The Proportional Reasoning Abilities of Preschool Children

by

Martine Seiden
Class of 2011

A thesis submitted to the faculty of Wesleyan University in partial fulfillment of the requirements for the Degree of Bachelor of Arts with Departmental Honors in Psychology

Middletown, Connecticut April, 2011
Acknowledgements

To...

Hilary Barth, my thesis adviser: Thank you for giving me the opportunity to conduct research this summer and school year, for your advice and guidance, and for all your help putting this project together. It has been an amazing experience.

Andrea Patalano: Thank you for being my second reader, and for inspiring me to enter the field of cognitive science my freshmen year.

My family: Dad, thank you for all of your encouraging words, Mom, I'm happy to say that I am a live donkey, not a dead doctor, and Sonyah, thanks for the late night hellos

Jenn- Thank you for being the best LC and all the cheerios

Anima, Rachel, and Shipra (fine, and Mattie too): Thank you for the entertainment

Emily Slusser: Thank you for all your willing help and advice throughout this process

Tom Castelli: Thank you for your help in designing such great apparatuses

Tali - Thank you for all your help with creating and running BBLB.

Ariana, Emily, Jae, and Qianqian: Thank you for the butterflies

Adele: Thank you for keeping me sane

Thank you to the Hughes Institute for a summer research grant

Finally, A huge 'thank you' to all of the preschools, families, and friends who participated in these studies and made this research possible!
Abstract

Proportional reasoning is important both in mathematics learning and in everyday life. Past studies of children's abilities to reason about proportions have shown both success at young ages (e.g. Acredolo, O'Connor, Banks, Horobin, 1989; Sophian, 2000) but also failure at older ages (e.g. Falk & Wilkening, 1998; Fujimura, 2001; Piaget & Inhelder, 1951/1975). This suggests that some proportional reasoning tasks are easier for children to succeed at then others. The current body of research was reviewed to identify which task variables elicit success versus failure. Two experiments assessed preschool age children's abilities to produce proportion estimates under favorable variables. Experiment 1 tested four- and five-year-olds on an intuitive estimation tasks that involved part-part reasoning. Children produced highly accurate estimates and showed a pattern of bias similar to how adults and older children estimate proportion. Experiment 2 tested four-year-olds and adults on a different implicit estimation task, which required part-whole reasoning. Children again made accurate estimates, but did not display a pattern of bias similar to adults. These results suggest that the current body of work underestimates children's considerable abilities to reason about proportions, and that further research is needed to explore how task variables affect success and how proportional reasoning skills develop.
## Table of Contents

Acknowledgements .............................................. 2  
Abstract ......................................................... 3  
Introduction ..................................................... 5  
Experiment 1: Part-Part Relation Judgments ............... 27  
  Methods .......................................................... 29  
  Results ........................................................... 32  
  Discussion ....................................................... 37  
Experiment 2: Part-Whole Relation Judgments ............. 40  
  Methods .......................................................... 41  
  Results ........................................................... 45  
  Discussion ....................................................... 48  
General Discussion ............................................. 50  
References ......................................................... 57  
Figure Captions ................................................ 64  
Figures ............................................................. 66
General Introduction

The ability to understand and reason about proportions is an important part of cognitive development and essential to both academic achievements and everyday life. A proportion is an equality between two ratios that can be represented symbolically as $a/b = c/d$. Proportional reasoning involves, "detecting, expressing, analyzing, explaining, and providing evidence in support of assertions about, proportional relationships" (Lamon, 2005). Success in proportional reasoning tasks relies on the ability to understand a relationship as a single quantity and then to be able to operate within it (Lamon, 2005). Put more simply, proportional reasoning involves thinking about the relations among relations. Reasoning about proportions has been called the “capstone of elementary school arithmetic,” (National Research Council, 2001) and is central to a variety of advanced academic subjects, such as mathematics, physics, and chemistry. It is also incorporated into many practical activities, such as reading graphs, deviating from a baking recipe, or determining costs after discounts, and is "a pervasive activity that transcends topical barriers in adult life" (Ahl, Moore, & Dixon, 1992).

A large and growing number of studies have explored how and when proportional reasoning skills develop. However, they have provided conflicting evidence. Several studies have shown that children have great difficulty with proportional reasoning until they are older (Chapman, 1975; Falk & Wilkening, 1998; Fujimura, 2001; Hunting & Sharples, 1988; Lunzer & Pumfrey, 1966; Piaget &
Inhelder, 1951/1975; Noelting, 1980; Siegler & Vago, 1987). Piaget and Inhelder (1951/1975) were the first to formally report on children’s proportional reasoning skills. Many of their tasks involved probabilities, which require proportional reasoning because they involve understanding a part (number of successes) compared to a whole (total number of attempts). One specific study involved predicting the color of a marble randomly selected from a known set. Based on the results, the researchers concluded that children had no abilities to make probabilistic choices before age seven, and very limited abilities until after age ten (Piaget & Inhelder, 1951/1975). Other studies have suggested that children are unable to make comparison judgments about mixtures based on ratios. Fujimura (2001) had nine and ten year olds solve proportional reasoning problems about the strength of juice created from mixing concentrate and water. Even after an intervention to, children had great difficulty solving the problems. These studies suggest that proportional reasoning is acquired late in childhood, and is relatively inaccessible to preschoolers and elementary school children.

In stark contrast, many more recent studies have shown that children as young as four or five years old can perform successfully on tasks that require some forms of proportional reasoning (Acredolo, et. al., 1989; Boyer, Levine, & Huttenlocher, 2008; Ginsburg & Rapoport; 1967; Huttenlocher, Newcombe, & Vasilyeva, 1999; Jeong, Levine, & Huttenlocher, 2007; Sophian, 2000; Sophian, Garyantes, & Change, 1997; Sophian & Wood, 1997; Spinillo & Bryant, 1991, 1999; Yost, Siegel, Andrews, 1962). For example, children as young as four were able to select which of two jars would produce a greater likelihood of selecting a winning
color marble when the absolute numbers differed (Yost, et. al., 1962). Young children can also match a mix of water and juice to its proportional equivalence of a smaller aggregate amount of liquid (Jeong et. al., 2008). These studies and others show that children do have some access to proportional reasoning skills as early as preschool years.

Evidence of early developing proportional reasoning abilities is bolstered by relatively new findings of numerical competence in preverbal infants and nonverbal animals. Many studies show that infants and animals have rudimentary systems for understanding numbers, and are sensitive to both magnitudes and to changes in amount (Brannon, 2000; Cordes & Brannon, 2009; Lipton & Spelke, 2004; McCrink & Wynne, 2007; Xu & Spelke, 2000; Xu, Spelke, & Goddard, 2005; Vallentin & Nieder, 2008). Six-month-old infants shown arrays of dots distinguished them from displays of other amounts when the amounts differ by at least a 1:2 ratio, such as 8 to 16 (Xu & Spelke, 2000). Infants and animals also seem able to manipulate these amounts to some extent, and studies have shown that they have a basic sense of addition and subtraction (Boysen & Berntson, 1989; Hauser, MacNeilage, & Ware, 1996; McCrink & Wynn, 2004; Olthof, Iden, & Robert, 1997; Wynn, 1992). By 11 months, infants can discriminate sequences of ascending numerosites from descending ones, suggesting that they can appreciate greater-than and less-than relations (Brannon, 2002). This is a precursor to proportional reasoning because it forms the basis for qualitative comparisons of amounts.

Infants also appear to have basic abilities to encode and compare extents, which is involved in making proportional matches. Infants as young as six months old
can extract and compare sets of two types of items when the ratios of the two types of items differ by a factor of at least 2 (McCrink and Wynn, 2007). It also appears that in at least some situations, infants encode the relations between the extents of the two objects rather than their individual absolute extents (Duffy et. al., 2005). In one revealing study, five-month-old infants were presented with a dowel inside of a clear glass jar. During test trials, infants saw this dowel and jar alternated with either another dowel-jar pairs of the same size ratio but different total size (same ratio condition), or with an identical dowel in a different size container (same extent condition). Infants distinguished between dowel-jar pairs that had different dowel to jar ratios, but not between dowel-jar pairs with the same dowel to jar ratio even though they differed in absolute size (Duffy, et. al., 2005). This implies that infants can extract and compare ratios. A strong interpretation of these results suggests that when comparing sets of items or objects, infants consider proportional matches to be an indication of sameness more so than the absolute size of any individual part of the set.

It appears that there is a relationship between recognizing ratio relations and how preverbal infants and non-verbal animals interpret and form expectation about the world around them. For example, when 6.5-month-old infants saw a hand push a box to the edge of a platform they expected it to fall when only 15% of its surface remained on the platform (an unbelievable outcome), but not when 70% remained (a believable outcome). If a hand continued to grasp the box when it was moved to the edge of the platform so that it was never unsupported, infants did not treat either condition as unbelievable (Baillargeon, et. al., 1992). Infants also exert force
proportional to the predicted weight of an object they are attempting to pick up (Mounoud & Bower, 1974). Foraging behavior by animals implies that they are also sensitive to proportions in their approach to their environment. For example, ducks are able to distribute their foraging time between two locations proportionate to the amount of food available, controlling for both size and frequency of food distribution (Harper, 1982). Even more strikingly, ducks distribute themselves across two locations in a proportion that directly reflects the distribution of food in the two locations, even before they have started to feed (Gallistel, et. al., 1991). Taken together, these suggest that emerging competencies for recognizing and comparing amounts and ratios are an important part of cognitive development. These competencies appear to inform some of our most basic intuitions and interactions with our environment. In sum, there is strong evidence that certain rudimentary understandings of numbers, amounts, and relations between amounts have highly intuitive bases that develop before the cognitive skills to reflect upon or verbalize them.

