical magnitude that can be represented on a continuum with other numerical magnitudes. As such, there has been a recent emphasis in both research (Siegler, Fazio, Bailey, & Zhou, 2013) and educational practices (National Governors Association Center for Best Practices, 2010) on children's understanding of fraction *magnitude* in particular.

Although there has been more recent attention to fractions, the majority of work investigating how children and adults think about numerical magnitude has been in the context of whole numbers. This research suggests that children and adults represent whole numbers as approximate, ordered magnitudes. This evidence often comes from performance on symbolic numerical comparison tasks in which the participant is asked to decide which of two numbers is largest as accurately and quickly as possible. These studies reveal ratio-dependent responding, such that the speed and accuracy with which a person responds is dependent upon the distance between the numbers and size of the numbers (together resulting in a ratio effect; Moyer & Landauer, 1967, 1973; Sekuler & Mierkiewcz, 1977). Critically, this pattern of responding suggests that these representations are both noisy (i.e., the representation of 6 is not exact, but instead also bleeds into the representations of other numbers) and ordered (i.e., 6 overlaps more with 7 than with 10; Moyer & Landauer, 1967, 1973).

This finding of ratio-dependence is robust for whole number comparisons, but has only recently been demonstrated in comparisons involving other kinds of rational numbers including fractions and decimals. This research suggests that adults and children are able to represent approximate magnitudes when comparing two fractions, two decimals, and when comparing across notations (i.e., fraction vs. decimal, fraction vs. whole number, decimal vs. whole number; DeWolf, Grounds, Bassok, & Holyoak, 2014; Faulkenberry & Pierce, 2011; Ganor-Stern, 2013, 2012; Hurst & Cordes, 2016; Hurst & Cordes, under review - B; Meert, Gregoire, & Noel, 2010; Kallai & Tzlegov, 2012; Schneider & Siegler, 2010; Sprute & Temple, 2011; Varma & Karl, 2013; Wang & Siegler, 2013; although there are some contexts where accessing the magnitude represented by a decimal or fraction can be more difficult, e.g., Bonato, Fabbri, Umilta, & Zori, 2007; Kallai & Tzelgov, 2012). Thus, these findings lead researchers to conclude that adults and children are able to represent the magnitudes of fractions, decimals, and whole numbers in an integrated way.

This evidence of an approximate and ordered "mental number line" for whole numbers has also prompted research in how whole numbers are mapped onto space. For example, behavioral evidence revealing people are faster to make magnitude judgments when small numbers are on the left and large numbers are on the right (e.g., Dehaene, Bossini, & Giraux, 1993; Gevers, Lammertyn, Notebaert, Verguts, & Fias, 2006; Nuerk, Wood, & Willmes, 2005) are consistent with the idea of an ordered mental number line.

Building on these findings, studies have found that children's numerical abilities benefit from experience using visual displays involving a linear spatial representation to learn about numbers. That is, researchers have shown that playing linear board games can improve children's numerical abilities, beyond that of playing circular board games, emphasizing the importance of a linear external representation for teaching and learning numbers (Ramani & Siegler, 2008; Siegler & Ramani, 2009). Less work, however, has investigated the potential benefits or weaknesses of various spatial representations of fraction magnitudes. Given that area models (e.g., pie charts) and number lines are frequently used in the classroom (National Governors Association for Best Practices, 2010), understanding the impact of these visual representations on children's understanding of fraction concepts would be important for promoting theoretical understanding of how children think about these numbers as well as for educational practices.

Some recent evidence suggests that fraction instruction using number lines may be better than instruction using more traditional area models, including pie charts (Cramer et al., 2002; Hamdan & Gunderson, 2017; Keijzer & Terwel, 2003; Saxe et al., 2013; Wang & Siegler, 2013). One of the reasons number lines may be more beneficial for understanding fraction magnitudes is that they draw connections between fractions, whole numbers, and decimals (Cramer et al., 2002; Saxe et al., 2013), something that is very difficult for children to do (e.g., Vamvakoussi & Vosniadou, 2010). On the other hand, some evidence investigating different kinds of graphical displays (e.g., bar charts versus pie charts) suggest that pie charts may be particularly useful for conveying *relative* amounts, but not absolute amounts (Shah & Hoeffner, 2002). Other work, however, suggests that number lines and area models (e.g., pie charts) may show distinct advantages and disadvantages. Research suggests that fraction notation is most beneficial for providing discrete, part-whole information, but not magnitude information (DeWolf et al., 2015; Rapp et al., 2015). As such, adults are more likely to represent a fraction using a pie chart than a number line (Hurst, Santry, Relander, & Cordes, in prep) and are more accurate when mapping between fraction symbols and pie charts than between fractions and number lines (Hurst, Relander, & Cordes, 2016). Decimals, on the other hand, are generally preferred for representing continuous numerical magnitudes (DeWolf et al., 2015; Hurst & Cordes, 2016) and are more likely to be associated with number lines (Hurst, Santry, Relander, & Cordes, in prep). Together, these findings suggest that number lines (which may be easily aligned with decimals, but not fractions; Hurst et al., in prep) may be particularly useful for representing numerical magnitude, whereas pie charts (aligned with fractions, but not decimals; Hurst et al., in prep) may be more useful for representing information about the components or the part-whole structure.

Given that fractions can represent both magnitude information and information about the discrete components and that number lines may more readily convey magnitude information, in the current study we explored whether thinking about fractions on a number line would lead adults and children to more readily think about the *magnitudes* associated with symbolic fractions. We investigated this central question by first assessing adults' spontaneous strategy use during a magnitude comparison task (Experiment 1) and then by providing children and adults with short mapping activities involving either number lines or pie charts before completing a symbolic magnitude task (Experiments 2 and 3). Thus, across three experiments we investigated whether adults' spontaneous use of number line based visualization strategies is associated with better magnitude performance, and whether priming children and adults to think about fractions as values on a number line may lead to better symbolic magnitude comparison than priming children and adults to think about the part-whole nature of fractions through an area model, specifically a pie chart.

Experiment 1

We investigated what visualization strategies adults spontaneously engage during a symbolic magnitude comparison task and how these strategies may be related to symbolic magnitude understanding. In particular, this study addressed the following research questions: How do adults spontaneously visualize fraction magnitudes? Does the type of visualization strategy relate to abilities to process the magnitude of symbolic fractions? In particular, do those individuals who visualize fractions as magnitudes falling along a number line perform better on a fraction magnitude comparison task than those who think about fractions as part of a whole using an area model (e.g., on a pie chart)?

Method

Participants. Fifty adults (M_{age} = 19.2 years, Range 18 to 24 years, 39 females) participated for partial course credit.

Measures. Adults completed all tasks in the same order on a 13-inch MacBook laptop using Xojo programming software (formerly named REALBasic): (1) magnitude comparison task and (2) strategy questionnaire. Adults received an additional speeded magnitude addition task with fractions in between these two tasks, however this task will not be reported here. Adults were tested one-on-one in a quiet room in our laboratory.

The experimenter remained in the room for the magnitude comparison task but left during the strategy questionnaire.

Magnitude Comparison Task. All adults received three blocks of trials that differed in the notation of the stimuli being compared: two fractions (FvF; e.g. 3/5 vs 2/9), one fraction and one decimal (DvF; e.g. 3/5 vs 0.22), and one fraction and one whole number (NvF; e.g. 4/3 vs 2). In each trial, two numbers were presented on the screen, one on the left and one on the right, and participants were asked to choose which of the two numbers was largest as quickly and accurately as they could by pressing the corresponding key on the keyboard (right or left arrow). The stimuli remained on the screen until the participant made a response and then a small fixation cross appeared in the middle of the screen for 1000ms until the trial began. At the beginning of each block, participants were given one notation-specific practice problem (the same problem across participants) with computerized feedback as to their accuracy. After the practice problem, participants were invited to ask additional questions or clarify the task. On test trials, no feedback was given.

The order of trials within each block and the order of the blocks were randomized across participants. Stimulus pairs on each trial came from one of two ratio bins (ratio = larger stimulus / smaller stimulus): small ratio bin (ratios ranged from 1.35 to 1.51) and large ratio bin (ratios ranged from 2.2 to 2.9). For each Ratio (2) X Block (3) combination there were four unique comparisons, shown twice (once with the largest number on the left and once with the largest on the right). Thus, there were a total of 48 trials (4 unique comparisons X 2 ratios X 2 (shown twice) X 3 blocks).

The decimal stimuli (used in the DvF block) were presented to the hundredths digit, with a whole number before each decimal point (e.g. 0.15; 1.36). The fraction stimuli (used in all blocks: DvF, NvF, and FvF) had numerator and denominator values each less than or equal to 10. The two fractions for each comparison in the FvF block were made up of four different integers (i.e., a/b vs. c/d where a, b, c, d were all different positive integers) in order to prevent the use of whole number strategies (e.g., Schneider & Siegler, 2010). In the DvF block, fractions ranged from 1/5 to 5/3 and decimals ranged from 0.22 to 3.5. In the NvF block, fractions ranged from 6/5 to 9/2 and whole numbers ranged from 1 to 6. In the FvF block, fractions ranged from 1/5 to 7/2. See the Supplementary Appendix 2 for a complete set of stimuli values. Decimals were approximately 2cm high x 5.5cm wide, Fractions were approximately 5.5 cm high and 2.7 cm wide, and Whole Numbers were approximately 2 cm high and 1.2 cm wide. The fixation cross between trials was approximately 0.5cm high x 0.5cm wide, in the center of the screen.

Strategy Questionnaire. Lastly, adults were presented with the three questions: (1) "In what way do you think about fractions? In other words, when you think about a fraction (for example, 1/2 or 4/5) how do you visualize it?" (2) "If you were explaining fractions to someone, which visual references would be best to use?" and (3) "What kind of visual references do you remember learning fractions with most?" For questions 2 and 3, participants were provided with the options of a) Pie Chart, b) Number line, or c) No visual aid. For question 3, adults were provided with the additional option of "Other", with a space to provide more details. Adults were asked to respond as honestly as possible.

Data Coding. Accuracy (proportion correct) and reaction time (RT) were measured and analyzed on the magnitude comparison task and the speeded addition task. Only RTs from correct responses and those within three standard deviations of the individual's average RT were included in the analyses. At the individual level, in order for average RT for each participant to accurately represent the speed in which they processed symbolic magnitudes, adults who performed at or below chance (4/8 or below) or who had fewer than three included trials (i.e., that were correct and within three SDs of their average RT) were excluded (similar criteria to those used in Hurst & Cordes, 2016). At the group level, values more than three standard deviations away from the group mean were considered outliers and were replaced with the next higher value within the acceptable range (3 SDs of the mean). This resulted in the replacement of less than 1% of the data and a final sample of only 38 (out of 50) participants having complete and useable RT data. Given the lower samples of useable RT data, we also analyzed adults' accuracy data (proportion of trials correct).

The responses from the visualization question on the strategy questionnaire (question 1) were coded based on three major themes: (1) visual area model strategies, which included any description that involved estimating the magnitude using an object or image involving parts of a whole shape or object, such as a pizza, a pie chart, or a rectangle with shaded in sections, (2) number line strategies, which included any strategy visualizing a number line or continuum, and (3) symbolic estimation strategies, which included any responses that involved estimating the magnitude using only symbolic methods, for example estimating the proximity of the value to anchors like 1/2 or 1 and/or converting to an approximate magnitude in decimal form. Some participants reported other strategies that did not fit into any of these categories, for example "one number on top of another" or "as a ratio". Participants' responses may have fallen into more than one of these categories, and those were coded as such. Two independent coders coded all responses and disagreed on 8/50 of the responses. Disagreements were discussed and settled by a third coder.

Results

Magnitude Comparison Task. Accuracy and RT on the magnitude comparison task were analyzed using two different repeated measures ANOVA with notation (3: FvF, NvF, and DvF) and ratio (2: small, larger) as within subject factors (see Figure 2.1).

For accuracy, there was a main effect of notation, $F(2, 98) = 6.7$, $p = 0.002$, *partial* $\eta^2 = 0.12$, with performance on FvF ($M = 0.83$) being significantly lower than NvF, $M = 0.89$, $t(49) = 3.2$, $p = 0.002$, and DvF, $M = 0.88$, $t(49) = 2.5$, $p = 0.016$, which were not significantly different from each other, $t(49) = 1.3$, $p = 0.2$. There was also a main effect of ratio, $F(1, 49) = 50.4$, $p < 0.001$, *partial* $\eta^2 = 0.5$, with higher accuracy on the large ratio, $M = 0.92$, than the small ratio, $M = 0.81$. Furthermore, there was no significant interaction between notation and ratio, $F(2, 98) = 1.4$, $p < 0.25$, *partial* $\eta^2 = 0.03$.

When looking at RT (with only n = 38), the pattern of results was identical. There was a main effect of notation $F(2, 74) = 13.1$, $p < 0.001$, *partial* $\eta^2 = 0.26$, with RT on FvF trials, $M = 2013$ms, taking significantly longer than both NvF, $M = 1702$ms, $t(37) = 4.9$, $p<0.001$, and DvF, $M = 1804$ms, $t(37) = 4.1$, $p<0.001$, which were not significantly different from each other, $t(37) = 1.5$, $p<0.15$. Again, there was also a main effect of ratio, $F(1, 37) = 19.7$, $p < 0.001$, *partial* $\eta^2 = 0.35$, with the small ratio, $M = 1941$ms, taking significantly longer than the large ratio, M = 1738ms. Furthermore, there was not a significant interaction between notation and ratio, $F(2, 74) = 1.14$, $p = 0.3$, *partial* η^2 = 0.03.

Figure 2.1: Adult's Accuracy and Reaction Time on the Experiment 1 Magnitude Comparison Task

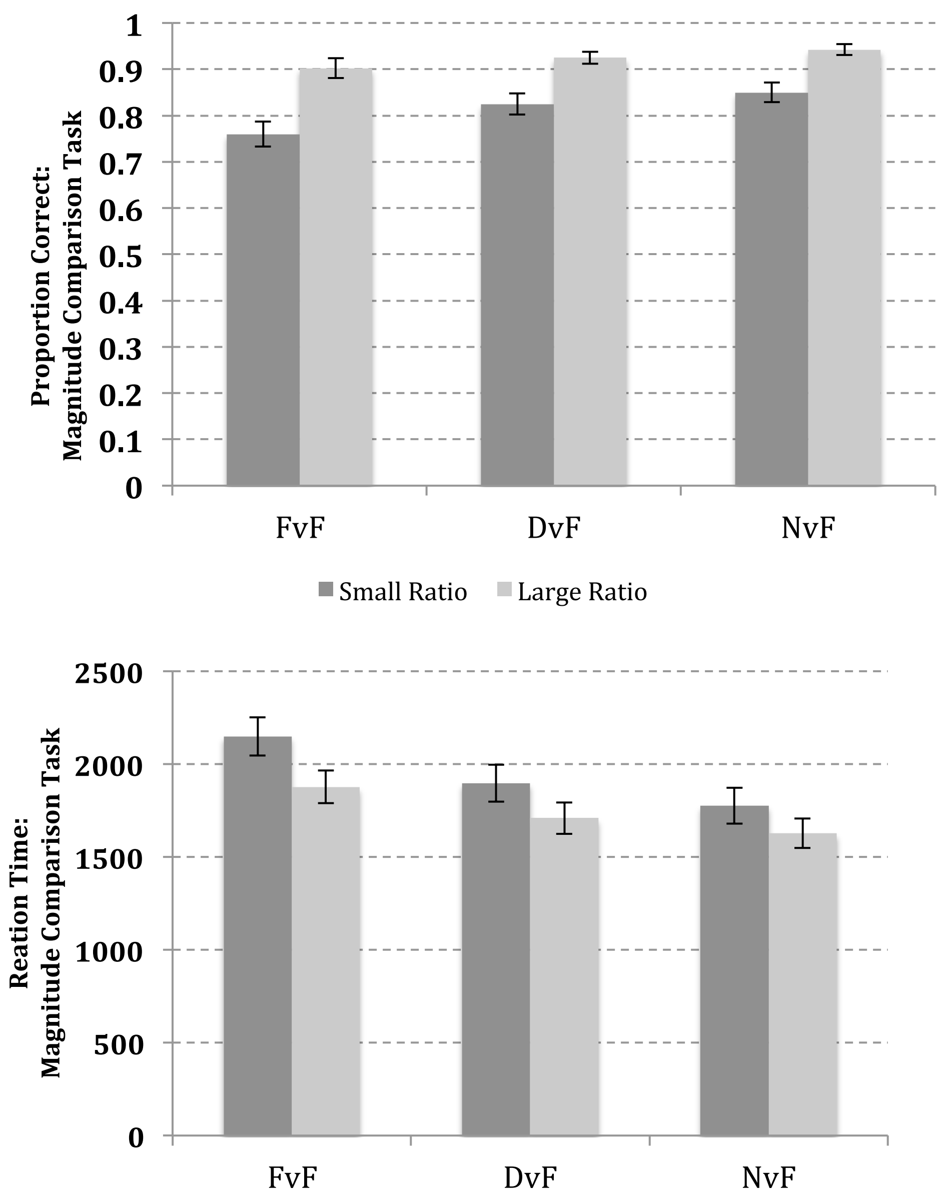


Figure 2.1: Performance on the magnitude comparison task of Experiment 1, broken down by ratio and notation for both dependent variables: Accuracy (upper panel) and Reaction Time (lower panel)

Strategies. In line with work suggesting that aligning fractions and number lines is particularly difficult (Hurst, Santry, Relander, & Cordes, in prep), only *one* individual (2%) reported using a number line in response to our first strategy question. In contrast, 64% of individuals (n=32) reported using a visual area model (e.g., imagining a pie chart) and 46% (n=23) reported using a symbolic estimation strategy (e.g., estimating decimal form; note, however that 10 adults are included in both groups as they reported both strategies). In addition, 8% (n=4) reported using a strategy that did not fit into the categories (e.g., "one number on top of another", "as a ratio").

Responses to the second question aligned with this pattern of results. That is, when asked to pick (of three options) which would be the *best* visual reference to use to explain fractions to someone else, almost all participants reported a pie chart: 84% (n=42), with fewer reporting a number line (14%; n=7) or using no visual reference (2%; n=1). Similarly, when asked which visual references they remember learning with (note that adults could select more than one option for this question), 82% (n=41) said they remember using a pie chart and only 22% (n=11) reported using a number line (note, again, that n=7 of these adults fall in both categories). In addition, 2% (n=1) reported not remembering any visual reference and 14% (n=7) selected "other". Three of the "other" respondents reported other strategies in addition to the pie charts: real world examples, shaded objects, or sets of discrete objects, two other respondents reported using discrete objects only and/or money, and two other respondents reported using base-10 blocks (sets of stackable blocks that can be organized into sets of 10 to easily communicate place-value).

Performance Differences Across Strategies. Originally, our primary interest was to investigate whether there were individual differences in magnitude understanding between people who opt to use area models (e.g., pie chart) versus number line visual references to think about fractions. However, almost none of the participants reported a number line visualization strategy and instead area models were the primary visual figure strategy adults reported, making it impossible to make any meaningful inferences from the data between these two strategy choices. Yet many adults did report using a symbolic estimation strategy. Given that it has been suggested that the primary benefit of the number line is that it provides a transparent way to integrate rational numbers of different notations (i.e., decimals, fractions, and whole numbers), it was hypothesized that adults who engaged in symbolic estimation strategies would be at an advantage compared to those who used part-whole visualization strategies, despite the fact that they do not involve a visual figure aid.

To explore this hypothesis, we compared performance on the magnitude comparison task between those who spontaneously reported using only area model visualization strategies and those who spontaneously reported using only symbolic estimation strategies. Notably, given that 20% of adults (n=10) reported using both of these strategies, we isolated this analysis to only those adults who reported *only one* of these strategies. In line with our predictions, proportion correct on the magnitude comparison task was higher for those who reported using a symbolic estimation strategy (n=13), M=0.9, compared to a visual area model strategy (n=22), M=0.82, $t(23.3)$=2.49, p=0.02. Given the already low samples in each of these groups, we opted not to look at

RT differences as that would result in an even lower sample (12/35 of these participants did not meet the criteria for inclusion in RT analyses).

Discussion

Results of our comparison task replicate previous work showing ratio-dependent responding, in terms of both accuracy and RT, when adults compared fractions and whole numbers, fractions and decimals, and two fractions (e.g., Ganor-Stern, 2013; Hurst & Cordes, 2016; Schneider & Siegler, 2010; Faulkenberry & Pierce, 2011). Together, these findings are consistent with claims that adults consider fractions as falling along an integrated continuum with whole numbers and decimals.

In addition, performance when comparing two fractions was significantly less accurate and slower than when comparing one fraction with a whole number or decimal, again replicating patterns in other work (Hurst & Cordes, 2016). This finding suggests that the difficulties encountered when thinking about rational number magnitudes may be particularly tied to the fraction notation itself. That is, comparing two fractions within the same notation was more difficult than comparing a fraction to a value in a different notation. Thus, the potential difficulty caused by comparing across notations was less than the difficulty caused by the fraction notation itself.

Moreover, adults' strategies and preferences for thinking about and learning fractions were consistent with findings from previous studies indicating an alignment between fractions and area models. Self-reports revealed that many of the adults preferred thinking about fractions via pie charts – with most also reporting pie charts as being the best way to teach someone about fractions and the way they remember learning about fractions. It may not be surprising that pie charts were chosen so frequently when it was provided as an option in our multiple choice question, especially since many people likely think of pie charts as the iconic fraction representation. More striking, however, is adults' responses on the open-ended, spontaneous question about the way in which they "think about fractions" (which came before the multiple-choice questions). Even without prompting, most adults reported using a visual strategy involving an area model or image, like a pie chart or a shaded object, highlighting the overall preference of this type of representation amongst our adult sample.

Notably, in contrast, only *one* person spontaneously reported using a number line. Even when number lines were provided as a multiple-choice option, markedly few participants selected them, suggesting that number lines are not perceived as being a particularly useful representation for thinking about fraction magnitudes, at least by the college students tested. Instead, the second most common category of strategies reported in our open ended question involved symbolic estimation strategies, such as converting the fraction magnitude into a decimal or approximating its relationship to other numbers like ½ or 1. Importantly, when comparing performance on the symbolic magnitude task between those who spontaneously reported using only a visual area model and those who reported using only a symbolic estimation strategy, the symbolic estimators performed better on the magnitude task. This may suggest that symbolic estimation strategies reflect more advanced knowledge of fractions and/or that these strategies allow for more accurate estimates of magnitude than part-whole visualization strategies. This is in line with recent work suggesting that fractions are useful for providing part-whole information, but are not as useful (compared to decimals) for providing magnitude information (DeWolf et al., 2014; Hurst & Cordes, 2016). Thus, it may be that those individuals who estimated fraction magnitudes relative to other symbols had a better understanding of the relations among magnitudes in fraction, decimals, and whole number notation. Whereas those individuals who used area model visualization strategies may have been more focused on the part-whole components of fractions specifically and less tuned to the numerical magnitude associated with the fraction, which is shared across other notations as well.

Although the adults in the current sample did not provide much evidence of thinking about fractions using a number line, this may be attributable to the age of our sample, as the emphasis on using number lines has been relatively recent (National Mathematics Advisory Panel, 2008). In line with this hypothesis, only a small proportion of participants reported having been taught fractions with number lines. Thus, in Experiment 2 we further investigated the relation between visual and symbolic representations of fractions by explicitly providing children and adults with different representations before completing the symbolic comparison task. By allowing people to engage in a visual mapping task directly, we investigated whether priming adults and children to think about fractions using either number lines or pie charts would impact subsequent performance on a symbolic magnitude comparison task. Furthermore, by testing children, in addition to adults, we can more thoroughly investigate the impact of and implications for education. In particular, given that current adults (based on Experiment 1) mostly show evidence of learning with area models and not number lines, they may be particularly resistant to number line priming, as these may be particularly unfamiliar representations. Children who are in the process of learning about fraction and decimal magnitudes and who are more likely to have encountered fractions on number lines, however, may have more malleable visual representations of fractions and thus may be more open to adapting the way they think about fractions.

Experiment 2

In Experiment 2, adults and 9- to 12-year-old children were randomly assigned to map fractions to either pie charts or to number lines. Following this task, participants engaged in the fraction comparison task (as in Experiment 1). In doing so, we investigated three specific research questions: (1) Would practice mapping fractions with number lines result in better symbolic magnitude performance than practice mapping fractions with pie charts? (2) How does performance on the visual-spatial mapping task, with either number lines or pie charts, relate to symbolic magnitude performance? In particular, it may be that people's ability to map fractions to number lines is more predictive of symbolic magnitude understanding than their ability to map fractions with pie charts. (3) What kinds of overt strategies (i.e., partitioning into sections versus estimating approximately) do people use when dealing with number lines and pie charts? Work with adults suggests adults are more error-prone when translating between fractions and number lines than pie charts (Hurst, Relander, & Cordes, 2016). Is this due to a difference in the types of overt strategies adults engage in when encountering a number line compared to a pie chart? Alternatively, greater error may arise from the inaccurate application of the same overt strategy. For example, partitioning strategies may simply be more accurate when applied to pie charts than to number lines (given the symmetric nature of pie charts).

Lastly, 9- through 12-year-old children and adults were included in Experiment 2 in order to investigate these questions in both a group of educated adults who have substantial experience with fractions and decimals and a group of children who are currently in the process of learning these concepts. This age group is older than children in other recent studies investigating the efficacy of number lines (e.g., Hamdan & Gunderson, 2017) so that we could investigate whether the way children approach symbolic fractions could be impacted using a brief practice rather than teaching an entirely new approach. Thus, we used children who were already familiar with fractions and had begun instruction on these topics.

Method

Participants. Forty-eight 9-12-year-old children and 100 adults were included in the analyses. Twenty-four children were randomly assigned to one of two between-subject conditions: Number Line (NL) Condition (M_{age} = 10.3 years, Range: 9 to 12.3 years, 15 females), or the Pie Chart (PC) Condition (M_{age} = 10.5 years, Range: 9.2 to 12.8 years, 14 females). Fifty adults were randomly assigned to each condition: Number Line (NL) Condition (M_{age} = 18.6 years, Range: 17 to 21 years, 40 females), or the Pie Chart (PC) Condition (M_{age} = 19.0 years, Range: 18 to 23 years, 36 females). Children were tested at local after school programs, summer camps, and public parks, as well as in our laboratory on the Chestnut Hill campus of Boston College or in their homes. Written consent was obtained from parents or legal guardians of all children and children provided written assent for their own participation. Children received a small prize, sticker, or $10 for their participation, depending on the regulations of the specific testing facility. Adults were tested in our laboratory on the Chestnut Hill campus of Boston College and received partial course credit for participation.

Design and Measures. Participants were randomly assigned to either the Number Line condition or the Pie Chart condition. All participants completed the visual representation activities and the magnitude comparison task. The visual representation activities used either Number Lines or Pie Charts, depending on the participant's condition. The magnitude comparison task was identical across the two conditions and to the task used in Experiment 1. For both adults and children, the experimenter remained quietly in the room for the duration of the study. The entire experiment took approximately 15 minutes for adults and 20 minutes for children.

Visual Representation Activities. All participants received two separate 21.5 cm by 14 cm paper booklets: one booklet for Number-to-Position (NP) trials and one for Position-to-Number (PN; adapted from Siegler & Opfer, 2003) trials. Although these terms (NP and PN) have typically been used for number line mappings only, for ease in communication we will be using them for mapping with both number lines and pie charts. Each booklet contained seven trials displaying a different target magnitude, presented in the same order for each participant. In the NP task the target magnitude was presented as a symbolic fraction: 1/2, 3/4, 1/3, 7/8, 3/5, 5/6, and 1/5. For the PN task the target magnitude was presented as a visual representation (dependent upon the condition): 1/2, 3/4, 1/4, 2/3, 7/9, 3/7, and 1/8.

All participants completed the NP booklet first, followed by the PN booklet. In the NP task, participants were presented a symbolic fraction above an empty pie chart or number line (depending on condition) and asked to place a mark on the number line or fill in the part of the circle that went with the number at the top. On each trial, after the participant made his or her mark, the experimenter showed the participant the correct response by making a mark on the line or on the pie chart with a yellow highlighter, using a premeasured cut out. The experimenter then gave the participant feedback by commenting on the child's response relative to the correct response (e.g., "You were pretty close!" or "Not quite, it's actually smaller"). The participant completed seven of these problems, each presented on a separate page.

In the PN task, participants were given a 0 to 1 number line with a mark on it or a pie chart with a shaded portion (depending on condition) and a fraction with a visible denominator but an empty box in the numerator above the image. Participants were asked to fill in the numerator of the fraction, using the denominator provided. The experimenter corrected responses on each trial by showing the participant a card with a number line or pie chart divided up into the relevant units and indicating the correct answer. Participants completed seven trials with each problem presented on a separate page.

Number Line Condition: For both the NP and PN tasks, a 13.8cm line with the end points marked with 0 and 1 was displayed in the center of the page. Approximately 3.5 cm above the center of the line was a symbolic fraction (about 1.75cm by 0.75 cm). In the PN task, there was a 1 cm long hatch mark located somewhere along the number line and the numerator of the symbolic fraction was replaced with an empty textbox (however, the appropriate denominator was provided on every trial).

Pie Chart Condition: For both the NP and PN tasks, a circle with a radius of approximately 4 cm was centered on each page. Within each circle, a vertical radial line extended from the center of the circle to the top. This was intended to be a reference line for participants (however, many chose not to use it and drew two of their own boundary lines). Approximately 1 cm above the top of the circle was a symbolic fraction. In the PN trials, the circle contained a shaded region (substantially darker gray than the other portion, which was left white) with one boundary edge extending from the top point directly down to the center and the shaded region extending to the right in clockwise direction. As in the Number Line condition, the numerator of the symbolic fraction was replaced with an empty textbox (for the PN task).

Data Coding. For the NP task, percent absolute error (PAE) was used as a measure of performance accuracy. PAE was measured as the average difference (in cm for the Number Line condition and degrees for the Pie Chart condition) between the participant's response and the correct location divided by the length of the line (Number Line condition) or divided by 360 degrees (Pie Chart condition), times 100 (resulting in a *percent* error). Two independent coders measured responses on each booklet and reliability was measured using the interclass correlation (ICC; modeled using consistency with a two-way model using R package "irr" by Gamer, Lemon, Fellows, and Singh (2012)). Reliability was excellent for all magnitudes in each condition (ICCs's > 0.85) and the average value given by the two coders was used in the analyses.

Booklets were also coded for evidence of overt, external strategies. Two independent coders determined whether, for each magnitude, there was evidence of an overt partitioning strategy (additional lines on the number line or pie chart beyond the response) or if there was no evidence of partitioning and participants simply estimated their answer (only the answer was written anywhere). Each participant was then categorized as either consistently using an overt partitioning strategy, consistently not using an overt partitioning strategy, or inconsistently applying strategies. Inter-rater reliability on these overall categorizations (measured using Cohen's Kappa with the R package “irr” (Gamer et al., 2012)) was excellent (Cohen Kappa = 0.9) and the codes from the first coder were used in the analyses.

For both the PAE analyses and the strategy categorization, performance on the magnitude of ½ was not included because overall performance on ½ was very high, the use of partitioning strategies beyond ½ would not be a relevant strategy, and because the visual aspects of our pie chart display may have made ½ easier for pie charts than for number lines. For both conditions, accuracy on the PN task was computed as the proportion of trials in which a correct numerator response was provided.

On the magnitude comparison task, accuracy was used as the primary dependent variable for children and both accuracy and RT (using the same inclusion criteria as Experiment 1, resulting in replacing ~ 1.7% of the data and a final sample of n=46 in the PC condition and n=43 in the NL condition for the RT analyses) are reported for adults (again, given the slightly lower samples in the RT analyses). For children, we did not analyze RT as it would have resulted in the exclusion of a substantial number of children who performed relatively poorly and RT for poor performers (who may have guessed or been confidence in their incorrect responses) and high performers (who accurately applied strategies) cannot be reliably compared.

