# The Nature of Representations of Number in Early Childhood: Numerical Comparison as a Case Study

by

Pierina Cheung

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Doctor of Philosophy

in

Psychology

Waterloo, Ontario, Canada, 2015

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

## Abstract

What is the nature of non-verbal representations of number? Broadly speaking, non-verbal representations of number can be divided into two categories: representations of particular numerosities and representations of unspecified numerosities. However, studies of representations of number do not only call for investigations into the representations of numerosities. As philosophers of mathematics (i.e., structuralists) have pointed out, the essence of number lies in numerical relations. Using numerical comparison as a case study, this dissertation asks questions about the nature of the non-verbal representations of particular numerosities and unspecified numerosities.

Research in the last few decades has found evidence for a non-verbal representation of particular numerosities – the approximate number system (ANS). The ANS encodes number as approximate numerical magnitudes, and is a dedicated system for representing number. While there is much evidence that the ANS can be used to represent numerical relations, little is known about whether a separate system for representing small sets of individual objects – parallel individuation – can also be used to represent number. In Chapter 2, I ask whether the parallel individuation system can support numerical comparison. In two experiments, children between the ages of 2 ½ and 4 ½ years old were asked to compare either exclusively small sets (< 4) or exclusively large sets (> 6) on the basis of number. The results of these studies suggest that parallel individuation supports numerical comparison prior to the acquisition of numerical language.

In addition to representations of particular numerosities, humans are also capable of representing unspecified number. For example, we understand that the numerical statement ‘x + 1 > x’ is true without representing the particular value of x. But when does this representation develop? In Chapter 3, I ask at what age children begin to show the capacity to reason about unspecified numerosities. In three experiments, children between the ages of 3 and 6 years were

asked to reason about the effects of numerical and non-numerical transformations on the numerosity of a set. These sets were large enough to be outside the range of parallel individuation and involved comparisons that are not computable by the ANS (Experiment 3), or were hidden so that the specific numerosity was unavailable (Experiments 4 and 5). After the transformation, children were asked whether there were more objects in the set. The results of these studies suggest that the ability to represent unspecified numerosities emerges at around age 4, and is fully in place by ages 5 to 6.

Together, these studies provide evidence that 1) in addition to the ANS, parallel individuation is one of the developmental roots of the representation of numerical relations and 2) the numerical reasoning principles that operate over representations of unspecified numerosities may develop later than computations that operate over representations of particular numerosities.

## Acknowledgements

I am especially grateful to four people who, in their own ways, have shaped who I am today. I am greatly indebted to all of them for what they have done to develop my life as a scientist. First and foremost, I am grateful to have Mathieu Le Corre as my advisor. His exceptionally clear thinking, careful analysis of empirical data, and his eye towards details of argumentation as well as presentation have had a very important influence on this dissertation and my thinking. I only hope that one day my thinking will be as sharp as his. I am also incredibly grateful to Katherine White. One may think that she only acts as my advisor 'on paper', but she has supported me in so many different ways that I don't think she or other people realize. I wish there were another word that could mean "100 times Thank you" (or you don't need another word because we can combine different words…) Thank you for the insightful comments on various drafts of this dissertation. Thank you also for the constant encouragement and support. Finishing this dissertation without her would have been very difficult. I also want to thank Peggy Li, who always asks the right questions that make me think deeply about the language and thought relation. Thank you for being such a wonderful and supportive mentor over the past years. Lastly, I thank David Barner, who introduced me to the notion of individuation and issues about language and human cognition that brought me to graduate school. I would not be interested in number had I not gone to Japan that year. Something magical happened on that trip. And I now perceive the world very differently because of my experience in Ladlab 1.0. The world is full of "individuals".

I would also like to thank my committee members, Stephanie Denison and Ori Friedman, for reading and commenting on my dissertation. Their comments and suggestions have improved the quality of this work.

Thank you to Bridjet Lee, Alex Butti, Elie O'hana, Brent Wettlaufer, and Miranda Sollychin for their help with data collection. Thank you also to all the children, families, daycare centres and schools for participating in the experiments.

Several people have made graduate school more fun: Angela Nyhout, Jesse Langstaff, Meghan Dale, Emanuela Yeung, and Pylin Chuatpatcharasopon. In particular, I thank Angela Nyhout for the continual support and conversations, and Jesse Langstaff for chatting with me about statistics and human decision making (and games). I'm so glad to have spent my time at Waterloo with such intelligent, supportive and humorous people. 2014 has been a particularly difficult year for me on many fronts. I want to thank Katie Wagner and Katie Kimura for being there at one of my lowest points in life. I also want to thank Ladlab 2.0 for the intellectual stimulation, jokes, and parties: KW, KK, Lara, Katharine, Iris, Jess, Dr. Drave (and the expresso machine). Doing research was a lot of fun at UCSD.

Brent Wettlaufer, who has seen the worst of me, deserves some special thanks. Thank you for always being there and for listening to my crazy research ideas. Thank you for making sure that I eat regularly, that I get out of my chair, and that I do not do work in the evening. Nobody has said "You can do it!" more times than you did. And I finally finished this PhD. Thank you also to the Boersen and the Wettlaufer clans for accepting me to their families, and for making life in Southwestern Ontario more fun. Heartfelt thanks to my family in Toronto: Aunt Grace, Aunt Winnie, and Uncle Ming for their never-failing support and for making sure that I eat enough Asian food in Canada.

It was hard leaving HK for Canada 11 years ago. I am grateful to all my wonderful friends back home, especially Low Ti, 瓜敏, 馬, 賴彥融, and Catherine, who have constantly sent me positive vibes, messages, and love despite the long distance.

Finally, I thank my mom, my dad, and my sister for all their love and support. Ever since I was in graduate school, my dad told me every time when we were on the phone: 女仔唔曬讀咁多書 (women don't need too much education). I may not have made him understand the kind of emotional thrill I get from doing scientific research and the excitement from thinking about the structure of the human mind (*and* Chinese History). And although he may not have wanted me to pursue this PhD, he always ended the conversation by saying 你自己鐘意啦 (do what you like to do). I dedicate this dissertation to my parents for their unconditional love and support. I also dedicate this dissertation to 公公 (my maternal grandfather), who was not able to see me enter into graduate school, but I think he would be happy to know that I have decided to embark on this journey.

致爸爸和媽媽

感謝您們多年來的支持跟鼓勵

致公公

以身試教, 讓我體會到終身學習的真正意思

## Table of Contents

List of Figures

List of Tables

# Chapter One: General Introduction

Humans are numerical beings. Whether we are thinking of how many cups of coffee we drink in a day, how much we have saved this month, or how far we are going for a hike, we think about number. But what is a representation of number? A growing body of research has shown that number can be represented with various representational systems (Dehaene, 1997; Fuson, 1988; Gallistel & Gelman, 2000; Feigenson, Dehaene, & Spelke, 2004; Carey 2009; Hyde & Spelke, 2009, 2011; Wynn, 1990, 1992a). These systems can be broadly divided into two categories – representations of particular numerosities and representations of unspecified numerosities. For example, when asked how many cups of coffee we drink in a day, we may be thinking about the number of individual cups of coffee we had today, and this reflects our ability to represent particular numerosities. In contrast, we need not know the value of x to determine that '$x + 1 > x$' is true, and this reflects our ability to represent unspecified numerosities.

However, investigations into the representations of number cannot be limited to explaining how we represent numerosities, but they must also explain how we compute relations between numerosities. Randy Gallistel put the point quite forcefully:

> "If the only sense in which the brain represents number is that there is a sensory/perceptual mapping from numerosity to brain states (the activities of detectors for specific numerosities), which make possible simple numerical discriminations, then the brain's representation of number is a representation in name only. Only if the brain brings combinatorial processes to bear on the neural entities that represent numerosities may we say that the brain represents number in an interesting sense of the term representation." (1989, p. 159).

Therefore, it is crucial to explore the nature of representations of number in the context of numerical relations. The overarching goal of the research reported in this dissertation is to characterize the nature and development of the representations of particular numerosities and unspecified numerosities through studies of numerical comparison. Moreover, given that the goal is to explore the nature and development of representations of number, the present investigation focuses on non-verbal representations because they may be the building blocks for later numerical knowledge.

### ***Non-verbal representations of number***

#### *A non-verbal representation of particular numerosities*

Research on the psychology of number has identified a non-verbal representational system that can be used to represent particular numerosities. This system is powerful enough to represent a broad range of numbers, and is called the approximate number system (ANS). Akin to a 'number line', the ANS represents number as continuous mental magnitudes; the size of 'the line' is proportional to the objective magnitudes of the numerosities represented by the stimuli. ANS representations are approximate and exhibit scalar variability – i.e., error in the representation is proportional to the numerosity of the set. When adults are asked to estimate the number of objects in a set under conditions that suppress counting, the error in their estimates increases as a function of the size of the set of objects (e.g., Whalen, Gallistel, & Gelman, 1999). Another signature of the ANS is that numerical comparisons follow Weber's law – discriminability between sets depends on the ratio between them. For example, adults make more rapid judgments when asked to compare two numerals that are further away from each other (e.g., 5 vs. 9) than two numerals that are closer to each other (e.g., 5 vs. 6; Moyer & Landauer, 1967). Moreover, the larger the numerals, the more difficult it is for adults to compare them if the distance between the two numerals is constant (e.g., comparing 8 vs. 9 is harder than 5 vs. 6 and the distance is 1 in both cases). These behavioural findings demonstrate the ratio signature of Weber's law that characterizes the ANS (e.g., Dehaene, 1997; Whalen, et al., 1999; Cordes, Gelman, Gallistel, & Whalen, 2001).

Research has consistently shown that the ANS is present early in infancy (e.g., Xu & Spelke, 2000; Dehaene, 1997). For example, it has been shown that infants discriminate between numerosities as a function of Weber's law – a signature of the ANS. The minimal ratio required for discrimination has been shown to develop with age. As young as 6 months, infants can discriminate 8 vs. 16 dots (1:2 ratio), but not 8 vs. 12 dots (2:3 ratio); by 9 months, they succeed at a 2:3 ratio discrimination, and by preschool years, at a 3:4 ratio (Feigenson et al., 2004; Halberda & Feigenson, 2008; Xu, Spelke, & Goddard, 2005; Lipton & Spelke, 2003; Xu & Spelke, 2000). Studies with infants have also shown that the ANS specifically encodes number, and not other non-numerical dimensions of sets. For example, as the number of objects increases in a set, the amount of space the objects occupy also increases. Thus, numerical and non-numerical dimensions are often correlated in sets of objects. Researchers have found that when non-numerical variables that co-vary with number, such as total surface area, are controlled for, infants continue to discriminate between sets of objects (e.g., arrays of dots or squares) on the basis of number, and that their discriminations follow Weber's law (e.g., Xu & Spelke, 2000). Moreover, starting in infancy, there is evidence that the ANS represents relations between numerosities such as ordinal relations (Brannon, 2002; Suanda, Tompson, & Brannon, 2008), and addition and subtraction (McCrink & Wynn, 2004; see also Barth, LaMont, et al., 2005, 2006). Together, the evidence suggests that the ANS is a dedicated number system. Not only can it represent the numerosities of sets, but it can also represent numerical relations and operations.

*A non-verbal representation of individual objects*

In addition to the ANS, some have argued that there is another system for representing the numerosities of sets – parallel individuation (e.g., Carey, 2009; Le Corre & Carey, 2007, 2008).

However, unlike the ANS, this system is dedicated to representing individual objects. Specifically, under parallel individuation, each individual object is represented with a distinct mental symbol, and the limit on the system's representational capacity comes from the number of objects it can represent in parallel. This system can only represent up to 3 or 4 objects. One piece of evidence for this system is that for the representation of small numbers ($\leq 4$) does not always show scalar variability. For example, in 1871, when estimating the number of beans being thrown into a box, W. J. Jevons found that he was almost perfect when estimating up to four beans, but his performance worsened as the number of beans increased (Jevons, 1871; see also Trick & Pylyshyn, 1994 for similar evidence). This suggests that the ANS is not always recruited in the small number range. The fast and precise enumeration for small sets ($\leq 4$) indicates the role of parallel individuation – a system for representing individual objects.

Evidence for parallel individuation is also found in infants (e.g., Feigenson & Carey, 2003, 2005; Carey, 2009; Feigenson, Carey, & Hauser, 2002a; Feigenson, Carey, & Spelke, 2002). Using a manual search task, Feigenson and Carey (2003, 2005) showed that 14-month-old infants do not show ratio-dependent performance for small sets. In this task (Feigenson & Carey, 2003, 2005), infants watch as objects are hidden in an opaque box and are allowed to search for them. Infants succeed on this task if they search longer when the box is expected to contain objects, and if they stop searching when the box is expected to be empty. Using this paradigm, Feigenson and Carey found that if infants see two objects hidden in a box, and one of them is retrieved, they continue to reach into the box to look for the second object. In contrast, if they see two objects hidden, and both of them are retrieved, they stop searching. This suggests that they can distinguish 1 vs. 2. However, infants fail to search when they see four objects hidden in the box and two of them are retrieved. Thus, infants succeed at 1 vs. 2 but not 2 vs. 4. Given that the ratio of comparison is 1:2

for both discriminations, this suggests that infants' success on this task does not depend on the ratio of the number of retrieved balls to the total number of hidden balls (see Feigenson et al., 2004 for a review of similar evidence from other paradigms). Rather, it is determined by the total number of hidden balls, indicating that a capacity-limited system – i.e., parallel individuation, is recruited.