Although infant and animal research is promising, it does not necessarily prove that infants or animals can reason about proportions, per se. For example, infants clearly cannot perform on proportional reasoning tasks that require complicated computations, physical manipulations, or verbal responses. What the research does evidence is that they have the basic tools needed to solve least some types of proportional reasoning problems. But if these tools emerge so early in life, then why do children have such great difficulty on certain proportional reasoning tasks until late in childhood?
Researchers have proposed several theories that attempt to address this long period of development. Spinillo and Bryant (1999) suggested that there are different types of proportional reasoning that become more accessible as a child develops and is exposed to formal education. Fischbein (1975) differentiated between an intuitive and early-developing understanding about relative frequency, and a later developing capability to conceptualize this understanding in computational terms. Surber and Haines (1987) identified a number of methodological issues that might have affected success in past studies. For example, they report that later researchers criticized many of Piaget and Inhelder's studies for relying heavily on verbal skills and lacking reliable objective measures. More recently, a large number of researchers have individually identified particular characteristics of tasks that affect success (e.g. Boyer, et. al., 2008; Jeong, et. al., 2007; Singer & Resnick, 1992; Sophian & Wood, 1997; Spinillo & Bryant 1991; 1999). Taken together, these investigations point to the existence of one or more elements that affect success on proportional reasoning tasks. The theories described above are not necessarily incompatible or mutually exclusive, and possibly work in tandem. What is currently missing from the research is an integration and review of the variables that affect success.

This thesis aims to provide some structure to the study of proportional reasoning that will help to explain discrepancies across studies and provide a framework for future research. The first goal of this thesis is to review the experimental literature on proportional reasoning comprehensively in order to identify variables common to a variety of proportional reasoning tasks, identifying those that appear to lead to success vs. failure. Identifying favorable variables could
contribute to designing a task involving proportional reasoning at which very young children could succeed. The second goal of this thesis, therefore, is to determine whether favorable task variables might support children’s success at proportional reasoning tasks at much younger ages than were previously thought possible.

**Proportional reasoning task variables**

Past studies on proportional reasoning have used a wide variety of tasks to assess ability. Some examples include juice mixing tasks that compare the strength of juice made from different amount of juice and water (e.g. Boyer, et. al., 2008; Fujimura, 2001; Noelting, 1980), probability tasks that address the likelihood of choosing a target from a set of target and non-target items (e.g. Acredolo et. al, 1989; Falk & Wilkening, 1998) or scaling tasks that require an observer to match a target stimuli to its proportional equivalent (e.g. Huttenlocher, Newcombe, and Vasilyeva, 1999; Sophian, 2000). Across these and other tasks there are at least three important variables that appear to affect performance: *degree of explicitness*, *quantity type*, and *relation type*. *Degree of explicitness* refers to the extent that a task is explicit or implicit: does the task overtly measure proportional reasoning or does it use a more indirect method? *Quantity type* refers to the nature of the variables: are they presented as continuous amounts or in discrete units? *Relation type* refers to the nature of the relationship between the variables being compared: does the task require comparing a part to its whole, or two distinct parts to each other? These variables appear to affect the success rates of both children and adults.

*Degree of explicitness*. Tasks can be more or less explicit in at least three ways: how explicitly they attempt to elicit a response involving proportional
reasoning, whether or not they require knowledge of formal mathematical skills to be solved, or whether or not they require a justification of a response. Tasks can usually be categorized as more or less explicit by asking these questions, but it can be difficult to say with precision how explicit a task is.

Other domains of cognitive development, such as memory, language acquisition, and pattern recognition, have consistently shown that implicit abilities emerge before similar skills to deal with explicit problems, and can sometimes be applied where explicit ones cannot. For example, an implicit memory system appears to emerge long before an explicit one, and can be accessed in many situations where an explicit one cannot (Graf & Schacter, 1985; Kolb & Whishaw, 2003; Naito, 1990; Parkin, 1989).

An earlier ability to perform on implicit tasks is also strongly supported by evidence that preverbal infants and non-verbal animals have certain intuitive skills to deal with proportions. For obvious reasons, infants and animals cannot perform on explicit tasks. Instead, researchers use implicit measures, such as looking time (e.g. Baillargeon, et. al., 1992; Duffy et. al., 2005, McCrink & Wynn, 2007; Xu & Spelke, 2000) or foraging behavior (e.g. Harper, 1982; Mounoud & Bower, 1974) to show that they have certain skills to assess relations among relations. These studies suggest that there is an evolutionary-driven cognitive basis for recognizing, and perhaps even utilizing, proportions and ratios when interacting with the surrounding environment.

Given the above findings, it is unsurprising that proportional reasoning tasks with a high degree of explicitness have generally produced very little success in younger children (Chapman, 1975; Fujimura, 2001; Hunting & Sharpley, 1988;
Inhelder & Piaget, 1951/1975; Noelting, 1980; Siegler & Vago, 1987). Preschool children perform poorly when directly asked to divide amounts into halves, thirds, and quarters (Hunting & Sharpley, 1988). They also show little success when asked to predict the most likely color to be randomly chosen from a known set and justify their response (Chapman, 1975; Piaget & Inhelder, 1951/1975). However, younger children have succeeded at tasks presented in a more intuitive format (Acredolo, et. al., 1989; Boyer, et. al., 2008; Ginsburg & Rapoport, 1967; Sophian, 2000; Sophian et. al., 1997; Sophian & Wood, 1997). Children as young as four years old can select which of two pairs of circles have the same ratio-relations as a larger pair when asked to show which picture is, 'the same' as the larger picture (Sophian, 2000). By 7 years of age, children can predict the probability of how happy a child who wanted a particular color of jellybean would be when given bags of jellybeans with differing proportions of the favored color. It seems that children can perform at much young ages on implicit versions of explicit tasks that have shown failure by older children (Acredolo, et. al., 1989).

In reviewing the literature, Moore, Dixon, and colleagues emerged as the only researchers who have systematically tested intuitive and explicit conditions of the same task (Ahl, et. al., 1992; Dixon, et. al., 1996; Moore, et. al., 1991). They have consistently found that younger children succeed earlier, and older children show a greater rate of success, on intuitive conditions. However, their work may be of limited wider applicability because all of their studies have used a similar task that involves estimating the temperature of water resulting from mixing two containers of waters at different temperatures. In intuitive conditions, the labels use qualitative descriptions
(such as very hot) and an unmarked scale of response, while the explicit conditions used numbers and subdivided scales of response. One disadvantage of this task is that it fails to adequately distinguish between the degree of explicitness and another variable, quantity type (discussed in the next section) since the explicit condition had discrete units but the implicit condition did not. Nevertheless, these studies are a starting point for showing that intuitive versions of a task produce earlier and higher levels of success than equivalent explicit versions.

While it appears all tasks showing success by children under eight years old have been intuitive, not all intuitive tasks confer success at young ages. One relatively intuitive study that showed failure at younger ages used a probability task that that avoided directly asking children to produce probabilities and did not use vernal justifications as a marker of success (Falk & Wilkening, 1998). The task was imbedded in a competitive game in which randomly selecting beads of one color resulted in a prize, but selecting another color did not. In each trial, the child was shown one "full" container that already had its ratio of winning to losing beads set (for example, two blue and two yellow beads respectively) and an incomplete container that always had 6 losing beads (here, 6 yellow beads). The child then had to decide how many beads of the winning color (here, blue) to add to the incomplete container before the child and the experiment were given their container to make a random pick. The child was indirectly encouraged to make the probabilities equal because the experimenter was allowed to choose which container she wanted first, and because at the end of the game the person with the most total yellow beads won a special prize. Even by ten years of age, children only got half of the easiest problems
correct. Failure on some intuitive tasks indicates that a low level of explicitness is necessary but not sufficient for a task to be accessible to younger children.

Why are explicit tasks so much more difficult? One reason could be that they are often too confusing or call upon other skills that are undeveloped. For example, one Piagetian task that proved very difficult for younger children asked subjects to reproduce given geometric figures that were "similar" or "with the same shape" on a different scale and to verbally justify their reproductions. The researchers concluded that the children were unable to perform on this task because they lacked the proper proportional reasoning skills. However, it's possible that instead they did not understand the task as presented, or that they had not yet developed other skills to properly perform on the task, such as fine motor or verbal skills. Other explicit tasks can be confusing because of limited verbal understandings, mathematical jargon or use of mathematical equations that may not yet be familiar to children (e.g. Fujimura, 2001; Hunting & Sharpley, 1988; Piaget & Inhelder, 1951/1957). Unlike more explicit tasks, implicit tasks tend to be simpler and not require precise mathematical reasoning. Their simplicity might prompt children to access intuitive skills for proportional reasoning that appear in infancy (Brannon, 2002; Baillargeon, et.al., 1992; Gallistel, et. al., 1991; McCrink & Wynn, 2007; Mounoud & Bower, 1974; Xu & Spelke, 2000). These intuitive abilities might be hindered by more explicit tasks.

In sum, there is strong evidence that in order for children to show success on a task, it must be presented in a simple implicit format that allow them to access more intuitive skills. However, implicit does not appear to be sufficient to confer success,
which suggests that there are other variables that influence performance on proportional reasoning tasks.

*Quantity type.* This variable refers to whether elements of a task are presented in discrete or continuous amounts. Discrete stimuli are generally presented as sets, such as a bag of ten marbles. Continuous stimuli are inherently uncountable, such as the amount of water in a glass. They are typically portrayed with qualitative descriptions, such as 'this glass is very full' or "this glass has more water than that one." Quantity type typically plays a role in stimuli and can also be a factor of response mechanism for production tasks, such as whether a scale of response is subdivided or not. Tasks can combine these two quantity types, such as by having participants compare discrete to continuous quantities, or by eliciting an answer in a quantity type that differs from the quantity type of the stimuli presented. This distinction has been recognized since the beginning of research on proportional reasoning. Inhelder and Piaget (1958) distinguished between qualitative and quantitative proportions and theorized that skills to deal with the former preceded skills for the later. However, they performed little empirical work to explore this difference.