Given the differences in response variability across children and adults, we analyzed data from the two age groups separately.

Results

Visual representation task. First, we analyzed performance on the visual representation activities for children. On the NP trials, we looked at average PAE across all trials (except half; see Figure 2.2). There was not a significant difference in performance between the PC condition, $M = 6.1\%$, and the NL condition, $M = 8.7\%$, $t(46) = 1.5$, $p = 0.15$. On the PN trials, children did very well overall and there was not a significant difference between the two conditions on proportion correct, $M_{NL} = 0.86$, $M_{PC} = 0.82$, $t(46) = 0.78$, $p = 0.43$.

Adults also did not show a significant difference in PAE on the NP task (PC condition: $M = 3.0\%$; NL condition: $M = 3.4\%$), $t(98) = 0.96$, $p = 0.3$. However, adults, were significantly more accurate in the NL condition than in the PC condition on the PN trials, $M_{NL} = 0.99$, $M_{PC} = 0.93$, $t(69.5) = 4.16$, $p < 0.001$; however, performance was very high and approaching ceiling on this task, providing little variability.

Figure 2.2: Percent Absolute Error on Experiment 2 mapping task for adults and children

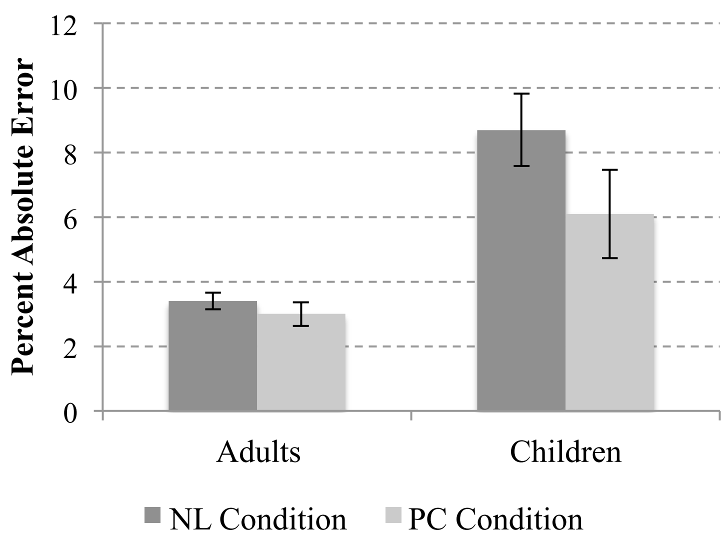


Figure 2.2: PAE on the NP task for adults and children in the number line (NL) and pie chart (PC) conditions.

Second, we looked at differences in the use of overt strategies across the two representations. In general, pie charts were associated with partitioning, whereas number lines were associated with general estimation (no evidence of overt partitioning). For children in the PC condition, 15 children consistently used a partitioning strategy, 4 children consistently used only an estimation strategy (i.e., no evidence of an alternative strategy), and 5 children were inconsistent in their use of strategies. For children in the NL condition, 18 children consistently used an estimation strategy, 4 children consistently used a partitioning strategy, and 2 children were inconsistent in their strategy use. When looking at only those children who were consistent in their strategy use, we find a significant difference between the PC and NL conditions, $\chi^2 = 15.1$, $p < 0.001$, with children in the PC condition much more likely to use a partitioning strategy and children in the NL condition much more likely to use an estimation strategy.

Overall, adults tended to use an estimation strategy regardless of the format of the representation, however number lines were more likely to be associated with estimation strategies than pie charts. In the NL condition, all 50 adults consistently used an estimation strategy – there was no evidence of partitioning on the number lines. In contrast, in the PC condition only 29 adults consistently used an estimation strategy, 7 adults consistently used a partitioning strategy, and 14 adults were inconsistent in their strategy use. Furthermore, when looking at just those participants using a consistent strategy, the types of strategy employed differed across conditions, $\chi^2 = 10.6$, $p = 0.001$, Fisher's Exact Test $p = 0.002$ (given the low expected values in some cells).

Magnitude comparison task. In order to investigate performance on the magnitude comparison task, including whether the number line versus pie chart mapping tasks impacted subsequent performance on the magnitude task, we analyzed the magnitude comparison task using an ANOVA on proportion correct with notation (3:

FvF, DvF, and NvF) and ratio (2: small and large) as within-subject factors, condition (2: NL and PC) as a between subject factor. Child data (see Figure 2.3a) did not show a main effect of notation, $F(2, 92) = 0.8$, $p = 0.5$, *partial* $\eta^2 = 0.02$. However, there was a significant main effect of ratio, $F(1, 46) = 30.2$, $p < 0.001$, *partial* $\eta^2 = 0.4$, and a ratio by notation interaction (reporting Huynh-Feldt correction for a violation of sphericity), $F(1.8, 83.6) = 4.2$, $p = 0.02$, *partial* $\eta^2 = 0.08$. Paired t-tests indicated that there was a significant ratio effect, with lower accuracy on the smaller ratio than the larger ratio, on the DvF trials: $M_{small} = 0.63$, $M_{large} = 0.74$, $t(47) = 3.5$, $p < 0.001$, and the FvF trials: $M_{small} = 0.60$, $M_{large} = 0.81$, $t(47) = 3.8$, $p < 0.001$. However, there was not a significant ratio effect on the NvF trials: $M_{small} = 0.69$, $M_{large} = 0.73$, $t(47) = 1.4$, $p = 0.16$, although the pattern of responding was consistent with a ratio-effect. Furthermore, there was a significant difference between the ratio effects in the FvF and NvF trials ($p < 0.007$), but not between NvF and DvF ($p = 0.15$) or DvF and FvF ($p = 0.13$). There was not a main effect of condition, $F(1, 46) = 1.9$, $p = 0.17$, *partial* $\eta^2 = 0.04$, nor any interactions involving condition, ps > 0.1, *partial* $\eta^2s < 0.05$. However, the means suggest that overall accuracy on the magnitude comparison task was slightly higher for children in the PC condition ($M = 0.73$) than children in the NL condition ($M = 0.66$).

Analyses of adult accuracy on the comparison task revealed a main effect of notation $F(1.9, 185.8) = 9.2$, $p < 0.001$, *partial* $\eta^2 = 0.09$, a main effect of ratio, $F(1, 98) = 57.5$, $p < 0.001$, *partial* $\eta^2 = 0.37$, and again, a ratio by notation interaction, $F(1.75, 171.8) = 8.9$, $p < 0.001$, *partial* $\eta^2 = 0.08$. Paired t-tests indicated there were significant ratio effects in each notation type: FvF $M_{small} = 0.83$, $M_{large} = 0.95$, $t(99) = 6.4$, $p < 0.001$, NvF $M_{small} = 0.91$, $M_{large} = 0.95$, $t(99) = 3.1$, $p = 0.003$, and DvF $M_{small} = 0.9$, $M_{large} =$

0.95, $t(99) = 4.2$, $p < 0.001$. However, accuracy on the FvF trials ($M = 0.89$) was significantly lower overall and resulted in significantly higher ratio effects than on the NvF trials ($M = 0.93$; performance difference: $p < 0.001$, ratio effect difference: $p = 0.001$) and the DvF trials ($M = 0.92$; performance difference: $p = 0.004$, ratio effect difference: $p = 0.005$). There were no significant differences between DvF and NvF on overall performance ($p = 0.35$) or ratio effects ($p = 0.38$). Again, there were no main or interaction effects involving condition (*p*s > 0.1).

The same analyses involving adults' RT (see Figure 2.3b) showed a very similar pattern: there was a main effect of notation $F(2, 174) = 19.0$, $p < 0.001$, *partial* $\eta^2 = 0.18$, a main effect of ratio, $F(1, 87) = 71.8$, $p < 0.001$, *partial* $\eta^2 = 0.5$, and a ratio by notation interaction, $F(1.73, 155.3) = 12.0$, $p < 0.001$, *partial* $\eta^2 = 0.1$. The largest ratio effects were on the DvF trials, $M_{small} = 2590$ms, $M_{large} = 2131$ms, $t(88) = 7.0$, $p < 0.001$, followed by the FvF trials, $M_{small} = 2603$ms, $M_{large} = 2365$ms, $t(88) = 3.9$, $p < 0.001$, and NvF trials, $M_{small} = 2080$ms, $M_{large} = 2016$ms, $t(88) = 1.9$, $p = 0.06$ (which were only marginally significant) and all three ratio-effects were significantly different from each other (*p*s < 0.001). Furthermore, FvF took the longest on average, followed by DvF (only marginally significantly different from FvF, $p = 0.06$), and NvF (significantly different from both, *ps* < 0.001). However, there were no main or interaction effects involving condition (*p*s > 0.4)

Figure 2.3A: Children's Accuracy on the Experiment 2 Magnitude Comparison Task

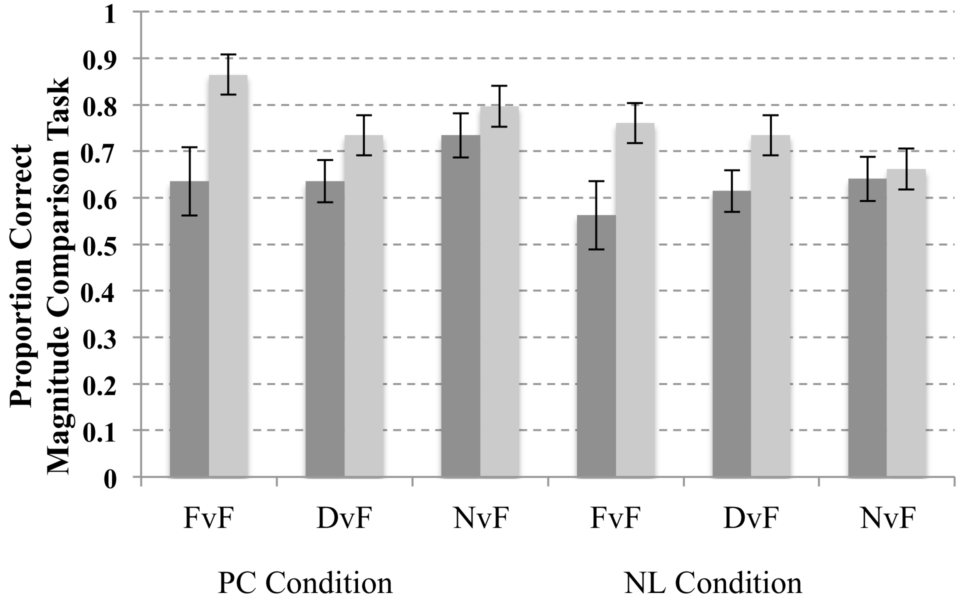


Figure 2.3B: Adults' Reaction Time on the Experiment 2 Magnitude Comparison Task

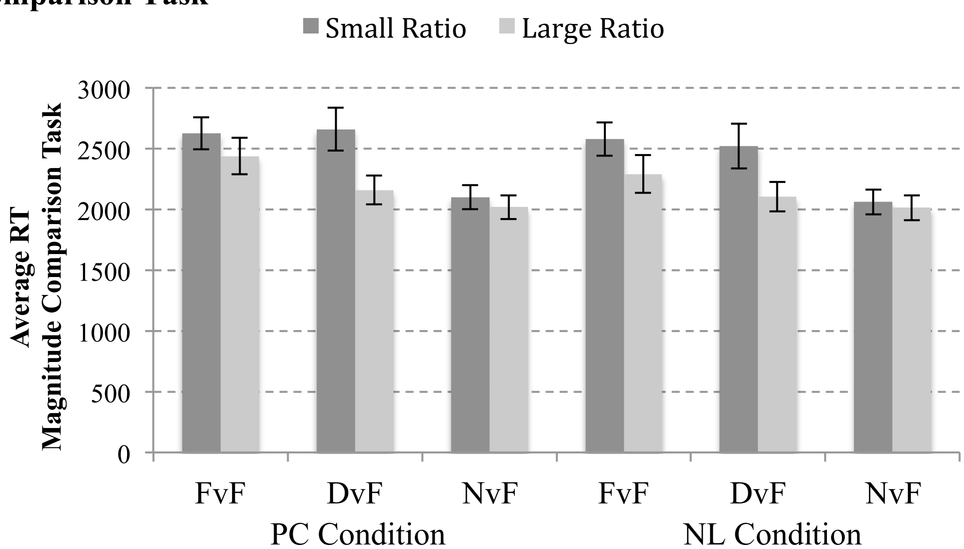


Figure 2.3: Performance on the magnitude comparison task in Experiment 2 for children (proportion correct, upper figure A) and for adults (average RT, lower figure B)

Relations between tasks. We next explored how performance on the visual-symbolic mapping task may be related to the symbolic comparison task for each of the two representations. For children in the PC condition, there was not a significant correlation between PAE on the PC mapping task and overall proportion correct on the symbolic magnitude comparison task, $r(23) = -0.18$, $p = 0.4$. However, for children in the NL condition, there was a significant correlation between PAE on the NL mapping task and proportion correct on the symbolic magnitude comparison task, $r(23) = -0.598$, $p = 0.002$. However, a direct comparison of these two correlations indicated that they were not significantly different from each other, $z = 1.64$, $p = 0.1$.

A similar pattern was found in the adult sample. Data from adults in the PC condition did not show a significant relation between PAE on the mapping task and proportion correct on the symbolic magnitude comparison task, $r(49) = -0.02$, $p = 0.9$. However, data from adults in the NL condition did show a significant relation, $r(49) = -0.285$, $p = 0.045$. Like the children, however, there was not a significant difference in these correlations, $z = 1.33$, $p = 0.18$. There was not a significant relation between average RT on the comparison task and PAE on the mapping task for data from adults in the NL condition, $r(43) = 0.05$, $p = 0.8$, or in the PC condition, $r(46) = -0.04$, $p = 0.8$.

Discussion

In Experiment 2, we investigated the relations between thinking about fraction notation within different visual-spatial representations and as symbolic magnitude in both children and adults. Overall, we did not find that a relatively short activity mapping fractions with number lines led to better symbolic magnitude performance than practice mapping fractions to pie charts. However, we did find evidence that both children and adults used the number lines and pie charts in different ways, although these differences were not accompanied by performance differences. Furthermore, we replicated findings from Experiment 1 and previous work suggesting that adults and children show evidence of ratio dependent responding across rational number notations, however this may be less consistent for fractions and whole numbers (Hurst & Cordes, 2016, under review - B).

Based on previous work suggesting that teaching with or providing short trainings with number lines lead to better fraction understanding than area model representations, specifically pie charts (Cramer et al., 2002; Hamdan & Gunderson, 2017; Keijzer & Terwel, 2003; Saxe et al., 2013; Wang & Siegler, 2013), our lack of finding (and if anything, a pattern in the opposite direction) may be surprising. It is possible that using a number line representation is not beneficial for symbolic magnitude comparisons. However, there are several other, non-mutually exclusive, possibilities for why these activities did not lead to differences in the subsequent symbolic magnitude task. First, the activities children and adults in the current study engaged in were not about instruction. Participants did receive feedback on the task, but they were not instructed as to how to use the number lines or pie charts. However, many of the recent studies showing an impact of number lines were with younger children who did not have much fraction understanding already, and as such, the primary manipulation involved instruction (e.g., Hamdan & Gunderson, 2017). It may be that the adults and older children in the current study (who already had some knowledge of fractions) may have been more likely to default to their learned or practiced way of thinking about fractions, such that our visualization activities may have not altered their approach to the tasks. That is, although number lines are becoming a larger part of the curriculum, other representations, including area model representations, are still a major component (e.g., National Governors Association Center for Best Practices, 2010) and so it may be that the participants required more explicit instruction and/or more extensive experience with number lines in order for these experiences to alter their response strategy. Lastly, some previous studies that showed a benefit for number lines included instruction with both fractions and decimals (e.g., Wang & Siegler, 2013). Considering this difference, it may be that the primary advantage offered by the number line model is that it provides a visual representation that emphasizes the relation between different symbolic representations, including decimals, fractions, and integers – relation which are not as readily represented via pie charts (e.g., Saxe et al., 2013). In line with this hypothesis, those adults in Experiment 1 who reported using symbolic strategies that integrated different magnitudes and different types of numbers performed better on the magnitude comparison task. However, children and adults in the current study only received number lines and pie charts with proper fractions (i.e., fractions between 0 and 1). Thus, it may be that the people are unable to reap the benefits of the number line model under these circumstances, where a pie chart may be just as effective. This idea was explored in Experiment 3.

Although we did not see significant performance differences across conditions on either the visual mapping tasks or the symbolic magnitude task, there was substantial evidence of strategy differences in the way children and adults worked with number lines versus pie charts. When working with number lines, both adults and children tended to show no evidence of engaging an overt, external strategy and thus may have directly estimated the location of the fraction on the number line. On the other hand, when working with pie charts, children tended to overtly partition the pie charts into pieces. Adults estimated more often than partitioned on pie charts, but this was less extreme than when using number lines. Although there were not obvious differences in accuracy in the current study, this pattern of strategy use may be why other studies have shown that adults are generally more accurate mapping symbols with pie charts, but have particular difficultly mapping fractions with number lines (e.g., Hurst, Santry, Relander, & Cordes, in prep). Given that decimals provide a more direct representation of magnitude than fractions for both children and adults (e.g., DeWolf et al., 2014; Hurst & Cordes, 2016, under review - B), this previous work combined with the strategy differences in the current experiment suggest that number lines may be more likely to be associated with numeric magnitudes than pie charts, which may, in turn, be more likely to be associated with information about the discrete components. Remarkably, this finding holds despite the fact that participants only mapped fractions. Thus, it is not that fractions always lead people to use partitioning strategies. Rather, children and adults may be more inclined to adapt their strategy based on the visual representation and thus, attempt to treat fractions like an approximate magnitude in the case of number line representations (by not using an overt partitioning strategy, but instead potentially engaging a magnitude estimation strategy) and treat fractions like a part-whole structure with specific components in the case of pie charts (by using a partitioning strategy).

In Experiment 3, we want to further extend these findings by looking at rational number magnitude understanding of notation beyond fractions. In particular, some research with number lines suggests that the benefit of number lines is because of the way it can integrate different notations and extend beyond just proper fractions (Cramer et al., 2002). Thus, in Experiment 3 we included fractions, decimals, and whole numbers in a larger range of values (0 to 5 number line or pie charts, rather than just 0 to 1) in order to investigate whether mapping different notations to a number line would result in better symbolic magnitude ability.

Experiment 3

In Experiment 3, we used a very similar paradigm to Experiment 2, but extended our visual mapping task to include symbolic magnitudes beyond just fractions. In so doing, we addressed three specific research questions, as in Experiment 2: (1) Does practice mapping fractions, decimals, and whole numbers to number lines result in better symbolic magnitude performance than practice mapping the same values to pie charts? (2) How does performance on the visual-spatial mapping task, with either number lines or pie charts, relate to symbolic magnitude performance? (3) What kinds of overt strategies do people use with number lines and pie charts and do these strategies differ for fractions and decimals? In particular, are the strategies that people engage during the mapping task more dependent on the notation they are given (decimals vs. fractions) or on the visual-spatial representation (number lines vs. pie charts). On the one hand, in Experiment 2, adults approached number lines and pie charts very differently even though they were given the same fraction magnitudes. Thus, we might expect that the same patterns would be true for decimals. On the other hand, decimals are more directly mapped to magnitude for children and adults (e.g., Hurst & Cordes, 2016, under review - B) and do not directly represent information about the components. Thus, we might expect that people will always use an estimation strategy with decimals, even when given a pie chart.

Furthermore, by including an increased number of trials involving distinct notations we can directly compare performance and strategies for fractions and decimals.

Method

Participants. Forty-nine 9-12-year-old children participated in the study, separated across two conditions: Integrated Number Line (Int-NL; N = 25, M_{age} = 10.3 years, Range: 9 to 12.25 years, 14 females) and Integrated Pie Chart (Int-PC; N = 24, M_{age} = 10.4 years, Range: 9.75 to 12.25 years, 14 females). Forty-eight adults also participated in the experiment: Integrated Number Line (Int-NL; N = 23, M_{age} = 19.4 years, Range: 18 to 22 years, 15 females) and Integrated Pie Chart (Int-PC; N = 25, M_{age} = 19.6 years, Range: 18 to 22 years, 29 females).

As in Experiment 2, children were tested at local after school programs, summer camps, and public parks, as well as in our laboratory on the Chestnut Hill campus of Boston College or in their homes. Written consent was obtained from parents or legal guardians of all children and children provided written assent for their own participation. Children received a small prize, sticker, or $10 for their participation, depending on the regulations of the specific testing facility. Adults were tested in our laboratory on the Chestnut Hill campus of Boston College and received partial course credit for participation.

Design and measures. Participants were randomly assigned to one of two conditions: Integrated Number Line condition or Integrated Pie Chart condition. All participants completed the visual representation activities and the magnitude comparison task. The visual representation activities included either number lines or pie charts, depending on condition and only contained NP trials (given the overall high accuracy of children and adults on the PN trials in Experiment 2). The magnitude comparison task was identical between the two conditions and to those used in Experiments 1 and 2.

Visual Representation Activities. Participants received a booklet (21.5 cm by 14 cm) that contained 14 NP trials. On each page was a symbolic number (fraction, decimal, or whole number) presented at the top of the page, in the same order across participants: 4, 6/4, 1/5, 1.4, 3.8, 3/4, 2.3, 0.2, 8/3, 9/2, 2, 10/3, 4.1, and 2.7. In the Integrated Number Line condition, a 14 cm line was in the center of the page with endpoints labeled 0 and 5 and 0.7cm vertical hatch marks located at the whole number units on the line (i.e., 1, 2, 3, and 4). In the Integrated Pie Chart condition, five circles (radius = 1.5cm) were aligned horizontally across the center of the page with 0.3mm between each circle and each circle only containing a 1.5cm line extending from the center of circle to the top.

As in Experiment 2, participants were asked to represent the number at the top by making a mark on the line or filling in the circles (depending on the condition). After the participant completed their response, the experimenter use a premade cut out of the correct answer to put a yellow hatch-mark in the correct spot on the line or fill in the correct amount in the circles by outlining the full circles and drawing boundary lines in partial circles (depending on condition). Unlike Experiment 2, participants were not given evaluative verbal feedback by the experimenter but instead the experimenter simply said: “This shows the number”, pointing to their correct yellow answer.

Data Coding. Coding of responses and strategies on the visual representation activities mimicked that of Experiment 2. Again, two independent coders coded for strategy use (estimation versus partitioning strategy for decimals and fractions separately; all Cohen’s Kappas > 0.75) and codes for the first coder are used in the analysis. Two independent coders also measured the accuracy of responses (using the same model and methods as Experiment 2; all IRRs > 0.68) and the average value between the two coders was used in the analyses. In addition, as in Experiment 2, we did not look at strategies or accuracy on the fractions involving 2 in the denominator (9/2), since a partitioning strategy is not applicable and almost all participants provided the answer with no overt, written strategy. Five children and three adults in the Int-PC condition and one child in the Int-NL condition responded to the booklet in an atypical fashion, making it impossible to score the accuracy of these participants in a way that is comparable to the others (e.g., coloring in the pie charts using other shapes). Thus, their data were not included in analyses involving performance on the visual mapping tasks, but they were included in analyses involving the other tasks and strategies on the visual mapping task.

Performance on the magnitude comparison task was treated the same way as Experiments 1 and 2. Adult RT analyses included a final sample of 20 in the Int-NL condition and 23 in the Int-PC condition and ~2.3% of the data were considered outliers and replaced.

Results

Visual representation task. First, we analyzed performance on the visual representation activities for decimals and fractions separately (see Figure 2.4). For children, there was not a significant difference in PAE on the number lines (N=24) versus pie charts (N=19) for the decimals: $M_{Int\text{-}NL} = 3.8$, $M_{Int\text{-}PC} = 10.0$, $t(41) = 0.6$, $p = 0.5$, or fractions $M_{Int\text{-}NL} = 10.0$, $M_{Int\text{-}PC} = 10.4$, $t(41) = 0.1$, $p = 0.9$.

Adults' performance, however, showed significantly less error when mapping to the number line (N = 23) than the pie chart (N=22) for decimals ($M_{Int\text{-}NL}$ = 1.1, $M_{Int\text{-}PC}$ = 2.7, $t(43)$ =4.2, $p < 0.001$), but did not show a significant difference on fractions ($M_{Int\text{-}NL}$ = 3.6, $M_{Int\text{-}PC}$ = 5.1, $t(43) = 0.9$, $p = 0.4$).

Figure 2.4: Percent Absolute Error on Experiment 3 mapping task for adults and children

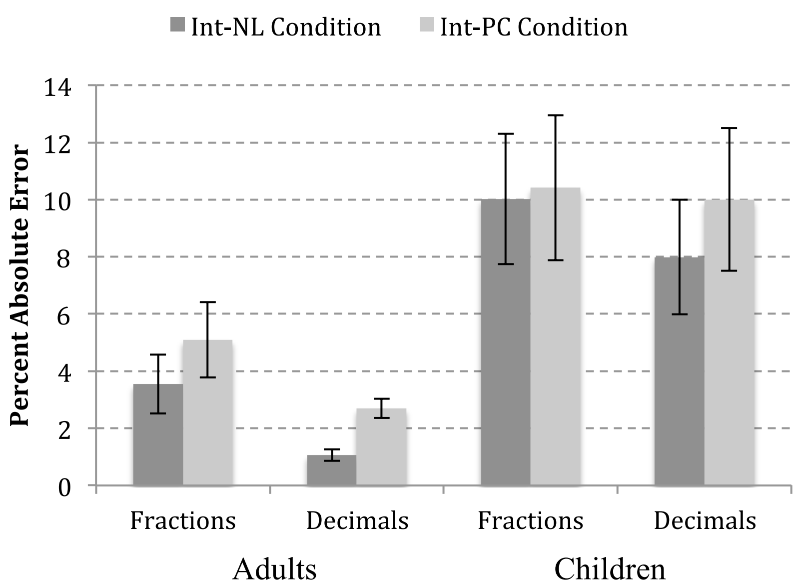


Figure 2.4: Percent Absolute Error on the visual mapping tasks for adults and children in the Integrated Number Line (Int-NL) and Integrated Pie Chart (Int-PC) conditions of Experiment 3

Second, we looked at strategy differences across the two representations for fractions (although, as in Experiment 2, excluding 9/2 given that almost every participant used an estimation strategy) and decimals separately. The full counts broken down by condition and age group are presented in Table 2.1.

		Int-NL condition		Int-PC Condition	
		Fractions	Decimals	Fractions	Decimals
Children	Estimation	20	21	8	7
	Partitioning	3	2	5	7
	Inconsistent	2	2	11	10
Adults	Estimation	21	22	17	13
	Partitioning	0	1	1	2
	Inconsistent	2	0	7	10

Table 2.1: The number of participants who used each strategy based on age group, type of trial, and condition

Most children in the Int-NL condition consistently used an estimation strategy for fractions and decimals. In the Int-PC condition, however, children showed much more variability in their strategies and were more evenly split across strategy types and trial types (although, numerically more still used an estimation strategy). Furthermore, when looking at just those children who consistently applied the same strategy (either estimation or partitioning, ignoring children who were inconsistent) the pattern of responses between the Int-NL and Int-PC conditions were significantly different for decimals ($\chi^2 = 8.1$, $p = 0.005$, Fisher's Exact Test: $p = 0.014$) and marginally for fractions ($\chi^2 = 3.1$, $p = 0.08$, Fisher's Exact Test: $p = 0.1$). Thus, although children in both conditions were more likely to use an estimation strategy, this was, in general, more extreme in the Int-NL condition than in the Int-PC condition.

When looking at the adult data, we see a similar pattern to the children above and as in Experiment 2. Most adults, in both conditions, tended to consistently apply an estimation strategy. However, this pattern is slightly less strong for decimals, where a substantial number of adults were inconsistent in their strategy use. When looking at just those adults who consistently used a single strategy (either estimation or partitioning, ignoring inconsistent adults), there was not a significant difference in strategy use across Int-NL and Int-PC for fractions ($\chi^2 = 1.2$, $p = 0.27$, Fisher's Exact Test: $p = 0.5$) or for decimals ($\chi^2 = 3.1$, $p = 0.08$, Fisher's Exact Test: $p = 0.16$; although this is marginal).

Magnitude comparison task. As in Experiment 3, we analyzed the magnitude comparison task using an ANOVA with notation (3: FvF, DvF, and NvF) and ratio (2: small and large) as within-subject factors and condition (2: Int-NL and Int-PC) as a between subject factor on children and adults separately (see Figures 2.5a and 2.5b). With the children, we replicated the main effects of Experiment 2: there was a main effect of ratio, $F(1, 47) = 38.0$, $p < 0.001$, *partial* $\eta^2 = 0.4$, and a ratio by notation interaction, $F(2, 94) = 3.4$, $p = 0.04$, *partial* $\eta^2 = 0.07$. In addition, however, there was a main effect of notation, $F(2, 94) = 3.3$, $p = 0.04$, *partial* $\eta^2 = 0.07$. Within each notation, performance was higher on the large ratio than the small ratio, FvF: $M_{small} = 0.57$, $M_{large} = 0.76$, $t(48) = 3.9$, $p < 0.001$, and DvF: $M_{small} = 0.59$, $M_{large} = 0.73$, $t(48) = 4.8$, $p < 0.001$, although this was only marginal for NvF comparisons: $M_{small} = 0.69$, $M_{large} = 0.75$, $t(48) = 1.99$, p =0.053 Furthermore, NvF showed significantly smaller ratio effects than both FvF (p = 0.02) and DvF ($p < 0.05$), which were not significantly different from each other (p = 0.4). Lastly, overall performance was significantly higher on NvF trials than both DvF (p = 0.019) and FvF trials ($p = 0.03$), which were not significantly different from each other ($p = 0.96$). Again, however, there were no significant main or interaction effects involving condition (ps > 0.4).

The ANOVA on proportion correct in the adult data revealed a significant main effect of ratio, $F(1, 46) = 20.9$, $p < 0.001$, *partial* $\eta^2 = 0.3$, with adults performing better on the larger ratio (M=0.96) than the small ratio (M=0.9). However, there were no other main effects or interactions (all *ps* > 0.05). Importantly, condition did not interact with any other variables.

Given the near ceiling performance of adults on this task, we also looked at RT. For RT, there was a main effect of notation $F(2, 82) = 22.8$, $p < 0.001$, *partial* $\eta^2 = 0.4$, a main effect of ratio, $F(1, 41) = 32.1$, $p < 0.001$, *partial* $\eta^2 = 0.4$, and a ratio by notation interaction $F(2, 82) = 6.4$, $p = 0.003$, *partial* $\eta^2 = 0.1$.

Additional analyses revealed there was a significant ratio effect in each of the three notation types, with slower performance on the small ratio compared to the large ratio (paired t-tests in each notation separately: NvF: $M_{Small} = 1982$ms, $M_{large} = 1815$ms, $p < 0.001$; DvF: $M_{Small} = 2612$ms, $M_{large} = 2075$ms, $p < 0.001$; FvF: $M_{Small} = 2642$ms, $M_{large} = 2406$ms $p = 0.002$). However, this ratio effect was largest for DvF (*p*s < 0.05) and the NvF and FvF ratio effects were not significantly different from each other ($p = 0.4$).