On the basis of this finding, some have proposed that, although the primary function of parallel individuation is to track objects, it can also be used to represent the numerosity of sets (Carey, 2009; Le Corre & Carey, 2007, 2008). Specifically, they suggest that symbols for individual objects can enter into computations of one-to-one correspondence. Recall that in the manual search task, infants succeed as long as no more than three objects are hidden in the box (Feigenson & Carey, 2003, 2005). Critically, in contrast to previous studies (e.g., Clearfield & Mix, 1999, 2001; Feigenson et al., 2002a), infants in this task appear to be tracking number, rather than continuous extent, such as the cumulative physical size of objects. For example, when infants retrieve a single large car whose volume is equal to the volume of the two cars hidden in the box, they expect to find another car in the box (Feigenson & Carey, 2003). Feigenson and Carey (2003, 2005; Feigenson, 2005; Carey, 2009) proposed that infants succeed by comparing the number of hidden objects to the number of retrieved objects on the basis of one-to-one correspondence. When the two numbers match (i.e., when there is a one-to-one correspondence between hidden objects and retrieved objects), infants stop looking for objects in the box.

Moreover, researchers have found that symbols for individuals can be "chunked" into sets. Feigenson and Halberda (2004) showed that infants can represent more than three hidden objects if the objects are presented as spatially distinct groups. For example, when infants are presented with two objects on the left side of the box, and two on the right, they represent all four objects as two chunks of two. Feigenson and Halberda suggest that, in this context, infants successfully represent more than three objects by mentally chunking the objects in each spatial group into representations of "sets" (Feigenson & Halberda, 2004; Rosenberg & Feigenson, 2013).

Given infants' ability to compute one-to-one correspondence over representations of individual objects, and their ability to chunk objects into sets, Carey and colleagues have proposed that parallel individuation is powerful enough to represent the numerosity of sets. However, on closer inspection, the evidence is in fact limited. In the manual search task, infants create a working memory model of *one* set of objects, and compare a visible set to a set of objects stored in working memory. Infants are thus not representing two distinct sets. To date, there is no evidence that parallel individuation can support the numerical computation of two sets. A system that can only represent the numerosity of a single set is not a system that is powerful enough to represent numerical relations. This motivates the research reported in Chapter 2. Specifically, I ask whether parallel individuation can be used to compare distinct sets to examine the extent to which parallel individuation is a system for representing the numerosity of sets.

*A non-verbal representation of unspecified numerosities*

As adults, we are not limited to representing particular numerosities and can make generalizations over all possible numbers. For example, we know that when 1 is added to a very large set, the set contains more elements. We need not know how many objects are in this set; we understand that this is a general fact about number that holds for all numbers. This numerical fact can be expressed using a general statement, '$x + 1 > x$'. Therefore, as adults, we can reason about number without representing particular numerosities. I call this a representation of unspecified numerosities because the numerical statement holds true regardless of what numerosity x refers to.

The representation of unspecified numerosities contrasts with that of particular numbers –

e.g., statements such as 2 + 1 > 2 – because knowledge about particular numbers does not generalize to other numbers. Knowing that 2 + 1 > 2 does not grant one the knowledge that ‘for any number x, x + 1 > x’, nor does it entail the knowledge of other specific number facts such as 3 + 1 > 3.

As adults, we can represent unspecified numerosities, but when and how do such representations develop? There has been a long-standing debate with regard to the origin of the representations of unspecified numerosities. On the constructivist view, children do not have a representation of unspecified numerosities until after they understand numerical equality (Piaget, 1952). Drawing on data from his classic study of conservation of number, Piaget concluded that preschoolers fail to recognize that two sets are numerically equal when they are in one-to-one correspondence. In Piaget’s task, children are shown two sets that have the same number of objects, and they are placed in one-to-one correspondence (e.g., two rows of beads, with one bead above another bead). The experimenter asks if both sets have the same number. After the child judges that the two sets are the same, the experimenter transforms one of the sets spatially by making it longer. The child is then asked again about whether the two sets have the same number of objects. Piaget (1952) found that children younger than 6 or 7 years fail to recognize that the spatial cue (e.g., length) is irrelevant to judgments about the numerosity of sets. Note that, to succeed on the conservation of number task, one need not represent the specific numerosities. Success on Piaget’s conservation of number task requires knowledge that if two sets of objects are in one-to-one correspondence, the numerosity of one set of objects, x, is the same as the numerosity of another set of objects, y. This sort of general knowledge necessitates the ability to represent unspecified numerosities, because x and y can take any values. Moreover, it is not necessary to know the specific numerosity to understand that stretching a row of objects does not change its numerosity.

On Piaget's theory, children do not have a representation of unspecified numerosities until after they are at the stage of concrete operations, at around ages 6 to 7 (Piaget, 1952).

On the other side of the debate is the view that the representation of unspecified numerosities is innate. The specific proposals differ with respect to the innate components that constitute the representation of unspecified numerosities. For example, Leslie and colleagues (Leslie, Gallistel, & Gelman, 2007; Leslie, Gelman, & Gallistel, 2008) propose that numerical primitives such as the concept of ONE and the successor function (i.e., for each numeral N, there is a unique successor, N + 1) are the innate building blocks of natural numbers: "there is an innately given recursive rule S(x) = x + ONE" where S refers to the successor function, and ONE is an "innately given symbol with an integer value" (p. 132; Leslie et al., 2007). Thus, given an innate knowledge of the successor function that can take on any particular value, their account assumes an innate representation of unspecified numerosities.

A similar proposal put forth by Rips and colleagues argues that humans have a powerful domain-general innate logic that allows one to derive knowledge that underlies the Peano axioms[1], which define the natural numbers (Rips, Bloomfield, & Asmuth, 2008). Given that a representation of unspecified numerosities is required to represent the Peano axioms, Rips et al.'s account is committed to an innate representation of unspecified numerosities.

Three sets of studies bear on the debate about children's representation of unspecified numerosities. First, as mentioned earlier, Piaget investigated the issue using the conservation of number task that assessed children's understanding of one-to-one correspondence. He found that children younger than 6 or 7 do not seem to understand that when two sets are in one-to-one

[1] The Peano axioms state that (1) 1 is a natural number; (2) 1 is not the successor of any natural number; (3) if n is a natural number, then n + 1 is a natural number; (4) for natural numbers n and m, if n = m, then n + 1 = m + 1; (5) for every natural number n, S(n) is a natural number.

correspondence, they have the same numerosity irrespective of length. Based on this finding, he concluded that children who fail the conservation of number task do not understand that one-to-one correspondence is a logically necessary equivalence that does not depend on perception. Thus, on his view, children who fail the conservation of number task do not have a representation of unspecified numerosities.

Second, a recent study used a different paradigm to examine if children can use one-to-one correspondence to determine numerical equality (Izard, Streri, & Spelke, 2014). Using a manual search task, the researchers provided one-to-one correspondence cues for 2 ½-year-olds and acted out stories to test if they could make use of such cues to make quantity discriminations. Crucially, they used sets (e.g., 5 or 6) that were outside of children's known number range. Thus, to succeed on the task, children have to reason about unknown numerosities using one-to-one correspondence. They found that children fail to use one-to-one correspondence to judge whether a set of retrieved objects matches a set of objects stored in memory.

Finally, researchers have also investigated children's representation of unspecified numerosities using tasks that require children to reason about the meanings of unknown number words (Sarnecka & Gelman, 2004; Condry & Spelke, 2008; Lipton & Spelke, 2006; Brooks, Audet, & Barner, 2013). Results are, however, inconsistent across these studies, and different interpretations have been proposed to explain children's performance (e.g., Brooks et al., 2013). In particular, in some studies, 2- and 3-year-old children appear to understand the conditions under which a change in number word is licensed (Sarnecka & Gelman, 2004; Brooks et al., 2013; Lipton & Spelke, 2006), whereas in some other studies with a slightly different paradigm, researchers have noted failure (Condry & Spelke, 2008; Brooks et al., 2013).

In sum, these three sets of studies suggest that children may not be able to represent unspecified numerosities until age 6. However, they have methodological concerns that may make it difficult to generalize their findings. First, in Piaget's conservation of number task, children's failure may be explained by a perceptual conflict between length and the numerosity of sets. It remains a possibility that children's performance would improve if perceptual conflict were removed. Second, in Izard et al.'s study, young children were required to remember the content of the stories, which were fairly complex. Finally, in studies on children's knowledge of unknown number words, their failure may not reflect a lack of numerical knowledge, but a lack of knowledge of how number words represent number.

Given these methodological concerns, in Chapter 3, I explore when children begin to be able to reason about unspecified numerosities using a paradigm that removes (1) the perceptual conflict that may be present in the Piagetian conservation of number task, (2) the memory demands in Izard et al.'s study, and (3) the need to interpret number word meanings.

In this paradigm, children were shown sets of objects, and observed some types of transformations performed on the set. Then the experimenter asked if there were more objects in the set after the transformation. Investigating children's understanding of numerical transformations is one way to tap their knowledge of numerical relations. It also provides important insights into the development of mathematical thinking because knowledge of numerical transformations (e.g., that the addition of one element affects all numbers in the same way) reflects our ability to think about symbolic operations, and one can combine symbolic operations such as those studied in Chapter 3 to form more complex numerical thought.

***The present research***

The goal of the present research is to investigate the nature of non-verbal representations of number in early childhood. I chose to study non-verbal representations because they may be the fundamental building blocks for later numerical knowledge.

In Chapter 2, I investigate the nature of representations of particular numerosities. While the ANS can represent relations between numerosities starting in infancy, little is known about whether parallel individuation can also be used to represent numerical relations between sets of objects. To test this, in two experiments, I asked preschoolers between the ages of 2 ½ and 4 ½ years to compare either exclusively small numerosities (< 4) or exclusively large numerosities (> 6). Crucially, the ratios of comparisons were equated between small and large numerosities. This allows me to use the accuracy of performance to address the question of whether parallel individuation can support numerical comparison. Moreover, to investigate whether parallel individuation is one of the developmental roots of numerical relations, I examined whether children who had not acquired number word meanings could use parallel individuation to compare numerosities.

In Chapter 3, I investigate the developmental trajectory of the representation of unspecified numerosities. In three experiments, I asked when children between the ages of 3 and 5 years begin to reason that the addition of one element increases the numerosity of a set regardless of set size. Using visible sets and hidden sets of objects, children were asked to reason about the effects of various types of transformations (e.g., addition, subtraction, and rearranging objects) on numerosity.

In the final chapter, I highlight some contributions that the present research makes to the field of numerical cognition, and discuss the implications for verbal representations of number, along with the similarities and differences between the computations of numerical comparison reported in Chapters 2 and 3.

## Chapter Two: The role of non-verbal representations of sets in numerical comparison

According to structuralism, an influential theory of the nature of mathematical objects, the natural numbers are nothing but a system of relations (e.g., Shapiro, 2000). Therefore, accounts of our knowledge of the natural numbers and of its developmental origins cannot be limited to explaining how we represent individual numbers and how we acquire these representations. They must also explain how we represent relations between them and how these representations develop.

One of the possible developmental roots of our representations of relations between numbers is the Approximate Number System (ANS). The ANS represents number approximately, and does not discriminate all pairs of numerosities with equal precision (Dehaene, 1997; Gallistel & Gelman, 2000; Meck & Church, 1983). Rather, the discrimination precision of the ANS is a function of the ratio of numerosities (e.g., Moyer & Landauer, 1967; Lipton & Spelke, 2003; Xu & Spelke, 2000; Xu, Spelke, & Goddard, 2005; Xu & Arriaga, 2007; Coubart, Izard, Spelke, Marie & Streri, 2014). For example, 9-month-olds can discriminate numerosities that differ by a 2:3 ratio (e.g., 6 vs. 9), but not a 3:4 ratio (e.g., 12 vs. 16). Starting in infancy, the ANS also represents relations between numerosities such as ordinal relations (Brannon, 2002; Suanda, Tompson, & Brannon, 2008; see Barth, LaMont, et al., 2005, 2006 for evidence from preschoolers), addition and subtraction (McCrink & Wynn, 2004; see also Barth, LaMont, et al., 2005, 2006), and proportions of discrete quantities (McCrink & Wynn, 2007). Although this is controversial, some studies suggest that later numerical reasoning about numerical relations is at least partially rooted in the ANS, even when the reasoning involves manipulating mathematical symbols (e.g., Halberda, Mazzocco, & Feigenson, 2008; Libertus, Feigenson & Halberda, 2011).

However, multiple studies have shown that, in various contexts, infants and adults often do not use the ANS to represent the numerosity of sets when they include 4 or fewer objects. Instead, they use a distinct system whose capacity is limited to sets of up to 3 or 4 objects (e.g., Feigenson, Carey, & Hauser, 2002a; Feigenson & Carey, 2003, 2005; Feigenson, Dehaene, & Spelke, 2004; Lipton & Spelke, 2004; Revkin, Piazza, Izard, Cohen, & Dehaene, 2008; Trick & Pylyshyn, 1994; Xu, 2003). Most agree that this system represents individual objects and their features and that the limit on its capacity comes from the number of objects it can represent in parallel. It is thus often referred to as "parallel individuation." There is also growing agreement that this system develops early in infancy (e.g., Feigenson, et al., 2002a; Hyde & Spelke, 2011; Coubart et al., 2014).

Some have argued that, in addition to representing objects, parallel individuation provides another developmental root of our representations of numerosities and of the relations between them (Carey, 2009; Feigenson & Carey, 2003, 2005; Le Corre & Carey, 2007, 2008). However, to date there is no evidence that parallel individuation can be used to *compare* the numerosities of *distinct* sets. As made clear by structuralist philosophers of mathematics, a system that can encode the numerosity of individual sets but that cannot compare the numerosities of distinct sets cannot support the acquisition of much numerical knowledge. Therefore, to further investigate the extent to which parallel individuation can be thought of as a developmental root of our numerical knowledge, the experiments in Chapter 2 ask whether children can use this system to compute relative numerosity– i.e., to determine which of two distinct sets contains more objects – prior to going to school, and prior to acquiring numerical language.