A higher sensitivity (and perhaps preference) for continuous quantities materializes as early as infancy (Duffy, et. al., 2005; Feigenson et. al., 2002; Mix, et. al. 2002). When comparing items or sets, infants appear to respond to multiple dimensions of continuous extent, but fail to respond to number (Feigenson, et. al., 2002). This difference is also apparent in childhood: children perform more accurately on tasks with continuous quantities than identical tasks with discrete units.
Children as young as six could accurately select which of two spinners had a higher chance of landing on a winning color when the winning and losing color sections were presented continuously. However, children as old as ten had great difficulty when the regions were subdivided. This difficulty persisted even when the subdivided regions were adjacent, implying that distinct units have a strong influence that even older children with some mathematical training have difficulty ignoring. (Jeong, et. al., 2007). Similarly, children as young as six years of age made more accurate proportional matches between targets and choice alternatives that were continuous whole pies, then pies that had been divided into slices (Spinillo & Bryant, 1999).

It should be noted that there is some limited evidence that younger children can perform on tasks involving discrete units. On highly implicit tasks, children as young as seven (Acredolo et. al., 1989) and even four (Yost, et. al., 1962) made proportional reasoning judgments that involved discrete units. In a review, only one study emerged showing younger children achieving success on a task that used both discrete stimuli and a discrete method of response. Ginsburg and Rapoport (1967) tested six- and 11-year-olds on a relatively implicit task that used discrete stimuli (marbles) and a scale to produce estimates made of a line of marbles. Children said the color of individual marbles out loud as they were transferred from one container to another (there were 40 marbles total in each trial). Subjects then used a special apparatus with lines of 40 marbles for each color presented during experiment to show their estimates of how marbles of each color they had seen in a trial. While 11-year-olds did slightly better than six-year-olds, both produced relatively accurate
estimates. If the children tested in this task used proportional reasoning to make their estimates, then continuous quantities might aid success, but are not necessary to elicit successful proportional reasoning from younger children.

However, this task should be cautiously categorized as testing proportional reasoning. Proportional reasoning involves the relations between relations, which would have been tested if children used their estimates of one color to inform their estimates of another color, or made their estimates of each color's total in relation to the total number of marbles. However, children could have also just been estimating the number each color of marble that they saw, without ever relating it to the entire set or to the amount of the other color, a possibility that the authors ignore. If that was true, this task would be better categorized as testing a sense of numerosities and amount, not proportional reasoning. A better indicator of children's proportional reasoning skills would have required the child to compare each marble color's proportion to the whole or the other color's proportion. For example, children could have been asked to replicate the ratios of color by composing a set of the same color marbles on a smaller scale. Therefore, we can only tentatively say that continuity is not a necessary task variable to elicit success in younger children. Further research is needed to determine the complete influence of quantity type on success.

Why might discrete units be more difficult to reason about? One reason, briefly touched upon in the last section, is that discrete tasks may tend to be more explicit, and continuous tasks more implicit. Like implicit tasks, continuous tasks might tap into intuitive skills more easily, while discrete tasks invoke competing conceptualizations. Several researchers have suggested that discrete units lead to
erroneous strategies based on counting target elements, while ignoring the relative amounts of target to non-target elements (Boyer, et. al., 2008; Jeong, et. al, 2007). This might lead children to believe that a jar of 3 blue marbles and 3 yellow marbles is more likely to produce a blue marble in a random pick than a jar of 2 blue marbles and 1 yellow marble. Furthermore, children do worse then when both target and choice stimuli are in discrete units than when just one of them (Jeong, et. al., 2007). This implies that children make mistakes of counting when a numerical match is possible, and mistakenly conceptualize the task as one of numerical equivalence rather than proportions. This raises the possibility that children in Ginsburg and Rapoport's (1967) study might have performed better than expected because the task design made counting difficult. Marbles were only presented individually and working memory was otherwise engaged in saying each marble's color aloud.

In sum, continuous amounts appear to confer success more often than discrete units, but younger children may be able to succeed at some discrete tasks, particularly if they dissuade counting. It seems that continuous amounts are more likely to invoke intuitive abilities, while discrete amounts may prompt children to use strategies that are not conducive to proportional reasoning, such as counting. Further work needs to investigate the links between implicitness and continuous amounts.

*Relations type* This variable refers to how the relationships of entities are expressed. Some tasks involve reasoning about the relations between two distinct entities or sets of entities, which are part-part relations. Other tasks involve comparing a section of an entity or set of entities to the entire entity or set, which are part-whole relations. Part-part and part-whole relations both express a constant
relationship of two quantities, but approach this relationship from different perspectives. A part-part expression conveys the relationship between two separate quantities while a part-whole relation approaches it as being between a portion of a quantity and the entire quantity. Comparing the sizes of two slices of pie would involve part-part relations, while comparing the part of a pie that had been eaten to the entire pie would be a part-whole relation. This difference is reflected in distinct mathematical expressions: part-part relations are generally presented as ratios while part-whole relations are presented as fractions, percentages, or decimals. These relations potentially involve different processes of reasoning that could develop independently or sequentially. One factor that greatly complicates designating the relation-type of a task is that it can be difficult to say with certainty whether a task is being approached as a part-part or a part-whole problem. For example, a bag of ten white and five black marbles could be represented as a part-part relation (10:5 black to white marbles) or as a part-whole relation that compared the number of one color of marbles to the entire bag (5 out of 15 marbles are black).

Singer and Resnick (1992) developed a classification system to assist in identifying whether studies assess part-part relations or part-whole relations. According to their system, problems encourage part-part reasoning if they involve two quantities that are part of a larger common set, and are set up in correspondence to each other. These include problems that present For Every relationships, such as reasoning about the number of oranges in a bowl for every apple in the bowl. Specific types of these situations include Juice Mixture problem that involve mixing concentrate and water (e.g. Jeong, et. al., 2008; Noelting, 1980) or Marbles problems
that involves comparing the proportions of two colors of marbles (e.g. Siegler, 1980; Ginsburg & Rapoport, 1967). Part-whole reasoning is used for problems in which the quantities are defined on the same measure space, so that the part and the whole must be like quantities. This includes Out Of situations, such as the number of apples out of a bowl of fruit. Special types of Out Of problems include Fullness problems that involve comparing different containers filled with different amounts. The authors also propose that proportion tasks that involve rate situations (such as miles per hour) are neither part-part nor part-whole situations since the variables (miles and hours) are not from the same measure space. This type of relationship is not used in any of the studies examined in this paper and is generally absent from the literature. Although these classifications are useful for determining how a task is intended to be represented, they do not necessarily reflect how elements are perceived, understood, or used in formulating a response. For example, an Out Of situation meant to be represented as five apples out of a bowl of ten fruits could instead be understood as a For Every situation (there is one apple for every one non-apple in the bowl).

There is some evidence that older children do better at part-part tasks. Singer and Resnick (1992) tested sixth to eighth graders on a highly explicit task using marbles. Children estimated which of two jars of black and red marbles was more likely to provide a red marble from a random pick. They also provided a verbal explanation for their choice. Some of the problems encouraged part-part reasoning by having the same total number of marbles in both jars (children had to compare the portions of one color across two jars). Other problems lent themselves to part-whole reasoning by varying the total number of marbles in each jar (children had to
represent the proportion of each color to the total of each jar in order to make an accurate response). Based on analyses of performance of different problem types and children's verbal justifications, the authors concluded that older children prefer to represent relations based on parts.

There is a very limited amount of empirical evidence for which strategy develops earlier. What evidence does exist generally comes from observations of behavior and analyses of justifications that children given. For example, Noelting (1980) recorded justifications given on a juice-mixing task that asked children to choose which of two mixtures was more concentrated. He reported that younger children tried to solve problems using part-part relations, by directly comparing how many glasses of water and of concentrate there were across the two choices. In comparison, older children attempted to use part-whole relations to compare the amount of concentrate to the entire amount of liquid. Spinillo and Bryant (1991) proposed that observing how children dealt with a halfway boundary could explain what relation-type they were using. If children can only deal with more-than or less-than relations, which are generally what is required of part-part relations, then comparing two amounts that are both greater than or less than half (e.g. comparing 4/6 and 5/6) should be much more difficult than comparing two amounts in which they are on different sides of the halfway mark (e.g. comparing 1/6 and 5/6). This is because children would only be encoding whether the amount of water was more or less than the amount of no water.

In order to test this theory, Spinillo and Bryant (1991) had four- to seven-year-olds select which of two boxes of blue and white bricks was a proportional match to a
picture of a rectangle divided into blue and white parts. Some trials had both choice alternatives be in the same half as the target (e.g. a target was 1/6 blue and the two choices were 2 blue bricks out of 12 total or 4 blue bricks out of 12 total). On other trials, the choices were in different halves (e.g. a target was 1/6 blue and the choices were 2 blue bricks out of 12 total or 8 blue bricks out of 12 total). A follow-up study tested children on a similar task using only continuous amounts (Spinillo & Bryant, 1999). Both of these studies showed that children as young as six could perform on the trials in which the two choice alternatives were in different halves. In addition, verbal justifications of the younger children almost exclusively referenced part-part relations ("half blue half white") rather than part-whole relations ("half of the box is blue"). These results propose the possibility that the tools for part-part relations are more accessible at younger ages than similar tools for part-whole relations.