In addition, the overall ANOVA revealed a marginal ratio by notation by condition interaction, $F(2, 82) = 2.7$, $p = 0.07$, *partial* $\eta^2 = 0.06$. To follow up on this marginal three-way interaction we examined the impact of ratio and notation in each condition separately. For adults in the Int-NL condition (N=20), there was a main effect of notation, $F(2, 38) = 12.9$, $p < 0.001$, *partial* $\eta^2 = 0.4$, a main effect of ratio, $F(1, 19) = 19.3$, $p < 0.001$, *partial* $\eta^2 = 0.5$, and a ratio by notation interaction, $F(2, 38) = 6.4$, $p = 0.004$, *partial* $\eta^2 = 0.3$. For adults in the Int-PC condition (N=23), there was a main effect of notation, $F(2, 44) = 10.1$, $p < 0.001$, *partial* $\eta^2 = 0.3$ and a main effect of ratio, $F(1, 22) = 11.7$, $p = 0.002$, *partial* $\eta^2 = 0.4$, but the interaction between ratio and notation was not significant, $F(2, 44) = 0.6$, $p = 0.5$, *partial* $\eta^2 = 0.03$. Thus, we compared the ratio effect in each notation across the two conditions. Adults in both conditions had comparable ratio effects (RT large – RT small) in the NvF (Int-NL $M_{\text{ratio effect}}$ = 176.1; Int-PC $M_{\text{ratio effect}}$ = 157.8; p =0.8) and in the FvF (Int-NL $M_{\text{ratio effect}}$ = 252.4; Int-PC $M_{\text{ratio effect}}$ = 219.8; p =0.8) comparisons. However, in the DvF comparisons, adults in the Int-NL condition ($M_{\text{ratio effect}}$ = 772) had marginally significantly higher ratio effects than adults in the Int-PC condition ($M_{\text{ratio effect}}$ = 301.9, p = 0.057).

Figure 2.5A: Children's Accuracy on the Experiment 3 Magnitude Comparison Task

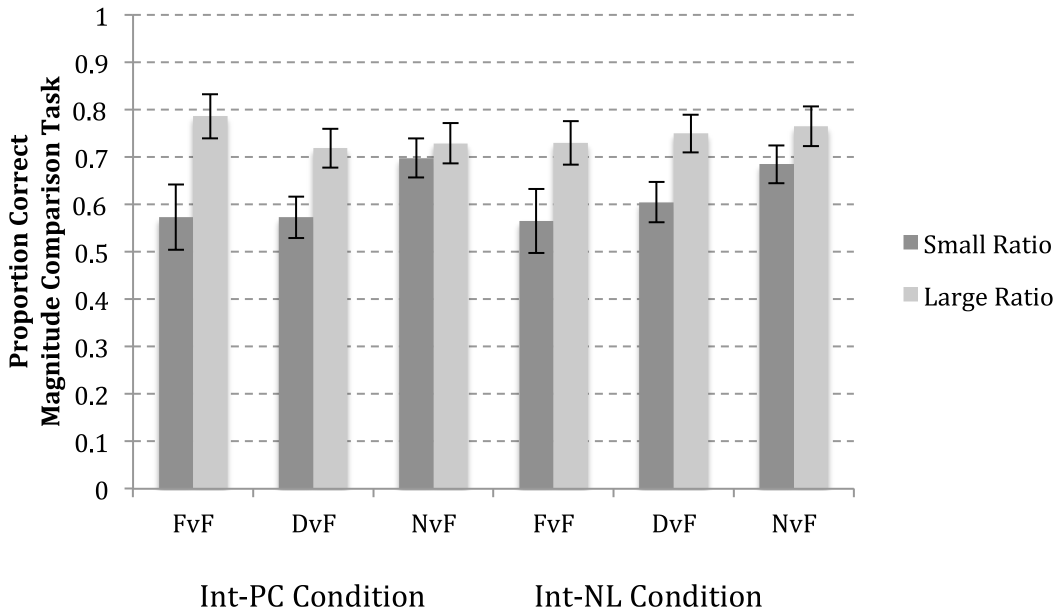


Figure 2.5B: Adults' Reaction Time on the Experiment 3 Magnitude Comparison Task

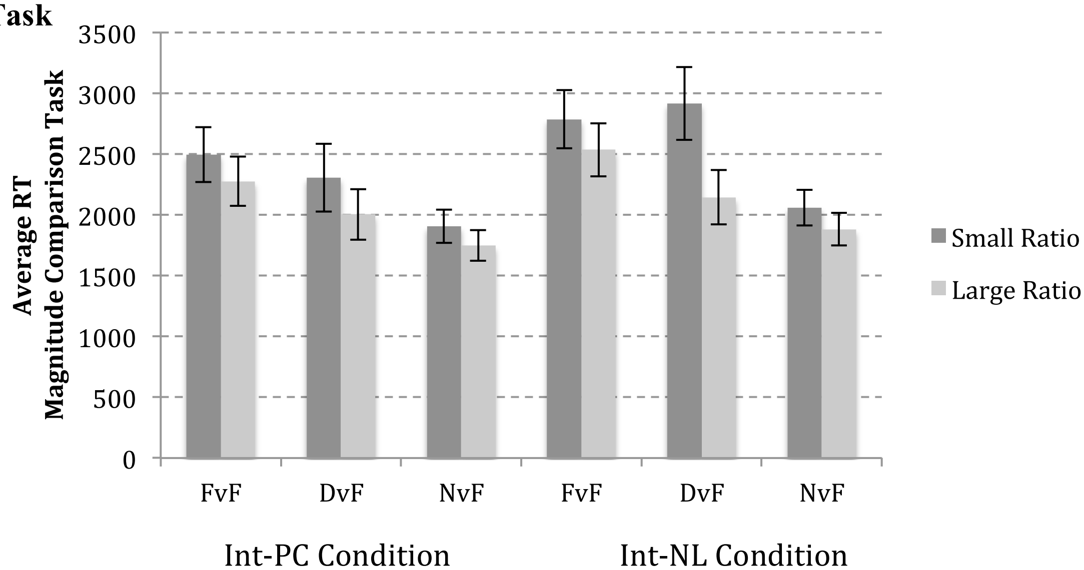


Figure 2.5: Performance on the magnitude comparison task in Experiment 3 for children (proportion correct, upper figure A) and for adults (average RT, lower figure B) , separated by condition: Int-PC (left bars) and Int-NL (right bars)

Relations between tasks. We next explored how performance on the spatial-symbolic mapping task may be related to the symbolic comparison ask for each of the two representations. For children in the Int-NL condition, PAE on the number line mapping task was significantly correlated with overall performance on the symbolic magnitude task, $r(23) = -0.795$, $p < 0.001$. However, in contrast to findings of Experiment 2, a correlation was also seen with performance on the Int-PC mapping task, $r(18) = -0.584$, $p = 0.009$. These correlations were not significantly different from each other, $z = -0.86$, $p = 0.39$.

This pattern of relation was different for adults, compared to the children in Experiment 3 or the adults in Experiment 2. There was not a significant relation between PAE on the mapping task and either overall accuracy or overall average RT on the comparison task in the Int-NL condition, Accuracy: $r(23) = 0.02$, $p = 0.9$; RT: $r(20) = 0.31$, $p = 0.18$. However, for adults in the Int-PC condition, there was a significant relation between PAE and accuracy on the comparison task, $r(22) = -0.47$, $p = 0.03$, but not between PAE and RT on the comparison task: $r(20) = 0.09$, $p = 0.67$. Furthermore, the relations between PAE and accuracy in the two conditions were marginally different from each other ($z = 1.65$, $p = 0.099$), but the relations between PAE and RT were not significantly different from each other across conditions ($z = 0.76$, $p = 0.4$).

Discussion

In Experiment 3, we replicated and extended the findings in Experiments 1 and 2 using the same symbolic magnitude task and extending the visual mapping task to include alternative notations in a wider range of magnitudes. Again, we did not find significant evidence that children performed better on the symbolic magnitude task after receiving practice mapping symbolic rational numbers to number lines, compared to pie charts. Rather, children in the two conditions did not perform significantly different from each other on symbolic magnitude task. Adults, on the other hand, show some small evidence that the number line tasks may have impacted performance on the magnitude comparison task when comparing decimals and fractions in particular. However, this was not evident in overall performance (i.e., adults did not get better after receiving the number lines), but rather adult data showed higher ratio-effects after participating in the number line, compared to the pie chart, priming activities. Thus, this may suggest that rather than number lines and pie charts leading to better or worse performance, they may impact the way in which adults approach the task. Given that ratio dependent responding is generally taken as evidence for attending to magnitude during the task (e.g.,Moyer & Landauer, 1967, 1973), it may be that the higher ratio effects in decimal vs. fraction comparisons are reflective of a more approximate, magnitude based strategy. Furthermore, it may be surprising that this finding arose specifically for the decimal versus fraction comparisons and not the other comparisons. This may be because fractions and decimals were both estimated using the same representational format, providing adults with a stronger basis for aligning these notations in ways they otherwise would not. The differences across fractions and decimals will be further discussed in the general discussion.

Although the number line and pie chart mapping tasks did not consistently prime subsequent performance on the comparison task, there were significant differences in performance with the number lines and pie charts themselves. Children did not show differences in their error on the mapping task for any notation. However, adults in the number line condition had lower error mapping decimals than adults in the pie chart condition (but this was not the case for fractions). These findings with decimals are in line with other work suggesting that decimals are more direct representations of magnitude and more directly aligned with number lines (e.g., Hurst & Cordes, 2016; Hurst, Santry, Relander, & Cordes, in prep), but further extend these findings by suggesting that adults may be more accurate mapping decimals with number lines than decimals with pie charts.

Although children did not show differences in their error between number lines and pie charts, they do show significant differences in their strategies. Although children and adults tended to rely on estimation more than partitioning in both conditions, for children this was less extreme in the pie chart condition. Thus, like Experiment 2, this suggests that children may be inclined to adapt their strategy based on the visual-spatial representation, using a magnitude estimation strategy for number lines and a part-whole partitioning strategy for pie charts. However, the results of Experiment 3 combined with Experiment 2 highlight that this was true for both fractions and decimals.

General Discussion

Across three experiments involving children (approximately 9-12 years old) and adults, we investigated the relation between symbolic magnitude understanding and the use of number lines and pie charts for thinking about the *magnitudes* of symbolic rational numbers. Overall, we do not find that our number line mapping tasks primed adults or children to better think about symbolic fraction magnitudes, regardless of whether the mapping task involved proper fractions only or a mix of notations and magnitudes. However, we did find consistent evidence across all three studies that although number lines may not be spontaneously used to represent fraction magnitudes (Experiment 1), adults and children did tend to apply more magnitude strategies with number lines than with pie charts (Experiment 2) across different magnitudes and notations (Experiment 3) and there is some evidence that these magnitude-based strategies are related to symbolic magnitude performance. Furthermore, we replicated previous work with children and adults (Ganor-Stern, 2012, 2013; Hurst & Cordes, 2016, under review - B) suggesting that adults and children can represent the magnitudes of fractions, decimals, and whole numbers in an integrated way (although, ratio dependent responding for fractions and whole numbers may be less robust than other comparisons).

However, unlike previous work suggesting that adults show particularly poor performance when mapping between fractions and number lines (Hurst, Relander, & Cordes, 2016), we did not see consistent performance benefits for mapping with number lines or pie charts. This may be due to the nature of the tasks. Hurst, Relander, and Cordes (2016) used a computerized mental estimation task that prevented adults from engaging in any kind of overt, written strategy and instead required estimation on the computer screen using the mouse. The current study, however, allowed people to use explicit physical strategies like partitioning (seen more frequently with pie charts than numbers lines) or other physical strategies like aligning their fingers on the page. It may be that the ability to use these strategies on paper and pencil or the increased noise from using physical drawings (i.e., lines that are not perfectly straight) evened out some of the performance differences seen in Hurst et al. (2016).

These findings have implications for the way we think about fraction concepts and for fraction education. In particular, these findings highlight that the magnitude component and the part-whole component of fraction symbols are not always tied together, but instead may be best conceptualized using distinct spatial representations. In particular, although fractions may not be spontaneously visualized using number lines (Experiment 1; Hurst, Santry, Relander, & Cordes, in prep) and adults show particular difficulty mapping between fractions and number lines, at least in some contexts (Hurst, Relander, & Cordes, 2016), when adults and children were forced to map between fractions and number lines they did not, in general, try to engage in a part-whole partitioning strategy, as they did with pie charts, but instead attempted to map the fraction to the number line with a holistic magnitude approach.

In addition, children and adults in Experiment 3 tended to rely on estimation strategies for both number lines and pie charts (although they still relied on estimation more for number lines), which was not the case for children in Experiment 2, who relied almost exclusively on part-whole partitioning strategies for pie charts and estimation for number lines. This may suggest that the inclusion of different notations on the same representation prompted children to engage estimation strategies, rather than partitioning strategies, even for pie charts. Some support for this comes from the relations among the tasks: although children's performance on the number line and integrated number line tasks was significantly related with symbolic magnitude comparison, only performance on the integrated pie charts and not the single pie charts was predictive of symbolic magnitude performance. Thus, there may be some evidence suggesting that in order to teach fraction *magnitudes* it may be important to emphasize the relations between different magnitudes, in different notations, for values beyond just zero-to-one using a number line spatial representation. Children and adults appear to use number lines with an approximate, estimation based approach and so using this representation for magnitude may more closely align with the way number lines are used for whole numbers as well. On the other hand, children tended to approach pie charts in a much more strategic, part-whole way. Thus, it may be that pie charts provide a benefit for conceptualizing the part-whole components of fractions.

Limitations and Future Directions

There are several limitations of the current study, resulting in many open questions that the current study was unable to address.

First, when coding for the strategies that adults and children engaged, we only looked at children and adults' use of overt, external strategies based on what they physically drew on the representations. It is of course possible that many children and adults engaged in *mental* partitioning strategies or physical strategies involving their fingers or hands, both of which we would not have been able to capture within the current data. Thus, it may be that some of the participants who we grouped as using "estimation" strategies simply used *mental* partitioning strategies, rather than overt written partitioning strategies. Despite this limitation, the current findings are still able to speak to the extreme differences in how children and adults approach number lines versus pie charts in terms of their use of overt strategies. That is, it is still important to acknowledge that the overt strategies that children and adults believe to be relevant, acceptable, and/or expected of them are dramatically different between pie charts and number lines. Future work should investigate whether these overt written differences reflect actual differences in the cognitive strategies individuals engage in when presented with number lines and pie charts, and whether the differential pattern observed in overt written strategies reflect differences inherent in the structures of number lines and pie charts or whether these differential strategies reflect learned implicit expectations set by our current educational practices.

Second, the current study focused on pie charts as a commonly used area model in contrast to number lines. However, pie charts are structurally different from number lines in other ways, beyond just being an area representation. In particular, number lines only contain information in a one-dimensional way; that is, magnitude information is conveyed via a horizontal representation of distance. Pie charts, on the other hand, involve both horizontal and vertical changes in space (i.e., changes in two dimensions) that must be integrated in order to interpret the relevant information (Spence, 1990). Thus, the differences seen in the current study for number lines and pie charts may not solely reflect different approaches to number line and area models but also reflect inherent differences in the way that people approach representations with one-dimensional vs. two-dimensional spatial information. Other area models, for example rectangles, may be more similar to number lines in that they can (in some cases) only involve a one-dimensional change (e.g., only be divided with horizontal lines), but are still area models that also involve a second, irrelevant dimension (e.g., the height of the rectangle). Research that has investigated differences in graphical formats (e.g., bar graphs vs. pie charts) suggests that these differences may impact how they are interpreted (Shah & Hoeffner, 2002). Thus, future work should investigate how adults and children may use other area models, such as rectangle models, which may be more similar to number lines in terms of the number of spatial dimensions that change.

Additionally, there are several factors that may have impacted the lack of impact of our priming activities. First, it may be that our priming activities were simply not strong enough to impact the strategies that adults and children use with symbolic magnitudes in the absence of overtly available physical references. For example, given that the magnitude task was on the computer and participants did not have the opportunity to work out their answers on paper, it may be that children were unable to make the connection between the explicit mapping task (which involved paper and pencil) and visualization strategies in the subsequent task. Alternatively, the priming task may simply have not been intensive enough to alter participant's approach to the subsequent magnitude comparison task. Future studies may consider explicitly drawing participant's attention to the use of mental visualization strategies, even in symbolic magnitude tasks where these are not explicitly available. Second, although we tested children who are actively in the process of learning fractions and decimals and adults who have already completed instruction on these topics, we did not systematically collect information about the educational experience of these participants. In particular, it may be that some participants already had substantial experience with number lines, used number lines for whole numbers but not fractions, or had not used number lines in any capacity. Thus, the older age of these children may have limited the impact of our priming manipulation. Future research should explicitly compare the malleability of fraction visualization across middle childhood to determine whether younger children may have more malleable fraction visualization strategies. If so, future work may want to investigate whether introducing the magnitudes of non-whole number quantities onto number lines before young children are introduced to the formal part-whole aspect of fraction notation can improve children's learning of rational number concepts more generally. More generally, it may be worthwhile in future work to manipulate various aspects of fraction instruction, including (1) visual referents of number lines vs. pie charts, (2) explicitly translating and mapping between representations and encouraging children to use these strategies, and (3) the timing of instruction on part-whole versus fraction magnitude concepts on the acquisition of different kinds of fraction knowledge, including magnitude, part-whole components, or arithmetic.

Lastly, the sample sizes used in Experiments 2 and 3 were relatively small. Given the variability in how children respond on these fraction tasks between 9- and 12-years of age, it may be that we did not have enough statistical power to truly investigate some of our effects of interest. When designing these studies, the majority of published studies investigating the impact of number line instruction involved large educational interventions that simultaneously manipulated the representations used (e.g., number lines, pie charts) and the teaching methods (e.g., encouraging problem-based learning; more feedback, etc.). As such, it was difficult to isolate a predicted effect size for the differential impact of number lines and pie charts on fraction understanding in particular. Thus, we based our sample sizes on rules of thumb. However, it is now clear that the current study is highly underpowered for the effects we are investigating. Recently, new studies have been published involving a more precise manipulation of number lines vs. pie charts, and their results suggest a medium to large effect (simple effect Cohen's d = 0.75 for the critical comparison) when the representations are paired with instruction in a younger group. Thus, for 80% power, we would need a minimum of approximately 30-35 in each group, whereas our sample ranged from 20-25 in each group, depending on the specific analysis. As such, the lack of a priming effect found in our data should be interpreted with caution. It is critical to investigate these questions with a larger sample.

Conclusions

In sum, the current study reports three experiments providing convergent evidence that adults and children are able to mentally represent the magnitudes of fractions, decimals, and whole numbers in an integrated, approximate, and ordered way, and that number lines may be better aligned with magnitude based strategies, whereas pie charts may be better aligned with part-whole based strategies. This provides some insight into the use of visual-spatial representations for fraction education and suggests directions for future research to further investigate how these representations may be related.

Supplementary Appendix 2

Stimuli values for the magnitude comparison task across all three experiments

Fraction vs. Fraction		
Fraction	Fraction	Ratio Bin
1/3	2/9	1.5
10/9	5/3	1.5
2/3	4/9	1.5
4/3	8/9	1.5
1/2	5/4	2.5
1/5	4/9	2.5
2/9	3/5	2.5
7/2	6/5	2.5

Fraction vs. Decimal		
Decimal	Fraction	Ratio Bin
0.3	2/9	1.5
0.44	2/3	1.5
1.11	5/3	1.5
1.3	8/9	1.5
0.22	3/5	2.5
0.44	1/5	2.5
0.5	5/4	2.5
3.5	6/5	2.5

Fraction vs. Whole Number		
Fraction	Whole Number	Ratio Bin
3/2	1	1.5
9/2	3	1.5
4/3	2	1.5
10/3	5	1.5
9/4	1	2.5
9/2	2	2.5
8/3	6	2.5
6/5	3	2.5

CHAPTER 3: ATTENDING TO RELATIONS: PROPORTIONAL REASONING IN 3- TO 6-YEAR-OLD CHILDREN

Abstract

When proportional information is pit against whole number numerical information, children often attend to the whole number information at the expense of proportional information (e.g., indicating 4/9 is greater than 3/5 because 4 > 3). In the current study, we presented younger (3-4-year-olds) and older (5-6-year-olds) children a task in which the proportional information was presented either continuously (units cannot be counted) or discretely (countable units; numerical information available). In the discrete conditions, older children showed numerical interference: responding based on the number of pieces instead of the proportion of pieces. However, older children easily overcame this poor strategy selection on discrete trials if they first had some experience with continuous, proportional strategies, suggesting this prevalent reliance on numerical information may be malleable. Younger children, on the other hand, showed difficulty with the proportion task, but were successful in a simplified estimation-style task, suggesting that younger children may still be developing their proportional and numerical skills in task-dependent ways. Lastly, across both age groups, performance on the proportional reasoning task in continuous contexts, but not discrete contexts, was related to more general analogical reasoning skills. Findings suggest that children's proportional reasoning abilities are actively developing between the ages of 3 and 6 and may depend on domain general reasoning skills. We discuss the implications for this work for both cognitive development and education.

Learning fraction concepts can be a difficult task for many students (e.g., National Mathematics Advisory Panel, 2008). Children make both procedural and conceptual errors that remain pervasive even in later school grades and into adulthood (e.g., Christou & Vosniadou, 2012; National Mathematics Advisory Panel, 2008; Ni & Zhou, 2005; Lorti-Forges, Tian, & Siegler, 2015; Vamvakoussi & Vosniadou, 2010). However, fraction concepts have been shown to be an important gatekeeper to many later math skills (Booth & Newton, 2012; Siegler et al., 2012), making it important to understand the specific difficulties children face and how they may be rectified.

These difficulties with formal fraction understanding are surprising given that intuitive reasoning about visually presented proportional information seems to be an early developing skill, with evidence suggesting that even preverbal infants (e.g., Denison & Xu, 2010; McCrink & Wynn, 2007) can process proportional information when presented non-symbolically. Infants as young as 6-months-old have been shown to keep track of the ratio between discrete items (McCrink & Wynn, 2007) and slightly older infants use this ratio information to make probabilistic inferences (Denison, Reed, & Xu, 2013; Denison & Xu, 2010). For example, infants are surprised when an experimenter randomly pulls a red ball from a bin that contains mostly white balls (but some red balls), suggesting that infants track the likelihood of probabilistic outcomes (e.g., Denison, Reed, & Xu, 2013).

Despite this early developing proportional reasoning, children show systematic errors in their processing of visual proportional information. For example, Jeong, Levine, and Huttenlocher, (2007) taught 6-10-year-old children a probability game in which they spun a spinner with two colored sections (red and blue). If the spinner landed on red, they would win stickers but if it landed on blue, they would lose stickers. Children were then shown pairs of spinners and were asked to judge which spinner was more likely to result in a winning spin. When the probabilistic information presented on the spinners was continuous (i.e., the spinner contained only one large red portion and one large blue portion that were not divided up into smaller discrete pieces), children performed above chance on these judgments. However, when the red and blue sections on the spinner were divided into discrete, countable pieces (see Figures 3.1A and 3.1B for examples of discrete and continuous spinners), children performed at chance levels (Jeong et al., 2007). Based on these and similar findings, researchers posit that when discrete countable information is available, children ignore the relation between the number of red and blue pieces, focusing instead on the more salient "numerator"—the total number of winning red pieces, turning a proportional task into a counting one (Boyer et al., 2008; Jeong et al., 2007). Importantly, since children succeed (i.e., select the spinner with the greater proportion of red) when presented with continuous spinners, but not when presented with discrete spinners, this points to poor strategy selection in the presence of specific perceptual features, rather than an inability to process proportional information.

Over-attention to whole number information at the expense of proportional information has also been noted when children and adults process symbolic fractions – a phenomenon referred to as a "whole number bias" (e.g., Alibali & Sidney, 2015; Ni & Zhou, 2005). For example, when judging the relative magnitude of two fractions, people often attend to the relative magnitudes of the whole number components (DeWolf & Vosnaido, 2015), while ignoring the relationship between the numerator and denominator (Bonato, Fabbri, Umilta, & Zorzi, 2007), resulting in poor fraction processing.

Thus, interference of discrete numerical[1] information with relational information processing is pervasive across both symbolic and non-symbolic representations of proportion in children. Yet, importantly, children's spontaneous attention to relational information has been shown to be an important predictor of formal fraction ability (e.g., McMullen, Hannula-Sormunen, Laakkonen, & Lehtinen, 2016), making it critical that we understand what factors impact whether a child is likely to focus on relational information or instead rely upon discrete numerical information when engaging in proportion tasks. This knowledge can inform our understanding of the difficulties children may encounter when first learning fraction notation, including early precursors to the whole number bias.

In the current study, we aimed to investigate what factors may impact children's strategy selection (i.e., relational reasoning vs. counting) when engaging in discrete proportional tasks. To do this, we explored the impact of four distinct factors thought to impact proportional thinking: (1) direct instruction, (2) prior experience, (3) age (as a proxy for numerical ability), and (4) analogical reasoning with visual patterns.

We hypothesized that children's focus on the absolute number of red pieces, instead of on the relation between the number of red to blue pieces, may be the result of an over-attention to the salient "winning" pieces, leading them to ignore other relevant aspects of the display (the blue pieces, total number of pieces, and/or the relative sizes of the pieces). Thus, in our study we explored whether instructing children to pay attention

[1] Throughout this paper, we will use the term "number" or "numerical" to specifically refer to whole-number quantities. Although fractions and proportional amount are also defined as numerical quantities, in this context we will refer to these quantities as "proportion" or "proportional" and whole number quantities as being "number" or "numerical".

to the total amount of both the red *and* the blue pieces makes children more likely to focus on the relation between these two amounts.

We also explored whether experience focusing on proportional information, in the absence of conflicting whole number information in continuous trials, may promote the use of relational strategies on subsequent discrete trials. Thus, we varied the order of continuous and discrete spinner blocks in our task, allowing us to assess whether experience on continuous trials promotes performance on subsequent discrete trials. One recent study suggests some promise for this approach in older children; 4th graders (but not 2nd graders or kindergarteners) were less likely to rely on numerical strategies on discrete proportional tasks if initially exposed to practice with continuous proportion (Boyer & Levine, 2015).

A focus on number in the context of a proportional task may be driven by an overall increased focus on whole numbers and counting. If so, then younger children (e.g., preschoolers), for whom numerical information may be less salient because they are still in the process of mastering the counting procedure and have yet to encounter formal mathematical tasks in school, may be less likely to be swayed by whole-number numerical information. Some evidence suggests preschoolers prefer relational information when put in conflict with whole number information (Sophian et al., 1995), however other studies report that 3-4-year-olds do not use proportional information to make inferences about probabilistic outcomes (Girotto, Fontanari, Gonzalez, Vallortigara, & Blaye, 2016). Thus, there are still several open questions about how (and whether) children this age process proportional information. In the current study, we used an identical paradigm with 3-4- and 5-6-year-old children in order to explore strategy selection in a discrete proportion task across varying levels of numerical abilities. In addition, since the proportional reasoning abilities of these younger children are still unclear, we wanted to investigate children's ability to think about probability in two contexts: comparing probabilities and predicting outcomes. In this way, we can better look at the development of proportional reasoning through both decision-making type-tasks (e.g., which spinner is better?) and in estimation type tasks (e.g., what do you think will happen and how sure are you?) across these distinct age groups.

Lastly, given the importance of early proportional reasoning in predicting formal fraction ability (McMullen et al., 2016), we were also interested in factors that may contribute to individual differences in this ability. Little is known about the cognitive correlates of proportional understanding in young children before they have been introduced to formal fractions. Given that proportional thinking requires the consideration of the *relation* between two quantities, it may be that a domain-general ability to abstract relational information is a necessary component of engaging in proportional reasoning. In fact, a proportion is a form of analogy by definition, requiring an ability to abstract the relation between two quantities (e.g., the relation 20:25 is the same as the relation 4:5), much like analogical reasoning requires an ability to abstract the relation between two entities (e.g., cat:meow as dog:?). Furthermore, analogical reasoning emerges very early in development (e.g., Ferry, Hespos, & Gentner, 2015) and has been found to be a critical component of learning across many domains (English, 2004; Richland & Simms, 2015). We hypothesized that general analogical reasoning (outside the domain of proportions) may be a critical predictor of proportional processing. Thus in the current study, we explored whether children's general analogical reasoning abilities was related to children's ability to reason proportionally across contexts.

In the current study we investigated 3-6-year-old children's proportional processing using a spinner paradigm modeled after Jeong et al. (2007). The focus of our study was to explore factors that contribute to and/or diminish the use of whole-number strategies when engaging in a proportional task. Our study addressed four open questions: (1) Does drawing children's attention to both parts of a proportional display through direct instruction promote proportional thinking? (2) Does experience attending to proportional information diminish the engagement of numerical strategies in a proportional task? (3) Are younger children (3-4 year olds), who have limited numerical abilities and thus may find numerical information less salient, more likely to succeed in a discrete proportional task? (4) Are individual differences in proportional reasoning correlated with general analogical reasoning skills?

Method

Three-to-six year old children participated in three distinct tasks – the spinner comparison task, a single spinner task (requiring children to judge the outcome of a spin on a given spinner and indicate their confidence in that outcome), and an analogical reasoning task. During the spinner comparison task, half of the participants received direct instruction, highlighting attention to the relative size of the red and blue portions of the spinner, whereas the other half of participants did not. Moreover, the order of presentation of the continuous and discrete blocks of trials was counterbalanced across participants, allowing an exploration of the effect of order on performance.

Participants

Two groups of children: Younger Children (3-4 year-olds; N=91; Range: 3.2 to 4.9 years, *M*=4.2 years, 45 females) and Older Children (5-6-year-olds; N=89; Range: 5.1 to 6.8 years, *M*=5.9 years, 42 females) were included in the study. An additional seven children participated, but were excluded because of experimenter or computer error (N=6) or an inability to differentiate the colors used in the task (N=1).

Children were recruited from the greater Boston, MA area and were tested at local museums (Museum of Science, Boston and Boston Children's Museum), day cares, preschools, and after school programs, as well as in the laboratory. In accordance with the guidelines of each testing facility, children received a sticker and/or small prize for participating. Demographic information was not systematically collected for children tested outside of the laboratory. For children whose parents reported demographic information within the laboratory, the sample was predominantly white and educated, approximately as follows: 72% White, 7% Asian, 2% Native Hawaiian/Pacific Islander, 2% Black or African American, and 17% mixed race. In addition, about 15% reported being Hispanic. Of those who reported their education, all mothers had at least a Bachelor's degree and 68% had a Master's degree or higher. Although demographic information was not systematically collected, the demographics of children participating outside of the laboratory are expected to be comparable to those collected in our laboratory.

In addition, a sample of 7-8 year-old children (N=88; Range: 7.1 to 8.9 years, *M*=7.9 years, 51 females) was collected, but these children performed at ceiling on the tasks (75% of children responded correctly on at least 14/16 trials) and so are not included in the analyses (see the Supplementay Appendix 3 for more information about these children).

Design

The Boston College Institutional Review Board approved all study procedures (protocol 10.064, Development of Quantity Concepts) and all parents provided informed consent for their child's participation. Children over the age of 7 also provided written assent.

This study involved a 2 (Training Condition) x 2 (Block Order) between-subjects design, such that children were randomly assigned to one of four between-subject conditions: Training & Continuous First (N=26 Younger; N=23 Older); Training & Discrete First (N=20 Younger; N=21 Older); No Training & Continuous First (N=22 Younger; N=23 Older); and No Training & Discrete First (N=23 Younger; N=22 Older). These sample sizes are in line with other studies using the same or similar paradigm and investigating similar effects (Boyer & Levine, 2015; Jeong et al., 2007). Based on this prior work, we expected medium to large effects of numerical interference (manipulated within subject) and of block order (manipulated between subject). Thus, these sample sizes within each simple comparison of interest are sufficient to detect medium effects with approximately 70% to 90% power. We were unsure of the expected effect sizes of the training and the relation between proportion and analogy.

All task components except the spinner orientation were presented on a Mac Laptop with a 13-inch screen using Xojo programming software. The experimenter recorded all responses using the laptop by pressing the corresponding keys on the keyboard. Each child received two blocks of spinner trials. Each block consisted of: (1)

Spinner orientation, (2) Training (if applicable), and (3) Spinner Comparison. Following the completion of both blocks (one discrete and one continuous), children then participated in two additional tasks. Thus, children participated in the following tasks in a single experimental session lasting approximately 15 minutes: (1) Spinner Comparison Task: Block 1 (consisting of: Spinner orientation, Training (if applicable), Spinner Comparisons) and Block 2 (consisting of: Spinner orientation, Training (if applicable), Spinner Comparisons), (2) Single Spinner Task, and (3) Pattern Analogy Task.