There are some reasons to think that parallel individuation is one of the developmental roots of our knowledge of relative numerosity. First, Carey and colleagues (Feigenson & Carey, 2003, 2005; Carey, 2009; see also Uller, Carey, Huntley-Fenner, & Klatt, 1999) have argued that, from infancy on, computations of one-to-one correspondence are defined over the representations of objects created by the parallel individuation system. For example, Feigenson and Carey (2003) suggested that infants keep track of the number of objects hidden in a box by representing each hidden object with a unique symbol that is held in working memory, where each symbol functions as a mental tally mark of sorts. Every time they retrieve an object, they match it to an active tally mark in working memory. When each tally mark has been matched to an object, infants stop reaching. Second, infants can use parallel individuation to represent more than one set at a time. That is, under some conditions, representations of objects can be grouped into two or three "chunks" of up to three objects (Moher, Tuerk, & Feigenson, 2012; Feigenson & Halberda, 2004; see Frank & Barner, 2011 for evidence of object-chunking in school-age children).

Given that parallel individuation can compute one-to-one correspondence within a set, and can represent more than one set at a time, I hypothesize that parallel individuation may also make it possible to determine which of two sets contains more objects. If it can do so prior to the acquisition of numerical language, then it could be one of the developmental roots of our knowledge of relations between numerosities. To my knowledge, no study has directly tested whether these two hypotheses are true. To test these hypotheses directly, a study must show that the basis of the comparison is numerosity and not any other physical property that co-varies with it (e.g., total physical size). It must also show that the comparison is not carried out with the ANS, the other non-verbal system of numerical representation that is available to infants.

To accomplish the latter goal, a study must (1) include numerosities that can be compared with parallel individuation (henceforth, "small numerosities") *and* numerosities that can only be compared with the ANS (henceforth, "large numerosities"); (2) use the same ratios for comparisons of small numerosities and comparisons of large numerosities; and (3) show that performance on comparisons of small numerosities is significantly different from performance on comparisons of large numerosities. Pairs of sets that straddle the boundary between parallel individuation and the ANS must be avoided because the pattern of performance to be expected in such situations is not well known.

Therefore, the right design crucially depends on the upper limit on the capacity of parallel individuation. Unfortunately, it is somewhat unclear. Some studies suggest that it can hold up to 4 objects, but many others suggest that it cannot hold more than 3 (e.g., Starkey & Cooper, 1980; Feigenson, et al., 2002a; Feigenson & Carey, 2005). The bulk of the evidence that suggests it can hold up to 4 objects comes from studies of adult humans or rhesus monkeys (e.g., Luck & Vogel, 1997; Pylyshyn & Storm, 1988; Hauser, Carey, & Hauser, 2000); only one out of many infant studies suggests that it can hold up to 4 objects (Ross-Sheehy, Oakes, & Luck, 2003). Moreover, a study of non-verbal numerical comparisons in human adults has shown that they use parallel individuation when sets are comprised of 3 or fewer objects, and that they rely on the ANS when they are comprised of 4 or more objects (Choo & Franconeri, 2013). Therefore, I suggest that comparisons should not include sets of 4 objects because the way such sets are represented is unclear. In other words, small numerosities should be limited to sets of up to 3 objects and large numerosities should consist of sets of at least 5 objects.

Most studies of non-verbal numerical comparisons fail to meet at least one of the above demands. Some have showed clear evidence for the use of parallel individuation in comparisons of sets but also showed that the comparisons were based on physical attributes of the sets (e.g., their total physical size) rather than on their numerosity (Feigenson, et al., 2002a; Feigenson, et al., 2002b; Clearfield & Mix, 1999, 2001; Xu, et al., 2005). Other studies have showed clear evidence of comparisons based on numerosity but did not include comparisons where both sets contained fewer than 4 objects (Odic, Pietroski, Hunter, Lidz, & Halberda, 2012; Odic, Libertus, Feigenson, & Halberda, 2013). Studies that have included comparisons where both sets contained fewer than 4 objects did not also include comparisons where both sets included at least 5 objects (Barner & Snedeker, 2005, 2006; Brannon & Van de Walle, 2001; Feigenson, 2005; Feigenson et al., 2002a, 2002b). Therefore, none of these studies are able to distinguish whether children recruit parallel individuation or the ANS to compare two small sets of objects.

To my knowledge, only two studies clearly met all of the demands laid out above (Choo and Franconeri, 2013; Wagner & Johnson, 2011).[2] However, neither of them included children who had not yet begun to learn number word meanings. One of the studies was conducted with children who had all learned the meaning of at least some number words (Wagner & Johnson, 2011). The other was conducted with adults (Choo & Franconeri, 2013). Therefore, these studies cannot tell us whether parallel individuation is one of the developmental roots of our knowledge of relative numerosity, in the sense of being available prior to the acquisition of numerical language.

I aimed to fill this gap with a study of non-verbal numerical comparisons that met all the conditions listed above, namely: (1) participants included children who had not learned the meaning of any of the number words that denote the numerosities used in this study; (2) I included pairs of comparisons where both sets were unquestionably within the range of parallel

[2]Two additional studies compared preschoolers' performance on non-verbal numerical comparisons of what they called small numerosities to their performance on comparisons of larger numerosities where all comparisons had the same ratio (Cantlon, Safford, and Brannon, 2010; Rousselle, Palmers and Noel, 2004). Neither of these studies found a difference between performance on small and large comparisons. However, on closer inspection, we noticed that there were problems with each study. Rousselle et al.'s study included three number ranges: small (1 to 4), medium (4 to 8) and large (8 to 16). Both the small and the medium number ranges included comparisons that potentially straddle the boundary between parallel individuation and the ANS (e.g., 3 vs. 4 and 4 vs. 6). Therefore, their results are difficult to interpret. Cantlon et al's study had only one pair of comparisons that actually had the right design features – namely, a pair comprised of comparisons of 1 vs. 2 and 6 vs. 12. Their results are also difficult to interpret. They found that children are better at finding the numerical match for 1 object (out of 1 and 2 objects) than at finding the match for 6 or 12 objects (out of 6 and 12). However, they also found that children are *not* significantly better at finding a match for 2 (out of 1 and 2) than at finding a match for 6 or 12 (out of 6 and 12).

individuation – i.e,. both contained no more than 3 objects – and pairs where both sets were unquestionably within the range of the ANS – i.e., both contained at least 5 objects; and (3) the ratios of the small numerosities were the same as the ratios of the large numerosities. Given that ratios were equated, if children recruit the ANS to compare small and large numerosities, I should observe equal performance. If participants are not equally accurate on the comparisons of small numerosities and on the comparisons of large numerosities, this would suggest that different systems underlie performance for these two number ranges. Given that the ANS is required for large number comparisons, this would suggest that parallel individuation can be used to compare small numerosities.[3] Moreover, if I find that children who have not yet acquired number words for the numerosities used in the study show a difference in comparing small and large numerosities, this would suggest that parallel individuation is one of the developmental roots of our knowledge of numerical relations.

---

[3] There are no strong a-priori reasons to make a directional prediction with regard to children's performance on small numerosity comparisons and large numerosity comparisons. On the one hand, one could predict that children's performance for small numerosity comparison would be *worse* than for large numerosity comparison. This is because in previous studies, infants have been shown to attend to continuous cues for small sets under parallel individuation (Clearfield & Mix, 1999, 2001; Feigenson et al., 2002a), so it is possible that preschoolers may fail to compare two small sets of objects on the basis of number in the current study. Specifically, one could predict that they would demonstrate above chance performance on trials in which continuous cues co-vary with number (i.e., size-number congruent trials, see Methods), but would perform at chance on trials in which continuous cues are in conflict with number (i.e., size-number incongruent trials, see Methods). On the other hand, one could predict that children's performance for small numerosity comparison would be *better* than for large numerosity comparison, because if infants can compute one-to-one correspondence on a single set of objects (Feigenson & Carey, 2003, 2005), and if they can chunk small sets of objects (e.g., Feigenson & Halberda, 2004), then there is reason to suggest that they can also compare two sets of objects on the basis of number. Given that it is unclear from previous research whether children will perform better or worse on small numerosity comparison than large numerosity comparison, a directional prediction was not made.

## Experiment 1 - Comparisons of small vs. large numerosities

### Method

Participants

A total of 99 2½ - to 4½-year-olds participated in this study. Fifty of them were tested on comparisons of small sets only (average age of 3 years 7 months; range: 2 years 7 months – 4 years 7 months; 22 males), and 49 were tested on comparisons of large sets only (average age of 3 years 7 months; range: 2 years 6 months – 4 years 7 months; 30 males). All of the children were recruited in Southwestern Ontario in Canada, and were predominantly monolingual speakers of English.

Design and procedure

Children were tested on a numerical comparison task and on Give-N, a standard assessment of number word knowledge (Wynn 1990). The Give-N task was always administered at the end of the testing session.

*Numerical comparison.* The numerical comparison task always started with an experimenter introducing two puppets to the children - a frog and a duck. Then, the familiarization trial began. Children were shown a picture of a rectangle and asked to name it, e.g., "Do you know what this is?" If children failed to provide a label, the experimenter suggested one (e.g., block, rectangle), and encouraged children to repeat it. After familiarization, the experimenter showed children two pictures of blocks, placed one in front of each of the two puppets, and said, "Froggie has some [blocks], duckie has some [blocks], who has more [blocks]?" This instruction was repeated for each trial.

The numerical comparison task had two between-subject conditions that differed in the range of numbers tested: small numbers (<4) and large numbers (>4). In each condition, sets that differed by a 1:3 and 2:3 ratio were shown. Children in the small numerosity condition were asked to compare 1 vs. 3 (4 trials) and 2 vs. 3 (4 trials). Children in the large numerosity condition were asked to compare 6 vs. 18 (4 trials) and 6 vs. 9 (4 trials). All sets consisted of red rectangles printed on letter-sized paper. The rectangles in each individual set were all of the same physical size. However, the relation between the physical size of individual rectangles, their total perimeter and area, and numerosity varied across trials (see Figures 1a, 1b, 1c and 1d). On half of the trials, the two sets had the same cumulative perimeter and surface area, so that the physical size of the individual rectangles in the sets conflicted with numerosity (i.e., the rectangles in the numerically smaller set were physically larger). Thus, I refer to these trials as "size-number incongruent." (Figures 1b and d). On the other half of the trials, the physical size of the rectangles was the same in both sets and thus did not conflict with numerosity. Consequently, the cumulative perimeter and surface area of the sets was confounded with numerosity (i.e., the set with a greater number of objects also had a larger cumulative perimeter and a larger surface area; Figures 1a and c). I refer to these trials as "size-number congruent."

In the small numerosity condition, for size-number congruent trials, the cumulative surface area of the sets ranged from 1.5 $cm^2$ to 12 $cm^2$, and their cumulative perimeter ranged from 5.32 cm to 30cm. For size-number incongruent trials, the cumulative surface area was always 6 $cm^2$, and cumulative perimeter was 18 cm. In the large numerosity condition, for size-number congruent trials, the cumulative surface area of the sets ranged from 18 $cm^2$ to 72 $cm^2$, and their cumulative perimeter ranged from 31.92 cm to 144 cm. For size-number incongruent trials, cumulative surface area ranged from 18 $cm^2$ to 36 $cm^2$, and cumulative perimeter ranged from 54 cm to 108 cm.

Pairs of sets of rectangles were presented in one of two item orders. For each order, the correct side (left or right), ratio presented (1:3 or 2:3), numerosity pair (2 vs. 3 or 1 vs. 3; 6 vs. 9 or 6 vs. 18), and trial type (size-number congruent or size-number incongruent) was randomized.

In addition, no two consecutive trials were of the same trial type and numerosity pair. No feedback was given. If the child did not respond, the experimenter repeated the question. No prompts were given.

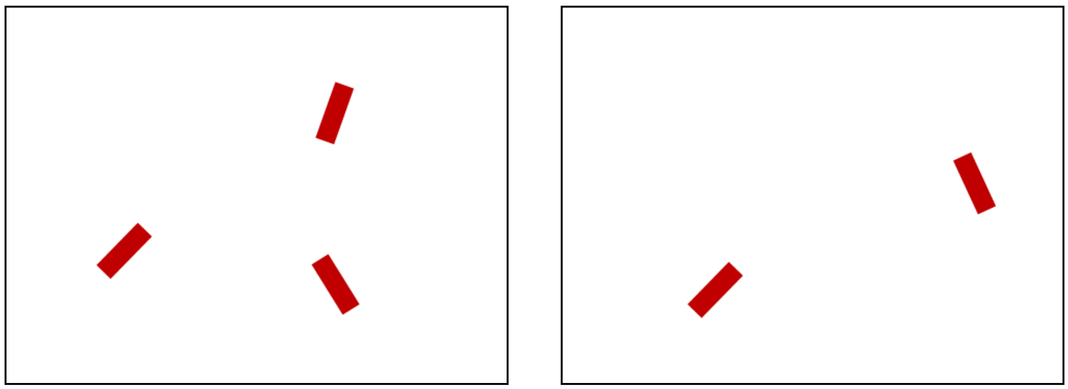

*Figure 1a*. A size-number ***congruent*** trial in the small numerosity condition.

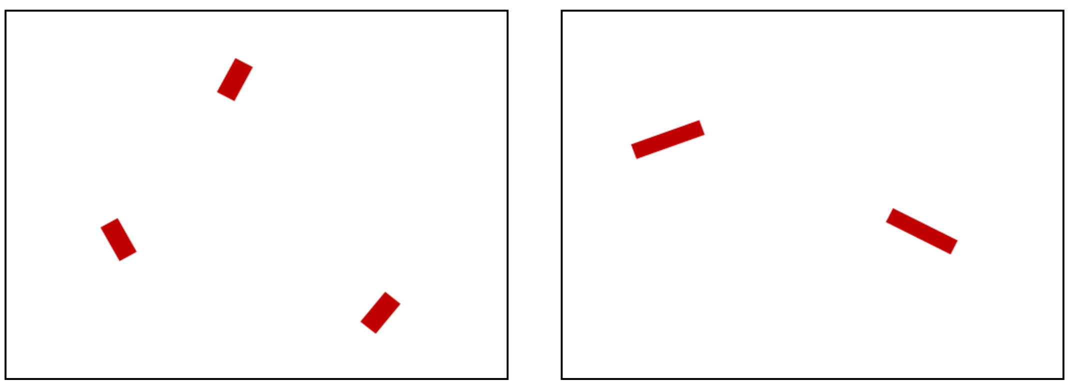

*Figure 1b*. A size-number ***incongruent*** trial in the small numerosity condition.