Part-part relations may be easier to represent because they require simpler reflection (Schwartz & Moore, 1998). For example, in representing the relationship of 20 white marbles and 20 black marbles, it takes one less step to simplify the relation into one of parts (20 to 20 equals a 1 to 1 relationship) than as a part to a whole (20 and 20 combined is 40, 20 out of 40 equals a 1 out of 2 relationship). Part-part relations become even easier for harder problems, such as a 5:7 ratio. Typically, these problems are represented as part-whole relations by translating them into alternative representations, such as a percent (42%). However, younger children may not possess the skills to perform these transformations. Similarly, Spinillo and Bryant (1991) proposed that part-part relations are easier for young children because they typically involve comparing which of two amounts is larger or smaller, which is
typically a simple and achievable task for younger children, and a skill that appears to be somewhat apparent even in infancy (Brannon, 2002).

At least one study has suggested that some skills to deal with part-whole relations may be apparent in very young children, and that there may not always be a preference for part-part relations (Sophian & Wood, 1997). Five- to seven-year-olds chose which of two cards of continuous or discrete amounts displayed a troll's "fair share" as determined by its proportional match of color to two examples of what a "fair share" for a particular troll was. Some problems had the stimuli divided into equal parts, which made part-part reasoning easier, while other problems divided stimuli into unequal parts, which made part-part reasoning very difficult. The children did not perform at significant rates until seven, and then did so only on the problems that were presented as necessitating part-whole reasoning (Sophian & Wood, 1997). It is not immediately clear why children performed as they did, but it is also interesting to note that the researchers found no differences on performance of the problems using discrete units from those using continuous amounts. The findings of this study show that more research is needed to confirm whether or not part-part relations are easier and develop sooner than part-whole relations. The later may be accessible to very young children under at least some circumstances.

In sum, it appears that relation-type plays a role in conferring success, but more work is needed to confirm the extent of this variable's influence. Part-part relations appear to be easier, perhaps because they generally call on more intuitive skills. Part-part relations may generally involve simpler representations, such as greater-than or less-than, that very young children have the capacity to reason about.
In contrast, part-whole relations may develop later from part-part relations and involve more formally taught reasoning.

_A fourth variable?_

An additional way that proportional reasoning tasks differ is by response mechanism; tasks can require selection or production. Selection tasks require selecting a correct response from one of two or more choice alternatives (e.g. Jeong, et. al., 2008; Sophian, 2000). Production tasks present stimuli and record perceived judgments in response (e.g. Acredolo, et. al., 1989; Ginsburg & Rapoport, 1967). Tasks that require the child to choose the correct response from two or more options have generally provided higher levels of success, even from younger children (e.g. Boyer et. al., 2008; Jeong, et. al, 2007; Sophian, 2000; Sophian & Wood, 1997; Spinillo & Bryant, 1991). Studies that instead have required the production of a response, either as an answer or explanation, are typically much more difficult (e.g. Fujimura, 2001; Hunting & Sharpley, 1988; Piaget & Inhelder, 1951/1975). However, this pattern does not necessarily mean that production tasks are inherently more difficult since very few studies with younger children have used a production mechanism, making direct comparison difficult. In addition, production tasks that have shown failure have generally also been high in explicitness (Fujimura, 2001; Hunting & Sharpley, 1988; Piaget & Inhelder, 1951/1975).

A selection paradigm can be disadvantageous in studying children's proportional reasoning skills for at least three reasons. 1) It fails to show that children can produce judgments using proportional reasoning, which is generally more complex 2) It may suggest that children do not have the skills to integrate information
across dimensions that they actually do possess (Acredolo, et. al. 1989; Anderson, 1980; Wilkening & Anderson, 1982). 3) It doesn't provide a pattern of responses that a production task does. A production task can show the relationship between true and judged proportions across the range of proportions (i.e. from 0% through 100%). This in turn provides a pattern of biases that can help inform what cues were used to form judgments (e.g. areas versus lengths). More research on how young children produce estimates also provides a comparison to estimation patterns of adults and older children, providing an idea of how the relationship between true and judged perceptions of stimuli change over age.

In an extensive review of the literature, only two tasks of proportional estimation with young children emerged. Ginburg and Rapoport (1967) tested children as young as six years old, who observed as 40 marbles were individually from one container to another, and then used a sliding scale to show how many of 40 marbles there had been of each color used (there was a two-color and a three-color condition). However, there are questions regarding to what extent this task tested proportional reasoning abilities. In Acredolo, et. al., (1989) children as young as seven years old used a continuous scale to estimate the happiness of another child who wanted a certain color jelly bean based on sets of jelly beans with different proportions of the favored color. Both of these studies demonstrate that young children can produce estimates on proportional reasoning tasks, and perhaps even tasks with discrete units, which are usually more difficult for children to reason about than continuous amounts (Boyer, et. al., 2008; Jeong, et. al., 2007). This leaves open the possibility that children of even younger ages can succeed at producing estimates.
Current study
An extensive review of the literature revealed that the most favorable conditions that allow for success in proportional reasoning have been a low-level of explicitness, continuous quantities, and part-part relations. When these conditions are applied to a task they should allow for relatively accurate proportional reasoning, even on a production task, at very young ages. The current research attempts to validate this hypothesis, as well as to explore the particular effects of the least studied variable, relation-type. The first experiment tests preschool age children's abilities to produce estimates on a task that uses all three favorable variables. Based on the success of children in Experiment 1, the second experiment attempts to address if preschoolers can also succeed at part-whole relations under other favorable variables.

Experiment 1
The major focus of Experiment 1 was to determine whether very young children could produce accurate estimates on a proportional reasoning task designed under favorable conditions. The task used had a low level of explicitness, continuous stimuli, and assessed reasoning about part-part relations. These variables have individually shown success in past studies with both older and younger children (Ahl, et. al, 1992; Boyer, et. al., 2008; Dixon, et. al., 1996; Jeong, et. al., 2007; Moore, et. al., 1991; Singer & Resnick, 1992; Spinillo & Bryant, 1999). The combination of all three of these favorable variables might allow children to succeed at a younger age than they would on a task with just one or two. Furthermore, favorable variables might allow younger child to actually produce estimates. This is more complex than
recognizing a proportional match, which is what the majority of studies with younger children have done.

A review of the literature revealed only one other study that presented children as young as four years old with a proportional reasoning task that used all three of these conditions, although it did not address using these variables or specifically seek to explore their effects (Sophian, 2000). Four- and five-year-olds were presented with a geometric shape or pair of shapes and had to choose which one of two choice, "was the same," as the target. One of these choice alternatives was proportionally equivalent to the target and the other was not. Even the four-year-olds chose correctly at a significant rate. However, since experiment used a forced choice paradigm it showed only that young children could recognize correct proportional matches, but did not speak to whether or not they could produce them.

To explore whether children could achieve at an estimation task, Experiment 1 tested 4- and 5-year olds on a production task that had several characteristics favorable to success in its design. In order to accommodate favorable variables in a production paradigm, a special apparatus was used that allowed children to manipulate the color of a line on a sliding scale by controlling its proportion of red to blue. Children used the sliding scale to simultaneously represent the size of each of one red and one blue circle.

The sliding scale response mechanism was based on the constant-sum method proposed by Metfessel (1947) and later developed by Comrey (1950). This method measures reasoning about ratio-judgments by having a subject divide an item of set length (e.g. a line) into parts that are proportionally equivalent to two or more stimuli.
This process conveys the believed ratio relationship between the stimuli. For example, dividing a line into a 25% segment and a 75% segment to represent the areas of two stimuli indicates that one stimuli is perceived as three times the size of the other stimuli. This method has been successfully used in explicit studies of proportional reasoning ability with both older children and adults (Hollands & Dyre, 2000; Spence, 1990; Spence & Krizel, 1994). It has also been suggested to be easier for older children than traditional and numeric magnitude estimations (Spence & Krizel, 1994). The apparatus used in Experiment 1 adapted this method for use as an intuitive estimation device for very young children.

Experiment 1 provided an implicit way to measure proportional reasoning ability because it did not explicitly attempt to elicit a response of proportional reasoning, did not call for knowledge of any formal mathematical skills, and did not require justifications of responses. Since neither the circles nor the scale that the child used were segmented, the nature of the design was unambiguously continuous. Accurate performance necessitated the use of part-part proportional reasoning by requiring the child to estimate the ratio of sizes of the two circles in comparison to each other and to then reproduce the ratio on a different scale.

**Method**

**Subjects** Participants were 26 children ranging in age from 4 years, 0 months to 5 years, 10 months \( (M = 4\text{ years }10\text{ months}, SD = 7\text{ months}) \). All participants were recruited from local families in Middletown, Connecticut, or the surrounding area, and were individually tested in a laboratory setting or their own preschool.
Apparatus and stimuli A special apparatus was developed to allow the child to manipulate the color of a rod on a continuous scale (see Figure 1). The apparatus was a red rod (100 cm in length) that was held up at either end by 2 black rectangular pieces of wood, and remained in place throughout the experiment. A blue tube, 50 cm in length, could be slid up and down the rod using a small grey handle. Half of the red rod was obscured by a black box that extended from one end of the rod to the middle of the rod, so that the visible portion of the red rod was 50 cm long. The box had a keyhole opening to allow the blue tube to slide through it. The blue tube was completely occluded when slid inside the box. When it was slid all the way out of the box, it covered the entire visible portion of the red rod. This allowed the proportion of red to blue of a 50 cm section to be precisely manipulated. Above both ends of the visible portion of the rod were card slots to hold laminated cards (the stimuli). The side of the rod facing the experimenter was segmented into 100 equal parts.

The visual stimuli were red and blue circles on individual laminated cards (11 cm x 4 cm). The circles were presented in pairs of one red and one blue circle whose combined areas always totaled 8 in$^2$. For example, a blue circle that was 2.02 in$^2$ was paired with a red circle that was 5.98 in$^2$, representing a relationship that could be expressed as 25% and 75% of the combined area, respectively. Nineteen pairs of circles were used in test trial, whose areas ranged from 5% through 95% of 8 in$^2$ in increments of 5% (see Figure 2). The order of presentation of the circle pairs alternated across children between two randomly made sets.