Measures

1. Spinner Comparison Task

Orientation: Prior to each set of comparison trials, children were oriented to an actual cardboard spinner (radius of 7.8cm), half of which was red and half was blue. The discrete spinner was broken into quarters, with the diagonally opposite quarters (i.e., non-consecutive) being the same color. Each circle had a black arrow (6.8cm long) that could be spun around the center of the circle. Children were initially given three stickers and were told that the experimenter would spin the arrow around the spinner and if the spinner landed on red the child would win another sticker and if the spinner landed on blue then the experimenter would take one of the child's stickers away. Then, the experimenter spun the spinner twice, each time behaving in accordance with where the spinner landed (i.e., either giving the child another sticker or taking one away).

Training: Following the orientation, children in the Training group participated in a brief instructional phase during which they were shown a single donut shaped spinner on the computer and were told that they were going to "figure out how to decide what is a good spinner and what is a bad spinner". Children were then asked how much red there was on the spinner, how much blue was on the spinner, and whether that meant it was a "good" spinner or a "bad" spinner. Children were encouraged to use the words "a little" (meaning, less than 50%) or "a lot" (meaning, more than 50%) when describing the amount of each color. If children gave specific number words (e.g., "two pieces"), they were asked whether it was a little or a lot in order to promote continuous, relational information as opposed to numerical information. If children used the words "a little" or "a lot" differently than intended or were incorrect about the spinner being "good" or "bad" (meaning, more or less than 50% chance of "winning", respectively) they were corrected. Children were then provided with an explanation as to why it was a good spinner or bad spinner (e.g., "It's a good spinner, because there is a lot of red and only a little bit of blue!").

Children in the Training group went through two training trials before each block, one with a good spinner (more red than blue) and one with a bad spinner (more blue than red). Children in the No Training group did not have any additional training and instead proceeded straight to the comparison trials.

Comparison Trials: During the comparison trials, children were shown two spinners on the computer screen. The spinners were presented as the same donut-shape as the actual spinners used during orientation, but did not include an arrow and were not spun or acted upon. Instead, children were asked to indicate which spinner would help them win more stickers (i.e., the spinner with the greater likelihood of landing on red). Children were prompted to point to the spinner (either on the left or on the right) and the experimenter recorded the child's responses by pressing corresponding keys on the keyboard. The spinners remained visible until the experimenter recorded the child's response.

The discrete spinners and continuous-matched spinners were presented in two separate blocks of 8 trials each. Between the discrete and continuous blocks (order counterbalanced across children), children were also told "this time, the spinners look a little different, but they work the same way: if the spinner lands on red you win a sticker and if it lands on blue I take a sticker away".

The same proportion values were used for both the discrete and continuous blocks. So, for example, if a discrete trial involved a comparison of 2/5 vs. 5/9, then a continuous trial involved the same comparison magnitudes. Within the discrete block, half the trials were numerically consistent (meaning that the spinner with *proportionally more* red also had *numerically more* red pieces; e.g., 2/5 vs. 5/9) and half of the trials were numerically misleading (meaning that the spinner with *proportionally more* red pieces had *numerically fewer* red pieces, such that comparing the number of red pieces and comparing the proportion of red across spinners would provide different answers; e.g., 4/9 vs. 2/3). Thus, we manipulated the consistency of whole number and proportional information during the discrete spinner blocks. The continuous blocks did not provide whole number information, but trials in the continuous blocks were matched to the specific proportions provided in the discrete block. Thus, continuous trials were divided into continuous consistent-matched trials and continuous misleading-matched trials based upon the specific proportions presented.

The numerically consistent comparisons presented were: 2/6 vs. 5/8; 5/7 vs. 8/9; 4/9 vs. 1/5; 3/6 vs. 5/8. The numerically misleading comparisons presented were: 2/3 vs.

3/9; 1/3 vs. 2/9; 5/10 vs. 4/5; 3/5 vs. 4/9. The ratios between the two proportions presented across trials ranged from 1.25 to 2.2, with an average ratio of 1.65 for the magnitudes used in consistent trials and 1.61 for magnitudes used in the misleading trials. In order to prevent surface area from being a relevant cue for discrimination, we used three different sizes of spinners: Small (diameter of 6.1cm), Medium (diameter of 8.8cm), and Large (diameter of 11.4cm). Within each block of eight trials, two trials involved a Small-Large comparison, two trials involved a Medium-Large comparison, and four trials involved a Small-Medium comparison. On half the trials, the larger spinner was the correct answer and on the other half the smaller spinner depicted the greater proportion of red, making size an unreliable cue for responding. The order of the trials within each block was randomized across participants.

2. Single Spinner Estimation

Following both spinner comparison blocks, children were presented a single donut shaped spinner and asked to decide whether they thought the spinner would land on blue or land on red when spun (the spinners were never actually spun). Following each decision, children were asked how sure they were that the spinner would land on that color by selecting one of five faces representing "Not sure at all/just guessing" to being "Really, really sure" (see Figure 3.1C).

There were six single spinner trials. Three spinners were discrete and three were continuous. Two of the spinners were small, three of the spinners were medium-sized, and one was large. The order of the trials was randomized across participants. The specific magnitudes presented were: 8/9, 2/6, 3/5 (discrete) and 1/5, 4/9, 2/3 (continuous).

3. Pattern Analogy Task

Lastly, children participated in the Pattern Analogy task (modeled after Goswami, 1989). The task consisted of two practice trials followed by six test trials, all of the form A:B → C:?. Children were presented with three pictures at the top of the computer screen, two on the left (A:B; e.g., Yellow Diamond: Yellow Circle) and one following a black arrow and followed by a question mark (→C:?; e.g., Red Diamond: ?). At the bottom of the screen were five picture options, which differed from the given analogy in specific ways: (1) the same as B (e.g., Yellow Circle), (2) the correct shape but incorrect color (e.g., Blue Circle), (3) the correct color but incorrect shape (e.g., Red Square), (4) the same as C (e.g., Red Diamond), and (5) the correct answer (e.g., Red Circle; see Figure 3.1D for an example). During the practice trials, the experimenter covered up the options with a piece of paper and helped the child through the trial by asking the child what was the same and what was different in the A:B pair (e.g., yellow circle, yellow square: both are yellow, but one is a circle and one is a square). Then, children were instructed to pick the picture that went in the spot with the question mark in order to make the two sides match. After going through the pattern, the experimenter revealed the options and asked the child to choose. After the child selected, they were given feedback and an explanation (e.g., "The red square goes with the red circle, so that they're both red but one is a circle and one is a square just like the other pictures"). Following the two practice trials, children were presented six test trials in which they did not receive any feedback.

The order of the five options was randomized across trials, but was the same for each child. The order of the practice trials was identical across children, but the order of the test trials was randomized. All shapes were inside a small white box (27.0 cm^2) and approximately the same size.

Figure 3.1: Example Stimuli

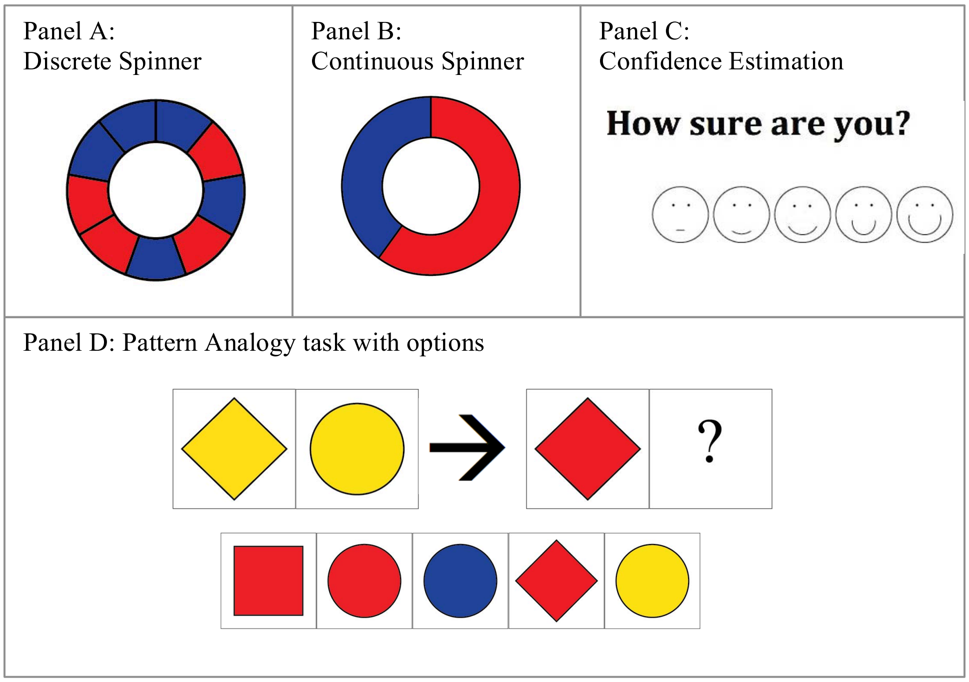


Figure 3.1: Example stimuli from each task: An example of a discrete spinner (Panel A) and a continuous spinner (Panel B) like those used in the comparison task and in the single spinner task; the scale children used to respond about their confidence level of the single spinner estimation trials (Panel C; note that the experimenter verbally read the question and explained the response scale); and an analogy problem with the options children were given to choose from along the bottom (Panel D).

Data Scoring

Accuracy was the primary dependent variable on all tasks. All children included in the analyses completed both spinner comparison blocks (discrete and continuous). One child in the younger group did not complete the Single Spinner task because of time constraints and that child's data were excluded from those analyses. Twenty-seven younger children and nine older children did not participate in the Pattern Analogy task because of time constraints, loss of attention during the task, or lost data due to computer error (N=1). Thus, a total of 142 children (N=62 younger and N=80 older) had complete data across all tasks. Notably, since these data are likely not missing at random (children with lower attention spans probably had a higher likelihood of not completing the Pattern Analogy task), regression analyses should be interpreted with this limitation in mind.

Results

Spinner Comparison Task

First, children in both age groups performed significantly above chance in the continuous block of the spinner comparison task (Younger: *M* (SE)=59% (2.6), $t(89)=3.4$, $p<0.005$; Older: *M* (SE)=80% (2.2), $t(88)=13.4$, $p<0.001$), revealing that on average children were able to engage in proportional reasoning. However, as a group, the younger children were only slightly above chance and a closer look at these children suggests that many of them did not perform above chance. Differences between these children are discussed after the main analyses.

Accuracy scores on the spinner comparison task were subjected to an ANOVA with Block type (2: Discrete, Continuous-Matched) and Trial Type (2: Misleading, Consistent) as repeated measures and Age Group (2: Younger, Older), Training (2: Yes, No), and Order (2: Continuous First, Discrete First) as between-subject factors. Analyses revealed a small main effect of Training ($F(1,172)=5.1$, $p = 0.026$, *partial* $\eta^2 = 0.03$), with those who received training (*M* (SE)=72% (2)) outperforming those who did not (*M* (SE)=65% (2)). However, training did not significantly interact with any other variables

(ps>0.05). In line with previous studies, analyses revealed a numerical interference effect, as indicated by a significant Block Type X Trial Type interaction ($F(1,172) = 13.03$, $p < 0.001$, *partial* $\eta^2 = 0.07$), revealing that children relied upon numerical information when it was available, performing better in discrete trials (compared to magnitude matched continuous trials) when numerical information was consistent ($M_{Continuous}$ (SE)=66% (2.1), $M_{Discrete}$(SE)=71% (2.4); $p = 0.04$) but performing worse in discrete trials (relative to magnitude matched continuous trials) when numerical information was misleading ($M_{Continuous}$ (SE)=72% (2.1), $M_{Discrete}$ (SE)=65% (2.1); $p = 0.01$)[2].

Furthermore, these factors (Block Type X Trial Type) significantly interacted with age group (three way interaction: $F(1,172)=11.9$, $p < 0.001$, *partial* $\eta^2 = 0.07$) and Age Group X Order (four way interaction: $F(1,172)=4.5$, $p = 0.034$, *partial* $\eta^2 = 0.03$). Given that the interference effect (seen in the Block Type x Trial Type interaction) differed across both age and block order, we looked at the impact of order, block type, and trial type in the younger and older groups separately using two 2 x 2 x 2 ANOVAs.

Analyses of data from the older children revealed both a Block Type X Trial Type interaction ($F(1, 87)=25.1$, $p<0.001$, partial $\eta^2=0.2$) and an Order X Block Type X Trial Type interaction ($F(1, 87)=7.3$, $p=0.008$, partial $\eta^2=0.08$). These interactions indicated

---

[2] It is worth noting that although it appears as though performance on the continuous "misleading" trials is higher than continuous "consistent" trials, this difference does not have a meaningful interpretation within the context of our manipulation (since the "misleading" and "consistent" manipulation is non-existent in the continuous trials). Rather, any differences here may reflect unintended differences between the specific stimuli and ratio comparisons used for these trials. However, we matched the specific magnitudes across the continuous and discrete trials specifically to control for this possibility of ratio differences. Moreover, the direction of this effect is the opposite of the numerical interference effect. Thus, this unintended difference in stimuli is not the cause of the numerical interference effect, and if anything may be muting the numerical interference effect reported here.

that performance in the older group was significantly impacted by whole number information on the discrete trials, but this interference differed depending on which block of trials children received first. Specifically, older children who received the discrete block prior to the continuous block demonstrated the expected numerical interference effect (Block Type X Trial Type interaction: $F(1, 42)=22.97$, $p<0.001$, partial $\eta^2=0.4$; see Figure 3.2, left panel). That is, performance was higher on discrete trials (compared to the equivalent continuous trials) when whole number information was consistent ($M_{Discrete}$ (SE)=87% (3.5) vs. $M_{Continuous}$ (SE) = 80% (3.2); $t(42) = -2.23$, $p = 0.03$, Cohen's d = 0.34) and lower (compared to the equivalent continuous trials) when whole number information was misleading ($M_{Discrete}$ (SE)= 62% (4.9) vs. $M_{Continuous}$ (SE)=89% (2.8); $t(42) = 4.48$, $p <0.001$, Cohen's d = 0.68). On the other hand, data from older children who received the continuous block first only showed a marginal Block Type X Trial Type interaction $F(1, 45)=3.6$, $p=0.06$, partial $\eta^2=0.08$ (see Figure 3.2, right panel), with no statistically significant difference in performance on discrete and continuous-matched trials when numerical information was misleading ($M_{Discrete}$ (SE)=76% (4.46), $M_{Continuous}$ (SE) =80% (4.0); $t(45)=0.94$, $p=0.35$, Cohen's d = 0.14) or when whole number information was consistent ($M_{Discrete}$ (SE) = 78% (4.0), $M_{Continuous}$ (SE) = 72% (4.2); $t(45)=1.41$, $p=0.17$, Cohen's d = 0.2).

Analyses of data from the younger group, on the other hand, revealed no main effects or interactions (*ps*>0.1; see Figure 3.3). That is, data from the younger group, in contrast to their older peers, did not reveal evidence of a statistically significant numerical interference effect.

Figure 3.2: Older children's Performance on Spinner Comparison Task

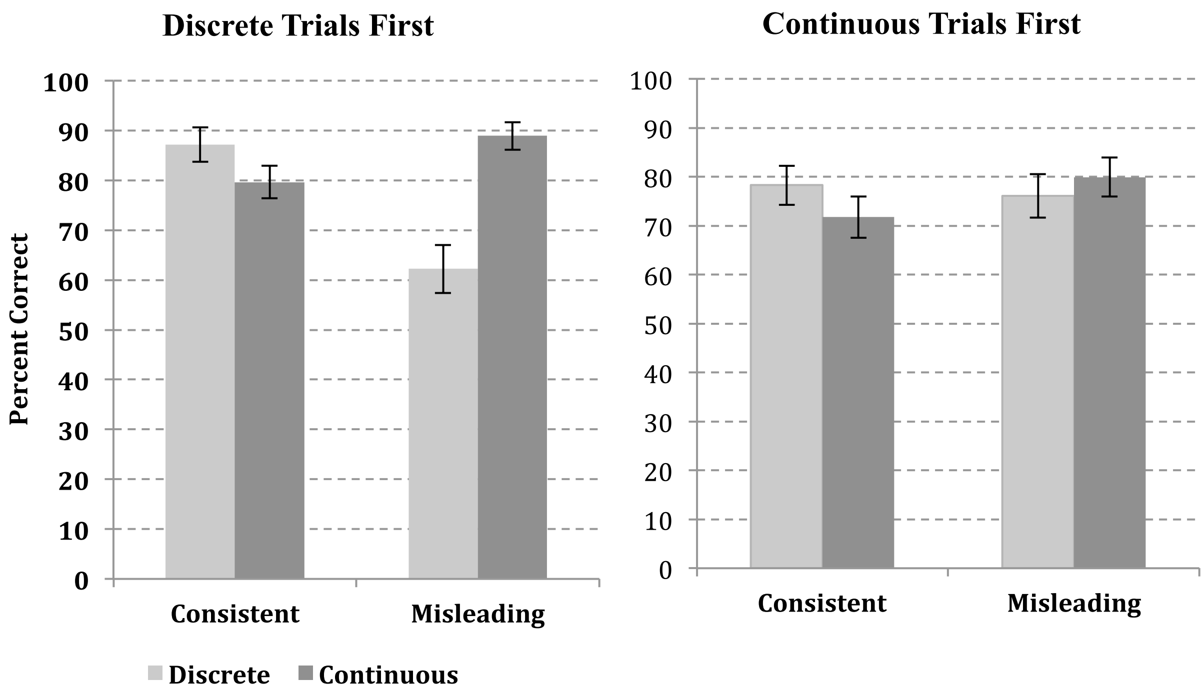


Figure 3.2: Older children's performance on the spinner comparison task separated into children who received the discrete block first (left) and children who received the continuous block first (right).

Figure 3.3: Younger Children's Performance on Spinner Comparison Task

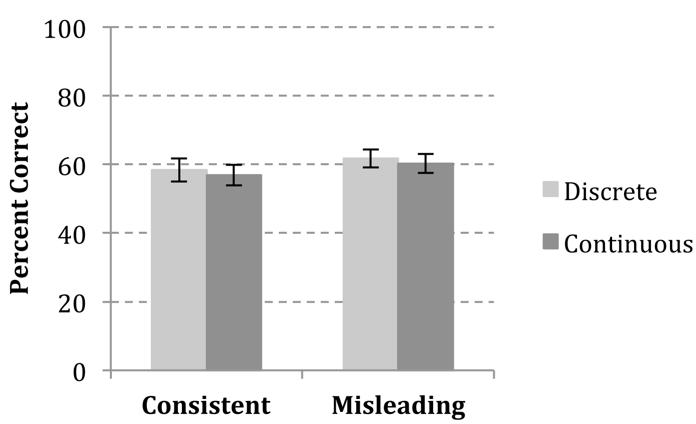


Figure 3.3: Younger children's performance on the spinner comparison task, across all children regardless of the order in which they received the blocks.

However, this younger group also performed much worse than the older group. Although, on average, they performed significantly above chance (*M*=59% on continuous trials), a closer look at these children reveals stark individual differences with many children performing below 50% correct. In order to investigate what might be happening with these younger children, we restricted additional analyses to data from only those children who scored above chance on the continuous trials (>50% correct), who demonstrated some understanding of the task demands and of continuous proportional information (N=49; M_{age}=4.2 years; Range 3.3-4.9 years; $M_{Continuous}$=77%; Range: 63% to 100% correct). Using a Block Type (2: Discrete, Continuous) X Trial Type (2: Consistent, Misleading) repeated measures ANVOA with Order (2: Continuous First, Discrete First) as a between subject comparison on only this subset of data, there was a main effect of block type, $F(1, 47) = 7.45$, p=0.009, partial η^2=0.137, with lower performance on the discrete trials (M=70%) than the continuous trials (M=77%). However, there were no other significant main effects or interactions (*p*'s>0.1), revealing that these younger children performed worse on *both* numerically misleading trials and numerically consistent trials, relative to performance on continuous-matched trials (Misleading: $M_{Continuous}$ (SE)= 77% (2.6), $M_{Discrete}$ (SE) =66% (3.8); Consistent: $M_{Continuous}$ (SE) = 78% (2.5), $M_{Discrete}$ (SE) = 74% (3.9)). Thus, data from younger children who performed above chance on the continuous trials still did not provide evidence of a numerical interference effect, but rather performed worse on discrete trials overall.

Single Spinner Trials

In addition to children's ability to compare probabilities (in the comparison task), we were interested in investigating children's abilities to predict the outcome of single probability events and whether the uncertainty of these probabilities was reflected in children's confidence judgments about the outcome. We did this by looking at performance on the single spinner trials in which children were asked to judge which color they thought the spinner would land on (i.e., what color had the highest probability). On average, both younger, M(SE) =70.1% (2.4), and older, M(SE)=89.8% (1.6), children performed above chance (50%; Younger: $t(89) = 8.29$, $p<0.001$; Older: $t(88) = 25.2$, $p<0.001$), but the older group outperformed the younger group ($t(152.98) = 6.83$, $p<0.001$)[3]. Given the poor performance of the younger group on the comparison task, it is worth noting that even those younger children who performed at or below chance on the continuous trials of the comparison task (N = 41) also performed significantly above chance on the single spinner trials, M (SE) = 67% (3.7), $t(40) = 4.5$, $p<0.001$.

Children's confidence judgments were analyzed by computing a slope between confidence judgments and how far the ratio of the spinner was from 50% (i.e., the highest level of uncertainty). A positive slope would indicate that the child adjusted their confidence judgments based on the degree of uncertainty in the spinner. That is, the further the probability was from 0.5, the higher confidence the child reported. Thirty-six children are excluded from these analyses: one younger child for experimenter error, as well as 15 younger and 20 older children for responding with the same confidence rating

[3] Note that although both continuous and discrete trials were included in this task, they were not perfectly matched on magnitude, making it impossible to make direct comparisons of performance on these two trial types.

on every trial (generally highest confidence). On average, data from younger children did not reveal statistically significant positive slopes, either on all trials (M_{slope}(SE) = 0.09 (0.1), $t(73)=0.85$, $p=0.4$, Cohen's d = 0.1) or on correct trials only (M_{slope}(SE)=0.08 (0.18), $t(71)=0.65$, $p=0.65$, Cohen's d = 0.05). Older children, however, did produce confidence judgments with significantly positive slopes on both correct trials (M_{slope}(SE)=0.59 (0.13), $t(68)=4.6$, $p<0.001$, Cohen's d =0.56) and across all trials (M_{slope}(SE)=0.60 (0.12), $t(68)=4.9$, $p<0.001$, Cohen's d =0.59).

Predicting Analogical Reasoning

Our primary interest was whether children's performance on proportion-based tasks was related to their more general ability to reason about patterns analogically and whether this depended on the context of the proportional reasoning. Thus, we computed partial correlations between pattern analogy scores (calculated as proportion correct, although see the Supplementary Appendix 3 for an alternate scoring system that takes into account the type of errors children made) and comparison scores on discrete trials, comparison scores on continuous trials, and a measure of the interference effect, controlling for age[4] and the other measures (see Table 3.1 for means and standard deviations of each measure). Our interference measure was calculated as: *performance on numerically consistent discrete trials – performance on numerically misleading discrete trials*. If children consistently relied upon numerical information, then this value should be close to 1 (near ceiling on consistent trials and near floor on misleading trials), whereas if children did not rely upon numerical information when making their

[4] Age was treated categorically for consistency across analyses. However, an identical pattern of results is found when age is treated continuously.

judgments, this value should be near zero (about equal performance, regardless of the nature of the discrete information).

Table 3.1: Descriptive Statistics for Correlational Analyses

		Mean Proportion Correct (SE)			
	N	Pattern Analogy	Discrete Trials	Continuous Trials	Numerical Interference
Younger Children	62	0.22 (0.03)	0.64 (0.03)	0.59 (0.03)	-0.024 (0.04)
Older Children	80	0.58 (0.03)	0.75 (0.02)	0.80 (0.02)	0.13 (0.05)

Accuracy on the Pattern Analogy task (when controlling for age and the other measures) was significantly correlated with performance on continuous proportion trials ($r = 0.2$, $p = 0.019$), but not with performance on discrete proportion trials ($r = 0.006$, $p = 0.79$) or numerical interference scores ($r = -0.08$, $p = 0.33$).

Discussion

The current study investigated 3-6-year-old children's attention to relational versus whole number information in discrete probabilistic contexts. In particular, we aimed to investigate whether specific context effects and individual differences may impact children's strategy selection when proportional information is presented discretely, as opposed to continuously. In doing so, the current data reveal substantial insights into the malleability of children's attention to whole number information when making proportional judgments and how individual differences in proportional reasoning may be related to more domain general skills.

The current study was the first to investigate numerical interference using this paradigm with 3- and 4-year-olds. Although as a group they performed slightly (and significantly) above chance on the comparison task, many children (approximately 50%) performed below chance on the continuous trials, which did not involve an opportunity for numerical interference. On the other hand, even those younger children who performed at or below chance on the comparison task were successful on the single spinner trials (judging whether they thought it would land on blue or red). Thus, it may be that deciding the probability of a single outcome is an easier task than comparing two outcomes (when the total number of pieces differ). Given that correctly judging a single spinner requires assessing the relationship between only two amounts (red vs. blue), while comparing two spinners requires comparing the relations among four different amounts (red vs. blue in each spinner, then comparing across spinners), it may not be surprising that children performed better on the single spinner task. Firstly, excessive cognitive demands (e.g., working memory, strategy selection, comparing multiple outcomes) of the comparison task set up may have limited their ability to demonstrate proportional understanding in the comparison task. Moreover, it could be argued that performance on the single spinner task may not necessarily require proportional reasoning abilities, but just an ability to compare the amount of blue to the amount of red without encoding this relation in any real way. Thus, although our findings point to the possibility that 3-4-year-old children understood the probabilistic outcomes tied to the relative amounts of red and blue, they do not rule out the possibility that children simply engaged a heuristic, such as "is there more red than blue" or "is there more red than 1/2".,

Interestingly, even those younger children who performed above chance on the comparison task did not produce data revealing a statistically significant numerical interference effect. Rather, these younger children generally performed worse on the discrete trials overall than the continuous trials, regardless of whether or not numerical information conflicted or aligned with proportional information. Given that the discrete trials involved non-consecutive red pieces, using a proportional strategy on these trials may have required more mental spatial manipulation than the continuous trials, potentially leading to their lower performance. Although we cannot claim that the younger group did not use whole number information, the findings that children performed worse overall on discrete trials than continuous trials, but not in a way consistent with whole number information, hints that maybe young children were less likely to use numerical strategies. In other words, given the young age of these children, they may not have seen the discrete trials as being a numerical task at all. This may be because these children likely had lower numerical abilities (e.g., might still be in the process of learning to count) and/or because numerical information is simply less salient to them (consistent with Sophian et al., 1995). Since we did not measure the numerical abilities of the children in the current study, our data cannot speak to this explicitly. However, given that between the ages of 3- and 6-years of age children are gaining a substantial amount of formal whole number knowledge (e.g., Hurst, Anderson, & Cordes, 2016; Le Corre & Carey, 2007; Wagner & Johnson, 2011) as well as proportional reasoning ability (e.g., Girotto et al., 2016; Sophian, Garyantes, & Chang, 1997), this may be a key period for investigating the developmental progression of these two concepts and how they may rely (or not) upon each other.

In contrast to the younger group, 5-6 year-olds performed quite well on both the comparison task and the single spinner task. Replicating previous findings (e.g., Boyer et al., 2008; Jeong et al., 2007), when whole number information was made available (i.e., the comparisons were discrete with different denominators) these older children relied on this information—performing better when whole number information was helpful and performing worse when whole number information was misleading.

This numerical interference effect, however, was diminished when children were first exposed to continuous trials prior to the discrete block. That is, children primed to use a proportional strategy (in the continuous block) were less likely to switch to a counting strategy, even when numerical information became available. It is notable that this transfer occurred despite quite an obvious stopping point between blocks that was highlighted by the experimenter (e.g., "now the spinners look different"). This finding is consistent with studies investigating formal fraction instruction, which suggest that introducing discrete fraction content as an extension of more continuous instruction with decimals and percentages (e.g., Moss & Case, 1999), rather than the other way around, may promote rational number understanding. By introducing content in this order, children may be more likely to extend the continuous proportional strategies into these new discrete contexts and rely less upon discrete numerical information overall. The benefits of introducing proportional information in the context of continuous representations prior to discrete representations are one potential avenue for future research.

Recently, Boyer and Levine (2015) reported a similar finding using a slightly different paradigm (a match-to-sample paradigm, where a child is asked to pick which of two options is the proportional match to a target, while ignoring the numerical match). Boyer and Levine reported that 4th graders who were given prior experience with continuous trials were less likely to use a numerical strategy on the subsequent discrete trials. Extending these findings, our data reveal this pattern as early as 5-6 years of age (corresponding to approximately kindergarten), suggesting that even at this young age, children's strategies for engaging in proportional tasks may be fairly malleable. It should be noted, however, that although Boyer and Levine tested younger children, they did not find the prior-experience effect in 2nd graders or kindergartners. This contrast may be due to several minor differences across paradigms. For example, match-to-sample tasks (as those used in Boyer and Levine) may be more cognitively taxing than comparison tasks, as they require making comparisons across a greater number of stimuli (3 stimuli versus 2 stimuli). In fact, the kindergarteners and 2nd graders studied by Boyer and Levine were less accurate overall in their task than the 5-6-year-olds were in the current study (the average score for kindergarteners in Boyer and Levine was between 50%-65% correct, whereas the 5- to 6-year-olds in the current study had average scores between 62%-89% correct). Thus, it may be that 5- to 6-year-old children are able to inhibit their reliance on numerical strategies, but only in contexts that are otherwise relatively easy for them. That is, when the task is challenging, children may be more likely to resort to relying on whole number information and less likely to use proportional strategies. However, it remains unclear why children would rely on numerical information in challenging contexts, given evidence that proportional reasoning may be a particularly early developing tool for encoding information (Duffy, Huttenlocher, & Levine, 2005; Huttenlocher, Duffy, & Levine, 2002). It may be that for children this age, numerical strategies are more salient or easier to match exactly than proportion (since counting and comparing the two salient sets is seemingly less taxing than comparing the relative size of four quantities). In contrast, in relatively simple tasks, children's reliance upon numerical information may be more malleable, allowing them to switch back to the more approximate proportional strategy when given the right experience.

Although experience with continuous trials encouraged children to rely on a proportional, rather than whole number, strategy, we did not find an effect of our direct instructional training. Overall, children in the training group did out-perform their peers who did not receive training, suggesting that highlighting the relevant parts of the displays may have increased children's attention to the task at a more global level. However, the training did not have a specific effect on strategy selection (i.e., did not result in a reduction of numerical interference). It may be that our instructional training was simply not strong enough to help children overcome their tendency to focus on numerical information or that there may be a more effective way to highlight relevant information. For example, we aimed to highlight the parts (i.e., number of red and number of blue), but it may be that children would benefit more from highlighting attention to the total number of pieces (i.e., the denominator) in addition to the number of red pieces (i.e., the numerator).