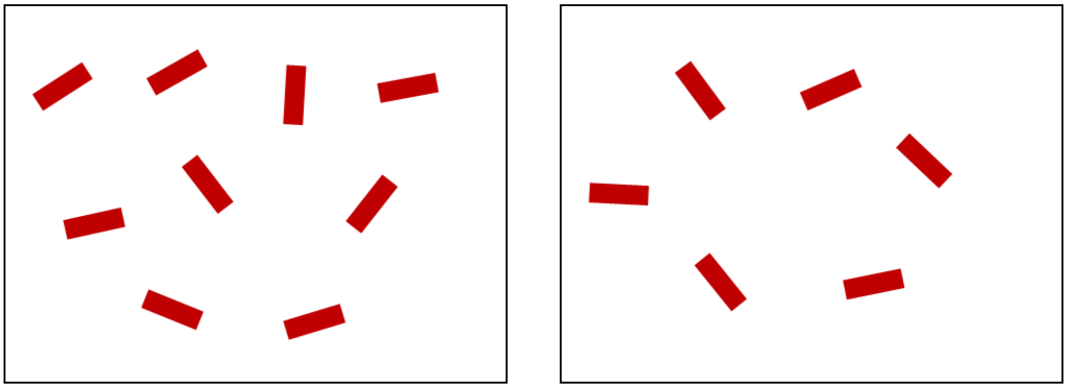

*Figure 1c*. A size-number ***congruent*** trial in the large numerosity condition.

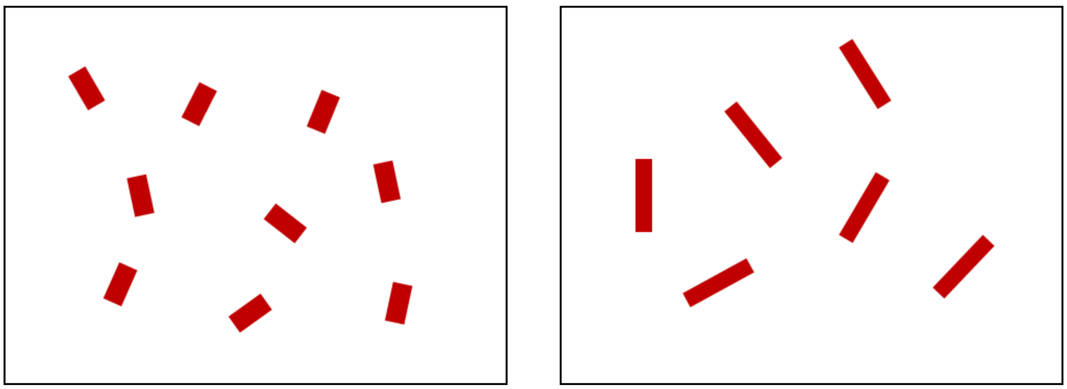

*Figure 1d*. A size-number ***incongruent*** trial in the large numerosity condition.

*Give-N*. The purpose of this task was to assess children's number word knowledge. Children were first introduced to a puppet, a tub of 10 fish, and a plate. They were then told that the puppet wanted to eat some fish, and the experimenter asked the child to give the puppet N fish (e.g., "Can you put one fish on the plate?"). Children were asked to give 1 and then 3 fish on the first two trials. If children succeeded on both, they were asked to give 5 fish. If children failed to correctly give 1 for "one" or 3 for "three", the experimenter asked for two fish. At this point, if children succeeded in response to a request for N, the next request was N+1; if they incorrectly responded to the request for N, the next request was for N-1. The highest numeral requested was "six".

Children were called ‘N-knowers’ (e.g., ‘1-knowers’) if they correctly gave N fish two out of three times when asked for N, but failed to give the correct number two out of three times for N+1. Children who failed to give one fish when asked for “one” were classified as ‘non-knowers’. Children who only knew a subset of the number words – i.e., ‘1-knowers’, ‘2-knowers’, ‘3-knowers’, and 4’-knowers’ – were called ‘subset-knowers’. Children who gave the correct number of fish for all numerals asked for (up to six) were called ‘Cardinal Principle-knowers’, ‘CP-knowers’ for short.

## Results

*Give-N*. The number of children and the mean age in each knower-level group are presented in Table 1. Age was significantly correlated with children’s knower-level, Pearson’s $r = .52$, $p < .001$.

| | **Small** | | **Large** | |
|---|---|---|---|---|
| | **N** | **Mean Age (Range)** | **N** | **Mean Age (Range)** |
| **Non-knowers** | 8 | 3;1 (2;9 - 3;8) | 7 | 3;5 (3;2 - 3;10) |
| **1-knowers** | 12 | 3;5 (2;7 - 4;0) | 10 | 3;0 (2;6 - 3;7) |
| **2-knowers** | 10 | 3;5 (2;8 - 3;11) | 12 | 3;9 (3;2 - 4;7) |
| **3-knowers** | 12 | 3;10 (3;1 - 4;7) | 11 | 3;8 (2;11 - 4;2) |
| **CP-knowers** | 8 | 3;11 (3;2 - 4;7) | 9 | 3;11 (2;11 - 4;7) |

Table 1. Age and number of children in each knower-level group in Experiment 1.

*Numerical Comparisons.* A linear mixed effects model was used to analyze proportion correct on the numerical comparison task. My first analysis found no order or gender effects (all $p$’s > .33), so these variables were not included in subsequent analyses. In the final model, I included Number Range (small vs. large), Size-Number Congruence (congruent vs. incongruent), Ratio (1:3 vs. 2:3), Knower-Level (non-knowers, subset-knowers, vs. CP-knowers) as fixed factors, centered Age as a covariate, and subject as a random factor. I also included all 2-way interaction terms involving Number Range.[4]

There was an effect of Number Range, $\beta = -.76$, $SE = .34$, $p = .015$. Children were better at comparing small numerosities ($M = .79$, $SE = .031$) than large numerosities ($M = .66$, $SE = .032$). There was also a main effect of Size-Number Congruence, $\beta = .095$, $SE = .042$, $p < .001$, with better performance on congruent ($M = .80$, $SE = .025$) than incongruent trials ($M = .65$, $SE = .030$). I also found an effect of Age, $\beta = .13$, $SE = .058$, $p < .001$. In addition, there was a main effect of Knower-Level, $\beta = -.18$, $SE = .10$, $p = .003$. CP-knowers ($N = 17$, $M = .93$, $SE = .039$) were significantly more accurate than both non-knowers ($N = 15$, $M = .60$, $SE = .061$, $t(30) = -4.72$, $p < .001$) and subset-knowers ($N = 67$, $M = .72$, $SE = .030$, $t(82) = -3.36$, p < .001). Non-knowers did not differ from subset-knowers ($t(80) = -1.7$, $p = .093$). I also found a Number Range x Age interaction, $\beta = .17$, $SE = .083$, $p =.040$. No other interactions were found.

I analyzed the Number Range by Age interaction effect by splitting the sample into two age groups using the sample mean: younger preschoolers (2 years 6 months to 3 years 6 months) and older preschoolers (3 years 7 months to 4 years 7 months). Independent samples t-tests revealed that there was an effect of Number Range for younger preschoolers (small: $N = 23$; large: $N = 21$), $t(42) = 2.27$, $p = .028$, but the effect was not significant for older preschoolers (small: $N$

[4] I chose linear mixed models because they allow for interactions between factors and covariates. I first fit a linear mixed model with all main effects and interaction terms and found that adding interaction terms did not improve the fit of the model compared to the model with only main effects. Also, none of the interactions except for Number Range x Age was a significant predictor. Thus, I only included 2-way interaction terms involving Number Range because I was primarily interested in whether small and large comparisons differed.

= 27; large: $N$ = 28), $t(53) = 1.37$, $p = .18$ (see Figure 2).

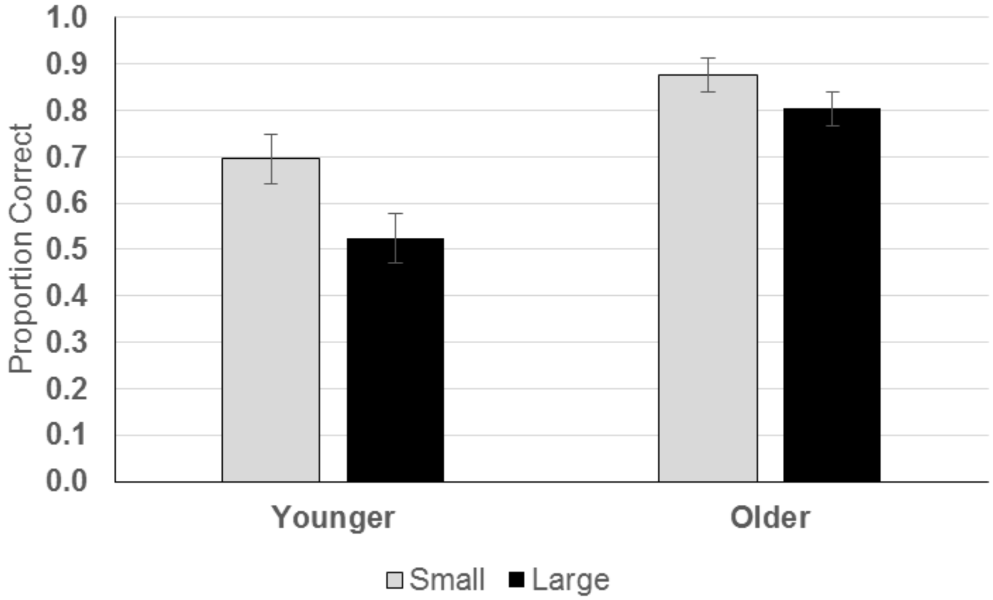


*Figure 2*. Overall proportion correct on small and large numerosity comparisons for younger and older preschoolers in Experiment 1. Error bars represent standard error of the mean.

I also tested whether children performed above chance for small and large numerosity comparisons. Older preschoolers performed above chance on both small, $t(26) = 10.26$, $p < .001$, and large comparisons, $t(27) = 8.17$, $p < .001$. However, younger preschoolers performed significantly above chance for small comparisons, $t(22) = 3.68$, $p < .001$, but not for large comparisons, $t(20) = .44$, $p > .66$.

Given the main effect of Number Range, I asked whether better performance on comparisons of small numerosities was driven by better performance on the 1 vs. 3 comparisons only. Indeed, by physical necessity, whenever one of the choices in a comparison is one object, the other choice is always the right answer. Thus, comparisons where one of the choices is a single object may be easier than all other comparisons. Yet, independent samples t-tests revealed that the effect of Number Range was not specific to comparisons of 1 vs. 3. Instead, the effect of Number Range was significant for comparisons with a 2:3 ratio, $t(97) = 2.05$, $p = .043$ (*Means*: 2 vs. 3 = .81 and 6 vs. 9 = .69), but not for comparisons with a 1:3 ratio, $t(97) = 1.71$, $p = .091$ (*Means*: 1 vs. 3 = .78 and 6 vs. 18 = .68).

The difference in accuracy suggests that children used two different systems - parallel individuation and ANS – for the small and large comparisons. However, it is possible that this accuracy difference arose because children used counting to determine which of two small sets was more numerous. To address this possibility, I analyzed whether subset-knowers – i.e., children who had not acquired the cardinal principle – performed significantly more accurately on small comparisons than large comparisons. Results show that they did, $t(65) = 2.16$, $p = .035$, (Small: $M = .78$, $SE = .043$; Large: $M = .66$, $SE = .040$). I also found that subset-knowers performed significantly better than chance on small comparisons, $t(33) = 6.54$, $p < .001$. The fact that even children who did not understand counting performed well on small comparisons (and better than on large comparisons) suggests that strategies such as counting are not responsible for the difference in performance for small and large comparisons.

Given the main effect of knower level, I also compared children's performance in the small and large numerosity conditions against chance for each knower-level group (see Figure 3). I combined congruent and incongruent trials in this analysis. One-sample t-tests showed that all knower-level groups performed above chance (all $p$'s $< .05$) for small numerosity comparisons except for 1-knowers ($p = .17$). For large number comparisons, it was only after children learned the meaning of 'two' that they demonstrated above chance performance (all $p$'s $<.05$). These results suggest that children are able to compare small sets in the absence of number word knowledge (see more below), but the ability to compare large sets may change with the development of number word learning.