Procedure During the first part of the experiment, the child was familiarized with the apparatus and task. The child sat directly across from the experimenter with
the visible portion of the red rod centered in front of him. He was encouraged to
familiarize himself with the apparatus by sliding the blue tube in and out of the black
box. The experimenter told the child: "See how you can change the line from all red
to all blue? We're going to play a game where you use this line to show me how big
different size circles are using the line. Can I show you how?"

The experimenter then placed a card of a 4 in² red circle in the slot above the
right side of the rod and a card of a 4 in² blue circle above the slot on the left side.
The experimenter moved the blue tube so that it covered 50% of the visible portion of
the rod, making equal amounts of red rod and blue tube visible. The experimenter
said, "These circles are the same size, so I made the line the same amount of red and
blue. But the bigger the circle is, the more of the line we make that color. And the
smaller the circle is, we make not as much of the line that color. Do you want to try?"

In the second part of the experiment, the child produced estimates of the size
ratios of 19 pairs of circles. The child was presented with the pairs of cards, which
were placed in the slots above the rod. For each pair, he was asked to show how much
blue and how much red should be displayed. The apparatus was alternately reset
between each trial to display the visible portion as completely red or completely blue,
so that the child could not use his previous estimate as a reference. Every few trials,
or if the child expressed confusion, the experimenter reminded him, "The bigger the
circle, the more of the line we make that color and the smaller the circle, the less of
the line we make that color." Each trial was followed by mildly positive
reinforcement, such as, "thank you" or, "good job."
Results

Accuracy of estimates Figure 3 shows the relationship between the true percentages of area and group mean judgments of area. True percentage refers to the percentage of the total area of the circle pair that was contributed by the blue circle. Judged percentage refers to the percentage of the rod that the child covered with the blue tube. Estimates were strongly predicted by actual circle area, $R^2 = .98, p < .0001$. Linear regressions of individual responses revealed a significant relationship between judged and true percentages for all but one participant (age four years 0 months), with all $R^2$ values $>.23, p < .05$. Participants were also analyzed in two separate age groups of four-year-olds and five-year-olds, in order to observe possible effects of age. The 4-year-old age group contained 13 children who ranged in age from 4 years 0 months to 4 years 11 months ($M = 4$ years 4 months, $SD = 3$ months). The 5-year-old group contained 13 children who ranged in age from 5 years 0 months to 5 years 10 months ($M = 5$ years 4 months, $SD = 4$ months). Estimates were predicted by actual circle area for both the four-year-olds ($R^2 = .93, p < .0001$) and the five-year-olds ($R^2 = .96, p < .0001$).

Each child's mean percent absolute error was calculated to give an overall assessment of accuracy of area estimates. Percent absolute error was computed as in Siegler and Booth (2004):

$$\frac{| (\text{Estimate} - \text{Actual Quantity}) / \text{Scale of Estimates}|}{\text{Scale of Estimates}}$$

To illustrate, if a child covered 25% of the rod with the blue dowel to represent a blue circle that was actually 15% of the combined area of the red and blue circle pair presented, the absolute error would be 10%, $\left[ \frac{| 25 - 15 |}{100} \right]$. Individual percent
absolute error means were all below 25% ($M = 15\%, SD = 5\%$). There was no correlation between age in months and mean percent absolute error, $r(24) = -.27$, $p = .18$.

Although the above analyses assume that children produced their estimates using area, children were not explicitly instructed to use area as a comparison tool. It is possible that children compared the circles by diameter instead. This would still require proportional reasoning ability, as children would need to compare the ratio of the two diameter lengths and then translate that ratio to the proportion of red to blue on the rod. In order to determine what cue children were using, analyses were reproduced comparing group estimates to true diameter length instead of true area (see Figure 4). Again, in all analyses the true and judged percentages refer to the blue circle and percentage of the rod covered by the blue tube. Mean estimates were predicted by actual percentage of diameter, $R^2 = .95$, $p < .0001$. This was also true for the Age four group ($R^2 = .94$, $p < .0001$) and the Age five group ($R^2 = .90$, $p < .0001$). This correlation is to be expected since area varies directly as a function of diameter.

One way to infer what cue children used to produce their estimates is to compare their error rates across the two methods. The lower rate of error probably arises from the cue that children used. Mean percent absolute error was calculated for the relationship between each child's estimates and the true percentages based on diameter. Individual children's mean percent absolute errors were all under 30% ($M = 18\%, SD = 6\%$). Individual percent absolute error showed no correlation with age in months, $r(24) = -.06$, $p = .79$. There was a significant effect of cue (area versus diameter), $t(25) = 2.92$, $p < .001$, such that mean percent absolute error was lower
when judged percents were compared to true area than true diameter. This suggests that most children spontaneously used area in their judgments rather than diameter length. Further analyses examining patterns of response and bias assume estimates of area. It should be noted that both methods would produce similar patterns of estimation because of their correlation.

*Pattern of bias* Past studies on proportional reasoning have shown a strong and consistent tendency for both children and adults to overestimate lower numbers and underestimate larger numbers (e.g. Hollands & Dyre, 2000; Hollands, Tanaka, & Dyre, 2002; Spence, 1990; Spence & Krizel; 1994). Hollands and Dyre (2000) reviewed seemingly unrelated studies of proportion judgments, such as visual tasks, durations of time, and gambling chances, and found a similar tendency across tasks for participants to over-estimate proportions smaller than .5 and to under-estimate proportions larger than .5 in a one-cycle bias pattern (see Figure 5A). In a few studies, the opposite one-cycle pattern has been observed, with proportions smaller than .5 being under-estimated and those larger than .5 being over-estimated. Other studies have shown proportion estimates that follow a multiple-cycle pattern in which a one-cycle bias pattern is repeated. For example, several studies of adults and elementary age children have shown a two-cycle bias pattern (see Figure 5B) where proportions between 0 and .25 and between .50 and .75 were overestimated, while proportions between .25 and .5 and between .75 and 1 were underestimated (Hollands & Dyre, 2000; Spence and Krizel, 1994).

Hollands and Dyre (2000) have proposed the idea of a *cyclical power model* that accounts for bias in implicit and explicit proportion judgments, and explains
multiple cycles of bias. The basis of their theory is an extension of Stevens's power law (Stevens, 1957, 1975) and Spence's (1990) power function. Stevens's power law describes the relationship between the actual magnitude of a physical stimulus, $\Pi$, and perception of the stimulus, $\theta$, as a power function,

$$\theta = \alpha \Pi^\beta$$

such that $\alpha$ is a scaling factor and $\beta$ is the power function's exponent that is specific to the stimulus involved. Stevens (1957, 1975) identified $\beta$ for a variety of stimulus continuum by fitting estimates of stimuli that differed along a particular physical dimension, such as length or area, to the power function. When $\beta=1$, the result is a straight line, when $\beta < 1$, progressive increases in the magnitude of the stimulus result in smaller and smaller increases in the perception of the stimulus's magnitude, and when $\beta > 1$, the opposite effect is observed.

Using Stevens' law, Spence (1990) derived a power model to fit proportion judgments. If $\Pi$ and $\Omega$ represent two stimuli being compared, with $\Pi + \Omega$ combining to form a whole (termed unity), then the proportion of $\Pi$ to the total ($\Pi + \Omega$) is represented by $\Pi/(\Pi + \Omega)$. Therefore the estimated proportion, expressed as $P$, would be reduced to:

$$P(\Pi) = \Pi^\beta/[(\Pi^\beta) + (1 - \Pi)^\beta], \text{ if } 0 \leq \Pi \leq 1$$

This power model adequately explains the one-cycle bias patterns found in some studies, but fails to account for multiple-cycle bias patterns (see Hollands & Dyre, 2000, for a detailed explanation).
In order to account for multiple-cycle bias patterns, Hollands and Dyre modified Spence's (1990) power model into the cyclical power model. Instead of \( \Pi \) and \( \Omega \) combining to unity, they used \( T \) to represent a total stimulus magnitude, while \( \Pi \) and \( \Omega \) represent a part and its complement:

\[
P(\Pi) = \frac{\Pi^\beta}{[(\Pi^\beta) + (T + \Pi^\beta)]} \text{ if } 0 \leq \Pi \leq T
\]

While this equation is similar to Spence's power model, it can be applied to multiple cycles of bias by assuming that one or more intermediate reference points can be used to limit the range of estimates, resulting in greater accuracy. For example, if an observer had to estimate the 25% mark on a line, she would have the end points of 0% and 100%, where 0% is represented by 0, and 100% is represented by \( T \). However, intermediate reference points, such as a tick mark at 50% of the line, would result in a more precise judgment by allowing the observer to restrict her response to a smaller range (in this case, the area between 0% and 50% or 50% and 100%, rather than 0% and 100%). These reference points may not need to be explicitly shown.

Evidence that estimating proportions of 1/2 are easier than other proportions (Spinillo & Bryant, 1991) suggests that adults and possibly children may be able to use a halfway point as a reference even if it is not explicitly marked. When an observer uses a reference point, the one-cycle bias pattern repeats between the 0 endpoint and the midpoint and again between the midpoint and the other end point, producing a two-cycle bias pattern. More cycles are possible with additional reference points (Hollands & Dyre, 2000).

The mean results of all children, as well as of the Age four and Age five subgroups were assessed using AICc methods to perform a formal model comparison.
of a straight line, one-cycle proportion model, and two-cycle proportion model. These comparisons took both model fit and complexity into account in determining the strongest model for each set of data. Examining the patterns of bias produced helps to clarify if the children judged proportion in a pattern similar to how adults produce proportion estimates. The results of these comparisons are shown in Figure 3. The bias patterns are also shown for each group. Group means were best explained by a two-cycle proportion model ($R^2 = .98$). The means of just the four-year-olds were explained by a one-cycle model ($R^2 = .93$), while the means of just the five-year-olds were best explained by a two-cycle model ($R^2 = .96$).