In addition to generally succeeding on the proportion tasks, there was a significant, positive relationship between older children's judgments of how confident they were of a particular outcome (e.g., that the spinner would land on blue) and the actual probability of the outcome. That is, children reported being more confident in their response when the likelihood of the outcome was actually much higher than the likelihood of the other outcome. Thus, by 5-6-years-old, children already show relatively sophisticated probabilistic reasoning in both estimation-type (what color do you think it will land on) and comparison (which spinner is more likely to land on red) contexts in which they are unable to rely solely on numerical information. Younger children, on the other hand, did not show this same pattern with their confidence judgments. It may be that 3-4 year-olds simply had trouble understanding the question or response scale. Research looking at preschoolers' abilities to provide confidence judgments in other domains suggest that this skill is still developing between 3- and 4-years-old (Hembacher & Ghetti, 2014; Ghetti, Hembacher, & Coughlin, 2013), so it may be that our younger participants were not able to articulate their uncertainty in this context. Alternatively, our data may reflect a limit to 3-4-year-olds' understanding of probability even in these single estimation contexts. That is, this finding may further suggest the possibility that these younger children approached the task using heuristic based judgments of "which has more" without encoding or representing the precise proportion or probability. Additional studies should investigate how early understanding of probability in this preschool age range may depend on specific aspects of the task demands. In particular, given the varying cognitive skills (e.g., working memory) and specific knowledge (e.g., whole number knowledge, estimation skills) required across distinct tasks commonly used in the literature (e.g., single display estimation tasks, comparison tasks, match-to-sample tasks), these tasks are likely not interchangeable. However, performance differences across these tasks may shed important insight into how children's understanding of proportion and ability to reason about the relevant relational information develops.

Lastly, the current study suggests that children's more general ability to reason analogically using geometric patterns is significantly related to their proportional reasoning in continuous contexts, but not in discrete contexts. Given that children engaged in numerical strategies on discrete trials, this finding might suggest that relational reasoning is specifically related to proportional reasoning (used on continuous trials) more so than counting or numerical reasoning (presumably the dominant strategy used on discrete trials). However, analogical reasoning (as measured here) was not significantly correlated with the extent to which children relied upon numerical information, in particular. Furthermore, as noted in the Method, given the not at random missing data for this analysis, these findings should be interpreted with caution. Rather, our goal is to suggest some directions for future work, to further investigate the potentially unique role of analogical reasoning in proportional thinking over other quantitative and mathematical knowledge. In particular, it is critically important to investigate the causal nature of this relationship (if any) and in particular whether analogical training, aimed at facilitating children's analogical reasoning through structural alignment, can foster recognition of the relational nature of proportion contexts (as has been shown in other domains, e.g., Gentner, Levine, Ping, Isaia, Dhillon, Bradley, & Honke, 2016).

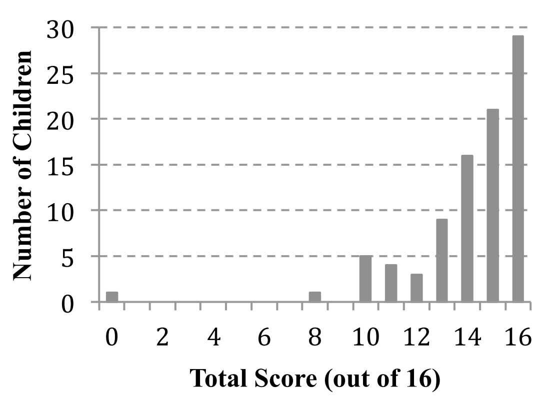
In sum, the current study provides substantial insight into the context effects and individual differences in children's use of numerical and proportional information in discrete and continuous probabilistic contexts. Future research should continue investigating the relationship between numerical and proportional information in order to further elucidate the factors that may lead to or prevent numerical interference (e.g., the saliency of numerical information in the fraction symbol system, global attention to the relation versus local attention to features during specific task demands), how the relationship may be leveraged to benefit learning rather than lead to interference (e.g., teaching continuous proportional information before discrete proportional information), and how domain general analogical reasoning skills may be leveraged to help children attend to relational information.

Supplementary Appendix 3

Additional Analyses

Additional 8-year-old Sample

An additional sample of 7-8 year old children also participated in this study. This group performed near ceiling levels on the task, leading to substantial skew in the distribution of their responses, violating assumptions of normality and making it difficult to analyze children's responses given the few number of questions per cell (4 trials per cell when broken down into discrete vs. continuous and misleading vs. consistent). On average, children performed at least 80% correct on all trial types: Consistent Trials $M_{Continuous}$ = 88% (SD=20%), $M_{Discrete}$ = 93% (SD=15%); Misleading Trials $M_{Continuous}$ = 93% (SD=18%), $M_{Discrete}$ = 80% (SD=30%). The figure below displays a histogram of overall performance (total possible score of 16, across the 4 trial types). Notably, almost a third of the children (29/88) got all the questions correct and more than half (50/88) got at most one incorrect. Below, we provide an analysis of this group's performance using analogue analyses as those used in the manuscript. It is worth noting that one child got all the problems incorrect (0/16). When this child is excluded from the analyses, the overall pattern of results reported below does not change.



Despite this extreme left skew, data from 7-8-year-olds also showed evidence of an interference effect. We conducted a 2 (Block Type) X 2 (Trial Type) repeated measures ANOVA, with Training (2) and Order (2) as between subject measures. There was a main effect of trial type, $F(1,84) = 4.22$, $p = 0.04$, partial $\eta^2 = 0.05$, and a marginal effect of block type, $F(1,84) = 3.9$, $p = 0.053$, partial $\eta^2 = 0.045$. However, these effects were qualified by a significant Block Type X Trial Type interaction, $F(1,84) = 24.27$, $p < 0.001$, partial $\eta^2 = 0.22$. In particular, as expected, children performed better in the discrete trials compared to the matched continuous trials when numerical information was consistent, $t(87) = 2.65$, $p = 0.01$, and worse in discrete trials compared to the matched continuous trials when numerical information was misleading, $t(87) = 4.3$, $p < 0.001$.

There was also a marginal Block Type X Training interaction, $F(1,84) = 3.9$, $p = 0.053$, partial $\eta^2 = 0.045$ and a significant Block Type X Order interaction, $F(1,84) = 12.7$, $p < 0.001$, partial $\eta^2 = 0.13$. In particular, those in the No Training condition performed similarly on the discrete and continuous blocks ($M_{Discrete} = 89\%$; $M_{Continuous} = 89\%$; $t(43)=0.198$, $p=0.8$) while those in the Training condition performed better on the continuous block (92%) than the discrete block (85%; $t(43)=2.86$, $p=0.007$). When looking at order effects, those who received the discrete block before the continuous performed better on the continuous block (94%) than the discrete block (84%; $t(45)=3.72$, $p=0.001$) and those who received the continuous block before the discrete, did not perform significantly different across the two blocks ($M_{Continuous} = 87\%$; $M_{Discrete} = 90\%$; $t(41)=1.1$, $p=0.28$). However, these interactions did not further interact with trial type and there were no other significant main or interaction effects (ps>0.1).

These findings are consistent with other work suggesting that this age group shows a numerical interference effect when evaluating discrete proportional displays. In addition, the effects are overall consistent with, but less nuanced than, the pattern shown by the 5-6-year-olds in the current manuscript. That is, children who received the discrete trials before the continuous trials performed worse on the discrete trials relative to the continuous trials, whereas those children who received the continuous trials before the discrete trials did not perform differently across the two blocks of trials. This gives some hint that having the continuous trials before the discrete trials may impact these children's strategy choice, and in particular lead to similar performance across the two blocks. However, these interactions did not significantly interact with trial type, suggesting that the pattern is less specified to particular trial types in this sample (as they were with the 5-6-year-olds), and may not be entirely explained by overcoming a numerical bias. Furthermore, the small (and marginal) Training x Block interaction suggests that our training led to a larger difference in performance between continuous and discrete blocks, relative to those in the no training condition. This further emphasizes that our training may not have been best suited to improving performance on the discrete trials specifically. It is possible that the less nuanced pattern of effects found for these 7-8-year olds (relative to the 5-6-year-olds) is because of lower variability in responses at the upper end of performance. However, given the extreme skew in the distribution suggesting ceiling level performance, both overall and within each trial type, the interpretation of these effects is limited. The analyses are provided here only for transparency and completeness, as well as to provide some potential insight into the performance of this oldest age group.

Alternate Scoring of Pattern Analogy

Many of the children, particularly the younger children, performed fairly poorly on the analogy task (for example, over a third of the children received less than 35% correct) based on using the dichotomous correct/incorrect scoring. However, this measure of accuracy treated all errors as equal, despite the fact that some of the "wrong" options were more related to the pattern than other options. Thus, in order to differentiate children who had some knowledge of how to make an analogy from those who did not, we also scored the analogy task using a slightly different method. In this error-based scoring scheme, the scores were not based on perceptual similarity to the correct answer, but rather whether the perceptual match was one that took into account the first half of the analogy (suggesting some understanding of needing to make a "match" *across* these relations) versus a perceptual match that ignored the first half of the analogy, which is likely a less difficult match (given the perceptual and temporal proximity). However, these distinctions were not tied to specific perceptual features (e.g., color vs. shape matches), but instead based on the location of the perceptual match in the analogy.

In our alternate scoring method, each trial of the Pattern Analogy task was scored with 0 to 3 possible points, so that children were given partial credit for making perceptual matches, even if they did not make the relational match. Selecting the correct analogical match was worth 3 points, reflecting a complete understanding of the analogical relation (e.g., the red circle in Figure 3.1D of the main text). Children who selected the same image as "B" (in A:B::C:?), were considered to have made the far-perceptual match (2 points; the yellow circle in Figure 3.1D). Meaning, they understood that they were required to reference the first half of the pattern (A:B), but selected the perceptual match rather than the relational match. Children who selected the same image as “C” were considered to have made a close-perceptual match (1 point; the red diamond in Figure 3.1D of the main text). Meaning, they understood the need to make a pattern, but selected the simple perceptual match on the second half of the analogy only (C:?) without referencing the first half, neither perceptually or relationally. The other two options were no directly related to the pattern, either perceptually or relationally (blue circle or red square in Figure 3.1D of the main text), and so selection of these options indicated no understanding of completing any clear pattern (0 points). Thus, the total possible score on the Pattern Analogy task ranged from 0 to 18 (across all six trials) and is reported as a proportion out of 18. This scoring method incorporated error-type in order to illuminate meaningful differences among low-scoring children. Thus, this scoring system creates a continuum from little to no understanding of patterning, to understanding basic aspects of patterning, to understanding relational and analogical patterning in particular.

Table: Descriptive Statistics for the alternate scoring of Pattern Analogy

	Pattern Analogy: Mean proportion correct (SD)
Younger Children (N=62)	0.48 (0.02)
Older Children (N=80)	0.74 (0.02)

Using this scoring scheme, pattern analogy score was significantly correlated with performance on the continuous trials ($r = 0.245$, $p = 0.004$), but not on discrete proportion trials ($r = 0.04$, $p = 0.618$), mirroring findings reported in the main text using the dichotomous scoring scheme. Furthermore, pattern analogy performance was significantly correlated with interference scores ($r = -0.17$, $p = 0.045$), suggesting that children who were more likely to use number during the discrete trials were also likely to have lower analogical reasoning scores.

Given that this relation is small, and not reproduced using an alternate scoring scheme, it should be interpreted with caution. However, it may point to the idea that children with lower analogical reasoning skills may be more likely to rely on numerical information and not proportional information.

CHAPTER 4: TALKING ABOUT PROPORTION: FRACTION LABELS IMPACT NUMERICAL INTERFERENCE IN NON-SYMBOLIC PROPORTIONAL REASONING

Abstract

Children's abilities to reason about proportional information are greatly influenced by the availability of discrete, whole number information. Unfortunately, because formal fraction labels emphasize whole number information (e.g., "*three*-fifths"), they serve to highlight differences across equivalent fractions (e.g., 3/4 and 6/8) while obscuring proportional magnitude. In two experiments, 5-7 year-old children were introduced to equivalent non-symbolic proportions labeled in one of three ways: (1) a single, categorical label for multiple fractions (both 3/4 and 6/8 referred to as "blick") which highlighted similarities, (2) labels that focused on the numerator (e.g., 3/4 labeled as "three blicks" or "three fourths"), or (3) labels that had a complete part-whole structure ("three-out-of-four"), with varying amounts of additional information. Children then completed measures of non-symbolic proportional reasoning that pitted whole-number information against proportional information. Across both experiments, children who heard the categorical labels were more likely to match non-symbolic displays based upon proportion than children in any of the other conditions, who demonstrated higher levels of numerical interference. These findings suggest that fraction labels impact children's attention to proportional information even in the context of part-whole displays and in particular, the use of whole number words within fraction labels may turn children's attention toward whole number information and away from proportional magnitude. We discuss the implications for these findings in terms of children's proportional reasoning

and fraction education. In addition, we discuss how individual differences in math ability and task differences may relate to numerical interference.

We regularly make informal judgments about visually presented proportional information in our environment, whether it be judgments of physical support (e.g., whether a book that is mostly, but not entirely, on a table is likely to fall off), chance and probability (e.g., if a bag of candy is mostly a disliked candy, you wouldn't want to reach in and grab one at random), or interpreting commonly presented formal information (e.g., interpreting a pie chart during a work meeting). Although this is common practice throughout our lives, these visual displays of proportional information are not always straightforward. In particular, they are made difficult by the fact that proportion is a relation between two quantities and the magnitude of these two quantities (e.g., numerical values) can interfere with our attention to the *relation* between the quantities (termed "numerical interference"; Boyer, Levine, & Huttenlocher, 2008; Hurst & Cordes, under review - A; Jeong, Levine, & Huttenlocher, 2007; Ni & Zhou, 2005).

Despite these potential difficulties, infants and young children show a fairly sophisticated ability to reason about proportional information (Denison & Xu, 2010; McCrink & Wynn, 2007; Sophian, Harley, & Martin, 1995; Spinillo, 2003). For example, 6-month-old infants attend to and keep track of the ratio between discrete items (McCrink & Wynn, 2007) and can use this ratio information to make probabilistic inferences (Denison, Reed, & Xu, 2013; Denison & Xu, 2010). For example, 12- to 14-month-old infants are surprised by an experimenter pulling a red ball from a bin that contains mostly white balls and only a few red balls, suggesting that they track the likelihood of probabilistic outcomes (e.g., Denison & Xu, 2010).

However, for both children and adults, proportional information is not always the most salient feature of these displays. When visual proportion is presented discretely (i.e., with distinct, divided units that can be explicitly counted, like a bag of candy or a rectangle divided into quarters), this countable numerical information may be particularly salient and can interfere with our ability to focus on the relation between the two components. The interference between absolute number and relative proportion has been at the center of a substantial amount of work with symbolic (e.g., Alibali & Sidney, 2015; Durkin & Rittle-Johnson, 2015; Ni & Zhou, 2005) and non-symbolic representations of proportion (e.g., Boyer et al., 2008; Boyer & Levine, 2015; Fabbri, Caviola, Tang, Zorzi, & Butterworth, 2012; Hurst & Cordes, under review - A; Jeong, et al., 2007). In particular, the availability of countable, discrete quantities has been shown to lead to systematic and predictable errors in the way children respond in proportion tasks. For example, children show success comparing and matching proportions when the proportions are presented continuously (i.e., in the absence of countable numerical information); but when proportional information is presented discretely, children will often ignore proportional information and instead respond in a way that is consistent with attending to the *number* of relevant pieces (i.e., the number of pieces in the "numerator") rather than the *relation* between the number of pieces and the total number of pieces (i.e., the overall proportion; Boyer et al., 2008; Boyer & Levine, 2015; Hurst & Cordes, under review - A; Jeong et al., 2007). Given recent evidence that individual differences in children's spontaneous attention to proportion is related to later arithmetic and fraction ability (McMullen, Hannula-Sormunen, Laakkonen, & Lehtinen, 2016; McMullen, Hannula-Sormunen, & Lehtinen, 2014), it is critical to understand factors that may highlight children's attention to numerical information at the expense of proportional reasoning.

Research has pointed to numerous factors that likely contribute to the prevalence of this numerical interference, such as teaching methods, previous experience, strategies, and individual differences in understanding fractions (e.g., Alibali & Sidney, 2015; Boyer & Levine, 2015; Empson, Junk, Dominguex, & Turner, 2006; Hurst & Cordes, under review - A; Spinillo, 2002). In the current study, we explore another factor that has been less explored – the words we use to talk about proportional information: fraction labels. Fraction labels are complex – involving multiple words, including whole-numbers labels (e.g., "*three*-fifths") to describe a single magnitude. Notably, one of the unique aspects of fractions is that there are infinitely many distinct fractions representing equivalent proportional magnitudes (e.g., 3/4 = 6/8). As such, there are infinitely many verbal labels representing the same proportional magnitude, possibly making it difficult for children to understand that very different labels (e.g., "three-fourths", "six-eighths") refer to the same proportion.

Substantial research suggests that whole number words (e.g., "one", "two", "three") are a key aspect of children's early number knowledge (Baroody & Price, 1983; Condry & Spelke, 2008; LeCorre & Carey, 2007; Odic, LeCorre, & Halberda, 2015). For example, children's understanding of number words may impact how they map symbols to quantities (e.g., Hurst, Anderson, & Cordes, 2017). Furthermore, the structure of number words may be particularly important for how children learn basic number concepts (e.g., LeCorre & Carey, 2007; Fuson & Kwon, 1992; Miller, Smith, Zhu, & Zhang, 1995). For example, children who learn number words in languages that use base-10 number word systems (such as Korean and Chinese where 11 may be referred to as "ten and one" and 23 is referred to as "two tens and three") learn to count faster and do better in number magnitude and arithmetic tasks than English speaking children, likely because of the transparency of the number words (Fuson & Kwon, 1992; Laski & Yu, 2014; Miller & Stigler, 1987).

Although there is substantially less work investigating the impact of fraction verbal labels, there is some evidence that the words we use to refer to proportion may also impact fraction learning (Paik & Mix, 2003; Mix & Paik, 2008). The formal English label system for fractions involves combining a whole-number label for the numerator with a number word variant for the denominator (e.g., 3/5 is labeled "three-fifths" and 4/6 is "four-sixths") and because there are infinitely many fractions with the same magnitude, equivalent fractions (e.g., three-quarters and six-eighths) are referred to using different labels. Because this labeling system emphasizes whole-number information and does not convey any information about the relation between the numerator and the denominator, it has the potential to be confusing for early fraction learners. In support of this, some research suggests that cross-linguistic differences in fraction labels may be related to differences in cross-cultural performance (although, are not the whole story; Paik & Mix, 2003). Paik and Mix (2003) taught US children about symbolic fractions using the phrasing "four of five parts" and "of five parts, four" (based on a translation from Korean fraction labels) and found that this significantly helped children's fraction reasoning over languages that just labeled the components (e.g., "four five") and traditional US labels ("four-fifths"). Importantly, these studies focused on how fraction labels impact learning of fraction symbols, where a transparent naming system may be more relevant for understanding how to read symbolic information.

In the current study, we investigate how the way we label and talk about fractions may impact reasoning about non-symbolic proportional information, in the context of discrete visually presented proportions. In particular, because fraction labels highlight the components that are different (the number of pieces) across equivalent fractions, while obscuring what is the same (the proportional magnitude), fraction labels may impede the learning of certain critical concepts, such as equivalent fractions, even before written symbols are introduced. If so, then turning children's attention toward the elements that are the same across equivalent fractions may help attend to the relevant features when working with equivalent fractions. Research on category learning more generally suggests that using a single noun to refer to multiple examples of a category helps children spontaneously attend to the similarities across the objects and extract the category-relevant features (e.g., Graham, Namy, Gentner, & Meagher, 2010; Plunkett, Hu, & Cohen, 2007). Thus, providing children with a single label for the entire equivalence class may help children form a category for equivalent fractions by focusing on the relational similarity across the distinct proportions (Singer-Freeman & Goswami, 2001).

In the current paper, we investigated how verbal labels impact children's attention to numerical versus proportional information while matching equivalent non-symbolic proportions. In two experiments, we introduced 6.5-year-old children to non-symbolic representations of equivalent proportions (3/4 and 6/8). Importantly, children were randomly assigned to different conditions in which we varied the words used to label the proportional information. In particular, we investigated whether labels that highlight a category of equivalent proportions (Categorical Condition in Experiments 1 and 2) result in less numerical interference than labels that emphasize the numerator (Numerator-Focused Condition in Experiment 1 and Traditional Condition in Experiment 2) or the part-whole nature of the proportion (Part-Whole Condition in Experiment 1 and Simplified Part-Whole Condition in Experiment 2). Although there is evidence that part-whole labels are beneficial for symbolic fraction understanding (Paik & Mix, 2003; Mix & Paik, 2008), part-whole labels also include whole number words that differ across equivalent proportions. Thus, if the use of distinct number words across equivalent proportion is the primary difficulty, then these part-whole labels may be just as likely to highlight numerical information (at the expense of proportional information) as labels that emphasize the numerator specifically.

Experiment 1

To explore the impact of verbal labels on children's proportional reasoning, children were randomly assigned to three different conditions in which we varied the words we used to label the proportional amount. In the Categorical Condition, multiple equivalent fractions were labeled with a single label (e.g., both 3/4 and 6/8 were referred to as "blick"), highlighting the categorical nature of these equivalent fractions. In the Numerator-Focused Condition, labels highlighted the numerator and obscured the denominator using a non-sense word (e.g., 3/4 referred to as "three-blicks" and 6/8 referred to as "six-daxes", akin to English fraction labels). Lastly, in the Part-Whole Condition, labels referred to both numerical components (e.g., 3/4 referred to as "three-out-of-four blicks"). We then explored the impact of these labels on children's abilities to match non-symbolic proportional information exactly and on their ability to compare non-symbolic proportional information. Moreover, we explored how the child's vocabulary may relate to how much the verbal labels used during training impacted their attention to numerical versus relational information.

Method

Participants. Our sample consisted of 125 children (M_{age} = 6.5 years, range: 5.25 years to 7.8 years, N_{female} = 55), separated into three between subject conditions: Numerator-Focused (N=41, M_{age} = 6.59 years, range: 5.41 years to 7.8 years, N_{female} = 18), Part-Whole (N=41, M_{age} = 6.55 years, range: 5.5 years to 7.8 years, N_{female} = 18), and Categorical (N=43, M_{age} = 6.46 years, range: 5.25 years to 7.8 years, N_{female} = 19). An additional 17 children participated, but were not included in any analyses because of interference from parents or other children (N=2) or experimenter error (N=15).

Children participated in our campus laboratory, at local child-care programs, or at local museums (the *Living Laboratory* at the Museum of Science, Boston, MA or the Acton Discovery Museum, Acton, MA) and received a small toy or sticker for participation, in accordance with the regulations set by each testing facility. Parents or legal guardians provided written consent for each child and children over 7-years-old provided written assent.

Measures. Participants completed five phases: (1) vocabulary assessment, (2) label training, (3) learning verification, (4) equivalence matching task, and (5) proportion magnitude comparison task. Children were randomly assigned to one of three conditions: *Categorical Label*, *Numerator-Focused Label*, and *Part-Whole Label*. The conditions only differed in the verbal information provided during the training phase, all other aspects of the tasks were identical. A trained experimenter administered all tasks and recorded the child's responses. The learning verification, equivalence matching, and comparison tasks were all presented on a 13-inch Mac Laptop using Xojo programming software. Children were not given specific feedback about their accuracy but were given general encouragement (e.g., “You’re doing great! Let’s keep going”).

Vocabulary assessment. The vocabulary task consisted of the Woodcock Johnson 3 assessment. Children were shown pictures and asked for the name of the image. The standard protocol for administering the Woodcock-Johnson III Picture Vocabulary Test (Woodcock, McGrew, & Mather, 2001) was followed. The experimenter started the test on the question recommended by the assessment based on the child’s grade and children were required to meet a basal criterion (first six questions administered correct) or to have completed the first question. The experimenter ended testing once children received six questions in a row incorrect. Children were live scored by the experimenter.

Some children did not have a useable vocabulary measurement because they did not receive the vocabulary assessment (the test was added after starting data collection; N=13), experimenter error during administration (N=6), or because they were classified as an outlier (N=1, more than three standard deviations outside the mean), resulting in data from 108 children (M_{age} = 6.57 years, range: 5.25 years to 7.8 years, N_{female} = 48) in analyses involving vocabulary (Numerator-Focus: N=36, M_{age} = 6.67 years, range: 5.66 years to 7.8 years, N_{female} = 16; Part-Whole: N=35, M_{age} = 6.6 years, range: 5.5 years to 7.8 years, N_{female} = 16; Categorical: N=37, M_{age} = 6.44 years, range: 5.25 years to 7.8 years, N_{female} = 16).

Training phase. In all conditions children were introduced to a toy animal character[5] (e.g., a plastic kangaroo named Roo) who "likes shapes that have just the right amount of color" (see the Supplementary Appendix 4 for the full script). The game was then introduced as a way to find out what amount of color the character liked. First, children were shown an empty grey circle (8 cm in diameter) divided into quarters (from the center, creating angular slices) and the experimenter colored in three adjacent pieces using a yellow marker saying: "I'm going to color [condition specific label]". After coloring the pieces, the experimenter said: "See, this is called [condition specific label], Roo likes this one" and then placed the shape on a tray in front of the toy character. The experimenter then brought out a rectangle (13.4 cm long by 6.3 cm tall, with vertical dividing lines along the shorter dimension into four equal pieces) that had three of four consecutive pieces already colored yellow and said "This is also [condition specific label], Roo likes this one too" and placed it on the tray. The experimenter then brought out two other pre-colored shapes (one circle and one rectangle, same dimensions as previously), with each one saying "But Roo won't like this one" and placing the shape off to the side away from the other shapes, but still visible. This was then repeated with new shapes to demonstrate that Roo also likes shapes with 6/8 colored yellow. At the end, there were four shapes (3/4 circle, 3/4 rectangle, 6/8 circle, and 6/8 rectangle) placed on the tray with the character, as exemplars of what Roo would like, and four counter-example shapes (3/8 circle (counter-example Roo did not like with same number of yellow pieces), 4/4 rectangle (counter-example Roo did not like with same total number of pieces), 6/12 circle (counter-example Roo did not like with same number of yellow

[5] The specific character varied, but for the purposes of clarity, we will refer to the character as "Roo" throughout.

pieces), and 2/8 rectangle (counter-example Roo did not like with same total number of pieces)) off to the side. The experimenter then drew the child's attention to the shapes on the tray with Roo and said, "See, these are all the ones that Roo likes". See Figure 4.1a for an example of the final set up.

In the *Categorical Label* condition, all to-be-learned proportions were paired with a single nonsense word ("This is called **blick**" for 3/4 and for 6/8). In the *Numerator-Focused* condition, the numerator was highlighted using a number word and the denominator was obscured using a non-sense word ("This is called **three-blicks**" for 3/4 and "This is called **six-daxes**" for 6/8). In the *Part-Whole* condition, the denominator was first highlighted, then the numerator was highlighted using "out of" terminology associated with parts and wholes ("Here are **four blicks** [highlighting the whole shape]. I'm going to color **three blicks**. See, this is called **three-out-of-four blicks**" for 3/4, and similarly for 6/8, except the pieces were labeled as "daxes" instead of "blicks"). Within each condition, the use of blick(s) and dax(es) with specific fractions was counter-balanced across children to avoid specific pairings between words and magnitudes.

Learning verification. The purpose of the learning verification task was to ensure that children had learned the specific proportions taught during the training phase. The learning verification task was performed on a laptop computer. Children were presented with two stimuli on each trial, one on the right and one on the left of the screen. On each trial, children were asked which of the shapes Roo would like. The stimuli remained on the screen until the child selected a response by pointing to one of the stimuli and the experimenter recorded the child's response by pushing the corresponding key on the keyboard. The next trial began as soon as the experimenter recorded the child's response.

Children received a total of six trials. The first trial included a direct comparison learned during training (3/4 vs. 3/8; circles 9.8 cm in diameter). The second trial included a similar comparison learned during training (6/8 vs. 4/8; rectangle with horizontal lines, 13.4 cm long by 8.8 cm tall). The remaining four trials involved a novel shape (square divided with vertical lines, 10 cm wide) and magnitudes that were learned proportions (3/4 and 6/8) as well as equivalent, non-learned proportions (12/16 and 9/12).

Equivalence task. The equivalence task was a match-to-sample task on the computer that assessed whether children could match non-symbolic visual displays based upon proportional information. On every trial, children were shown a sample stimulus (circle, 7.7 cm in diameter) presented in the upper half of the screen for 1000ms before two options (rectangles with horizontal lines, 13.4 cm long by 7.7 cm tall) were displayed on the bottom half of the screen. A thick black line divided the upper and lower halves of the screen to separate the sample from the options (see Figure 4.1b for example stimuli). Children were then asked to pick which of the bottom pictures "best matched" the top (sample) picture. All stimuli remained on the screen until the child gave a response and the experimenter recorded the child's response using the keyboard. The next trial began as soon as the experimenter recorded the child's response.

Children participated in 13 trials (presented in a random order) that varied in the make-up of the options (see the Supplementary Appendix 4 for a full set of magnitudes used). Five trials did not involve competing numerical information such that the correct answer was either an exact match (and thus, matched on number *and* proportion; 3 trials) or the correct answer was an equivalent proportion (matched on proportion, but not number) and the incorrect answer did not match the sample stimulus on either number or proportion (2 trials). The remaining 8 trials involved competing numerical information, in which the proportional response matched on proportion but not number (i.e., was an equivalent proportion) and the other option had the same denominator (i.e., was broken up into the same number of pieces; 4 trials) or had the same numerator (i.e., had the same number of yellow pieces; 4 trials) as the sample stimulus. Thus, both answers only matched on one feature (number OR proportion) and thus these two features were in clear competition. On half of these competing numerical trials, the proportional response had a greater number of pieces (e.g., target = 2/8, correct answer = 4/16) and on the other half the proportional response had fewer pieces (e.g., target = 2/6, correct answer = 1/3).

Figure 4.1: Example Stimuli

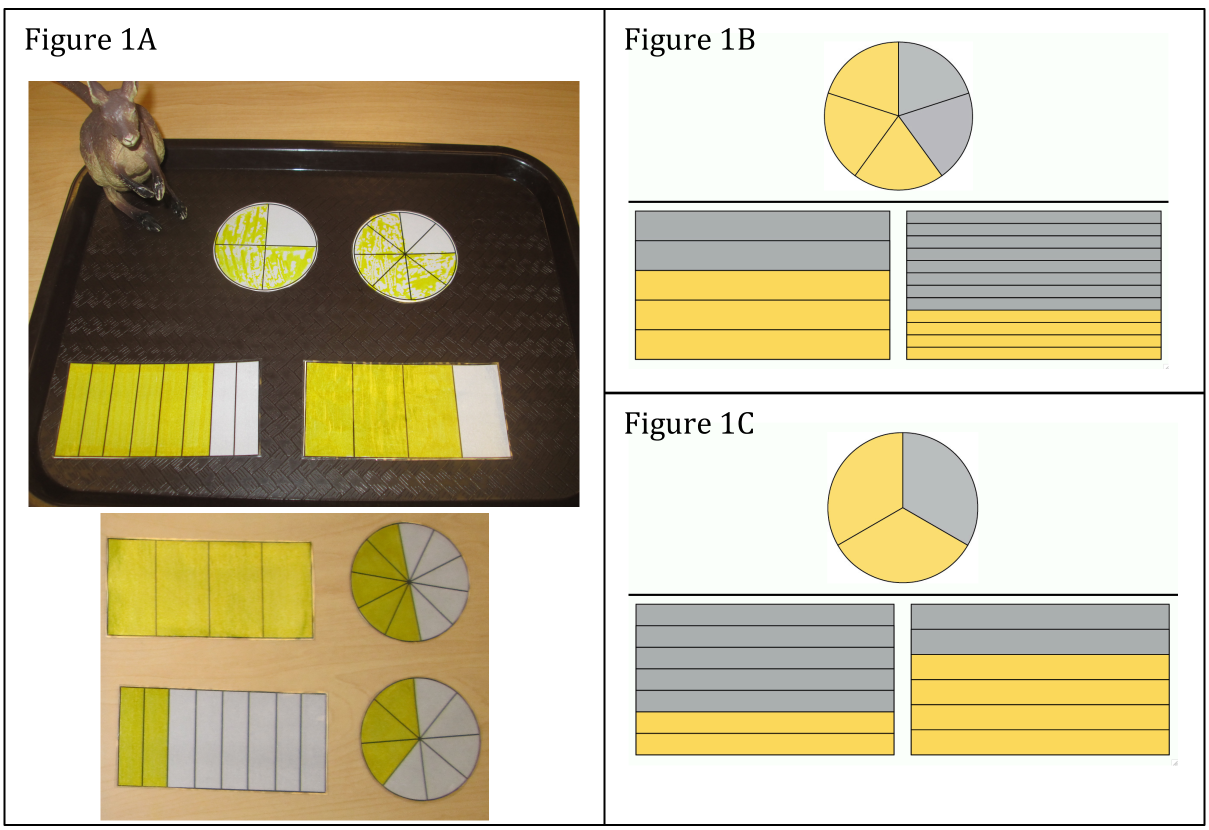


Figure 4.1: Stimuli used during the label training (Figure 1a, left) and the equivalence matching task (right). Figure 1a shows the shapes that the character (i.e., Roo) does like (upper figure, on the tray) and the shapes the character would not like (lower figure). On the right, Figure 1b shows an example of a trial without competing numerical information and Figure 1c shows a trial with competing numerical information.