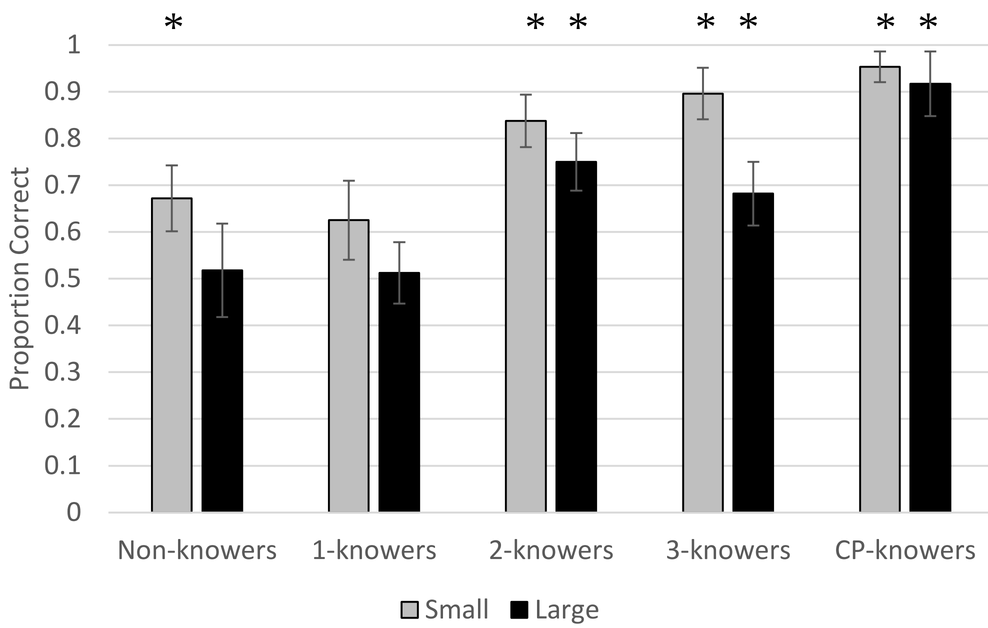


*Figure 3*. Overall proportion correct on small and large numerosity comparisons by knower-level group in Experiment 1. Error bars represent standard error of the mean. * $p < .05$

To further explore whether the ability to use parallel individuation to compare numerosities is available prior to the acquisition of any number word meanings, I analyzed whether non-knowers performed significantly above chance on size-number incongruent trials for small numerosities, and whether their performance for small numerosities was also better than for large numerosities. I combined the 2 vs. 3 and 1 vs. 3 comparisons and found that non-knowers performed significantly above chance, $M = .66$, $SD = .23$, $t(7) = 1.93$, $p = .048$ (1-tailed), d = .68 (90% CI: .0071, 1.31).[5] Moreover, given the possibility that non-knowers' performance might be

[5] One may argue that a two-tailed test is more appropriate and conservative. Under a two-tailed test, non-knowers' performance on small comparisons would almost reach significance, p = .096. However, it is also important to evaluate an effect based on its effect size. The effect size found here - d = .68 (90% CI:

carried by the comparison that involved a set of 1, I removed 1 vs. 3 and only analyzed their performance on 2 vs. 3. I found that non-knowers continued to perform marginally above chance, *M* = .63, *SD* = .35, *W* = 1.63, *p* = .10.[6] I also found that they performed marginally better on comparing small sets (1 vs. 3 and 2 vs. 3) than large sets (6 vs. 18, 6 vs. 9), *M* = .43, *SD* = .24, $t(13) = 1.89$, $p = .082$. These results suggest that parallel individuation can likely support numerical comparisons prior to the acquisition of numerical language.

## Discussion

Two- to four-year-old children were tested on a non-verbal numerical comparison task in one of two conditions: comparing sets of three or fewer objects or sets of six or more objects. The ratios of the numerosities in both the small and large numerosity comparisons were the same. Despite that, children did not perform equally well on all comparisons. Rather, they were more accurate on small than on large comparisons. This difference also held in children who had not learned the meaning of any number word – i.e., non-knowers. Non-knowers performed above chance when they compared small numerosities, and did better on small numerosities than on large ones. Finally, the results also indicate that the effect of Number Range only held in younger children (2;6 to 3;6).

These results suggest that children relied on distinct systems to compare small and large numerosities – i.e., parallel individuation for small ones and the ANS for large ones. Given that the same difference was observed in non-knowers and that non-knowers performed above chance on comparisons of small numerosities even when area and numerosity were not congruent, it may be that some of the operations that can be performed over representations of objects created with

.0071, 1.31) - is of medium size. Future studies should replicate the current study with a larger sample of non-knowers.

[6] A one-sample Wilcoxon signed rank test was performed for this analysis because of the non-parametric data structure for the comparison of 2 vs. 3.

parallel individuation provide one of the developmental roots of representations of relative numerosity.

However, at first sight, the absence of a ratio difference for the large comparisons may appear to suggest that children did not use the ANS to compare large numerosities. In particular, children were not more accurate on the large comparisons with a large ratio (6 vs. 18) than on the large comparisons with a smaller ratio (6 vs. 9), which may raise the concern that my task was not appropriately designed to test the question of interest because it may not have engaged the ANS.

I believe that this concern is unwarranted for two reasons. First, three other studies of non-verbal numerical comparisons in preschoolers with designs similar to the current study also failed to find accuracy differences for ratios similar to Experiment 1 – i.e., 1:2 and 2:3 (Abreu-Mendoza, Soto-Elba, & Arias-Trejo, 2013; Halberda & Feigenson, 2008; Rousselle & Noel, 2008). Importantly, two of these studies provided positive evidence that their task engaged the ANS – i.e., they found that children were better on comparisons with 1:2 and 2:3 ratios than on comparisons with harder-to-discriminate ratios (Abreu-Mendoza et al., 2013; Halberda & Feigenson, 2008). Moreover, children's average accuracy on the large numerosity comparisons in the present study was similar to that which has been reported in previous studies – i.e., near 70% for comparisons with ratios between 2:3 and 1:2 (Halberda & Feigenson, 2008; Rousselle & Noel, 2008). Therefore, our failure to find a difference between performance on comparisons of 6 vs. 9 (a 2:3 ratio) and of 6 vs. 18 (a 1:3 ratio) does not mean that our task did not engage the ANS. Rather, it is consistent with, perhaps even predicted by, what we know about how ratio affects children's performance when they use the ANS to compare numerosities.

Second, there is no other plausible explanation of children's performance on large comparisons in our study. The only other strategies that are available to *adults* are counting, and breaking up the large collections into smaller collections of 2 to 4 elements and adding these up. Neither of these strategies can explain why I find the same pattern of results when I restrict my analyses to subset-knowers. Subset-knowers do not know the cardinal principle. Therefore, they cannot have compared the collections by counting them. Moreover, it is highly unlikely that children who do not know the cardinal principle nonetheless know the sums required to determine the numerosity of collections of 6, 9 or 18 elements by breaking them up into small collections of 2 to 4. Therefore, it is unreasonable to assume that subset-knowers used this strategy. Then how did subset-knowers compare the large numerosities? I believe that, short of postulating new representational systems for which no evidence has been provided, the only explanation left is that they did so with the ANS.

Another possible concern is that children controlled how long the collections were presented to them, so that, consequently, they could have used counting instead of parallel individuation and/or the ANS to compare numerosities. Since small collections are easier to count than large ones, this could explain why children were more accurate on small comparisons than on large ones. However, as I argued above, the fact that subset-knowers showed this pattern of results rules out this alternative. There remain two alternatives that cannot be ruled out directly by the results of Experiment 1. First, it could be that children were, in fact, relying on the ANS for all comparisons but that they performed more poorly on large comparisons because these comparisons make greater demands on non-numerical aspects of processing. For example, larger sets necessarily require children to divide their attention over more objects than small sets, and this could cause a decrement in performance. Although possible in theory, this “processing demands” alternative is unlikely to be the right explanation. First, on this alternative, the difficulty of comparisons should increase continuously as a function of the absolute value of numerosities.

Contrary to this prediction, previous research provides strong evidence that when infants (Wood & Spelke, 2005) or adults (Barth, Kanwisher, & Spelke, 2003) use the ANS to compare numerosities, their level of performance for a given ratio of numerosities is the same regardless of the absolute value of the numerosities. Second, it may be that children performed better on comparisons of small numerosities because the correct answer in this condition was always the same – i.e., 3 – whereas the correct answer in the large numerosity condition alternated between 9 and 18. Both of these alternatives were tested in Experiment 2.

## Experiment 2 - Comparisons of small vs. large numerosities with a wider range of set sizes

Experiment 2 had two main goals. First, I sought to replicate the effect of Number Range in Experiment 1. Second, I aimed to address the two alternative explanations raised in Experiment 1. As in Experiment 1, one group of children compared small numerosities only and another group compared large numerosities only. All comparisons differed by a 2:3 ratio.

To address the processing demands alternative the large comparisons included a wider range of numerosities: in addition to 6 vs. 9, I also included 10 vs. 15 and 12 vs. 18. Evidence that children's accuracy decreases as numerosity increases would support this alternative. On the other hand, evidence that (1) children perform better when they compare small numerosities than when they compare large numerosities at the same ratio, *and* that (2) they perform equally on all comparisons of large numerosities would provide strong evidence against the processing demands alternative.

Experiment 2 also tested whether the children in Experiment 1 who compared small numerosities were more accurate than those who compared large ones because the correct answer was always the same numerosity in the former condition but not in the latter. To test this, the small numerosity condition of Experiment 2 included comparisons of 1 vs. 2 and 3 vs. 4 in addition to comparisons of 2 vs. 3. Children's performance on 3 vs. 4 will also shed light on whether a set of 4 objects can be represented with parallel individuation.

## Method

### Participants

A total of 86 2 ½ - to 4 ½-year-olds participated in this study. There were 45 children in the small numerosity condition, with an average age of 3 years 5 months (range: 2 years 6 months – 4 years 5 months; 24 males), and 41 children in the large numerosity condition, with an average age of 3 years 5 months (range: 2 years 6 months – 4 years 9 months; 21 males). A majority of the children were recruited in Southwestern Ontario in Canada (*N* = 63; 24 in the small numerosity condition), and the remaining children were recruited in the Greater Boston Area in the US (*N* = 23; 21 in the small numerosity condition). Participants were predominantly monolingual speakers of English.

### Design and procedure

The stimuli and procedure were similar to Experiment 1. All except one child completed Give-N after the numerical comparison task. The numerical comparison task had two between-subject conditions that differed on the range of numbers tested: small numbers (≤ 4) and large numbers (>4). As in Experiment 1, each individual set was composed of red rectangles of the same size, and all sets were printed on letter-sized paper. All pairs of numerosities differed by a 2:3 ratio: children in the small numerosity condition were asked to compare 2 vs. 3 (8 trials), and those in the large numerosity condition compared 6 vs. 9, 10 vs. 15, and 12 vs. 18 (4 trials each). Children in the small numerosity condition also compared 1 vs. 2 (2 trials) and 3 vs. 4 (2 trials), and thus children completed a total of 12 trials in each condition. Size and number were incongruent in all comparisons (as in Figures 1b and 1d). Pairs of sets of rectangles were presented in one of two item orders. In each order, the correct side (left or right) was counterbalanced and the order was pseudo-randomized such that no two consecutive trials were of the same pairs of numerosities. No feedback was given.

In the small numerosity condition, the cumulative surface area ranged from 6 to 8 $cm^2$ and cumulative perimeter ranged from 18 to 24 cm. In the large numerosity condition, cumulative surface area ranged from 18 to 36 $cm^2$, and cumulative perimeter ranged from 54 cm to 108 cm.

Results and Discussion

*Give-N.* The number of children and the mean age in each knower-level group are presented in Table 2.[7] Age was significantly correlated with children's knower-level, Pearson's $r = .48$, $p < .001$.

| | Small | | Large | |
|---|---|---|---|---|
| | **N** | **Mean Age (Range)** | **N** | **Mean Age (Range)** |
| **Non-knowers** | 3 | 3;0 (2;10 - 3;3) | 3 | 2;10 (2;8 - 3;0) |
| **1-knowers** | 8 | 3;0 (2;6 - 3;10) | 12 | 3;1 (2;7 - 4;9) |
| **2-knowers** | 10 | 3;2 (2;7 - 4;5) | 12 | 3;3 (2;6 - 4;0) |
| **3-knowers** | 9 | 3;5 (2;10 - 4;2) | 5 | 3;8 (2;8 - 4;3) |
| **CP-knowers** | 14 | 3;10 (3;3 - 4;5) | 9 | 4;1 (3;4 - 4;7) |

Table 2. Age and number of children in each knower-level group in Experiment 2.

*Numerical Comparisons.* Preliminary analyses showed that children recruited in Canada

[7] Across both small and large numerosity conditions, I found 1 4-knower and he was classified as a 3-knower in the analyses. One child in the small numerosity comparison did not complete the Give-N task but was included in analyses for the numerical comparison task.

and in the US performed similarly, $t(84) = 1.29$, $p = .20$. Given that almost all American children (21/23) were tested in the small numerosity condition, I also compared Americans (n = 21) to Canadians (n = 22) on comparisons of small numerosities only, and again, found no difference, $t(43) = .47$, $p = .64$. Thus, I combined data from the two locations in the subsequent analyses**.** Preliminary analyses also revealed that there were no order or gender effects, so I collapsed across these variables (all $p$'s > .36).

First, I asked if the effect of Number Range in Experiment 1 could be replicated. I constructed a generalized linear model predicting overall proportion correct using Number Range (small vs. large) and Knower-Level (subset-knowers vs. CP-knowers)[8] as predictors, and centered Age as a covariate. I also included two interaction terms: Number Range x Age and Number Range x Knower-Level. Results revealed a main effect of Number Range, $\beta = -.17$, $SE = .11$, $p = .015$. As in Experiment 1, children were better at comparing small numerosities ($M = .82$, $SE = .035$) than large numerosities ($M = .70$, $SE = .038$). There was also an effect of Age, $\beta = .091$, $SE = .07$, $p = .022$. In addition, I found a main effect of Knower-Level, $\beta = -.18$, $SE = .09$, $p = .022$. CP-knowers ($N = 23$, $M = .85$, $SE = .057$) performed significantly better than subset-knowers ($N = 62$, $M = .69$, $SE = .031$). No interactions were found.

To test the processing demands alternative, I compared the large numerosity comparisons to each other. No significant differences were found (all p's > .64). I also compared 2 vs. 3 to each of the large numerosity comparisons individually. A significant difference was found in each case: 6 vs. 9, $t(76.3) = 2.42$, $p = .019$; 10 vs. 15, $t(84) = 2.25$, $p = .027$; 12 vs. 18, $t(84) = 2.58$, $p = .012$

[8] I only had a total of six non-knowers in Experiment 2 ($n = 3$ in the small numerosity condition, $n = 3$ in the large numerosity condition), and thus they were included in the subset-knowers group.

(see Figure 4).[9] These results strongly suggest that the Number Range effect was categorical - i.e., children were more accurate on comparisons of small numerosities than on comparisons of large ones, but were equally accurate on all comparisons of large numerosities.