**Discussion**

The most notable finding of Experiment 1 is that children as young as four years old can produce accurate judgments of proportion under favorable conditions. This contrasts with numerous past studies that have concluded that children do not possess effective proportional reasoning skills until adolescence, even on selection tasks. For example, even by age ten, children have been shown to mistakenly refer to number rather than proportion when making probability estimates or comparisons (Chapman, 1975; Piaget & Inhelder, 1951/1975). They also have difficulty comparing juice mixtures made with different proportions of water and juice (Fujimura, 2001; Noelting, 1980). To my knowledge, no other task has shown that children this young can produce relatively accurate proportional estimates.

Children's mean percent absolute error for area were relatively low, and the bottom range shows that a few individual children made particularly accurate estimates. It is also noteworthy that most children appeared to produce judgments
based on area rather than diameter. Using area is requires a more complex judgment because it involves assessing and integrating information about two dimensions instead of just one.

Models of proportional reasoning best explained group responses, even by the Age four subgroup, suggesting that children perceived this task as one of proportional reasoning and treated it accordingly. This implies that there are similarities in how very young children perceive and report on intuitive tasks and how adults and older children do so on explicit tasks. The specific proportional model that best explained the group estimates was a two-cycle model. When divided by age, a difference emerged: the group means of four-year-olds were best explained by a one-cycle model, while the group means of five-year-olds were best explained by a two-cycle model. Holland and Dyre (2000) investigated numerous past studies of proportional reasoning and suggested that adults and older children produce a one-cycle pattern of estimation when they do not have access to a midpoint, and a two-cycle pattern when they do. In this case, a subject appears to be making estimates on a scale between the two endpoints. This pattern appears to duplicate itself when there is access to a midpoint. The authors suggested that when a midpoint is recognized it can be used as an additional reference point, so that judgments occur on two smaller scales: one between an endpoint and the midpoint, and other between the midpoint and the second endpoint.

It seems that by age five, but not age four, children are able to recognize and use a midpoint as an additional reference point to the two end points. While both ages produced estimates predicted by models of proportion estimation, the five-year-olds
appeared to access a more complex strategy that makes use of a halfway mark. It is unclear if this change is due to general development in cognitive skills over age, or if it might be related to more experience within an educational setting, since the older children probably had spent longer in preschool and a few had attended Kindergarten. It also appears that it does not greatly affect rates of error, since there was no significant correlation between age in months and mean percent absolute error.

What allowed some children to recognize a midpoint if they were not provided a reference point to use during trials? Although no explicit midpoint was given during experimental trials, children were shown the midpoint mark on the apparatus during the training period along with the corresponding circles. It's possible that children were able to remember and use this as a reference point during the experiment. The five-year-olds might have been better at recognizing the use of a midpoint as well as remembering which circles represented the 50-50 mark. The success of the five-year-olds could also be in part a result of more experience with a "half" concept, particularly in a preschool or kindergarten setting, where children are often encouraged to recognize and understand "half." Ultimately although suggestive, these results are based on relatively small samples and should be confirmed with a larger number of subjects.

Experiment 1 showed that when condition are favorable to success, preschool age children can produce proportional reasoning estimates at accuracy very much unpredicted by past studies. Can these results be repeated on a different proportional reasoning task that elicits estimates on part-whole relations? Experiment 2 explores
whether children can produce accurate part-whole estimates under favorable conditions.

**Experiment 2**

Experiment 2 looked to answer the same question as Experiment 1: can very young children succeed at a proportion judgment task if it is designed with favorable task variables that have tended to produce success in past studies? Experiment 1 showed that children as young as four years old are able to produce estimates about part-part relations, which some researchers have suggested are easier and can be reasoned about earlier than part-whole relations (Schwartz & Moore, 1998; Singer & Resnick, 1992; Spinillo & Bryant, 1991). However, there are at least three reasons to consider that children may be just as capable at part-whole relations. 1) At least one study has shown the opposite finding Sophian and Wood (1997) had children select which of two displays represented the same fair share (i.e proportion) as two sample stimuli. Only by age 7 were children able to select correctly at a significant rate, and then only on problems set up to be solved as part-whole problems. 2) There is little empirical evidence to show that part-part reasoning develops earlier and is easier for very young children; of the limited number of studies available, most focus on older children. It's possible that part-part relations are only easier in certain situations. 3) Because many proportional reasoning situations can be represented using either relation type, it can be difficult to know for certain which type is being used, even if a task is designed to elicit one or the other. Experiment 2 examines success in the context of a task under favorable conditions designed to elicit part-whole proportion judgments.
There were several major changes in the task design of Experiment 2. A different special apparatus was used, that allowed children to indicate a response by placing a magnetic marker along an unmarked rod. The task was embedded in a story about a puppet wanting sweaters that were "just big enough" for teddy bears of different sizes. Visual anchor points of a large bear and an extremely small bear flanked either end of the rod. To achieve success on the task, children needed to match the location on the bar between the two anchor points that was proportionally equivalent to the size of a stimuli bear compared to the anchor points.

This format invited part-whole reasoning because success required that the child be able to consider the relationship between each bear stimuli and the large anchor bear as a part to whole relation in order to accurately place the marker on the bar. The design was implicit because it did not explicitly ask for a proportional reasoning response, did not require formal mathematical knowledge, and correct justifications did not play a part in assessing success. The stimuli and scale of response were again both continuous, since neither the bears nor the rod were segmented. Due to the high rate of success in Experiment 1, Experiment 2 tested mostly four-year-olds with several additional older 3-year-olds and younger five-year-olds. Adults were also tested for use as a comparison group.

Method

Subjects Child participants were 26 children ranging in age from 3 years 6 months to 5 years 4 months ($M = 4$ years 6 months, $SD = 6$ months). All participants were recruited from local families in Middletown, Connecticut, and the surrounding area, and were tested in a laboratory setting or their own preschool. The adults were
14 undergraduates at Wesleyan University ranging in age from 19 to 23 year of age ($M = 21.1$ years, $SD = 0.86$ years), and were tested at Wesleyan University.

*Apparatus and Stimuli* The stimuli and materials consisted of a specially designed apparatus, magnet piece, and laminated picture cards (see Figure 6). The apparatus consisted of a platform with a rectangular rod (51 cm in length) held above the platform by two triangular pieces attached to either end of the rod. The section of the rod that faced the child was covered with an unmarked black magnetic strip. The magnet piece was a small rectangle that the child or experimenter could stick to the magnetic side of the rod in order to indicate a response. A straight line was drawn down the middle of the back of the magnet piece on both sides. This enabled the child to give a precise response and for the experimenter to record the child's exact estimates using a 50 cm measuring tape on the side of the rod facing the experimenter. The ends of the measuring tape were 0.5 cm from the ends of the rod, so that if the magnetic piece was placed at either absolute end of the rod, the corresponding recorded response would be 0 cm or 50 cm.

During experimental trials, the child was presented with visual stimuli of individual animated teddy bears on laminated picture cards ($11$ cm$^2$) with. The bears varied by area, which was determined by pique height and width of the bear. The child was presented with cards of a 40 cm$^2$ bear and a 0.4 cm$^2$ bear during familiarization trials. The card of the 0.4 cm$^2$ bear was attached to the right edge of the rod and the card of the 40 cm$^2$ bear was attached to the left edge of the rod at the start of the experiment. Both cards remained attached for the duration of the experiment. During training, children were presented with new cards of the 40 cm$^2$
bear and 0.4 cm$^2$ bear as well as a 20 cm$^2$ bear. The experimental trials consisted cards of nine different bears varying in area by 10% increments of 40 cm$^2$. The experimenter wore a lion hand puppet throughout the experiment.

Procedure A female experimenter individually tested each child. The child sat directly across from the experimenter with the apparatus in front of the child. During the first part of the experiment the child was introduced to the task and familiarized with the stimuli and apparatus. The experimenter introduced the child to a hand puppet, and said: "Mr. Lion needs your help today. See, Mr. Lion has teddies who are very cold, and he wants to give them each a sweater. This is how you can help." The experimenter gave the child the magnet piece and encouraged him to place the magnet piece at different places on the rod.

The child was shown a picture card with a 40 cm$^2$ bear and asked to identify the picture (all children correctly identified it as a teddy bear). The experimenter confirmed that the picture was of a teddy bear and attached it to the right edge of the rod. The child was then asked to identify the picture card of a 0.4 cm$^2$ bear. If the child correctly identified it as a bear, the experimenter confirmed: "That's right, it's an itty bitty bear." If the child could not identify the picture as a bear, the experimenter encouraged the child to look closely and said: "Actually, it's an itty bitty bear. Can you see it?" The experimenter told the child: "This picture belongs right here," and attached it to the left end of the rod where it remained for the duration of the experiment.

During the second part of the experiment the child was trained on how to complete the task. The experimenter told the child: "Mr. Lion is going to show you
some pictures, and you look at each picture and decide where the magnet goes to match the picture, so that Mr. Lion gets a sweater that's just big enough for each teddy. Let me show you how it works." The experimenter then showed the child new cards of the 40 cm² bear and the 0.4 cm² bear and said, "If Mr. Lion gives me this picture then I move the magnet right here." The experimenter then put the magnet at the edge of the rod with the attached picture of the bear identical to the one shown to the child. The experimenter then said, "That's where it matches the picture so Mr. Lion can get the sweater that's just big enough for the teddy." This was repeated for the second card as well. The order of the two cards was randomly selected for each child.

The child was reminded that the magnet could go anywhere on the rod and the experimenter moved the magnet to each end of the rod while telling the child that the magnet sometimes went next to each of those pictures, but could also go anywhere in-between. The experimenter placed the magnet at three different spots along the rod while telling the child, "It can go here, or here, or here. You just need to look at the picture and then decide." The child was shown a 20 cm² bear and was told that when Mr. Lion showed him that picture he would place the magnet, "right here." The experimenter then put the magnet piece exactly in the middle of the rod. The child was reminded that the placement of the magnet there would allow the teddy to get a sweater that was just big enough. This training also demonstrated to children where a stimuli representing 50% matched a 50% mark on the rod.