Comparison task *(adapted from Hurst & Cordes, under review - A)* The comparison task was also performed on a laptop computer and it assessed whether children were able to judge the relative magnitude of two proportional displays. Children were first introduced to a cardboard spinner (not on the computer, 8cm in diameter with a small black arrow, approximately 3cm long), such as those used in children's board games and were told that the yellow pieces were "winning pieces" and the grey pieces were "losing pieces". The experimenter then spun the arrow twice and asked the child what the outcome meant.

Following this familiarization, children were then presented with two pictures of spinners on the computer screen on each trial (on the right and left of the computer screen). On each trial, children were asked which of the two spinners (on the computer) was "better". The stimuli remained visible until the child selected an answer and the experimenter recorded the child's response using the corresponding keys on the keyboard. All spinners were circles divided into equal pieces around the center point so that numerical information was available, but the circles were presented in three different sizes in order to prevent direct perceptual comparisons (small circles 6 cm in diameter; medium circles 8.8 cm in diameter; large circles 11.5 cm in diameter).

Children received 8 trials (in a random order). On half the trials, number was consistent with proportion, such that the spinner with the greater proportion of yellow also had the greater number of yellow pieces ("numerator"; the most salient, "winning" pieces). On these trials, choosing the spinner with the higher number of yellow or the highest proportion of yellow would result in the same answer (e.g., 2/5 vs. 5/9; "consistent trials"). On the other half of trials, number was inconsistent with proportion, such that the spinner with the greater proportion of yellow had a lower absolute number of yellow pieces (e.g., ¾ vs. 5/11). Thus, choosing the spinner with the greater number of yellow pieces would result in an incorrect response (e.g., 4/9 vs. 2/3; “misleading trials”). The numerically consistent comparisons presented were: 2/6 vs. 5/8; 5/7 vs. 8/9; 4/9 vs. 1/5; 3/6 vs. 5/8. The numerically misleading comparisons presented were: 2/3 vs. 3/9; 1/3 vs. 2/9; 5/10 vs. 4/5; 3/5 vs. 4/9.

Results

Initial analyses involving our control variables revealed no differences across conditions on key variables, ensuring random assignment across the three conditions. Across conditions, children did not significantly differ in terms of their age, $F(2,122)= 0.44$, $p=0.64$, *partial* $\eta^2 < 0.01$, or vocabulary, $M_{\text{Categorical}} = 19.6$, $M_{\text{Numerator}} = 19.5$, $M_{\text{Part-Whole}} = 19.8$, $F(2,105)= 0.08$, $p=0.9$, *partial* $\eta^2 < 0.01$. Moreover, performance on the learning verification task was comparable across groups, $M_{\text{Categorical}} = 0.79$, $M_{\text{Numerator}} = 0.79$, $M_{\text{Part-Whole}} = 0.74$; $F(2,122)= 0.63$, $p=0.5$, *partial* $\eta^2 = 0.01$, revealing comparable levels of learning across the three conditions.

Equivalence Task. A mixed-measures ANCOVA with Numerical Interference (2: present or absent) as a repeated measure, Condition (3: Categorical, Numerator-Focused, Part-Whole) as a between subject variable, and age as a covariate (however, when age is not included as a covariate the pattern of results is identical) was conducted on the proportion of trials children selected the proportional match in the equivalence task. Critically, there was a significant Condition X Numerical Interference interaction, $F(2,121) = 6.5$, $p=0.002$, *partial* $\eta^2 = 0.097$ (see Figure 4.2). Thus, we looked at performance across the three conditions on trials with and without numerical interference separately.

Preference for the proportional match on trials that did not involve competing numerical information did not vary across conditions when controlling for age as a covariate, $M_{\text{Categorical}}$= 0.67, $M_{\text{Numerator}}$ =0.65, $M_{\text{Part-Whole}}$ = 0.71; $F(2,121)$= 1.4, p=0.25, *partial* η^2 = 0.02. In contrast, performance on numerical interference trials (i.e., in which one response matched on proportion only and the other response matched on the numerator or denominator of the sample) did vary as a function of condition (again controlling for age as a covariate), $F(2,121)$= 5.4, p=0.006, *partial* η^2 = 0.08. In particular, those in the Categorical Label condition, M = 0.60, picked the proportional match more than children in the Numerator-Focused, M = 0.51; $F(1,81)$= 4.12, p=0.046, *partial* η^2 =0.05, and the Part-Whole, M = 0.45; $F(1,81)$= 10.7, p=0.002, *partial* η^2 = 0.12, conditions. However, there was no significant difference between performance in the Part-Whole and Numerator-Focused conditions, $F(1,79)$= 1.83, p=0.18, *partial* η^2 = 0.02.

Furthermore, children in the Categorical Label condition selected the proportional match more often than chance levels (.50) on both interference, p = 0.001, and non-interference, p<0.001, trials, which were not significantly different from each other, $t(42)$ = 1.5, p=0.13. However, consistent with past studies, there were significant interference effects, showing a significantly lower preference for the proportional match on trials with competing numerical information than trials without competing numerical information for children in the Numerator-Focused, $t(40)$ = 3.3, p = 0.002, and the Part-Whole, $t(40)$ = 6.3, p < 0.001, conditions. In particular, children in both conditions selected the proportional match significantly more often than chance on the trials without competing numerical information ($ps < 0.001$), but not on trials with competing numerical information ($ps > 0.1$), suggesting they were sometimes matching on number and sometimes matching on proportion.

Figure 4.2: Performance on Equivalence-Matching Task for Experiment 1

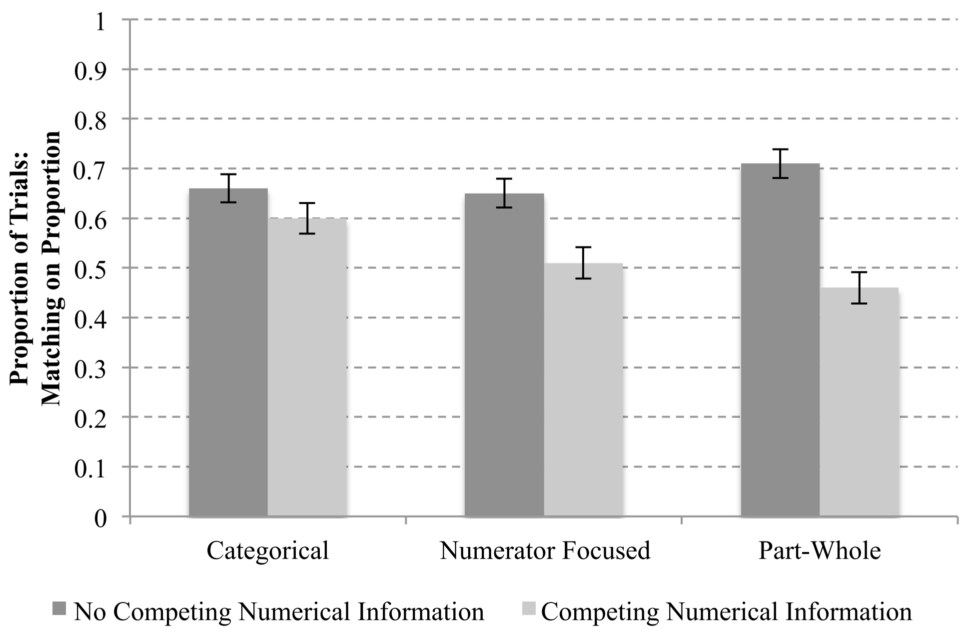


Figure 2: Adjusted means for children's preference for the proportional match on the equivalence-matching task in Experiment 1 across condition and trial type. Chance responding would be at 0.5.

Lastly, there were not significant correlations between performance on the equivalence task trials with and without competing numerical information in any of the three conditions, controlling for age: Categorical condition partial $r(40) = 0.19$, $p = 0.2$, Numerator-Focused partial $r(38) = 0.1$, $p = 0.5$, and Part-Whole partial $r(38) = 0.15$, $p = 0.4$.

Comparison task. Next, we explored whether the differential label training extended to a proportion comparison task (see Figure 4.3). First, we looked at overall performance on the task by conducting an ANCOVA on proportion correct with

Condition (3: Categorical, Numerator-Focused, and Part-Whole) as a between-subjective variable and age as a covariate. Overall, children performed 73% correct on average, regardless of trial type and this did not significantly differ across condition: $M_{\text{Categorical}}$ = 0.70, $M_{\text{Numerator}}$ =0.77, $M_{\text{Part-Whole}}$ = 0.73; $F(2,121)= 0.95$, p=0.39, *partial* $\eta^2 = 0.02$.

Next, and most importantly, given that numerical information was available in both trial types, we calculated an interference score to measure the *strength* of the interference (Interference Score = proportion correct on consistent trials – proportion correct on misleading trials), such that a high interference score would correspond with consistent use of numerical information (performing well when number is consistent but poorly when number is misleading) and a score around zero would correspond to no reliance upon numerical information (performing similarly across trial types, regardless of whether the numerical information available was consistent or misleading). We conducted an ANCOVA with condition (3: Categorical, Numerator-Focused, and Part-Whole) as a between subjects measure and age as a covariate on interference scores. There was not a significant main effect of condition, $M_{\text{Categorical}}$ = 0.09, $M_{\text{Numerator}}$ =0.17, $M_{\text{Part-Whole}}$ = 0.14; $F(2,121)= 0.46$, p=0.63, *partial* $\eta^2 < 0.01$. However, looking at the average interference effect, ignoring condition, there was a significant interference effect (interference score > 0), $M = 0.13$; $t(124) = 0.13$, $p < 0.001$, Cohen's d = 0.36.

Figure 4.3: Proportion Correct on Comparison Task

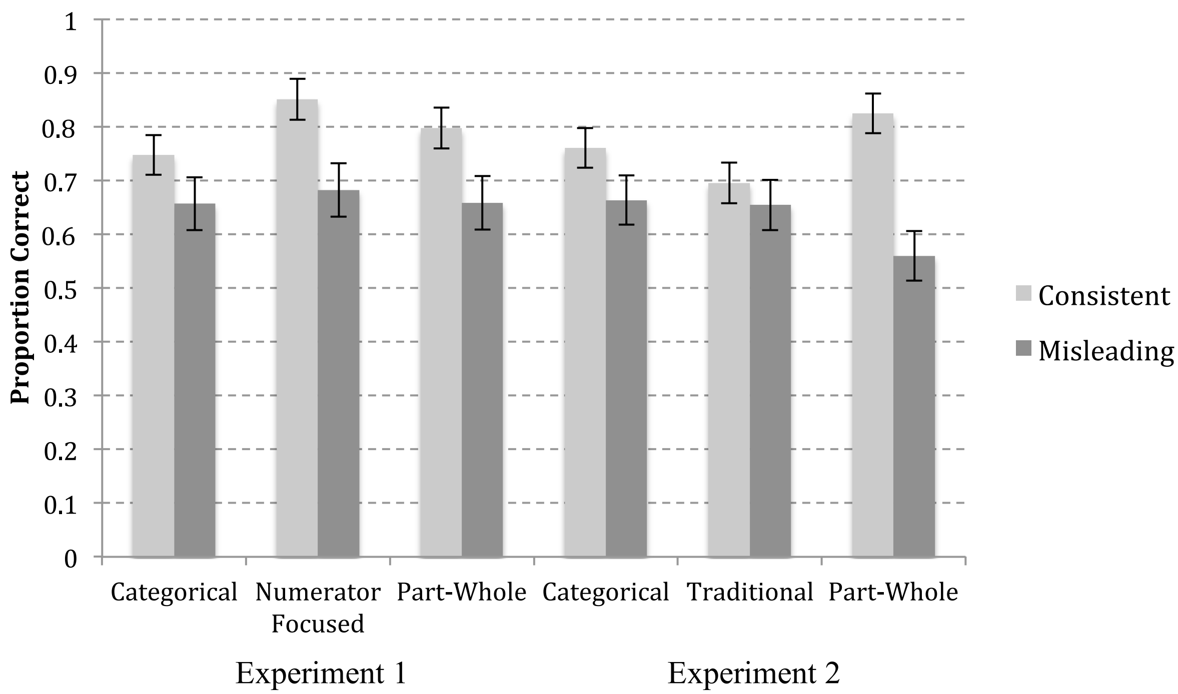


Figure 4.3: Adjusted means for children's performance on the spinner comparison task, controlling for age, broken down by condition for both Experiment 1 (left bars) and Experiment 2 (right bars). Performance is displayed for each trial type separately (consistent and misleading). Interference scores are the difference between these two trial types.

Relations among tasks. In order to get a better sense of children's proportional reasoning, we compared performance across the distinct measures of proportional reasoning and numerical interference. There was not a significant relation between interference on the comparison task and performance on the equivalence task trials without competing numerical information, partial $r(122) = -0.073$, $p = 0.4$, controlling for age, but there was a marginally significant relation with equivalence task trials with competing numerical information, partial $r(122) = -0.172$, $p = 0.057$, controlling for age.

Discussion

In Experiment 1, 5- to 7-year-old children learned about a set of equivalent proportions using one of three distinct labels: a categorical label, a numerator-focused label, or a part-whole label. Overall, we find that the labels used to reference proportional amounts impacted the relative saliency of different kinds of information children attended to in proportional reasoning tasks. In particular, children who received the categorical label were more likely to select the proportional match, and less likely to select a competing numerical match, on an equivalence matching task than children in either of the two other groups. Children who heard numerator-focused or part-whole labels, on the other hand, generally responded at chance, sometimes selecting the number match and sometimes selecting the proportion match. Thus, our data indicate that the format of fractions labels impacts early proportional reasoning, even in visual, non-symbolic contexts.

The impact of label training did not extend to the comparison task, however. There are several, non-mutually exclusive explanations for why this was the case. First, it may simply be that since the comparison task occurred after the equivalence task, our training manipulation may have been too short lasting to extend to this last task and by the time children reached the comparison task, they resorted back to their default manner of processing proportional information. Alternatively, the distinct nature of the tasks may have led to the different pattern observed. That is, our training was specific to understanding equivalent proportions and as such, it may be that children were less inclined to spontaneously extend what they learned during the exact matching training to the concept of comparison. In particular, it may not be that children generally learned to *always* attend more to proportion than to number, but instead grasped some more nuanced rule about attending to proportion when making equivalence judgments.

In Experiment 2, we aimed to replicate and extend these findings by exploring how performance in our conditions differed from that of how children typically respond when hearing traditional English labels for fractions. Although Experiment 1 included a numerator-focused condition that, in some ways, was conceptually similar to traditional labels, the non-sense word used for the denominator (e.g., blicks or daxes) was much more obscure than the denominator words used in traditional labels (e.g., fourths or eighths). Thus, in Experiment 2, we included a condition with traditional, formal labels to investigate how categorical labels compare to traditional fraction labels. In addition, the particularly number-focused performance by children in the part-whole condition may be surprising, given the other evidence that this phrasing is beneficial in other contexts (Paik & Mix, 2003). However, it may be that the difficulty shown by children in the part-whole condition was due to the additional phrase providing them with the denominator information, overwhelming them with verbal information. Thus, in Experiment 2 we modified the wording in our part-whole condition to be more simplified, resulting in a closer match of linguistic complexity across conditions.

Experiment 2

Experiment 2 had three primary purposes. First, we wanted to compare how performance in the key condition in Experiment 1 (the Categorical Label condition) differed from performance of children hearing formal labels actually used in American English. Thus, we replicated the *Categorical Label* condition and included a condition that used traditional fraction labels (e.g., “three-fourths”; *Traditional Label*). Second, we wanted to include a simplified Part-Whole label that did not have the confound of including an extra phrase during training (that did not have the phrase "Here are four blicks" [highlighting the whole shape]). Thus, we replicated the Part-Whole Label condition from Experiment 1, but simplified it to just include the label phrase (more parallel to the other conditions) and not the extra phrase that specifically highlighted the denominator (*Simplified Part-Whole* condition). Third, we included a measure of math knowledge outside of the domain of proportion in order to investigate how math ability may be related to numerical interference in proportional judgments.

Method

Participants. Our sample consisted of 127 children (M_{age} = 6.4 years, range: 5.25 years to 7.6 years, N_{female} = 70), separated into three between subject conditions: Categorical Label (N=43, M_{age} = 6.51 years, range: 5.42 years to 7.6 years, N_{female} = 24), Traditional Label (N=41, M_{age} = 6.49 years, range: 5.5 years to 7.6 years, N_{female} = 22), and Simplified Part-Whole Label (N=43, M_{age} = 6.28 years, range: 5.25 years to 7.5 years, N_{female} = 24). An additional 17 children participated, but were not included in any analyses because of interference from parents or other children (N=1), an inability to complete the tasks or follow the instructions (N=2), or experimenter (N=6) or computer (N=1) error. Children were recruited and consented as in Experiment 1.

Some children did not have a useable vocabulary measurement because of experimenter error resulting in them not receiving the test (N=1) or an error during administration (N=12), resulting in data from 114 children (M_{age} = 6.4 years, range: 5.25 years to 7.6 years, N_{female} = 61) in analyses involving vocabulary (Categorical: N=37, M_{age} = 6.44 years, range: 5.4 years to 7.6 years, N_{female} = 20; Traditional: N=37, M_{age} =

6.50 years, range: 5.5 years to 7.6 years, $N_{female} = 20$; Simplified Part-Whole: N=40, M_{age} = 6.3 years, range: 5.25 years to 7.5 years, $N_{female} = 21$).

Design. Children were randomly assigned to one of three conditions: *Categorical Label* (same as Experiment 1), *Simplified Part-Whole Label* (similar to *Part-Whole* condition in Experiment 1), and *Traditional Label* (new condition). As in Experiment 1, the conditions only differed in the verbal information provided during the training phase, all other aspects of the tasks were identical. Children completed the same tasks used in Experiment 1 as well as an additional math knowledge test (completed last).

Measures. Visual stimuli of all tasks were identical to Experiment 1. The procedures for all tasks, except the training conditions and number knowledge test (new task), were also identical to Experiment 1.

Training Phase. In the *Categorical Label* condition, the to-be-learned proportion was paired with a single nonsense word ("I'm going to color blick. See, this is called **blick**" for 3/4 and for 6/8), as in Experiment 1. In the *Simplified Part-Whole Label* condition, we used the same label as used in the *Part-Whole* condition from Experiment 1, but did not explicitly highlight the denominator in order to simplify the verbal information being provided ("I'm going to color three-out-of-four blicks. See, this is called **three-out-of-four blicks**" for 3/4). In the *Traditional Label* condition, we used formal fraction labels taught in school ("I'm going to color three-fourths. See, this is called three-fourths" for 3/4). For the full scripts, see the Supplementary Appendix 4.

Math knowledge test. We adapted questions from the Woodcock-Johnson III Applied Problems Test (Woodcock, McGrew, & Mather, 2001). Children received 10 questions that involved identifying non-symbolic quantities (3 questions), mental arithmetic based on non-symbolic quantities (5 questions) and word problems, read aloud by the experimenter (2 questions). Children completed all problems and were live scored by an experimenter.

Results

Overall, there was no differences in age, $F(2,124)=1.6$, $p=0.2$, *partial* $\eta^2 = 0.03$, vocabulary, $M_{\text{Categorical}} = 19.7$, $M_{\text{Traditional}} = 18.9$, $M_{\textit{out-of}} = 19.7$, $F(2,111)= 0.64$, $p=0.53$, *partial* $\eta^2 = 0.01$, or on performance on the learning verification task, $M_{\text{Categorical}} = 0.81$, $M_{\text{Traditional}} = 0.69$, $M_{\textit{S. Part-Whole}} = 0.74$; $F(2,124)= 2.36$, $p=0.1$, *partial* $\eta^2 = 0.01$, across conditions.

Equivalence matching task. As in Experiment 1, we used an ANCOVA with Numerical Interference (2: present or absent) as a repeated measure, Condition (3: Categorical, Traditional, Simplified Part-Whole) as a between subject variable, and age as a covariate (again, removing age as a covariate did not impact the pattern of results), on performance on the equivalence task (proportion of trials selecting a proportional response). Critically, there was again a significant Condition X Numerical Interference interaction, $F(2,123)= 3.28$, $p=0.04$, *partial* $\eta^2 = 0.05$ (see Figure 4.4). Thus, we looked at differences across the three conditions on trials with and without competing numerical information separately.

As in Experiment 1, there was no main effect of condition on trials that did not have competing numerical information, controlling for age as a covariate, $F(2,123)= 0.27$, $p=0.8$, *partial* $\eta^2 < 0.01$. Moreover, children in all three groups selected the proportional match more often than predicted by chance (higher than 0.5 or 50%): $M_{\text{Categorical}}= 0.68$, $M_{\text{Traditional}} = 0.70$, $M_{\textit{S. Part-Whole}} = 0.67$; all ps< 0.001).

In contrast, analyses of data from trials with competing numerical information (i.e., in which one response matched on proportion only and the other response matched on the numerator or denominator of the sample) revealed a significant main effect of condition (again controlling for age as a covariate), $F(2,123)= 3.58$, $p=0.03$, *partial* $\eta^2 = 0.06$. In particular, those in the Categorical Label condition, $M = 0.61$, selected the proportional match significantly more often than children in the Traditional condition, $M = 0.48$; $F(1,81)= 6.1$, $p=0.02$, *partial* $\eta^2 =0.07$, and marginally more often than children in the Simplified Part-Whole condition, $M = 0.50$; $F(1,83)= 3.9$, $p=0.051$, *partial* $\eta^2 = 0.05$. However, there was not a significant difference between children's performance in the Simplified Part-Whole and Traditional label conditions, $F(1,81) = 0.13$, $p=0.72$, *partial* $\eta^2 < 0.01$. Mirroring this pattern of findings, only children in the Categorical Label condition selected the proportional match more often than chance on these trials, $p = 0.004$, all other ps > 0.5.

Figure 4.4: Performance on Equivalence-Matching Task for Experiment 2

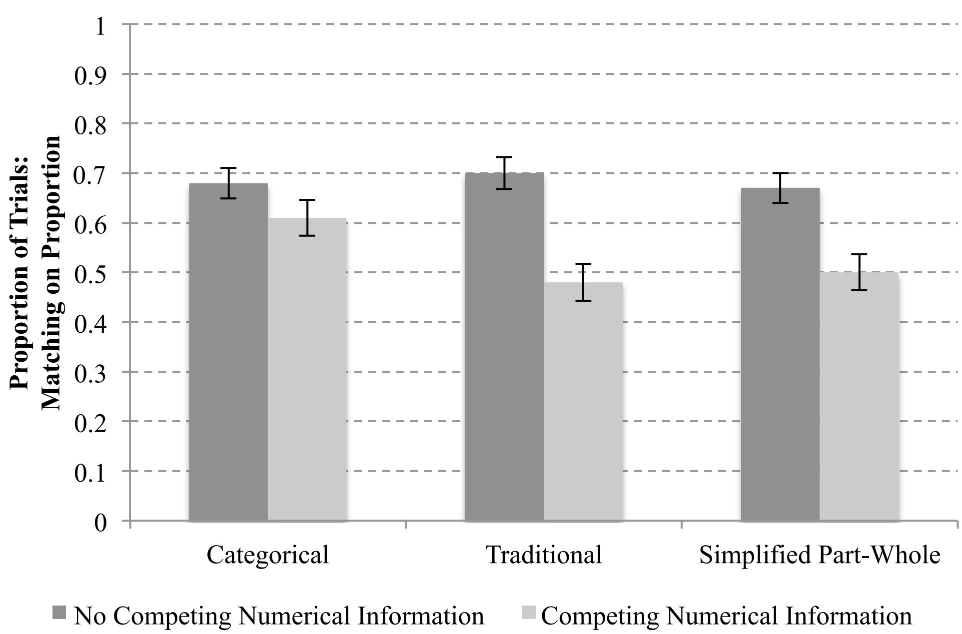


Figure 4.4: Adjusted means for children's preference for the proportional match on the equivalence-matching task in Experiment 2 across condition and trial type. Chance responding would be at 0.5.

Lastly, consistent with the idea that labels containing whole number words (Traditional and Part-whole conditions) detracted attention from proportional information, a significant positive correlation between preference for the proportional match on the equivalence task trials with and without competing numerical information was found only for children in the Categorical condition, controlling for age: Categorical condition partial $r(40) = 0.43$, $p = 0.004$, but not in the other two conditions: Traditional partial $r(38) = 0.18$, $p = 0.3$, and Part-Whole partial $r(40) = 0.09$, $p = 0.6$.

Comparison task. Next, we explored performance on the comparison task (see Figure 4.3). We conducted an ANCOVA on proportion correct with Condition (3: Categorical, Traditional, and Simplified Part-Whole) as a between-subjective variable and age as a covariate. Overall, children, performed 69% correct on average, regardless of trial type and this did not significantly differ across condition, $M_{\text{Categorical}} = 0.71$, $M_{\text{Traditional}}=0.68$, $M_{\text{S. Part-Whole}} = 0.69$; $F(2,127)= 0.42$, $p=0.7$, *partial* $\eta^2 < 0.01$.

However, because our hypothesis was that labels would influence attention to numerical information specifically, as in Experiment 1, we calculated an interference score to measure the *strength* of the interference (Interference Effect = consistent – misleading) and used an ANCOVA with condition (3: Categorical, Traditional, and Simplified Part-Whole) as a between subjects measure and age as a covariate. Here, unlike Experiment 1, there was a main effect of condition, $F(2,123)= 3.48$, $p=0.03$, *partial* $\eta^2 = 0.05$.

In particular, there was not a significant interference effect (i.e., different from 0) for children in the Categorical, $M = 0.1$; $t(42)=1.6$, $p = 0.1$, or the Traditional, $M = 0.04$; $t(40)=0.69$, $p = 0.5$, label conditions, which were not significantly different from each other ($p = 0.5$). However, there was a significant interference effect in the Simplified Part-Whole condition, $M = 0.26$; $t(42)=4.3$, $p < 0.001$, which was significantly higher than the interference effect in the Traditional condition, $p=0.01$, and marginally higher than the interference effect in the Categorical condition, $p=0.07$.

Relations among tasks. As in Experiment 1, we compared performance across the distinct measures of proportional reasoning and numerical interference. In addition, we looked at the relations between numerical interference and math knowledge in order to look at individual differences in children's attention to numerical information.

First, we looked at the relations between performance on the equivalence task and numerical interference on the comparison task. As in Experiment 1, there was not a significant relation between interference on the comparison task and preference for the proportional match on the equivalence task when proportion and number did not compete (partial $r(124) = -0.097$, $p = 0.3$, controlling for age). However, as expected, there was a significant negative relation between interference on the comparison task and preference for the proportional match in the equivalence task trials that did involve competing numerical information (partial $r(124) = -0.28$, $p = 0.001$, controlling for age), such that higher numerical interference in the comparison task was associated with a greater preference to match based on whole-number information in the equivalence task.

Next, in Experiment 2 we included a measure of math knowledge to investigate how math knowledge may predict children's numerical interference in each of the three conditions. We used linear regressions with age, math, condition (dummy coded), and math by condition (dummy coded) interactions in order to isolate the effects of each condition, and the variation across conditions (note: reported βs are standardized coefficients).

Math score did not significantly predict performance on the equivalence task trials with numerical information in the Traditional, $\beta = -0.12$, $t(105) = -0.7$, $p = 0.5$, or Simplified Part-Whole, $\beta = 0.09$, $t(105) = 0.4$, $p = 0.7$, conditions. However, math score was significantly related to performance in the Categorical, $\beta = 0.37$, $t(105) = 2.4$, $p = 0.02$, condition. In particular, children in the Categorical label condition with higher math knowledge were more likely to choose the number match. Furthermore, the relation in the Categorical condition was significantly different from the relation in the Traditional condition ($p = 0.03$), but not the Simplified Part-Whole condition ($p = 0.3$) and the Traditional and Simplified Part-Whole conditions were not different from each other ($p = 0.39$).

However, math score did not significantly predict numerical interference on the comparison task in the Categorical ($\beta = -0.12$, $t(105) = -0.7$, $p = 0.5$), Traditional ($\beta = 0.02$, $t(105) = 0.1$, $p = 0.9$), or Simplified Part-Whole ($\beta = -0.1$, $t(105) = -0.4$, $p = 0.7$) conditions, and these relations did not differ from each other (all dummy coded interaction effects ps > 0.5).

Discussion

Overall, the main results of Experiment 2 replicate those of Experiment 1. In particular, the labels children heard influenced performance on our proportional reasoning tasks. Children who heard a single categorical label used to refer to multiple examples of equivalent fractions were significantly more likely to match on proportion, rather than number, on a subsequent equivalence matching task than children who heard traditional fraction labels or children who heard the modified part-whole labels (although, this simple effect was only marginal). Furthermore, unlike Experiment 1, we also found our label training to have an impact on performance on the comparison task, such that children in the simplified part-whole condition showed a significant interference effect (unlike children in the categorical and traditional conditions), which was significantly higher than children in the traditional and categorical conditions (although, the differences between categorical and simplified part-whole conditions was marginal). This may suggest some evidence that labels that emphasize equivalent proportions may also impact reasoning on a comparison task, however, the strength and pattern of this effect remains unclear and are further discussed in the General Discussion.

Lastly, we did see a relation between math ability and attention to proportion (over numerical features) on the equivalence-matching task, but only for children in the categorical label condition. Since our categorical label condition did not involve any number words, it may be that individual differences in children's baseline attention to number (potentially related to their math knowledge) were more predictive of individual differences in numerical interference for children in this condition. However, in the conditions that involved number words, these individual differences may have been washed out by the fact that all children were prompted to attend to number during the training phase. However, these relations were only found for number matching during the equivalence match-to-sample task and not numerical interference during the comparison task. Given the differences across these two tasks throughout both experiments, reasons for this difference are further discussed in the General Discussion.

General Discussion

Across two experiments, 5- to 7-year-old children learned about a set of equivalent proportions using one of four distinct labels: a categorical label (Experiments 1 and 2), a numerator-focused label (Experiment 1), a traditional label (Experiment 2), and a part-whole label (Experiments 1 and 2). In both experiments, children who received the categorical label were more likely to select the proportional match, and less likely to select a competing numerical match, on an equivalence matching task than children who received labels that focused on the numerator (numerator-focused condition of Experiment 1 and traditional label condition of Experiment 2) or that focused on the numerator and denominator (complete and simplified part-whole conditions of Experiments 1 and 2).