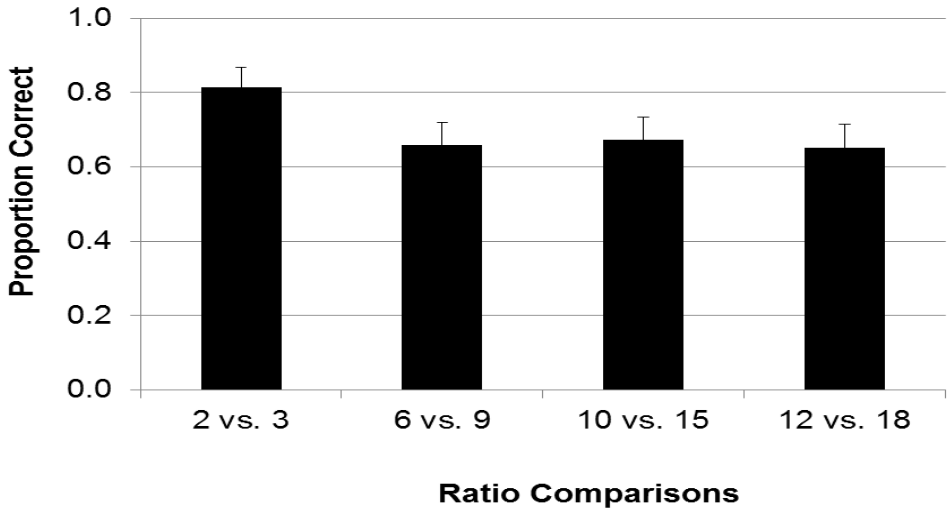


*Figure 4*. Proportion correct on small and each of the large numerosity comparisons in Experiment 2. Error bars represent standard error of the mean.

To examine whether children succeeded on 2 vs. 3 in Experiment 1 because they somehow defaulted to choosing 3 (or to avoiding 2) as the correct answer without doing the comparison. If children used the first strategy, they should have always chosen the wrong collection on comparisons of 3 vs. 4. If they used the second, they should have always chosen the wrong collection on comparisons of 1 vs. 2. This was not so. Rather, children performed significantly above chance on 1 vs. 2 ($M$ = .77, $SD$ = .38), $t(44) = 4.8$, $p < .001$, and on 3 vs. 4 ($M$ = .74, $SD$ = .33), $t(44) = 4.9$, $p < .001$.

Moreover, given a main effect of knower level, I compared children's performance on both

9 The statistics for equal variances not assumed was reported for the comparison between 2 vs. 3 and 6 vs. 9 because Levene's Test for Equality of Variances was significant ($p$=.045).

small and large numerosity conditions against chance as in Experiment 1 (see Figure 5). One-sample t-tests showed that only 1-knowers, 3-knowers, and CP-knowers performed significantly above chance on small numerosity comparisons (all *p*'s $< .05$). Non-knowers ($p = .46$) and 2-knowers ($p = .091$) did not. However, as shown in Table 2, there were only 3 non-knowers and, although 2-knowers' performance was not significantly above chance, it approaches significance. In the large numerosity comparison, only CP-knowers performed significantly above chance. Two-knowers' ($p = .089$) and 3-knowers' ($p = .094$) performance approaches significance, whereas non-knowers' and 1-knowers' did not (all *p*'s $> .45$).

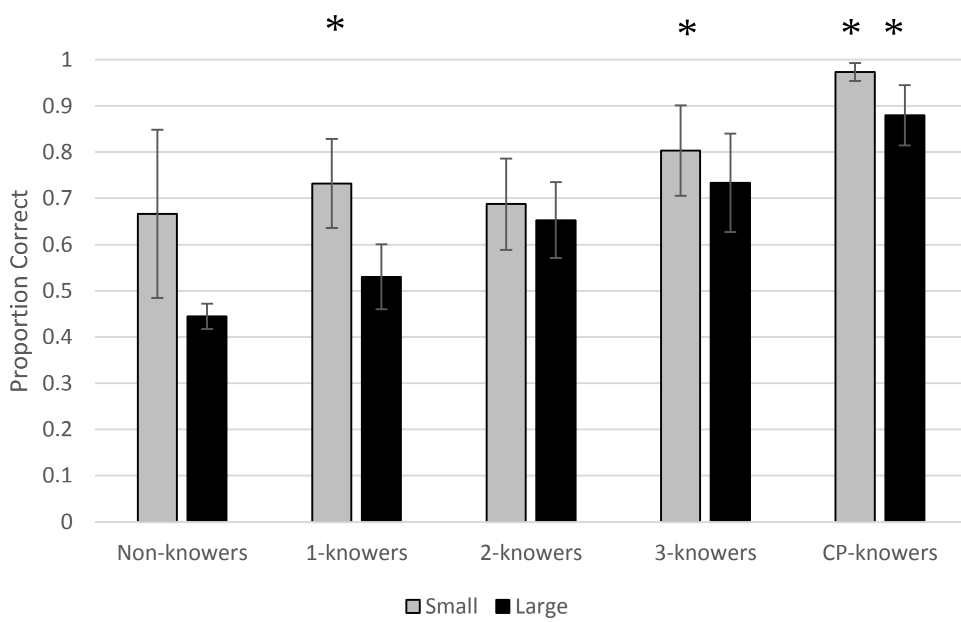


*Figure 5*. Overall proportion correct on small and large numerosity comparisons by knower-level group in Experiment 2. Error bars represent standard error of the mean. * $p < .05$

The evidence thus far suggests that preschoolers use parallel individuation to compare small numerosities. I next asked whether this ability is available prior to the acquisition of numerical language. Due to a small sample of non-knowers in the small numerosity condition (*N*

= 3), I combined non-knowers and 1-knowers in the current analysis. I included 1-knowers and analyzed their performance on 2 vs. 3 because they have not yet acquired number word meanings for the number words that designate the small numerosities I used in this analysis, namely 'two' and 'three.' I found that they performed significantly above chance, $t(10) = 2.64$, $p = .013$ (1-tailed; $M = .71$, $SD = .27$). I also examined whether non- and 1-knowers performed better on comparisons of small than large numerosities ($M = .51$, $SE = .057$), and found that they did, $t(24) = 2.10$, $p = .046$. This finding suggests that by the time children acquire the meaning of 'one', they can recruit parallel individuation to compare the numerosities of sets of up to four objects.

## Chapter Two Discussion

Two separate experiments showed that preschoolers were modestly, but reliably more accurate in comparing numerosities between 1 and 3 than numerosities larger than 5, despite the fact that the ratios of all comparisons were the same. The same pattern was found on comparisons where size and numerosity were incongruent – i.e., where the total area and the total perimeter of both sets of rectangles were equal so that individual rectangles in the small set were larger than the ones in the large set. Moreover, both experiments showed that this effect is also found in children who have not learned the meaning of any number word beyond "one."

In my view, the best explanation of these findings is that preschoolers used distinct systems to compare small and large numerosities: parallel individuation for 1 to 3 (and likely, 4) and the ANS for more than 5. Various aspects of my results allow me to rule out alternative explanations. First, Experiment 2 showed that, although children were equally accurate on comparisons of a relatively wide range of large numerosities (6 vs. 9, 10 vs. 15, and 12 vs. 18), they were more accurate on comparisons of small numerosities (2 vs. 3) than on any of the comparisons of large numerosities. This suggests that children did not perform more poorly on comparisons of large numerosities because forming representations of larger numerosities made greater demands on domain-general perceptual or attentional processes. Rather, it suggests that there was a categorical difference between performance on comparisons of the numerosities of sets that can be represented with parallel individuation and performance on comparisons of numerosities that can only be represented with the ANS.

Second, it cannot be that children were more accurate on comparisons of small numerosities because they could count small sets more easily than large ones. Indeed, children were discouraged from counting in this study, and, according to the experimenter's informal observations, very few children actually attempted to count (see also Sophian, 1987). Furthermore, the same difference was found in children who had not learned how to use counting to determine the number of objects in a set – i.e., children who were subset-knowers or non-knowers. Another possibility is that most children could name the small numerosities more easily than the large ones without counting, and that naming the numerosities made the comparisons easier (see Le Corre & Carey, 2007). However, this does not explain why children who had not learned the meaning of any number words beyond "one" were more accurate on comparisons of small numerosities than on comparisons of large ones.

While the current experiments found an effect of number range, two previous studies that tested performance on small and large non-verbal numerical comparisons in young children who had learned some numerical language did not find this effect (Abreu-Mendoza, Soto-Alba & Arias-Trejo, 2013; Wagner & Johnson, 2011). The results of Experiment 1 suggest a possible explanation for this discrepancy. Specifically, Experiment 1 suggests that the difference between small and large comparisons is more pronounced in children who are younger than 3 years 6 months than in older children. The participants in Abreu and colleagues' and in Wagner and Johnson's studies were older on average (the average age in these studies was 4 years plus or minus one month) than the ones in the present study (the average age in both Experiments was near 3 years 6 months). Moreover, neither group of researchers tested whether there was an interaction between age and number range. Therefore, they may have failed to find a difference between performance on small and on large comparisons because their analyses did not focus on children of the right age.

*Relation between knowledge of the cardinal principle and performance on non-verbal numerical comparisons*

Across both experiments, I found that CP-knowers were better at numerical comparisons than subset-knowers, even when age was partialled out. Why is acquiring the cardinal principle related to an improvement in children's performance on non-verbal numerical comparisons? One possibility is that acquiring the cardinal principle is related to changes in the non-verbal representations that are used to represent and compare numerosities in non-verbal numerical comparison tasks. Another possibility is that acquiring the principle is related to changes in other factors (e.g., domain-general processing) that affect performance on non-verbal numerical comparison tasks.

Ideally, if the acquisition of the cardinal principle is related to a change in non-verbal representations of numerosity, one would expect to find a correlation between knowledge of the cardinal principle and performance on non-verbal numerical comparisons regardless of the relation between the physical properties of the collections (e.g., the sum of the areas of their elements or "total area") and their numerosity. In accord with this, I find an effect of knowledge of the cardinal principle that is independent of age when trials where numerosity and total area are confounded are intermixed with trials where total area is equated across collections (Experiment 1). I also found this effect when total area is always equated across collections (Experiment 2). The latter finding converges with other studies that suggest that the effect of the acquisition of the cardinal principle is more robust in tasks where total area is always equated across collections. Indeed, Abreu-Mendoza and colleagues found that when the task only includes numerical comparisons where the collections are equated for total area, knowledge of the cardinal principle predicted performance on non-verbal numerical comparisons independently of age. However, they also found that, when the task includes at least some trials where relative numerosity and relative total area are confounded (i.e., where the more numerous collection also has a larger total area), knowledge of the cardinal principle did not predict performance on non-verbal numerical comparisons independently of age. Rousselle & Noel (2008) also obtained the latter result. Thus, it may be that the correlation between the acquisition of the cardinal principle and performance on non-verbal numerical comparisons is not due to changes in non-verbal representations of numerosity *per se*. Rather, it may be due to changes in factors that play a greater role in numerical comparisons when the total area of the collections is not congruent with their numerosity. For example, Abreu-Mendoza and colleagues propose that the correlation may be mediated by changes in children's ability to focus on numerosity as a dimension to be compared and/or in their executive function (see Negen & Sarnecka (under review) for a similar proposal).

*How does parallel individuation support numerical comparison of distinct sets?*

How does parallel individuation support the computation of numerical comparison? Previous studies of infants' memory for hidden objects suggest that a possible mechanism is via the computation of one-to-one correspondence. Feigenson and Carey (2003, 2005) found that infants use parallel individuation to represent up to three objects hidden in a box, and infants use this system to track and compare the number of hidden objects to the number of retrieved objects. They propose that infants succeed by performing a computation of one-to-one correspondence. That is, when there is a one-to-one correspondence between the set of hidden objects and the set of retrieved objects, infants stop looking for objects in the box. Moreover, Feigenson and Halberda (2004) found that when the objects to be hidden are presented as spatially separate groups, infants can represent at least two sets of up to three hidden objects. Therefore, it is possible that, in order to compare the numerosities of distinct sets of up to three objects, preschoolers represent the objects in each set with parallel individuation, and then compare the sets on the basis of one-to-one correspondence. If the representation of one of the sets contains an extra object, children judge that this set is more numerous. Future studies should investigate whether this is indeed the mechanism whereby parallel individuation supports numerical comparisons.

*Conclusions*

The experiments in Chapter 2 are the first to provide evidence that parallel individuation supports numerical comparisons, and it does so prior to the acquisition of numerical language. Thus, there may be two developmental roots of our knowledge of numerical relations: the comparisons of the size of the mental magnitudes created with the ANS, and, possibly, the one-to-one correspondence operations defined over representations of objects created with parallel individuation.

# Chapter Three: The development of representations of unspecified numerosities in numerical comparison

In the last few decades, research on the development of numerical cognition has focused on the development of representations of particular numbers, namely via parallel individuation, the approximate number system, and counting. However, some types of numerical knowledge and reasoning involve thinking about numbers in general – i.e., knowledge that applies to all possible numbers. For example, for any given set of objects with numerosity x, x changes whenever an object is added to or subtracted from the set; this knowledge applies regardless of what numerosity x refers to. Thus, whether the set contains 1, 5, or 100 objects, when an object is added to the set, the numerosity has increased. Likewise, whenever an object is removed from a set of 1, 5, or 100 objects, the numerosity has decreased. These examples illustrate that there are facts about number that necessarily hold true regardless of the number of objects in the set. These facts can be expressed using general statements with algebraic expressions (e.g., $x + 1 > x$) that are supported by representations of unspecified numerosities.

Knowledge about unspecified numerosities contrasts with knowledge about particular numerosities. For example, when an object is added to a set of 2, the set no longer has 2 objects; it has more than 2 objects. But this knowledge is only specific to '2', and knowing that $2 + 1 > 2$ does not necessarily grant you the knowledge that $50 + 1 > 50$. In this sense, the knowledge is number-specific, because knowing facts about '2' does not entail knowing facts about all possible numbers. The present study investigates when we begin to demonstrate general numerical knowledge as a first step to discovering how we construct such knowledge. In particular, I explore one of the most fundamental aspects of numerical knowledge – i.e., the knowledge that adding one element increases the numerosity of a set, expressed as $x + 1 > x$.