In the third part of the experiment the child was tested for understanding. The child was shown the 40 cm² bear and the 0.4 cm² bear in random order and asked
where he would put the magnet to match. If the child placed the magnet within 5% of
the correct endpoint, the experimenter moved on. If the child was incorrect, he was
again shown where the magnet would go to match the 40 cm$^2$ and the 0.4 cm$^2$ bears
and then retested. The majority of children did not need to be retrained. The child was
then reminded that the magnet could go anywhere on the rod and again shown the 20
cm$^2$ bear and its corresponding location on the bar.

The final part of the experiment was the test phase. The child produced
estimates of where the magnet would go on the rod to match different bears so that
each bear could get a sweater that was just big enough. The child was presented with
the magnet marker and a picture card of a bear for each of 9 randomized experimental
trials. The experimenter asked the child, "Where does the magnet go when Mr. Lion
shows you this picture?" The child placed the marker on the magnetic marker for
each trial and the experimenter recorded its position on a reverse 0-50 cm scale.
These responses were later coded on a 0-100 scale for easier analysis. Each trial was
followed by mildly positive reinforcement before removing the magnet and returning
it to the child for the next trial.

Results

Accuracy of Estimates: Unsurprisingly, adults produced highly accurate area
estimates that were strongly predicted by actual area, $R^2 = .98$, $p < .0001$ (See Figure
7). Group estimates for children were highly predicted by true percentages of area, $R^2$
= .94, $p < .0001$. Two children placed at least 9/10 of their estimates within 5% of the
same end point for each trial were eliminated from further analyses (ages 4 years 1
month and 4 years 8 months). The remaining 24 children ranged in age from 3 years 6
months to 5 years 4 months ($M = 4$ years 6 months, $SD = 6$ months). Mean estimates of area for the remaining children are shown in Figure 8A, and were highly predicted by true percentages of area, $R^2 = .96$, $p < .0001$.

Individual mean percent absolute errors were calculated to give an overall impression of accuracy. Mean percent absolute errors were determined by comparing true percentages of area represented by the stimuli presented during the test phase to the judged percentage. The true area was the area of the stimuli bear over $40 \text{ cm}^2$. Judged area was determined by the location of the marker on the rod, with the end that had the $0.4 \text{ cm}^2$ bear attached representing $0\%$, and the end that had the $40 \text{ cm}^2$ bear attached representing $100\%$. Adult mean percent absolute errors were all below $13\%$ ($M = 9\%$, $SD = 2\%$). The mean percent absolute errors for all children were below $37\%$ ($M = 20\%$, $SD = 6\%$). There was no significant correlation between percent absolute error of area estimates and age in months, $r(24) = -.05$, $p = .83$.

One possibility that must be considered is that children did produce judgments based on differences in area, a two-dimensional cue. They might have used a one-dimensional cue instead, such as height or width (which were equivalent). The children's estimates were highly predicted by actual height, $R^2 = .94$, $p < .0001$ (see Figure 8B), which is expected, since area varies as a direct result of height. Individual mean percent absolute errors for children were calculated using true percentages represented by height compared to judged percentages. The true percentage of height was determined by dividing the height of the stimuli by $6.32 \text{ cm}$, which was the height of the large anchor bear. The mean percent absolute errors were all below $45\%$ ($M = 21\%$, $SD = 8\%$). There was no significant correlation between percent absolute
error of area estimates and age in months, \( r(24) = -0.31, p = .14 \). There was no
significant effect of cue used on error, \( r(23) = .1, p < .01 \). It is still unclear which cue
children used when making their estimates, so further analyses have addressed both
possibilities.

**Pattern of bias:** Group means for children and adults were formally compared
using AICc methods for fits to a one-cycle proportion model, two-cycle proportion
model, and a straight line to determine which model provided the best explanation of
the estimates. For the children group mean estimates this is a straight line assuming
both use of area \( (R^2 = .96) \) and height \( (R^2 = .94) \), as shown in Figure 8. Group means
for children assuming area were additionally compared to a power model and
logarithmic model because of the unusual pattern of over-estimation for lower
percentage. A straight line better explained the data than either of these models. The
adult group means were best explained by a two-cycle proportion model, shown in
Figure 7 \( (R^2 = .99) \).

**Combined Analyses**

Experiments 1 and 2 used slightly different age groups in testing. In order to
provide a way to more accurately compare the two studies, this section will compare
performance of the four-year-olds in both Experiments. Thirteen children in
Experiment 1 were four-year-olds \( (M = 4 \text{ years 4 months}, SD = 3 \text{ months}) \). Of the 26
children in Experiment 2, 18 were four years old. However, two of these children
were eliminated from analyses for placing 9/10 of their estimates within 5% of an
endpoint. The remaining 16 children ranged in age from 4 years 1 month to 4 years
11 months \( (M = 4 \text{ years 6 months months}; SD = 3 \text{ months}) \). There was no significant
effect for task on mean percent absolute error scores $t(27) = 1.76, p = .09$. Group
mean estimates of the four-year-olds in Experiment 2 were formally compared for fits
to both cyclical proportion models and a straight line using AICc methods. Their
estimates were best explained by a straight line, $R^2 = .93$. The Age four group of
Experiment 1 had been best explained by a one-cycle proportion model.

**Discussion**

The results of Experiment 2 support the findings of Experiment 1 that even
preschoolers can succeed at producing proportion judgments under favorable
conditions. It extends these findings to show that preschoolers can also achieve
success on a part-whole reasoning task. Children produced relatively accurate linear
estimates with an average of only 11-12% more percent absolute error than adults,
depending on which cue was used.

One question that remains unanswered is what cue the children used to
produce judgments. Estimates compared to both true area and true height produced
patterns that do not have an immediate explanation. Assuming use of area produces a
linear pattern where smaller stimuli were overestimated but larger ones were not.
Assuming height produces a linear pattern where all estimates were underestimated.
One possible explanation for the difference between higher and lower proportions
when area estimates are used is that children were sensitive to ratio-difference
between sizes (for example, the difference from the 10% to 20% bear is a 100%
difference but the difference from the 20% to 30% bear is a 50% difference).
However, if this explanation were true, a logarithmic model, which is based on
exponential growth, should have best explained the pattern. Further testing with
additional points on the lower half of the scale would help to explore this possibility. Another possibility is that children used a combination of area and height for the smaller stimuli and used area for the larger stimuli, which could be explored by doing further statistical analyses and further research with a larger number of subjects and greater number of testing points.

Comparison of child and adult performance shows large differences in error and response pattern. Individual mean percent absolute error varied largely across child participants, while scores for adults were all relatively similar, showing ceiling effects of estimation skills by adulthood. This was not a result of age differences across the child participants, since there was no significant effect of age. Instead, this suggests that by adulthood, individual differences in implicit proportional reasoning abilities diminish.

Unlike adults tested, or the children in Experiment 1, children in Experiment 2 did not produce estimates that were best explained by a proportional model. This suggests that they did not conceptualize the task similarly to adults. This might have been a result of its complexity. Experiment 1 involved comparison of two circles, while Experiment 2 involved comparison of stimuli that were familiar objects to two other versions of the same object. Furthermore, Experiment 2 involved a background story, which Experiment 1 did not. The linear patterns additionally suggest that the children were unable to use the midpoint as an additional reference point, and did a poor job at estimating the 50% bear despite having seen it used in the training trials. However, this is not unexpected since the four-year-olds in Experiment 1 did not use a midpoint in their estimates either.
A direct comparison of the performance of four-year-olds in both Experiments provides tentative evidence that there are not significant differences in error across part-part and part-whole tasks. However, the different models shown by patterns of response suggests that preschoolers may be more accurate at assessing and using reference points on part-part relations. However, these results must be accepted with caution because of the small sample sizes and the numerous differences between the two tasks. What these apparent differences do demonstrate is that further research is needed to explore the differences of performance across these two types proportional reasoning relationships.

**General Discussion**

This thesis had two objectives. The first was to correctly identify variables that lead to success or failure. The three proportional reasoning variables that emerged from an extensive review of the literature were level of explicitness, quantity-type, and relation-type. The conditions of each of these variables that appear to be more favorable to success are a low-level of explicitness, continuity, and part-part relations, respectively. Although these variables have been assessed independently, more investigation and empirical research is needed to explore how they interact, and the extent to which they influence success of younger children.

The second objective was to see if very young children could produce estimates on a proportional reasoning task under favorable conditions. The results of Experiments 1 and 2 provide strong evidence that children as young as four years of age can produce accurate estimates. The results of these experiments contrast with past studies that have asserted that proportional reasoning is a skill that does not
emerge until late in childhood (e.g. Chapman, 1975; Fujimura, 2001; Lunzer & Pumfrey, 1966; Noelting, 1980; Piaget & Inhelder, 1951/1975). The results are particularly noteworthy because children did not just select a proportional match, but produced a series of judgments. In reviewing the literature, no other study emerged of children under age 6 accurately producing proportional reasoning estimates.

The favorable variables that were consistent across Experiments 1 and 2 were that both tasks were highly implicit, calling on intuitive rather than formally taught mathematical skills, and that both used continuous stimuli and unmarked scales of response. Both of these variables have been shown in past studies to result in success by younger children better than their alternatives (Ahl, et. al., 1992; Boyer et. al., 2008; Dixon, et. al., 1996; Jeong, et. al., 2007; Moore et. al., 1991; Spinillo & Bryant, 1999). Children performed significantly better and at earlier ages on selecting a spinner with a higher chance of landing on a certain color when the colors were not divided into subsections that when they were (Jeong et. al., 2007). They are also better at estimating the proportional happiness of selecting a certain color from a known set (Acredolo, 1989) then directly estimating the chances of selecting a certain color from a known set (Piaget & Inhelder, 1951/1975). There also appears to be overlap between these variables, so that quantitative amounts can increase the level of explicitness of a task (Ah. et. al., 1992; Dixon, et. al., 1996; Moore et. al., 1991). Future research is proportional reasoning should explore the mutual effects of these variables.