There are at least two non-mutually exclusive explanations of this pattern of findings. One possibility is that labeling multiple fractions within the same equivalence class (e.g., 3/4 and 6/8 in this case) with the same categorical label highlighted attention to commonalities across the shapes (i.e., common proportion), allowing children to ignore those components of the shapes that may have differed (i.e., the number of pieces). This is in line with other work suggesting that common labels can promote categorization in infancy and childhood by highlighting similarities across differing exemplars of the same category (e.g., Fulkerson & Waxman, 2007; Graham, Namy, Gentner, and Meagher, 2010). On the other hand, it may be that the inclusion of whole-number words as part of the label (e.g., “three” in “three blicks”), either in reference to the numerator and/or the denominator of the proportion, promoted attention to the *number* of pieces in each visual exemplar, at the expense of attending to proportion. As such, attending to numerical information, which varied across proportions within the same equivalence class (i.e., “three” in ¾ and “six” in 6/8), may have hindered attention to the commonality across these fractions – a common proportional magnitude. Given that we did not have a strict control group (i.e., a group that did not receive any labels) we do not have a baseline measure of how children would typically perform in this kind of task with these specific magnitudes. As such, we cannot distinguish between these two potential explanations with any certainty. However, given other work suggesting that children around this age tend to be more inclined to match on number than proportion, even without hearing any labels, (e.g., Boyer et al., 2008; Hurst & Cordes, under review- A; Jeong et al., 2007), our pattern of findings that children in the categorical label condition performed more similarly across trials with and without competing numerical information than children in any of the other conditions (who performed significantly worse when number and proportion were explicitly pit against each other) suggests the latter may be more likely. That is, the categorical label may have led children to treat the trials with competing numerical information the same way they approached the trials without competing numerical information. However, it is still an open question as to whether the number words used in the traditional, numerator-focused, and part-whole labels increased children’s attention to numerical components and/or whether the single category-based label used in the categorical label condition increased children’s attention to the proportional relation. Critically, our results do indicate that the labels used to talk about proportion can have a significant impact on the type of information that children attend to when matching equivalent proportions.

Importantly, we did not find that these labels solely impacted children's learning of the specific equivalence class of three-quarters. First, there was not a significant effect of condition on children's ability to learn which proportional amounts the character liked, as assessed by the learning verification task. Second, our matching task assessed matching across a variety of fraction equivalents, not just those equivalent to a three-quarter magnitude, revealing that our training generally impacted how children attended to proportional information in the face of conflicting numerical information. Furthermore, children across conditions did not perform significantly differently on proportion matching trials that did not involve competing numerical information, suggesting that the different labels did not impact children's *ability* to match on proportion more generally, but rather their strategy or attention in particular contexts. Thus, it is not that children were better or worse able to learn a particular set of equivalent fractions or to make proportional matches, but rather when we asked children to match proportions beyond those used during training and contrasted these matches with competing numerical information, children who received the categorical label were more likely to match based on equivalent proportion, rather than number, while those who received number based labels were at chance responding (sometimes choosing number, sometimes choosing proportion).

In contrast to other work suggesting that using part-whole labels may be beneficial for children's thinking about proportion (Paik & Mix, 2003), in the current study we did not find that the part-whole labels led to any significant differences from numerator-focused or traditional labels (Experiments 1 and 2, respectively). However, there may be several reasons for these seemingly contradictory findings. Notably, Paik and Mix (2003) worked with slightly older children (mean age of about 7-years-old in their youngest, first grade sample) and focused on children's understanding of symbolic fractions, rather than their proportional reasoning in the absence of symbols (as was the central topic of the current study). Thus, it may be that the part-whole label structure is more apt for allowing children to understand the structure of fraction symbols, but does not make the relationship between equivalent proportions any more transparent. This is reasonable to believe since the part-whole structure of Paik & Mix's "out of" label directly corresponds to the structure of the symbolic fraction. However, it does not make any reference to the proportional similarity across equivalent fractions and still uses distinct labels for equivalent proportions. Another potential (not-mutually exclusive) explanation is in the actual phrasing we used, compared to Paik and Mix (2003); there were seemingly minor, but potentially important, differences between our part-whole condition and the phrasing used by Paik and Mix (2003) that may provide some insight as to what is a critical component of understanding fractions labels. Our part-whole condition used a nonsense word in the place of the word "parts" or "pieces": "This is called three out of four <u>blicks</u>". On the other hand, Paik and Mix explicitly used the word parts: "of four <u>parts</u>, three" and "three of four <u>parts</u>". Importantly, they found that this explicit mention of the word parts was a critical component of the transparency of the label (Paik & Mix, 2003; Mix & Paik, 2008), which may explain the lower performance shown by children in the part-whole condition in our study. Our goal was to implicitly highlight the "parts" concept by referencing the "denominator" ("look at these four blicks") and using the phrase "out of", however it may be that children this young need the connection between proportional amount of pieces or parts (i.e., the unit) to be more explicit.

In addition to the equivalence-matching task, children performed a separate comparison task in which they had to judge which of two circles had a greater proportion of yellow (i.e., the spinner with the greater chance of landing on yellow). In Experiment 1 and some conditions of Experiment 2, we replicated other studies (Jeong et al., 2007; Hurst & Cordes, under review - A) in that children showed significant numerical interference in this task: performing better when numerical information was consistent with proportional information and performing worse when numerical information was misleading (i.e., contradicted proportional information). However, children who received traditional fraction labels and the single categorical label in Experiment 2 did not show a significant interference effect. This might give some hint that the verbal label training with equivalent proportions may also direct children's attention to proportional information in a very distinct comparison task, however, the inconsistency in this finding across experiments and conditions (i.e., we would expect traditional labels in Experiment 2 to be more like the part-whole labels than the categorical label) makes it unclear exactly what is happening in this task.

On the one hand, we did see a significant correlation between numerical interference on the comparison task and preference to match on number in the equivalence task, even when controlling for age, suggesting that these tasks may be tapping into a similar construct involving children's tendency to attend to numerical information. On the other hand, these differences (both in terms of differing pattern of relations and different effects of the training label) may suggest that the underlying process involved in the comparison task may differ from the equivalence task. For example, although the categorical label promoted equivalence matching, it may be that both the categorical label (only in Experiment 2) and the traditional label are helpful for prompting children to attend to proportion in probabilistic comparison contexts. In addition, individual differences in math ability were not predictive of individual differences in the comparison task, but were predictive of individual differences in numerical interference in the equivalence task. These findings may reflect the relative difficulty or familiarity of working with proportion in these contexts. In particular, it may be that children were more familiar with comparing probabilities or proportions than they were selecting equivalent matches. This is in line with some work suggesting that comparing probabilities may be easier and younger children may be more readily able to inhibit number based responding in these contexts than in equivalence tasks (Hurst & Cordes, under review - A). However, these potential similarities and differences between the comparison task and the equivalence task across conditions and experiments are not entirely clear or systematic (e.g., the categorical condition showing different patterns across Experiments 1 and 2), thus future work should systematically evaluate the relative similarities and differences in how children approach proportional information in these distinct conceptual contexts (comparison versus equivalence matching).

While our findings suggest that the way we label fractions in English may negatively impact children's ability to attend to non-symbolic proportional information, it is not our goal to suggest that we change the English language to have a different label system for fractions. However, there are still several implications of the current findings.

In particular, the current study suggests our current label system (which emphasizes the numerical components) may be inadvertently turning children's attention to particular aspects of fractions (i.e., the components) and not toward the magnitude as a whole (i.e., the proportion), especially for thinking about equivalent fractions. Although prior work has investigated how written symbols or labels impact children's understanding of symbolic fractions (e.g., Ni & Zhou, 2005; Paik & Mix, 2003), this work is the first to suggest that the way we talk about fractions can impact children's attention to proportional versus numerical features of non-symbolic proportion well before these are introduced formally. Given the regular use of fraction words in many informal contexts (e.g., baking, money, relative time or distance, etc.), it is important to recognize that the way we talk about fractions before children learn formal symbols may be providing an early basis for the whole-number bias. Thus, future work should investigate other methods for counteracting this negative influence of traditional fraction labels to encourage children to attend to the proportional amount when thinking about equivalent fractions. For example, we may be able to emphasize continuous features of proportional amount. Other work suggests that providing children with continuous proportional amounts can prevent children's attention to number (e.g., Boyer & Levine, 2015; Hurst & Cordes, under review - A) and children may be better at decimal magnitude than fractions, even early on (Hurst & Cordes, under review-B). Thus, emphasizing the continuous, equivalent nature of proportion through continuous features and/or decimal magnitudes may help children attend to proportional information when it is relevant.

In conclusion, the current study suggests that children who are exposed to fraction labels that highlight the categorical nature of equivalent visual proportions are less likely to attend to irrelevant whole number information when matching equivalent proportions, focusing more on the relevant proportional information itself. These findings highlight that the way we talk about proportion and fractions may be inadvertently turning children's attention to features that are not relevant within certain proportional reasoning contexts. Future work should continue to investigate what factors may be guiding children's attention toward or away from the relevant features within specific concepts (e.g., comparison versus equivalence) and formats (e.g., symbolic versus non-symbolic) in order to provide a more complete picture of children's proportional reasoning.

Supplementary Appendix 4

Label Training Scripts

Full Script: all conditions, both experiments, except Part-Whole Experiment 1:

"We're going to look at some shapes that have some parts that are colored and some parts that are not colored. This is Roo. She likes when her shapes have just the right amount of color and just the right amount with no color! Even if they're different shapes or the shapes are different sizes, the coloring amount has to be just right! Let's see what amount of color Roo likes!"

<u>Three-Fourths:</u>
Put the empty 4-quadrant circle on the table
"I'm going to color [condition specific label] *!"* – Color in 3⁄4 pieces with highlighter
"See, this is called [condition specific label]. *Roo likes this one"* – Place the colored circle on the tray next to Roo.

Put the filled in 3⁄4 rectangle on the table
"This is also called [condition specific label]. *Roo likes this one too!"*—Place the colored rectangle on the tray next to Roo.

Put the filled in 3/8 circle on the table: *"Roo won't like this one"*
Put the filled in 4/4 rectangle on the table: "*Roo won't like this one either"*
Move both of 3/8 circle (Shape 5) and the 4/4 rectangle (Shape 6) to the side, so that they're visible but not on the tray.

<u>Six-Eighths:</u>
Put the empty 8-quadrant circle on the table
"I'm going to color [condition specific label]*!"* – Color in 6/8 pieces with your highlighter
"See, this is called [condition specific label]. *Roo likes this one"* – Place the colored circle on the tray next to Roo.

Put the filled in 6/8 rectangle on the table
"This is also called [condition specific label]. *Roo likes this one too!"* – Place the colored rectangle on the tray next to Roo.

Put the filled in 6/12 circle on the table: *"Roo won't like this one"*
Put the filled in 2/8 rectangle on the table: "*Roo won't like this one either"*
Move both of 6/12 circle and the 2/8 rectangle to the side, so that they're visible but NOT on the tray.

"Roo likes all of these!" [With your hand, pointing to the shapes on the tray]

Condition Specific Labels:

Categorical (Experiment 1 and 2): *blick* OR *dax* (counterbalanced between participants)

Numerator-Focused (Experiment 1): *three-blicks* and *six-daxes* OR *three-daxes* and *six-blicks* (counterbalanced between participants)

Traditional (Experiment 2): *three-fourths* and *six-eighths*

Simplified Part-Whole (Experiment 2): three-out-of-four blicks and six-out-of-eight daxes OR three-out-of-four daxes and six-out-of-eight blicks

Full Script: Part-Whole Experiment 1

Note: as above, whether "blicks" or "daxes" was paired with sixths or eighths was counterbalanced across children. Here, we are using blicks then daxes for simplicity.

"We're going to look at some shapes that have some parts that are colored and some parts that are not colored. This is Roo. She likes when her shapes have just the right amount of color and just the right amount with no color! Even if they're different shapes or the shapes are different sizes, the coloring amount has to be just right! Let's see what amount of color Roo likes!"

Three-Fourths:
Put the empty 4-quadrant circle on the table
"Look at these four blicks"
"I'm going to color three blicks"– Color in 3⁄4 pieces with highlighter
"See, this is called three out of four blicks. Roo likes this one" – Place the colored circle on the tray next to Roo.

Put the filled in 3⁄4 rectangle on the table
"This is also called three out of four blicks. Roo likes this one too!"—Place the colored rectangle on the tray next to Roo.

Put the filled in 3/8 circle on the table: *"Roo won't like this one"*
Put the filled in 4/4 rectangle on the table: "*Roo won't like this one either"*
Move both of 3/8 circle (Shape 5) and the 4/4 rectangle (Shape 6) to the side, so that they're visible but not on the tray.

Six-Eighths:
Put the empty 8-quadrant circle on the table
"Look at these eight daxes"
"I'm going to color six daxes"– Color in 3⁄4 pieces with highlighter
"See, this is called six out of eight daxes. Roo likes this one" – Place the colored circle on the tray next to Roo.

Put the filled in 6/8 rectangle on the table

“This is also called six out of eight daxes. Roo likes this one too!” – Place the colored rectangle on the tray next to Roo.

Put the filled in 6/12 circle on the table: *“Roo won’t like this one”*
Put the filled in 2/8 rectangle on the table: “*Roo won’t like this one either”*
Move both of 6/12 circle and the 2/8 rectangle to the side, so that they’re visible but NOT on the tray.

“Roo likes all of these!” [With your hand, pointing to the shapes on the tray]

Stimuli Equivalence-Matching Task

	Sample Stimulus	Correct / Proportion Match	Incorrect / Number Match
No Competing Information	3/8	6/16	5/7
	2/6	1/3	4/7
	3/5	3/5	4/12
	4/7	4/7	4/14
	10/18	10/18	17/18
Numerical VS. Proportional Information	6/7	12/14	6/13
	2/3	4/6	2/7
	4/8	2/4	4/18
	4/10	2/5	4/6
	4/5	8/10	2/5
	5/7	10/14	3/7
	2/8	1/4	5/8
	6/18	3/9	10/18

CHAPTER 5: IMPLICATIONS AND FUTURE DIRECTIONS

Proportion is ubiquitous; from early-developing intuitions in infants (e.g., Denison & Xu, 2010; McCrink & Wynn, 2007) and non-human animals (e.g., Rugani, McCrink, de Hevia, Vallortigara, & Regolin, 2016; Tecwyn, Denison, Messer, & Buchsbaum, 2017) to formal systems that regulate our economic and social worlds. Proportional information takes on many forms, both conceptually (e.g., ratios, percentages, etc.) and in the format of the representation (e.g., 1/2 vs. 0.5), and these representations do not equally align with proportion concepts. For example, whether the use of fraction or decimal notation will be preferred depends on whether the task at hand relies upon understanding proportional magnitude, arithmetic, or part-whole relations (e.g., DeWolf et al., 2014; Hurst & Cordes, 2016, under review - B; Hurst, Relander, & Cordes, 2016; Iuculano & Butterworth, 2011). In order to understand the *conceptual* difficulties children face, and how to rectify them, we must focus on understanding how these distinct proportional *representations* align with proportional concepts. Nuances across these representations can focus attention to different kinds of information available in the representation, meanwhile shaping the concepts individuals believe to be relevant, the strategies individuals employ to solve a task involving proportions, and ultimately the decisions individuals make when solving a problem (whether that is deciding to take a loan with 18% interest or deciding to invert the fraction before multiplying).

In this dissertation, I presented three projects, involving six experiments, investigating how the format of various common proportional representations impact the way children and adults perform on proportion tasks. I have taken two approaches to these questions. In Project 1, I investigated whether there are specific benefits for representing magnitude information using number lines in particular and how fractions and pie charts (which contain information about both magnitude and the components) may represent magnitude in some contexts. Although we did not find that performance on the spatial mapping tasks primed performance on a subsequent symbolic magnitude task, we did find evidence that both adults and children employ distinct strategies when using pie charts and number lines to think about symbolic fractions and decimals.

Second, in Projects 2 and 3, I investigated those factors that direct children's attention toward proportion magnitude, and away from competing numerical components, even when component information may be particularly salient. Overall, our findings suggest that it was relatively easy to encourage children to attend to proportion over numerical features by first providing them with experience attending to proportion in the absence of competing numerical information or by labeling the magnitude with a single categorical label that highlighted common magnitudes across equivalent proportions. Our findings revealed that attention to proportion early in childhood may be particularly malleable to early input and experiences. In addition, the child's age and analogical reasoning ability may also predict the information children attend to such that younger children, who are less numerically adept, may be less likely to attend to numerical features and children with heightened relational reasoning skills may be better at proportional reasoning more generally.

Building on previous literature documenting numerical interference effects (Boyer et al., 2008; Jeong et al., 2007), these three projects extend this work by determining that numerical interference is malleable and by identifying specific factors and individual differences that influence whether individuals are likely to attend to proportional magnitude or the specific components.

Implications for Proportional Reasoning

Taken together, findings of these projects ultimately suggest that how children and adults use a given representation for proportional information depends upon what kind of information the representation is providing. In particular, in the case of proportion, representations that retain specific information about the components may be particularly difficult to use in magnitude contexts because the information about the components interferes with attention to proportional magnitude. However, importantly, children's attention to either magnitude or componential information appears to be fairly malleable, such that across contexts, their attention can be easily directed toward the relevant information.

Representational Specificity

Prior work has found that experience with number lines can improve symbolic magnitude abilities more so than area models or no representation (e.g., Hamden & Gunderson, 2017; Wang & Siegler, 2013). In Project 1, I attempted to replicate and extend these findings specifically within the context of a symbolic speeded numerical comparison task using a significantly shorter priming task; however we found no effect of our manipulation. While it is possible that whether one visualizes fractions as falling along a number line or as a pie chart does not have any impact on symbolic magnitude judgments, our findings can only be interpreted as inconclusive as post-hoc power analyses revealed that the study was underpowered. However, other more nuanced findings of the project indicated that number lines promote thinking about proportional magnitude information. In particular, adults and children almost exclusively used an estimation strategy with number lines, regardless of whether they were presented fraction or decimal notation. Thus, the ordered and linear nature of number lines may be particularly important for prompting people to think about magnitude, even with fraction symbols that also contain information about the numerical components. Pie charts, on the other hand, resulted in a greater tendency to invoke partitioning strategies, suggesting that they are more likely to cause individuals to focus on specific information about the discrete components. Thus, these findings suggest that number lines are very much associated with magnitude information and may encourage people to think about all symbolic representations as magnitudes. In this way, the current findings indirectly support evidence that number lines may be particularly useful for learning about fraction magnitudes.

However, this extreme specificity of number lines for magnitude information also highlights some important limitations of this representation. Although number lines may be particularly good at conveying magnitude, these representations may not readily be interpreted as providing information about the fraction components. Pie charts, however, were associated with more variable strategies, including both magnitude estimation strategies and partitioning strategies, which emphasize the discrete components. In particular, although mapping fractions to pie charts primarily involved a partitioning strategy for proper fractions (between 0 and 1), when multiple pie charts were used in a way that involved fractions, decimals, and whole numbers, beyond simple proper fractions, adults and children show much more variability in the way they used the pie charts. That is, some people relied exclusively on estimation strategies, some relied exclusively on partitioning strategies, while others engaged in both partitioning and estimation strategies depending on the specific context. Thus, these findings might suggest that the way individuals approach area models may be sometimes in terms of information about the components (i.e., with a partitioning strategy), and sometimes in terms of estimating magnitude directly (i.e., with an estimation strategy). Given that we did not find any significant differences in accuracy between mapping fractions to number lines and pie charts, we cannot speak to the benefits of perceiving pie charts in a more flexible manner—however, future research could further investigate whether this flexibility offered by area models may result in more adaptable responding in which individuals can focus on either magnitude or component information, depending on the context. This is line with other work revealing that when adults were forced to map between symbolic and spatial representations on a computer (without paper or pencil strategies being available) they were in general more accurate with pie charts than with number lines and were particularly inaccurate when mapping from a number line to a fraction (Hurst, Relander, & Cordes, 2016). Therefore, the benefits of number lines may be restricted to representing continuous magnitude information. Thus, in instances where the goal of instruction is to focus on information about the specific components (e.g., differentiating 4/8 from 8/16), then using a number line may not be as beneficial and instead come at the expense of readily interpreting the specific information about the components.

There are also open questions about the specificity of symbolic rational numbers. Fractions are particularly poor for representing magnitudes (e.g., DeWolf et al., 2014; Hurst & Cordes, 2016; Project 1, Experiment 1) but useful for conveying part-whole information (given the highly salient components), and decimals are particularly useful for representing magnitudes (e.g., DeWolf et al., 2014; Hurst & Cordes, 2016) but not information about the specific components (given that all equivalent fractions have the same decimal value). Given these complementary affordances, it may be worthwhile to explicitly communicate these specific benefits, as well as the potential drawbacks of each representation, to children. In other words, it may be that a critical, but currently missing, component of proportion instruction is explicitly teaching children how to choose which representation to use in a given context or for a specific purpose. That is, given that fractions, decimals, percent, ratios, number lines, pie charts, and many others, are variations of each other, children may be missing key information about those contexts in which each representation would be *best.* By focusing on the accessibility of magnitude information across all representations, students may not be gaining an understanding of the adaptive specificity of each of these representations. There is some evidence that this alignment between distinct spatial representations (discrete vs. continuous) and symbolic notation (fractions and decimals) is made implicitly in children's textbooks, with fractions being used in discrete word problems and decimals being used in both continuous and discrete problems (Rapp, Bassok, DeWolf, & Holyoak, 2015). Furthermore, 4th graders who have more explicit alignment preferences between different kinds of representations have higher fraction ability than children who do not have as strong associations between representations (Moseley & Okamoto, 2008). Thus, given that a skill critical for fraction proficiency (and mathematics more generally) is to know what information to use when (Fazio, DeWolf, & Siegler, 2016), making this alignment explicit for students and allowing for some notational specificity (i.e., acknowledging that fractions are better for part-whole and not great for magnitude) may lead to better proportional understanding and more flexible thinking overall.

Preventing Numerical Interference

In the current projects, I have been almost entirely focused on encouraging children and adults to think about *magnitude* and the components of the proportional information were considered almost exclusively in contrast to proportional magnitude. However, this tension between separate components and holistic magnitude may not be a necessary part of proportional reasoning. Rather, given that young children (in Projects 2 and 3, as young as 5- or 6-years-old) show flexibility between attending to magnitude versus numerical components and that children and adults do not show a consistent approach to working with pie charts (Project 1), it may be possible to maintain this flexibility throughout early education of fraction concepts and prevent and/or reduce the impact of these overt whole number biases early on. For example, it may be beneficial to provide some formal fraction instruction much earlier than current standards (typically in 3^{rd} or 4^{th} grade; National Governors Association Center for Best Practices, 2010) and introduce proportional concepts using continuous representations before discrete representations to prevent children from relying upon strong whole number biases when proportional reasoning.

One way to better integrate thinking about magnitude and part-whole representations is through other concepts children learn, such as measurement. Work with young children suggests that numerical interference in discrete area models arises because of an overt salience of counting strategies and whole-number concepts (e.g., Boyer et al., 2008). However, when children make this counting error, they also ignore important features of the items they are counting: the *size* of the unit. Numerical biases lead children to indicate that an area model depicting 4/9 has a greater number of relevant pieces (i.e., numerator) than one depicting 3/5 (that is, 4>3), however (given two identically sized objects), the size of each part in the former (each 1/9 piece) is smaller than that of the latter (each 1/5 piece). Although measurement is also not always an easy concept to learn (e.g., Clements, 1999; Solomon, Vasilyeva, Huttenlocher & Levine, 2015; Vasilyeva, Casey, Dearing, & Ganley, 2009), children spend a substantial amount of time learning about units when they are learning measurement as early as 1st grade (National Governors Association Center for Best Practices, 2010). As such, children's understandings of proportional concepts may benefit from incorporating these earlier taught ideas of measurement along with part-whole reasoning about fractions to allow children to better integrate their ideas of part-whole with their ideas of relative magnitude.

Relatedly, work comparing the use of continuous versus discrete area models (Boyer & Levine, 2015; current Project 2) and symbolic representations (Moss & Case, 1990) suggests that teaching with continuous magnitude representations (i.e., continuous area models; decimals) before introducing discrete representations or formal fractions may improve children's ability to attend to proportional information and learn about fractions. Incorporating these two ideas together, explicitly highlighting the relations between continuous representations and discrete representations by highlighting that discrete representations have a well-defined unit that continuous representations do not may help children make the connections between these representations. For example, when children engage in discretizing the same continuous representation into different discrete representations (e.g., dividing a circle into fourths and then another circle into eighths), in a way that highlights that the primary difference in these discrete representations is the size of the unit, they may be able to better attend to the information that is retained among these transformations, rather than superficial perceptual differences in the representations themselves.

Domain General Skills

Overall, findings from the current project involving analogical reasoning add to a substantial literature suggesting that proportional reasoning and fraction learning are impacted by domain general skills like working memory and executive functioning (Fuchs et al., 2014; Jordan, Hansen, Fuchs, Siegler, Gersten, & Micklos, 2013; Vukovic et al., 2014; Ye, Resnick, Hansen, Rodrigues, Rinne, & Jordan, 2016). In Project 2, I investigated the role that analogical reasoning may play in proportional reasoning specifically, since proportional reasoning is by definition interpreting the relation between two quantities. However, the specific role of the domain general skills may also be particularly relevant when thinking about numerical interference and different kinds of representations for proportional reasoning.

Given that proportional information is represented in many different ways, including different symbolic and spatial representations, the role of working memory, and in particular different components of working memory, may be particularly relevant for how these representations are used. For example, given the recent evidence of using number lines for teaching fraction magnitude (National Governors Association Center for Best Practices, 2010), we must also consider how this supposed benefit of number lines may differ for children who may have lower visuospatial working memory. In line with this, some evidence with adults has found that individual differences in the kinds of working memory strategies adults when comparing symbolic rational number magnitudes are related to their algebra ability (Hurst & Cordes, 2017). These findings suggest that different aspects of working memory (verbal or visual-spatial) may be differentially implicated in proportional reasoning, dependent upon both the format of the proportional representation, the particular strategies chosen to solve the task, and overall conceptual understanding.

The primary phenomenon investigated in the current dissertation was numerical interference in the context of proportional representations. Given that this preference for numerical information over proportional magnitude has often been attributed to an over-emphasis on discrete whole numbers prior to the introduction of formal proportions (Boyer et al., 2008; Durkin & Rittle-Johnson, 2015; Lortie-Forgues, Tian, & Siegler, 2015; Vamvakoussi, 2015), it may be important to consider the central role that inhibitory control may play. In particular, if discrete numerical components are more immediately salient such that attending to proportional magnitude requires additional effort in the presence of part-whole representations, then children with lower inhibitory control may be more likely to show numerical interference. Future work should continue directly investigate the potentially unique role of general relational reasoning, working memory, and inhibitory control in learning about formal proportional information, either in symbolic or spatial representational formats. This research may be able to shed light on not only individual differences leading to potential limitations for some individuals, but also ways to counteract these limitations.

In sum, although the current projects revealed that numerical interference is dependent on context, is malleable, and related to individual differences, additional research is needed to investigate additional factors relevant to numerical interference and whether numerical interference can be prevented more globally.

Future Directions: Viewing Proportional Reasoning Through Another Lens

The current projects focused on a specific phenomena of proportional reasoning: numerical interference, and to do so highlighted the role of multiple representations and interfering information in children's proportional reasoning. However, proportional reasoning is not the only domain where multiple representations and competing information are relevant to children's learning of a concept or success on a task. Rather, these ideas are relevant in other domains of science and math (e.g., Physics: Cheng, 1999) and for children's developing sense of symbolic representation more generally. Thus, future research further investigating children's use of multiple representations and interfering information in proportional reasoning may benefit from some of the insight gained in these other separate, but highly related, domains.

Symbolic Representation in General

Investigating how concepts, ideas, or objects are represented, both externally and in the mind, is a central study of many aspects of cognition, even beyond formal representations in mathematics. For example, research on children's use of written symbols (e.g., Bialystok, 1992, 2000; Bialystok & Martin, 2002), maps (e.g., Huang & Spelke, 2015), and symbolic artifacts (e.g., scale models; DeLoache, 2011) has provided substantial insight into the way children mentally represent things and communicate that representation with others (e.g., in the form of written symbols).

Children's ability to use representations in a symbolic way shows substantial development in early childhood, well before the introduction of formal mathematical symbols. This research highlights at least two key features that are critically important for symbolic reasoning: representational insight (DeLoach, 2002) and understanding the one-to-one relation between symbol and referent (e.g., Cohen, 1985; Eskritt & Olson, 2012). "Representational insight" refers to knowledge children have that there is a relation between the symbol and the referent in a given context (DeLoach, 2002). That is, children must know that the symbol is acting as a representation. However, this is not always obvious and things like perceptual similarity, prior experience, or individual differences can greatly impact whether children will gain representational insight in a particular context. However, children do show relatively early abilities to symbolically represent information, succeeding in using scale-model representations around 3-years-old (see DeLoach, 2002 for a review). In addition to children's use of given representations, understanding the one-to-one relation between symbol and referent is important for symbolic representation. Although children initially show difficulty with this one-to-one mapping, for example interpreting symbols differently from the way they produced them (e.g, Eskritt & Olson, 2012), by around 8-years-old children show a reliable preference for one-to-one mappings between symbol and referent (Cohen, 1985).

Although children are successful at more generally symbolic thinking fairly young (e.g., using symbolic representations around 3-years-old; producing and interpreting meaningful symbols around age 8-years), younger than the age at which they are introduced to formal symbolic proportional symbols, we may be able to interpret some of children's difficulty with formal proportion symbols within the mindset of their earlier difficulties with symbolic reasoning more generally. For example, one potential factor impacting children's ability to use varying mathematical representations (and proportional representations in particular) may be whether or not they have *representational insight*; in other words, whether they are even approaching the symbol as a representation, rather than just an object in and of itself. Given that a fraction, for example, can be a magnitude, a part-whole relation, or a division operation, considering these multifaceted interpretations may make it difficult for children to think about the fraction notation as a symbolic representation at all. Separately, some of the difficulty in proportional reasoning may also be driven by a lack of one-to-one mappings among symbols, concepts, and contexts. For example, multiple symbols can be used to represent the same or similar concept (e.g., fractions and decimals) and the same symbol may be reused across concepts and contexts (e.g., the dividing line for subtraction, division, and fractions). Thus, children must become fairly flexible in their use of symbols, while still maintaining the conventional notations used across many domains of math.

Thus, it may be that some of children's difficulty in symbolic proportional reasoning (e.g., fractions, decimals, percent, etc.) may lie in understanding the symbolic notation itself, rather than in the meaning of the specific symbols. Furthermore, by comparing the ways in which children reason symbolically in formal math with the way they learn to reason symbolically in early childhood we may be able to shed additional light on these difficulties.

Multiple Representations in Other Domains

The idea of using multiple representations is not unique to proportional information and has been discussed in educational guidelines, theoretical frameworks, and empirical lab studies about mathematical reasoning and problem solving across many math and science domains, in addition to fraction learning (e.g., Ainsworth, 1999; Braithwaite & Goldsone, 2013; Cheng, 1999; Gagatsis & Shiakalli, 2004; O'Halloran, 2014; Rau & Matthews, 2017; Weise, 2003; Zahner & Corter, 2010). However, the use of multiple representations is not always straightforward and there are several things to consider. For example, when the connections among multiple representations are unclear, children may not benefit from the use of multiple representations (Rau & Matthews, 2017). In addition, Ainsworth (1999) suggests that an important aspect of using multiple representations is to acknowledge the purpose of having multiple representations, whether it is to construct deeper meaning, to provide complementary information, or to constrain the interpretation of other representations. Thus, carefully considering the purpose of the representations used in fraction instruction and, as discussed earlier, the potential specificity of certain representations may help to clarify the relations among representations and, as a consequence, allow children to better develop their conceptual understanding of the topic more generally.

Furthermore, there is substantial evidence that being able to reason about rational number information is particularly important for learning algebra (e.g., Bailey, Hoard, Nugent, & Geary, 2012; DeWolf et al., 2016; Hurst & Cordes, 2017; Siegler et al., 2012). One factor that may be implicated in the relation between these two domains may be a more general ability to work with multiple representations and having flexibility in translating between representational formats. For example, although fraction and decimal arithmetic is particularly difficult for children (Lortie-Forgues et al., 2016), some evidence suggests that children's ability to easily translate among multiple representations of rational numbers may be particularly critical for solving fraction and decimal arithmetic (Deliyianni & Gagatsis, 2013; Panaoura, Gagatsis, Deliyianni, & Elia, 2009, 2010).