Previous research has shown that sometime between the preschool and kindergarten years, children begin to show signs of general knowledge that adding one element to a set increases its numerosity. In one study, 5-year-olds were shown a box of objects labelled with a number word. The experimenter then performed a transformation that changed the numerosity (e.g., taking away one object; a numerical transformation) or a transformation that did not change the numerosity (e.g., shaking the box; a non-numerical transformation), and asked whether there was still the same number of objects in the box (Lipton & Spelke, 2006). Crucially, the researchers used numbers that were beyond children's counting range (e.g., 127) to ensure that children could not recruit particular facts about number words to succeed on the task. They found that children correctly responded with the original number word after non-numerical transformations, but a different number word after numerical transformations, providing strong evidence that at age 5, children understand how number words label sets in general.

Some studies suggest that younger children have similar knowledge; however, the evidence is weaker. Sarnecka and Gelman (2004) presented older 2- and 3-year-olds with one set of objects placed in an opaque box and labeled it with a small (i.e., two, three) or a large (e.g., five, six) number word, e.g., "Here are [five] moons." Then, the experimenter performed a numerical (adding one object, subtracting one object) or a non-numerical (shaking or rotating) transformation. Children were asked whether the original or a new, alternative number word applied, e.g., "Now how many moons – is it five or six?" They found that regardless of whether small or large number words were used, children correctly chose the alternative number word after the numerical but not non-numerical transformations. However, in another study with a slightly different paradigm, the experimenter used two sets of objects rather than one, and labelled one of the two sets with a number word (e.g., "This tray has [five] sheep."; Condry & Spelke, 2008). Children between the ages of 3 and 3 ½ were tested, and number words that were outside of their known number word range (e.g., five, seven) were used. The experimenter then performed a numerical or non-numerical transformation on the labelled set, and asked children to point to the tray that was labelled by either the original number word (e.g., five) or a different number word (e.g., six). Using the two-set task, they found that regardless of whether an object was added to the set or whether the objects were rearranged, 3-year-olds were equally likely to choose the labelled and the unlabelled tray, suggesting that they may not understand how transformation affects number word meanings. Thus, children appear to behave differently on the one-set task and the two-set task.

While it is possible that the two-set task may be more cognitively demanding than the one-set task, Brooks, Audet and Barner (2013) argue that children may not be recruiting numerical knowledge to solve the one-set task. In the one-set task, children were presented with a labelled set and observed changes in numerosity, then a new number word was offered as an alternative that contrasted with the original number word. They argue that if children expect the experimenter to provide relevant information in communication, they may infer that the alternative number word is offered because it is relevant, and thus choose the alternative label for pragmatic reasons. On their view, children selected a new number word after numerical transformation not because of their knowledge about number word application, but because of pragmatic inference.[10] In support of their argument, they first replicated children's failure on the two-set task and their success on the one-set task in a within-subject design. Then they extended the same pattern of results to novel words denoting object kinds (e.g., "This is a dax"). They found that children who were presented

[10] However, it should be noted that on Brooks et al.'s account, children still possess the knowledge about what kind of transformation affects numerosity. This is because children only apply the alternative label in the context of numerical transformation, but not in the context of non-numerical transformation. I will revisit this point in the General Discussion section of this chapter.

with a single, labeled novel object preferred a new label after the experimenter added a piece to the novel object, but not when the experimenter shook the object. When two identical objects were presented and only one was labeled, children were equally likely to assign a label to its original referent regardless of whether the labeled object was transformed or shaken. Children's behaviour in the novel object tasks was similar to that in tasks involving number words, providing evidence that their success in previous studies of number word application may be explained by a domain-general pragmatic inference mechanism.

Regardless of whether children are using domain-specific numerical reasoning or domain-general pragmatic reasoning, most previous studies on children's acquisition of general or specific facts about number require children to interpret or use number words in their tasks. Children's apparent difficulty in mastering the logic of number words may therefore reflect a lack of knowledge of how number words represent number, and might not constitute evidence for a lack of general knowledge about number. Moreover, studies on children's verbal numerical comparison have found that 4-year-olds may not understand that "ten" is more than "six" (Le Corre, 2014; Le Corre & Carey, 2007; Schaeffer, Eggleston, & Scott, 1974). Thus, even if children apply a different, larger number word after the addition of one object as suggested by Sarnecka and Gelman (2004), this does not necessarily mean that they understand the numerosity has increased as a result of addition.

The evidence thus far suggests that by 5, children understand how numerical and non-numerical transformations affect the use of number words, but it is unclear at what age children begin to show such understanding. Also, children's difficulty with number word application may be due to problems specific to the acquisition of number word meanings, and not their cognitive ability to reason numerically. Thus, a non-verbal reasoning task that does not require children to reason about number word meanings would provide another test of children's knowledge of the numerical principle 'x + 1 > x'. I know of one study that assesses children's understanding of numerical relations such as greater than (>) or less than (<) without recruiting their number word knowledge. In this study, Gelman (1977) trained children between the ages of 2 and 4 years on the numerical relation between a set of two objects and a set of three objects, and tested whether they understood what type of transformation affected numerosity. Children were told that the set of three objects was the "winner", and the set of two objects was the "loser". Then the experimenter did one of three things to the "winner" set: surreptitiously removed an object, rearranged objects or substituted an individual for another individual of a different kind. Interestingly, children responded that there were no winners when the experimenter removed an object, and they chose the set of three objects as the "winner" in cases where the experimenter rearranged objects or substituted an individual. Nevertheless, this study used small sets (< 4), and thus, children could potentially recruit representations of particular numerosities to succeed in this non-verbal reasoning task. Children's ability to generalize the greater-than/less-than relation to larger set sizes remains unknown.

*The present experiments*

The experiments in this chapter examine children's representations of unspecified numerosities – specifically, the knowledge that when an object is added to or taken away from a set, it necessarily follows that the numerosity of a set changes. Across three experiments, children's reasoning about changes in numerosities of visible sets (Experiments 3), and hidden sets (Experiments 4 and 5) was investigated. Given that in previous studies, children demonstrate conflicting behaviour in tasks involving number words, Experiments 3 and 4 did not include number words in the instructions or ask questions about number word reference (e.g., “Can you point to the tray with five objects?”). This design allows me to ask whether children have general knowledge about number that applies to all set sizes without recruiting their knowledge about number words. In Experiment 5, I explored the possibility that highlighting the numerical comparison task that was required of the participants may improve their performance.

## Experiment 3: Children’s reasoning about numerosity changes with visible objects

To investigate when children begin to understand that adding one element to a set of objects increases its numerosity regardless of set size, I designed a paradigm that asked children to reason about numerical and non-numerical transformations of visible sets of objects that differed in set size – small sets and large sets. Children observed four kinds of transformation: addition of an object, subtraction of an object, addition and subtraction of the same object, and moving one object. After the transformation, children were asked whether the set contained more elements.

Given that children cannot rely on the ANS for reasoning about large sets because a numerical ratio of 15:14 (in the case of subtraction) or 15:16 (in the case of addition) is beyond the range of ratio discrimination that preschoolers can perform (see Halberda & Feigenson, 2008), their performance on large number trials is critical for determining whether children are capable of reasoning about unspecified numerosities. Specifically, if children demonstrate general understanding of number, they should recognize that addition and subtraction, but not moving one object, change the numerosity of a set. However, if they fail at reasoning about large sets, it could be due to task demands. To ensure that task demands are not an issue, small sets were also included. Given that children can recruit parallel individuation to reason about small sets, if they succeed with small sets, but fail with large sets, this provides evidence that they do not possess general numerical knowledge. However, if they fail with both small and large sets, then this suggests that their failure is due to a general difficulty in understanding the task.

Importantly, children were prevented from counting across all experiments reported in this chapter.

## Method

### Participants

Forty-eight children between the ages of 3 and 5 participated. There were 19 3-year-olds (range = 3 years 1 month to 3 years 11 months; mean = 3 years 6 months; 10 boys), 14 4-year-olds (range = 4 years 0 months to 4 years 11 months; mean = 4 years 3 months; 10 boys), and 15 5-year-olds (range = 5 years 0 months to 5 years 11 months; mean = 5 years 6 months; 10 boys). They were recruited at daycare centres in Kitchener-Waterloo and nearby areas.

### Procedures and stimuli

Each child completed the Transform Set task.[11] Each testing session lasted about 10 - 15 minutes. Two experimenters sat across from the child.

*Transform Set task*. At the beginning of the session, children were introduced to two puppets – Winnie the Pooh and Giraffe –and were told that they were friends. Each puppet was manipulated by a different experimenter. Children completed two familiarization trials before moving onto the experimental phase. The purpose of the familiarization trials was to test whether children can reason about changes made to a single object. During familiarization, Winnie the Pooh took out a ball that could be made bigger and smaller (Hoberman's Sphere), and asked

[11] Children in Experiments 3-5 also completed the Give-N task (as described in Chapter 2) at the end of the testing session to examine effects of counting knowledge. However, a majority of the children had already acquired the cardinal principle (i.e., all 4- and 5-year-olds, and more than half of the 3-year-olds were CP-knowers), making it difficult to study the effects of counting on numerical reasoning. Thus, results from the Give-N task were not analyzed.

Giraffe to make it 'bigger'. Giraffe either made it bigger or smaller. The direction of the change was counterbalanced across participants. Children were then asked, "Does Winnie the Pooh have a bigger ball now?" A 'yes' or 'no' response was recorded. If children answered incorrectly, feedback was given, and the puppets repeated the action and the experimenter asked the question again. Almost all children succeeded on the two familiarization trials with no feedback. One child failed one of the familiarization trials but succeeded on a second attempt, and was included in the analyses.

The experimental trials began with a dialogue between Winnie the Pooh and Giraffe.

*Winnie the Pooh*: Giraffe, I have some toys to show you. Do you want to see?

*Giraffe*: Yeah, sure!

*Winnie the Pooh*: Look, here are my [blocks] and I want more [blocks]. Can you help me?

*Giraffe*: Yeah, let me help you. Look, [child's name], look. (Giraffe performs the transformation).

Then the experimenter asked, "Does Winnie the Pooh have more [blocks] now?" Children gave a 'yes' or 'no' response. Stimuli included eight different kinds of objects – pom poms, bows, buttons, red Lego blocks, yellow Lego blocks, bells, rocks, and hearts. All objects were presented on a letter-sized piece of paper.

Children received both small and large number trials, which were conducted in blocks. For the small number trials, an object was always added to and subtracted from a set of two objects, such that in the case of addition, it resulted in three objects, and in the case of subtraction, it resulted in one object. For the large number trials, an object was always added to and subtracted from a set of 15 objects, such that in the case of addition, it resulted in 16 objects, and in the case of subtraction, it resulted in 14 objects. The order of the small and large number blocks was counterbalanced between children at all daycares such that at each daycare, there was an approximately equal number of children who received the small number block first and the large number block first. For each number block, the experimenter transformed the set in one of four ways: (a) by moving one object ('Move 1'); (b) by adding one object ('Plus 1'); (c) by removing one object ('Minus 1'), and (d) by moving one object outside of the sheet and then putting it back into the set of objects ('-A+A). There were two trials for each transformation, making a total of 16 trials for each child. There were two possible item orders, and the order of trial type was randomized such that no two consecutive trials were of the same trial type.

## Results and Discussion

If children can represent unspecified numerosities, then they should be able to distinguish when the numerosity of a set changes: 'Plus 1' and 'Minus 1' trials change the numerosity of a set, and 'Move 1' and '-A+A' trials do not. Critically, if children can demonstrate this knowledge in the large number trials, then this constitutes evidence for their ability to reason about unspecified numerosities because neither the ANS nor parallel individuation representations can support reasoning of large sets of the sort tested in this study. (The ANS cannot because of the ratios used.) If children only succeed on small number trials, then it suggests that they are relying on the representations of particular numbers, and may lack the cognitive capacity to reason algebraically.

### *Overall performance in Transform Set Task*

I computed a score for each participant in both small and large number trials. Participants received a score of 1 for each trial that they answered correctly, and thus, the maximum score is 8 for both the small and large number trials.

Preliminary analyses found no order or gender effects ($p$'s > .49), so these variables were not included in subsequent analyses. First, to examine the performance on small and large number trials, a repeated measures ANOVA using proportion correct as the dependent variable, with Number Range (Small and Large) as a within-subjects factor, and Block Order (Small first vs. Large first) and Age Group (3-year-olds vs. 4-year-olds vs. 5-year-olds) as between-subjects factors was conducted. There was a main effect of Age Group, $F(2,42) = 8.33$, $p < .001$, $\eta_p^2 = .28$. Tukey HSD revealed that 3-year-olds (57.6%) did not differ significantly from 4-year-olds (58.9%), but both groups performed significantly worse than the 5-year-olds (85.8%). There was also a main effect of Block Order, $F(1,42) = 8.25$, $p = .027$, $\eta_p^2 = .11$, indicating that children performed better overall when they were tested on small number trials first (small number first 74.6%; large number first: 60.4%). More importantly, there was a significant main effect of Number Range, $F(1,42) = 36.02$, $p < .001$, $\eta_p^2 = .46$. Children were better at reasoning about small sets (78.0%) than large sets (57.0%). No interactions were found.

To examine the effect of Number Range, I analyzed if children's performance on large number trials was above chance (50%). If children can reason about unspecified numerosities, they should perform above chance. Separate analyses were conducted for each of the three age groups. One-sample t-tests showed that only 5-year-olds (79.2%) performed significantly above chance on large number trials, $t(15) = 3.95$, $p < .001$. Neither the 3- nor the 4-year-olds were significantly different from chance, $M$ = 46.1%, $t(18) = -.71$, $p = .49$, $M$ = 45.5%, $t(13) = -.66$, $p = .52$, respectively. This suggests that the capacity to reason about unspecified numerosities may not appear until 5.

However, one may argue that 3-year-olds simply do not understand the task (e.g., difficulty of reasoning about the puppet's intentions in general). To address this, I computed one-sample t-tests for small number trials. Results showed that children from all age groups had above chance performance on small number trials: 3-year-olds: $M$ = 69.1%, $t(18) = 3.27$, $p = .004$; 4-year-olds: $M$ = 72.3%, $t(13) = 2.96$, $p = .011$; 5-year-olds: $M$ = 92.5%, $t(14)=9.38$, $p < .001$.