The main difference between Experiment 1 and 2 was that Experiment 1 assessed skills of reasoning about part-part relations, while Experiment 2 assessed
skills of reasoning about part-whole relations. Part-part relations involve comparing two separate but like entities (in Experiment 1, two circles). Part-whole relations involve comparing a subset of any entity to its entirety (in Experiment 2, the area of a smaller bear compared to a larger bear). However, the two tasks also varied in other ways, such as in the stimuli used and that the task of Experiment 2 was embedded in a story. This made direct comparison of performance on the two tasks difficult. The general success found on both tasks shows that the question over which relation-type is easier or can be assessed earlier is not yet resolved. A next step in this research is to design a task with comparable part-part and part-whole conditions. One way that this could be done would be to present children with red and blue circles and pie-slice pieces and having them represent the relationship of pie slice out of the pie using the red rod and blue tube of the apparatus used in Experiment 1.

The results of these two experiments might appear to validate the theory of some researchers that there are different types of proportional reasoning that develop at different ages (e.g. Spinillo & Bryant, 1999). However, the similarities in patterns of bias shown by adults and children across a wide variety of tasks (Hollands and Dyre, 2000) and also found in Experiment 1, introduce the possibility that children and adults are using similar strategies across different task types. An alternative hypothesis to Spinillo and Bryant (1999) that should be considered, is that children do not fail at certain tasks because they require a different type of proportional reasoning, but instead because they cannot conceptualize certain tasks as ones of proportional reasoning.
It's possible that the certain variables that tend to produce success do so because they more easily allow children to conceptualize the task, intuitively or reflectively, as one of proportional reasoning. In contrast, variables that tend to produce failure do so because they generally result in tasks that are more complicated or elicit erroneous alternative conceptualizations of the task that result in incorrect methods of solving a problem, such as counting (Jeong, et. al., 2007). Other researchers have proposed theories that support this hypothesis. Schwartz and Moore (1998) theorized that children's difficulty with proportional reasoning stems from the complexity of the problems, rather than a lack of conceptual structure. They proposed that mathematical instruction and experiences form mathematical tools allows older children and adults to simplify proportional reasoning problems so that they can be understood within the context of a known simpler conceptualization. In other words, children do not develop new concepts to deal with part-whole relations, but instead learn tools to transform them into simpler problems that can be solved. Resnick and Singer (1993) hypothesized that young children's performance on proportional reasoning tasks was constrained by knowledge about multiplicative relations rather than difficulties with conceptualizing relations of relations. Schwartz and Moore (1998) suggest that. These researchers provide strong theories for specific situations in which situations are reconceptualized rather than new concepts learned. However, none go so far as to say that this applies to the general acquisition of proportional reasoning.

If success in proportional reasoning tasks is a result of correctly conceptualizing the task, then improvement in proportional reasoning over childhood
could be the combination of two steps. First, children are able to conceptualize a wider array of problem types as belonging to the realm of proportional reasoning. Second, in the period of time after children begin correctly conceptualize a task their accuracy increases. The first of these improvements can be measured by assessing patterns of response, the second, by assessing levels of error. The first improvement might be suggested by the difference between the results of Experiments 1 and 2: the differences in the patterns of bias suggests that by age four children are better at treating part-part tasks as proportional reasoning tasks than part-whole tasks. The second improvement would be shown by the difference between error rates in children and adults, who would probably also display a similar pattern of bias but with a much lower rate of error. Currently, additional research to test adults on the same task as Experiment 1 is being planned and will help to clarify this theory.

The first improvement, expanded conceptualization, might be the result of experience and exposure to formal education. For example instruction in mathematics provides children with tools that allow them to deal with explicit proportional reasoning tasks, such as calculating numerical proportions. The second improvement, lower rates of error, is probably also the result of experience and education, as well as cognitive developments. In tasks that require formal computation, children make fewer mistakes with age. In tasks that involve more implicit comparison, such as the tasks used here, age confers increased concentration and reasoning that helps to produce more accurate estimates. For example, using an unmarked reference point in making proportional judgments on a scale increases accuracy, but requires that an observer recognize the potential use of a midpoint, as well estimate its location.
If all proportional reasoning tasks do in fact share the same basic conceptualization, future research should broaden its focus to include more studies and reviews that explore why children are or are not able to correctly conceptualize a task as one of proportional reasoning and what factors in development allow children to increasingly correctly conceptualize tasks of proportional reasoning as such. This theory and potential future research could also greatly benefit the field of education by helping to develop new methods for teaching complex and numerical proportional reasoning skills (such as fractions or ratios) that focus on making connections to existing reasoning skills, rather than teaching them as new concepts.

In summary, there are three main conclusions of these findings. 1) Children as young as four years of age have access to at least some proportional reasoning skills, and can apply those skills under the right circumstances. 2) Preschool children can go beyond recognition of proportional matches, and actually produce accurate estimates. 3) Young children producing estimates on implicit tasks appear to be subject to some of the same biases as adults on explicit tasks. Taken together, these findings suggest that the development of proportional reasoning skills is a complex process that builds on simple and intuitive concepts. It seems that certain task variables may affect how likely children are access these intuitive abilities, and further research should investigate and provide empirical research for what those variables are and how they affect success. In particular, further research is needed to explore the differences between reasoning about part-part and part-whole relations.

These findings also present the possibility that children do not develop new proportional reasoning concepts over time. Instead, it might be the case that as
children grow older and are exposed to formal mathematical education, they can conceptualize more types of tasks as proportional reasoning. This could have broader implications for educational settings, and suggests that instruction in proportions and ratios should focus on building from intuitive and existing concepts rather than trying to present new concepts. By helping children to access their existing capabilities, performance on explicit tasks, such as complex equations, might improve. This work provides a good starting point to investigating the above suggestions, but further study with larger samples is needed to confirm and expand these conclusions. What is clear is that our knowledge of proportional reasoning development is far from complete, and that the full extent of children's abilities has yet to be discovered.
References


Figure Captions

Figure 1. Apparatus and visual stimuli for Experiment 1. The blue circle shown is 2 \text{ in}^2 and the red circle is 6 \text{ in}^2. The rod as shown corresponds to the circle areas and is 25\% blue and 75\% red.

Figure 2. The 19 pairs of circles used in Experiment 1, shown at approximately 1/16 of their original diameters.

Figure 3. Group means and bias for all children, and for four- and five-year-olds separately, as a function of true area. Error bars represent standard errors of the mean. The dotted lines show the \( y=x \) lines. The solid lines show the fits of the model that best explains each group's pattern of estimation. The best fitting model was determined by a formal comparison using AICc. Group and five-year-olds' means were best explained by a two-cycle proportion model; four-year-olds' means were best explained by a one-cycle proportion model.

Figure 4. Comparison of mean estimates of error based on area and diameter length. The solid line shows mean estimates by area and the dashed line shows mean estimates by diameter. The dotted line shows the \( y=x \) line.

Figure 5. A. Predicted pattern of judgments and bias on a task of proportional reasoning when only the endpoints are used as reference points. Proportions under .5 are overestimated and proportions under .5 are underestimated (one-cycle model; Spence, 1990). B. Predicted pattern of judgments and bias on a task of proportional reasoning in which a midpoint is used as an additional reference point. Proportions from 0 to .25 and .5 to .75 are overestimated, and proportions from .25 to .5 and .75 to 1 are underestimated (two-cycle model,
Hollands & Dyre, 2000). For both models, patterns are shown for three values of $\beta$, the exponent of a power function that expresses the relationship between actual and judged magnitude of various stimuli. When $\beta = 1$, $y = 1$.

Figure 6. Apparatus and visual stimuli for Experiment 2. Attached to right side of the bar is a bear of area 40 cm$^2$, and attached to the left side is a bear of area 0.4 cm$^2$. These images were attached for the duration of the task.

Figure 7. Group means as a function of true area for adults in Experiment 2. Error bars represent standard errors of the mean. The solid line shows the model that best explains the pattern of results using AICc methods (two-cycle proportion model) and the dotted line shows the $y=x$ line.

Figure 8. For both graphs, error bars represent standard errors of the mean and the dotted line shows the $y=x$ line. AICc methods were used to compare the fits of various models to the response pattern. A. Group mean estimates as a function of true area for children in Experiment 2. The five models considered to explain the pattern of estimations were a straight line, one-cycle proportion model, two-cycle proportion model, power function, and a logarithmic function. The solid line shows the model that best explains the pattern of response (straight line). B, Group mean estimates as a function of true height for children in Experiment 2. A straight line explained the responses better than either proportion model it was compared to. Its fit is shown by the solid line.
Figure 1

Experiment 1: Apparatus and stimuli
### Experiment 1: Stimuli pairings

![Stimuli pairings diagram](image-url)
Figure 3

Experiment 1: Mean estimates

All Ages (n = 26)

Age 4 Group (n = 13)

Age 5 Group (n = 13)

$T_rue Pecentage$
Figure 4

Experiment 1:
Child group mean estimates by area versus by diameter

All ages (n=26)

Age 4 group (n=13)

Age 5 group (n=13)
Figure 5

A. One-cycle proportion judgment model:
Predicted estimation with no additional reference point

B. Two-cycle proportion judgment model:
Predicted estimation with additional reference point
Experiment 2:
Adult group mean estimates

$n = 14$

$r^2 = .99$
Experiment 2:
Child group mean estimates

$n = 24$

A. As a function of area

B. As a function of height