In addition, research on common errors in algebra indicates that people make representational errors because they rely on perceptual heuristics (e.g., Bernardo & Okagaki, 1994; Fisher, Borchert, & Bassok, 2011; Landy, Brookes, & Smout, 2014; Landy & Goldstone, 2007). For example, reversal errors occur when symbolic representations of a word problem are written in the same order in which the information is provided in the word problem, such as writing "there are four times as many students as professors" as 4*students = # of professors. However, this error can be prevented when adults are forced to use or think about different relational interpretations, for example using a division representation instead of a multiplication representation, preventing heuristic based alignment (Bernardo & Okagaki, 1994; Fisher et al., 2011).

Taking these findings together, the use of distinct representations is a critical component of algebra and rational number problem solving. Thus, it may be possible to promote flexible problem solving with multiple representations in a domain general way that would improve people's use of mathematical representations in multiple domains. Future research should investigate the use of multiple representations across domains, including algebra, in order to determine whether there are generalizable features of particular representations that may better promote attention to the relevant conceptual information.

Conclusion

In sum, in this dissertation I investigated how children and adults use different kinds of proportional representations for thinking about magnitude and whether their use of representations that contain both magnitude and information about the component parts may be malleable to different context effects and show individual differences across age and more domain-general skills. Overall, the current findings suggest that although the availability of information about discrete components does lead to numerical interference, this heightened attention to components, rather than magnitude, is highly malleable and dependent on individual differences as well as contextual features. In particular, for older children and adults, number lines may be particularly oriented toward a magnitude interpretation of fractions and decimals, but pie charts may show some flexibility in how they are used. In younger children, attention to proportional magnitude over numerical components may be promoted by both relevant experiences and input as well as impacted by individual differences. Together, these findings point to new ways to think about the role of numerical interference in proportional thinking and highlight novel directions for fraction education to promote children's abilities to think about fraction magnitudes.

References

Ainsworth, S. (1999). The functions of multiple representations. *Computers & Education, 33,* 131-152.

Alexander, P. A., Dumas, D., Grossnickle, E. M., List, A., & Firetto, C. M. (2016). Measuring Relational Reasoning. *The Journal of Experimental Education, 84,* 119-151. doi:10.1080/00220973.2014.963216

Alibali, M. W., & Sidney, P. G. (2015). Variability in the natural number bias: Who, when, how, and why. *Learning and Instruction, 37,* 56-61. doi:10.1016/j.learninstruc.2015.01.003

Atagi, N., DeWolf, M., Stigler, J., & Johnson, S. P. (2016). The Role of Visual Representations in College Students' Understanding of Mathematical Notation. *Journal of Experimental Psychology: Applied*. doi:10.1037/ xap0000090

Bailey, D. H., Hoard, M. K., Nugent, L., & Geary, D. C. (2012). Competence with fractions predicts gains in mathematics achievement. *Journal of Experimental Child Psychology, 113,* 447-455. doi:10.1016/j.jecp.2012.06.004

Baroody, A. J., & Price, J. (1983). The Development of the Number-Word Sequence in the Counting of Three-Year-Olds. *Journal of Research in Mathematics Education, 14,* 361-368.

Barth, H., Kanwisher, N., & Spelke, E. (2003). The construction of large number representations in adults. *Cognition*, *86*(3), 201-221.

Barth, H., La Mont, K., Lipton, J., and Spelke, E. (2005). Abstract number and arithmetic in preschool children. *Proceedings of the National Academy of Sciences, 102*, 14116-14121.

Bernardo, A. & Okagaki, L. (1994). Roles of Symbolic Knowledge and Problem-Information Context in Solving Word Problems. *Journal of Educational Psychology, 86,* . doi:10.1037/0022-0663.86.2.212

Berteletti, I., Lucangeli, D., & Zorzi, M. (2012). Representation of numerical and non-numerical order in children. *Cognition, 124,* 304-313. doi:10.1016/j.cognition.2012.05.015

Bialystok, E. (1992). Symbolic Representations of Letters and Numbers. *Cognitive Development, 7,* 301-316.

Bialystok, E. (2000). Symbolic Representation across Domains in Preschool Children. *Journal of Experimental Child Psychology, 76,* 173-189. doi:10.1006/jecp.1999.2548

Bialystok, E., & Martin, M. M. (2003). Notation to symbol: Development in children's understanding of print. *Journal of Experimental Child Psychology, 86,* 223-243. doi:10.1016/S0022-0965(03)00138-3

Bonato, M., Fabbri, S., Umilta, C., & Zorzi, M. (2007). The Mental Representations of Numerical Fractions: Real or Integer. *Journal of Experimental Psychology: Human Perception and Performance, 33,* 1410-1419. doi:10.1037/0096-1523.33.6.1410

Booth, J. L., & Newton, K. J (2012). Fractions: Could they really be the gatekeeper's doorman?. *Contemporary Educational Psychology, 37,* 247-253. doi:10.1016/j.cedpsych.2012.07.001

Booth, J. L., & Siegler, R. S. (2006). Developmental and Individual Differences in Pure Numerical Estimation. *Developmental Psychology, 41,* 189-201.

Boyer, T. W., & Levine, S. C. (2015). Prompting Children to Reason Proportionally: Processing Discrete Units as Continuous Amounts. *Developmental Psychology, 51,* 615-620. doi:10.1037/a0039010

Boyer, T. W., Levine, S. C., & Huttenlocher, J. (2008). Development of Proportional Reasoning: Where Young Children Go Wrong. *Developmental Psychology, 44,* 1478-1490. doi:10.1037/a0013110

Braithwaite, D. W., & Goldstone, R. L. (2013). Integrating Formal and Grounded Representations in Combinatorics Learning. *Journal of Educational Psychology, 105,* . doi:10.1037/a0032095

Chan, J. Y., & Mazzocco, M.M.M. (2017). Competing features influence children's attention to number. *Journal of Experimental Child Psychology, 156,* 62-81. doi:10.1016/j.jecp.2016.11.008

Cheng, P. C.-H. (1999). Unlocking conceptual learning in mathematics and science with effective representational systems. *Computers & Education, 33,* 109-130. doi:

Christou, K. P., & Vosniadou, S. (2012). What Kinds of Numbers Do Students Assign to Literal Symbols? Aspects of the Transition from Arithmetic to Algebra. *Mathematical Thinking and Learning, 14,* 1-27. doi:10.1080/10986065.2012.625074

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

Cohen, D. J., & Blanc-Goldhammer, D. (2011). Numerical bias in bounded and unbounded number line tasks. *Psychonomic Bulletin and Review, 18,* 331-338. doi:10.3758/s13423-011-0059-z

Cohen, S. (1985). The Development of Constraints on Symbol-Meaning Structure in Notation: Evidence from Production, Interpretation, and Forced-Choice Judgments. *Child Development, 56,* 177-195. doi:10.2307/1130184

Condry, K. F., & Spelke, E. S. (2008). The development of language and abstract concepts: The case of natural number. *Journal of Experimental Psychology, 137,* 22-38. doi:10.1037/0096-3445.137.1.22

Cordes, S., & Brannon, E. M. (2008). Quantitative competencies in infancy. *Developmental Science, 11,* 803-808. doi:10.1111/j.1467-7687.2008.00770.x

Cramer, K., & Wyberg, T. (2009). Efficacy of different concrete models for teaching the part-whole construct for fractions. *Mathematical thinking and learning*, *11*(4), 226-257.

Cramer, K. A., Post, T. R., & delMas, R. C. (2002). Initial Fraction Learning by Fourth- and Fifth-Grade Students: A Comparison of the Effects of Using Commercial Curricula with the Effects of Using the Rational Number Project Curriclum. *Journal for Research in Mathematics Education, 33,* 111-144.

Cramer, K., Wyberg, T., & Leavitt, S. (2008). The Role of Representations in Fraction Addition and Substraction. *Mathematics Teaching in the Middle School, 13,* 490-496.

Dehaene, S., Bossini, S., & Giraux, P. (1993). The Mental Representation of Parity and Number Magnitude. *Journal of Experimental Psychology: General, 122,* 371-396.

Deliyianni, E. & Gagatsis, A. (2013). Tracing the development of representational flexibility and problem solving in fraction addition: a longitudinal study. *Educational Psychology: An International Journal of Experimental Educational*

Psychology, 33, 427-442. doi:10.1080/01443410.2013.765540

DeLoache, J. (2002). Early Development of the Understanding and Use of Symbolic Artifacts. In U. Goswami (Ed.) *The Blackwell Handbook of Childhood Cognitive Development (pp.* 312-336). doi:10.1002/9781444325485.ch12

Denison, S., & Xu, F. (2010). Twelve- to 14-month-old infants can predict single-event probability with large set sizes. *Developmental Science, 13,* 798-803. doi:10.1111/j.1467-7687.2009.00943.x

Denison, S., Reed, C., & Xu, F. (2013). The Emergence of Probabilistic Reasoning in Very Young Infants: Evidence from 4.5- and 6-Month-Olds. *Developmental Psychology, 49,* 243-249. doi:10.1037/a0028278

Desmet, L., Gregoire, J., & Mussolin, C. (2010). Developmental changes in the comparison of decimal fractions. *Learning and Instruction, 20,* 521-532. doi:10.1016/j.learninstruc.2009.07.004

DeWolf, M., & Vosniadou, S. (2015). The representation of fraction magnitudes and the whole number bias reconsidered. *Learning and Instruction, 37,* 39-49. doi:10.1016/j.learninstruc.2014.07.002

DeWolf, M., Bassok, M., & Holyoak, K. J. (2015). Conceptual Structure and the Procedural Affordances of Rational Numbers: Relational Reasoning with Fractions and Decimals. *Journal of Experimental Psychology: General, 144,* 127-150. doi:10.1037/xge0000034

DeWolf, M., Grounds, M. A., Bassok, M., & Holyoak, K. J. (2014). Magnitude Comparison with Different Types of Rational Numbers. *Journal of Experimental Psychology: Human Perception and Performance, 40,* 71-82.

doi:10.1037/a0032916

Dowker, A., & Nuerk, H-C. (2016). Editorial: Linguistic Influences on Mathematics. *Frontiers in Psychology, 7,* 1-4. doi:10.3389/fpsyg.2016.01035

Duffy, S., Huttenlocher, J., & Levine, S. (2005). It Is All Relative: How Young Children Encode Extent. *Journal of Cognition and Development, 6,* 51-63. doi:10.1207/s15327647jcd0601_4

Empson, S. B., Junk, D., Dominguez, H., & Turner, E. (2006). Fractions as the Coordination of Multiplicatively Related Quantities: A Cross-Sectional Study of Children's Thinking. *Educational Studies in Mathematics, 63,* 1-28. doi:10.1007/sl0649-005-9000-6

English, L. (2004). Mathematical and Analogical Reasoning in Early Childhood. In L. English (Ed.),*Mathematical and Analogical Reasoning of Young Learner*s (pages 1-22). New York, NY: Routledge.

Eskritt, M., & Olson, D. (2012). From Depiction to Notation: How Children use Symbols to Represent Objects and Events. *Journal of Cognition and Development, 13,* 189-207. doi:10.1080/15248372.2011.590786

Fabbri, S., Caviola, S., Tang, J., Zorzi, M., & Butterworth, B. (2012). The role of numerosty in processing nonsymbolic proportions. *The Quarterly Journal of Experimental Psychology, 65,* 2435-2446. doi:10.1080/17470218.2012.694896

Faulkenberry, T. J., & Pierce, B. H. (2011). Mental Representations in Fraction Comparison: Holistic Versus Component-Based Strategies. *Experimental Psychology, 58,* 480-489. doi:10.1027/1618-3169/a000116

Fazio, L. K., DeWolf, M., & Siegler, R. S. (2016). Strategy Use and Strategy Choice in

Fraction Magnitude Comparison. *Journal of Experimental Psychology: Learning, Memory, and Cognition.*

Ferry, A. L., Hespos, S. J., & Gentner, D. (2015). Prelinguistic Relational Concepts: Investigating Analogical Processing in Infants. *Child Development, 86,* 1386-1405. doi:10.1111/cdev.12381

Fisher, K., Borchert, K. & Bassok, M. (2011). Following the standard form: Effects of equation format on algebraic modeling. *Memory & Cognition, 39,* 502-515. doi:10.3758/s13421-010-0031-6

Fuchs, L., Schumacher, R., Sterba, S., Long, J., Namkung, J., Malone, A., Hamlett, C., Jordan, N., Gersten, R., Siegler, R. & Changas, P. (2014). Does Working Memory Moderate the Effects of Fraction Intervention? An Aptitude–Treatment Interaction. *Journal of Educational Psychology, 106,* 499-514. doi:10.1037/a0034341

Fulkerson, A. L., & Waxman, S. R. (2007). Words (but not Tones) facilitate object categorization: Evidence from 6- and 12-month-olds. *Cognition, 105,* 218-228. doi:10.1016/j.cognition.2006.09.005

Fuson, K. C., & Kwon, Y. (1992). Korean Children's Understanding of Multidigit Addition and Substraction. *Child Development, 63,* 491-506.

Gagatsis, A., & Shiakalli, M. (2004). Ability to translate from one representation of the concept of function to another and mathematical problem solving. *Educational Psychology, 24.* doi:10.1080/0144341042000262953

Gamer, M., Lemon, J., Fellows, I., & Singh, P. (2012). irr: Various coefficients of interrater reliability and agreement. *R package version 0.84*, *137*.

Ganor-Stern, D. (2012). Fractions but not negative numbers are represented on the mental number line. *Acta Psychologica, 139,* 350-357. doi:10.1016/j.actpsy.2011.11.008

Ganor-Stern, D. (2013). Are 1/2 and 0.5 represented in the same way?. *Acta Psychologica, 142,* 299-307. doi:10.1016/j.actpsy.2013.01.003

Gentner, D., Levine, S. C., Ping, R., Isaia, A., Dhillon, S., Bradley, C., & Honke, G. (2016). Rapid Learning in a Children's Museum via Analogical Comparison. *Cognitive Science, 40,* 224-240. doi:10.1111/cogs.12248

Gevers, W., Lammertyn, J., Notebaert, W., Verguts, T., & Fias, W. (2006). Automatic response activation of implicit spatial information: Evidence from the SNARC effect. *Acta Psychologica, 122,* 221-233. doi:10.1016/j.actpsy.2005.11.004

Ghetti, S., Hembacher, E., & Coughlin, C. A. (2013). Feeling Uncertain and Acting on It During the Preschool years: A Metacognitive Approach. *Child Development Perspectives, 7,* 160-165. doi:10.1111/cdep.12035

Girotto, V., Lontanari, L., Gonzalez, M., Vallortigara, G., & Blaye, A. (2016). Young Children do not succeed in choice tasks that imply evaluating chances. *Cognition, 152,* 32-29. doi:10.1016/j.cognition.2016.03.010

Goswami, U. (1989). Relational Complexity and the Development of Analogical Reasoning. *Cognitive Development, 4,* 251-268.

Graham, S. A., Namy, L. L., Gentner, D., & Meagher, K. (2010). The role of comparison in preschoolers' novel object categorization. *Journal of Experimental Child Psychology, 107,* 280-290. doi:10.1016/j.jecp.2010.04.017

Gray, S. A., & Reeve, R. A. (2016). Number-specific and general cognitive markers of preschoolers’ math ability profiles. *Journal of experimental child* *psychology*, *147*, 1-21.

Halberda, J., & Feigenson, L. (2008). Developmental Change in the Acuity of the "Number Sense": Approximately Number System in 3-, 4-, 5-, and 6-Year-Olds and Adults. *Developmental Psychology, 44,* 1457-1465. doi:10.1037/a0012682

Hamdan, N., & Gunderson, E. A. (2017). The number line is a critical spatial-numerical representation: Evidence from a fraction intervention. *Developmental psychology*, *53*(3), 587.

Hembacher, E., & Ghetti, S. (2014). Don't Look at My Answer: Subjective Uncertainty Underlies Preschoolers' Exclusion of Their Least Accurate Memories. *Psychological Science, 25,* 1768-1776. doi:10.1177/0956797614542273

Huang, Y., & Spelke, E. S. (2015). Core Knowledge and the Emergence of Symbols: The Case of Maps. *Journal of Cognition and Development, 16,* 81-96. doi:10.1080/15248372.2013.784975

Hughes, M. (1983). Teaching Arithmetic to Pre-School Children. *Educational Review, 35,* 163-173.

Hurst, M., & Cordes, S. (2017). A systematic investigation of the link between rational number processing and algebra ability. *British Journal of Psychology*.

Hurst, M. A., & Cordes, S. (under review - A). Attending to Relations: Proportional Reasoning in 3- to 6-year-old children. [note: this is Project 2]

Hurst, M. A., & Cordes, S. (under review - B). Children's understanding of fraction and decimal symbolic magnitudes and its relationship to pre-algebra ability.

Hurst, M., & Cordes, S. (2016). Rational-number comparison across notation: Fractions, decimals, and whole numbers. *Journal of Experimental Psychology: Human*

Perception and Performance, 42, 281-293. doi:10.1037/xhp0000140

Hurst, M., Anderson, U., & Cordes, S. (2016). Mapping Among Number Words, Numerals, and Non-Symbolic Quantities in Preschoolers. *Journal of Cognition and Development, 18,* 41-62. doi:10.1080/15248372.2016.1228653

Hurst, M., Relander, C., & Cordes, S. (2016). Biases and benefits of number lines and pie charts in proportion representation. In *Proceedings of the 38th Annual Conference of the Cognitive Science Society* (pp. 586-591).

Hurst, M. A., Santry, M., Relander, C., & Cordes, S. (in preparation). lignment between Distinct Spatial and Symbolic Representations of Proportion

Huttenlocher, J., Duffy, S., & Levine, S. (2002). Infants and Toddlers Discriminate Amount: Are They Measuring?. *Psychological Science, 13,* 244-250.

Iuculano, T., & Butterworth, B. (2011). Understanding the real value of fractions and decimals. *The Quarterly Journal of Experimental Psychology, 64,* 2088-2098. doi:10.1080/17470218.2011.604785

Jeong, Y., Levine, S. C., & Huttenlocher, J. (2007). The Development of Proportional Reasoning: Effect of Continuous Versus Discrete Quantities. *Journal of Cognition and Development, 8,* 237-256. doi:10.1080/15248370701202471

Jordan, N. C., Hansen, N., Fuchs, L. S., Siegler, R. S., Gersten, R., & Micklos, D. (2013). Developmental predictors of fraction concepts and procedures. *Journal of Experimental Child Psychology*, *116*(1), 45-58.

Kallai, A. Y., & Tzelgov, J. (2012). When meaningful components interrupt the processing of the whole: The case of fractions. *Acta Psychologica*, *139*(2), 358-369.

Keijzer, R., & Terwel, J. (2003). Learning for mathematical insight: a longitudinal comparative study on modelling. *Learning and Instruction*, *13*(3), 285-304.

Landy, D., Brookes, D., & Smout, R. (2014). Abstract numeric relations and the visual structure of algebra. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(5), 1404.

Landy, D. & Goldstone, R. (2007). How Abstract Is Symbolic Thought?. *Journal of Experimental Psychology: Learning, Memory and Cognition, 33,* 720-733. doi:10.1037/0278-7393.33.4.720

Laski, E. V., & Yu, Q. (2014). Number line estimation and mental addition: Examining the potential roles of language and education. *Journal of Experimental Child Psychology, 117,* 29-44. doi:10.1016/j.jecp.2013.08.007

Le Corre, M., & Carey, S. (2007). One, Two, Three, Four, Nothing More: An Investigation of the Conceptual Sources of the Verbal Counting Principles. *Cognition, 105,* 395-438. doi:10.1016/j.cognition.2006.10.005

Lortie-Forgues, H., Tian, J., & Siegler, R.S. (2015). Why is learning fraction and decimal arithmetic so difficult?. *Developmental Review, 38,* 201-221. doi:10.1016/j.dr.2015.07.008

Matthews, P. G., & Chesney, D. L (2015). Fractions as percepts? Exploring cross-format distance effects for fractional magnitudes. *Cognitive Psychology, 78,* 28-56. doi:10.1016/j.cogpsych.2015.01.006

McCrink, K., & Wynn, K. (2007). Ratio Abstraction by 6-Month-Old Infants. *Psychological Science, 18,* 740-745. doi:10.1111/j.1467-9280.2007.01969.x

McMullen, J., Hannula-Sormunen, M. M., & Lehtinen, E. (2014). Spontaneous Focusing on Quantitative Relations in the Development of Children's Fraction Knowledge. *Cognition and Instruction, 32,* 198-218. doi:10.1080/07370008.2014.887085

McMullen, J., Hannula-Sormunen, M. M., Laakkonen, E., & Lehtinen, E. (2016). Spontaneous Focusing on Quantitative Relations as a Predictor of the Development of Rational Number Conceptual Knowledge. *Journal of Educational Psychology, 108,* 857-868. doi:10.1037/edu0000094

Meert, G., Gregoire, J., Noel, M-P. (2010). Comparing the magnitude of two fractions with common components: Which representations are used by 10- and 12-year-olds?. *Journal of Experimental Child Psychology, 107,* 244-259. doi:10.1016/j.jecp.2010.04.008

Miller, J. F., Smith, C. M., Zhu, J., & Zhang, H. (1995). Preschool Origins of Cross-National Differences in Mathematical Competence: The Role of Number-Naming Systems. *Psychological Science, 6,* 56-60.

Miller, K. F., & Stigler, J. W. (1987). Counting in Chinese: Cultural Variation in a Basic Cognitive Skill. *Cognitive Development, 2,* 279-305.

Mix, K. S., & Paik, J. H. (2998). Do Korean Fraction Names Promote Part-Whole Reasoning?. *Journal of Cognition and Development, 9,* 145-170. doi:10.1080/15248370802022605

Moseley, B., & Okamoto, Y. (2008). Identifying Fourth Graders' Understanding of Rational Number Representations: A Mixed Methods Approach. *School Science and Mathematics, 108,* 238-250.

Moss, J., & Case, R. (1999). Developing Children's Understanding of the Rational Numbers: A New Model and an Experimental Curriculum. *Journal for Research* *in Mathematics Education, 30,* 122-147.

Moyer, R. S., & Landauer, T. K. (1967). Time required for Judgements of Numerical Inequality. *Nature, 215,* 1519-1520. doi:10.1038/2151519a0

Moyer, R. S., & Landauer, T. K. (1973). Determinants of reaction time for digit inequality judgments. *Bulletin of the Psychonomic Society,* doi:10.3758/BF03334328.

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

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

Ni, Y., & Zhou, Y-D. (2005). Teaching and Learning Fractions and Rational Numbers: The Origins and Implications of the Whole Number Bias. *Educational Psychologist, 40,* 27-52. doi:10.1207/s15326985ep4001_3

Nuerk, H-C., Wood, G., & Willmes, K. (2005). The Universal SNARC Effect: The Association between Number Magnitude and Space is Amodal. *Experimental Psychology, 52,* 187-194. doi:10.1027/1618-3169.52.3.187

O'Halloran, K. L. (2014). The language of learning mathematics: A multimodal perspective, *The Journal of Mathematics Behavior.*

Odic, D., Le Corre, M., & Halberda, J. (2015). Children's mappings between number words and the approximate number system. *Cognition, 138,* 102-121. doi:10.1016/j.cognition.2015.01.008

Paik, J. H., & Mix, K. S. (2003). U.S. and Korean Children's Comprehension of Fraction Names: A Reexamination of Cross-National Differences. *Child Development, 74,* 144-154.

Panaoura, A., Gagatsis, A., Deliyianni, E., & Elia, I. (2009). The structure of students' beliefs about the use of representations and their performance on the learning of fractions. *Educational Psychology: An International Journal of Experimental Educational psychology, 29,* 713-728. doi:10.1080/01443410903229437

Panaoura, A., Gagatsis, A., Deliyianni, E., & Elia, I. (2010). A model on the cognitive and affective factors for the use of representations at the learning of decimals. *Educational Psychology: An International Journal of Experimental Educational Psychology, 30,* 713-734. doi:10.1080/01443410.2010.501103

Plunkett, K., Hu, J-F., & Cohen, L. B. (2008). Labels can override perceptual categories in early infancy. *Cognition, 106,* 665-681. doi:10.1016/j.cognition.2007.04.003

Rakoczy, H., Clüver, A., Saucke, L., Stoffregen, N., Gräbener, A., Migura, J., & Call, J. (2014). Apes are intuitive statisticians. *Cognition*, *131*(1), 60-68.

Ramani, G. B., & Siegler, R. S. (2008). Playing linear numerical board games promotes low-income children's' numerical development. *Developmental Science, 11*(5), 655-661.

Rapp, M., Bassok, M., DeWolf, M., & Holyoak, K. J. (2015). Modeling Discrete and Continuous Entities with Fractions and Decimals. *Journal of Experimental Psychology: Applied, 21,* 47-56. doi:10.1037/xap0000036

Rau, M. A., & Matthews, P. G. (2017). How to make "more" better? Principles for effective use of multiple representations to enhance students' learning about fractions. *ZDM Mathematics Education*. doi:10.1007/s11858-017-0846-8

Resnick, L. B., Nesher, P., Leonard, F., Magone, M., Omanson, S., Peled, I. (1).

Conceptual Bases of Arithmetic Errors: The Case of Decimal Fractions. *Journal for Research in Mathematics Education, 1989,* 8-27.

Richland, L. E., & Simms, N. (2015). Analogy, higher order thinking, and education. *WIREs Cognitive Science, 6*(2)*,* 1770192. doi:10.1002/wcs.1336

Rugani, R., McCrink, K., de Hevia, M-D., Vallortigara, G., & Regolin, L. (2016). Ratio abstraction over discrete magnitude by newly hatched domestic chicks (Gallus gallus). *Nature Scientific Reports, 6,* 1-8. doi:10.1038/srep30114

Saxe, G. B., Diakow, R., & Gearhart, M. (2013). Towards curricular coherence in integers and fractions: a study of the efficacy of a lesson sequence that uses the number line as the principle representational context. *ZDM Mathematics Education, 45,* 343-364. doi:10.1007/s11858-012-0466-2

Schneider, M., & Siegler, R. S. (2010). Representations of the Magnitudes of Fractions. *Journal of Experimental Psychology: Human Perception and Performance, 36,* 1227-1238.

Sekuler, R., & Mierkiewicz, D. (1977). Children's judgments of numerical inequality. *Child Development*, 630-633.

Shah, P., & Hoeffner, J. (2002). Review of graph comprehension research: Implications for instruction. *Educational Psychology Review*, *14*(1), 47-69.

Siegler, R. S., & Opfer, J. E. (2003). The Development of Numerical Estimation: Evidence for Multiple Representations of Numerical Quantity. *Psychological Science, 14,* 237-243.

Siegler, R. S., & Ramani, G. B. (2009). Playing Linear Number Board Games - but Not Circular Ones - Improves Low-Income Preschoolers' Numerical Understanding.

Journal of Educational Psychology, 101, 545-560. doi:10.1037/a0014239

Siegler, R. S., Duncan, G. J., Davis-Kean, P. E., Duckworth, K., Claessens, A., Engel, M., Susperreguy, M. I., & Chen, M. (2012). Early Predictors of High School Mathematics Achievement. *Psychological Science, 23,* 691 - 697. doi:10.1177/0956797612440101

Siegler, R. S., Fazio, L. K., Bailey, D. H., & Zhou, X. (2013). Fractions: the new frontier for theories of numerical development. *Trends in Cognitive Sciences, 17,* 13-19. doi:10.1016/j.tics.2012.11.004

Sinclair, A., Siegrist, F., & Sinclair, H. (1983). Young Children's Ideas about the Written Number System. *The Acquisition of Symbolic Skills,* 535-542.

Singer-Freeman, K. E., & Goswami, U. (2001). Does half a pizza equal half a box of chocolates? Proportion matching in an analogy task. *Cognitive Development, 16,* 811-829. doi:10.1016/S0885-2014(01)00066-1

Solomon, T. L., Vasilyeva, M., Huttenlocher, J., & Levine, S. C. (2015). Minding the gap: Children’s difficulty conceptualizing spatial intervals as linear measurement units. *Developmental psychology*, *51*(11), 1564.

Sophian, C., Garyantes, D., & Chang, C. (1997). When Three is Less Than Two: Early Development in Children's Understanding of Fractional Quantities. *Developmental Psychology, 33,* 731-744.

Sophian, C., Harley, H., & Martin, C. S. M. (1995). Relational and Representational Aspects of Early Number Development. *Cognition and Instruction, 13,* 253-268. doi:10.1207/s1532690xci1302_4

Sophian, C., Harley, H., & Martin, C. S. M. (1995). Relational and Representational

Aspects of Early Number Development. *Cognition and Instruction, 13,* 253-268. doi:10.1207/s1532690xci1302_4

Spence, I. (1990). Visual psychophysics of simple graphical elements. *Journal of Experimental Psychology: Human Perception and Performance*, *16*(4), 683.

Spinillo, A. G. (2002). Children's use of part-part comparisons to estimate probability. *Journal of Mathematical Behavior, 21,* 357-369. doi:10.1016/S0732-3123(02)00134-7

Spinillo, A. G. (2002). Children's use of part-part comparisons to estimate probability. *Journal of Mathematical Behavior, 21,* 357-369. doi:10.1016/S0732-3123(02)00134-7

Sprute, L., & Temple, E. (2011). Representations of Fractions: Evidence for Accessing the Whole Magnitude in Adults. *Mind, Brain, and Education, 5,* 42-47. doi:

Tecwyn, E. C., Denison, S., Messer, E. J. E., & Buchsbaum, D. (2016). Intuitive probabilistic inference in capuchin monkeys. *Animal Cognition,* doi:10.1007/s10071-016-1043-9

Vamvakoussi, X. (2015). The development of rational number knowledge: Old topic, new insights. *Learning and Instruction, 37,* 50-55. doi:10.1016/j.learninstruc.2015.01.002

Vamvakoussi, X., & Vosniadou, S. (2010). How Many Decimals Are There Between Two Fractions? Aspects of Secondary School Students' Understanding of Rational Numbers and Their Notation. *Cognition and Instruction, 28,* 181-209. doi:10.1080/07370001003676603

Varma, S., & Karl, S. R. (2013). Understanding decimal proportions: Discrete representations, parallel access, and privileged processing of zero. *Cognitive Psychology, 66,* 283-301. doi:10.1016/j.cogpsych.2013.01.002

Vasilyeva, M., Casey, B. M., Dearing, E., & Ganley, C. M. (2009). Measurement skills in low-income elementary school students: Exploring the nature of gender differences. *Cognition and Instruction*, *27*(4), 401-428.

Vukovic, R. K., Geary, D. C., Gersten, R., Fuchs, L. S., Jordan, N. C., & Siegler, R. S. (2014). Sources of Individual Differences in Children's Understanding of Fractions. *Child Development, 85,* 1461-1476. doi:10.1111/cdev.12218

Wagner, J. B., & Johnson, S. C. (2011). As association between understanding cardinality and analog magnitude representations in preschoolers. *Cognition, 119,* 10-22. doi:10.1016/j.cognition.2010.11.014

Wang, Y. & Siegler, R. S. (2013). Representations of and translations between common fractions and decimal fractions. *Psychological and Cognitive Sciences, 58,* 4630-4640. doi:10.1007/s11434-013-6035-4

Wiese, H. (2003). Iconic and non-iconic stages of number development: the role of language. *TRENDS in Cognitive Science, 7,* 385-390. doi:10.1016/S1364-6613(03)00192-X

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). Woodcock-Johnson III Tests of Achievement. Itasca, IL: Riverside Publishing

Ye, A., Resnick, I., Hansen, N., Rodrigues, J., Rinne, L., & Jordan, N. C. (2016). Pathways to fraction learning: Numerical abilities mediate the relation between early cognitive competencies and later fraction knowledge. *Journal of Experimental Child Psychology, 152,* 242-263. doi:10.1016/j.jecp.2016.08.001

Zahner, D., & Corter, J. E. (2010). The Process of Probability Problem Solving: Use of External Visual Representations. *Mathematical Thinking and Learning, 12,* 177-204. doi:10.1080/10986061003654240