*Relationship between age and trial type*

To further explore how children's reasoning about unspecified numerosities differs by age, I analyzed whether children's responses differed on the four different transformation types. If children can represent unspecified numerosities, then for both small and large sets, they should understand that 'Plus 1' is the only transformation that increases numerosity, and respond 'yes' to the question about whether the puppet has more objects. Friedman's ANOVA with the four transformation types as within-subject variables was computed separately for small and large number trials and for each age group with proportion of 'yes' responses as the dependent variable. Figures 6-8 display children's proportion of 'yes' responses for each transformation type in both small and large number trials.

For 3-year-olds, responses varied across the four transformation types for small number trials ($\chi 2$ (3) = 14.0, $p = .003$) but not for large number trials ($\chi 2$ (3) = 7.34, $p = .062$; see Figure 6). Wilcoxon signed rank tests revealed that 3-year-olds were able to differentiate between the transformation types for small sets but not for large sets (see Table 3). To adjust for multiple comparisons, a Bonferroni corrected alpha level of .0083 (.05/6 comparisons for each age group) was applied to all pairwise comparisons involving the transformation types.[12]

[12] Across the three experiments reported in this chapter, all trial type analyses were adjusted for Bonferroni correction. It is important to note that Bonferroni is often overly conservative. Significant and marginally significant values are marked with asterisks.

For 4-year-olds, responses varied across the four transformation types for both small ($\chi 2$ (3) = 24.3, $p$ < .001) and large number trials ($\chi 2$ (3) = 15.6, $p$ = .001; see Figure 7). Wilcoxon signed rank tests revealed that 4-year-olds were sensitive to the differences among the transformation types for small sets; for large sets, they appear to be able to differentiate between addition and subtraction, but not between addition and the no-numerosity change trials (i.e., '-A+A' and 'Move 1'; see Table 4).

For 5-year-olds, responses varied across the four transformation types for both small ($\chi 2$ (3) = 38.1, $p$ < .001) and large number trials ($\chi 2$ (3) = 30.2 $p$ < .001; see Figure 8). Wilcoxon signed rank tests revealed that 5-year-olds were able to differentiate between all the transformation types for both small sets and large sets (see Table 5).

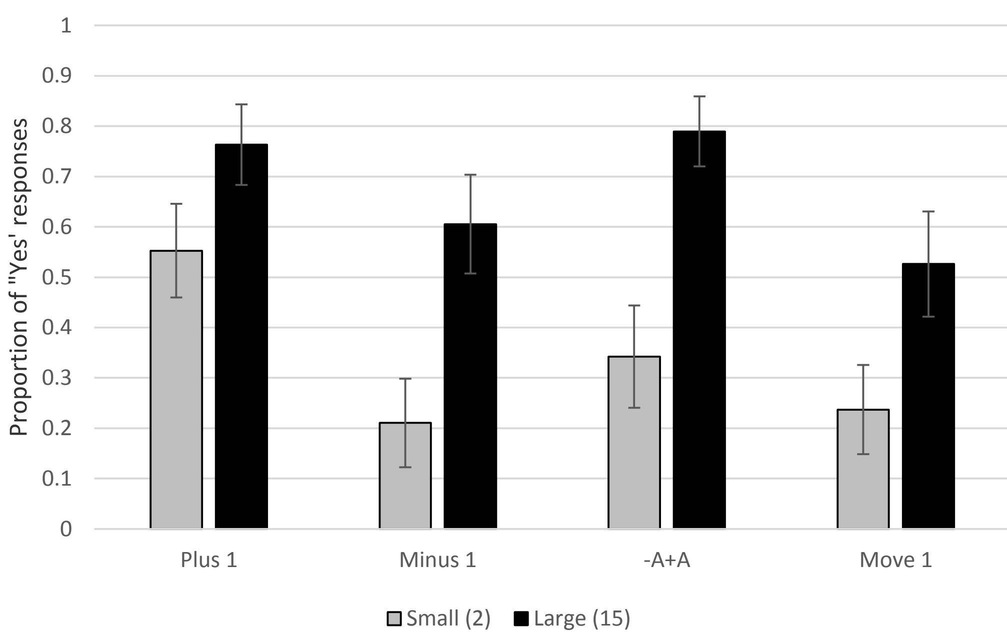


Figure 6. Three-year-olds' proportion of 'yes' responses in both small and large number trials. Error bars represent standard errors of the mean.

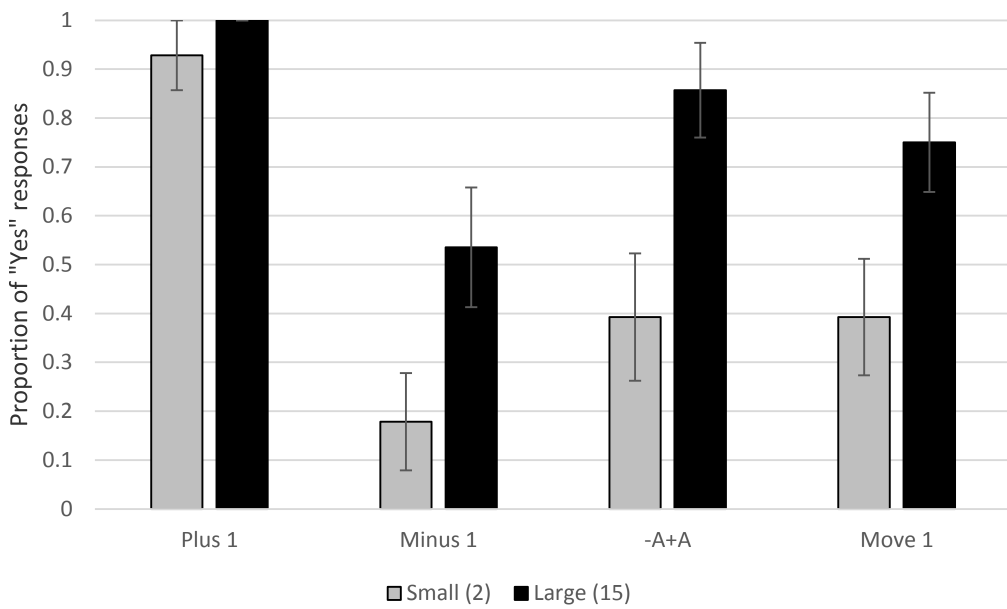


Figure 7. Four-year-olds' proportion of 'yes' responses in both small and large number trials. Error bars represent standard errors of the mean.

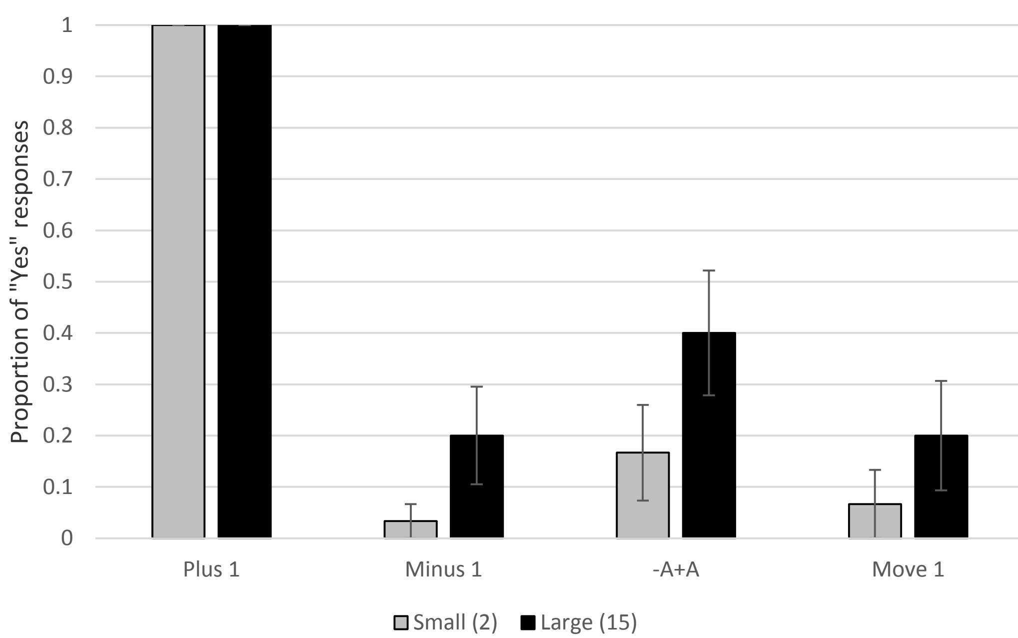


Figure 8. Five-year-olds' proportion of 'yes' responses in both small and large number trials. Error bars represent standard errors of the mean.

| | Small | Large |
|---|---|---|
| **Plus 1 – Minus 1** | -2.67 (.008)** | 1.19 (.233) |
| **Plus 1 – Minus A Plus A** | -1.90 (.057) | -.33 (.739) |
| **Plus 1 – Move 1** | -2.44 (.015)** | -2.01 (.045) |

Table 3. Z-scores (p-values) from Wilcoxon signed rank tests for 3-year-olds. *** significant difference with Bonferroni correction (adjusted alpha = .0083). ** marginally significant difference with Bonferroni correction.

| | Small | Large |
|---|---|---|
| **Plus 1 – Minus 1** | -3.21 (.001)*** | -2.60 (.009)** |
| **Plus 1 – Minus A Plus A** | -2.71 (.007)** | -1.41 (.157) |
| **Plus 1 – Move 1** | -2.76 (.006)** | -2.07 (.038) |

Table 4. Z-scores (p-values) from Wilcoxon signed rank tests for 4-year-olds. *** significant difference with Bonferroni correction (adjusted alpha = .0083). ** marginally significant difference with Bonferroni correction.

| | Small | Large |
|---|---|---|
| **Plus 1 – Minus 1** | -3.77 (< .001)*** | -3.42 (.001)*** |
| **Plus 1 – Minus A Plus A** | -3.50 (< .001)*** | -2.97 (.003)*** |
| **Plus 1 – Move 1** | -3.74 (< .001)*** | -3.46 (.001)*** |

Table 5. Z-scores (p-values) from Wilcoxon signed rank tests for 5-year-olds. *** significant difference with Bonferroni correction (adjusted alpha = .0083).

Results from this experiment show that the ability to reason about unspecified numerosities begins to emerge by age 4, and is fully in place by age 5. Using a Transform-Set task, I found that 3-year-olds can reason about the effects of transformations on small sets, but not on large sets, suggesting that while they can recruit parallel individuation for small number trials, they lack an ability to reason about unspecified numerosities. By age 4, children are capable of differentiating between addition and subtraction, but they fail to recognize that taking away an individual and returning it back to the set does not result in a change of numerosity of the set. Five-year-olds are able to recognize that addition increases the numerosity of the set, but other transformations do not.

The current results raise questions about why 4-year-olds understand the effects of addition and subtraction, but fail to understand that the addition of one element and subtraction of one element cancels each other out. There are two possible explanations. First, children may succeed on the ‘Plus 1’ and ‘Minus 1’ trials, but fail on the ‘-A+A’ trials because the latter type of transformation involves two steps – i.e., the subtraction *and* addition of the same individual, whereas the former transformations only involve one step. It is thus possible that reasoning about the effect of ‘-A+A’ is computationally more complex than ‘Plus 1’ and ‘Minus 1’, and makes greater processing demands. Second, 4-year-olds may be responding based on a ‘last-action’ heuristic that does not require them to reason about the effects of transformation on the numerosity of a set. For example, whenever the last action is adding an object, one gives a positive response and answers that there are ‘more’, and whenever the last action is taking away an object, one gives a negative response and answers that there are ‘not more’. To examine these two possibilities, I added other types of transformations in the following experiments. Specifically, I included transformations that involved the substitution of an individual (e.g., replacing a black rock with a white rock), and the removal of 3 individuals and addition of 1 individual. If children's difficulty with the '-A+A' trial is due to the fact that this transformation involves two steps, then they should fail on both the substitution and 'Minus 3 Plus 1' transformations. If they are responding based on a 'last-action' heuristic, then they should respond that there are more on all transformations except for 'Minus 1' and 'Knock' trials.

### Experiment 4: Children's reasoning about numerosity changes with hidden objects

Experiment 3 suggests that the capacity to reason about numerical transformations without representing any specific numerosities may begin to emerge around age 4, and is fully in place at age 5. However, the experiments with visible sets could underestimate children's ability to reason about unspecified numerosities because some of the transformations required children to resolve a conflict between their perception and their abstract representation of the set of objects. For example, in the case of moving one object, children may experience conflict between their perceptual reasoning (i.e., the set of objects takes up more space) and conceptual reasoning (i.e., the numerosity has not changed even though objects are moved). And this may explain why 3- and 4-year-olds responded that there were more elements around 50-70% of the time after one object was moved. Nevertheless, children's performance on the '-A+A' trial in Experiment 3 suggests that the reasoning conflict explanation does not fully explain the results. On a '-A+A' trial, the pre- and post-transformation sets look perceptually the same, yet 3- and 4-year-olds said there were more elements post-transformation 80% of the time. Moreover, 3-year-olds perform poorly on the 'Minus 1' trials, despite the fact that perceptual and conceptual reasoning coincide. This makes it unlikely that such a conflict can account for all of the results in the previous experiment. Nevertheless, to ensure that the perceptual aspect of sets does not interfere with children's numerical judgments, Experiment 4 presented sets that were hidden. This also provides another test for children's reasoning about representations of unspecified numerosities.

In Experiment 4, objects were presented in an opaque box and children could not see how many objects were in the box. Then a numerical or a non-numerical transformation was performed, and the experimenter asked if there were more objects in the box after the transformation. If older children's success on large number trials in Experiment 3 was explained by their ability to represent unspecified numerosities, then the same pattern of results should be observed in Experiment 4. That is, 5-year-olds should demonstrate an ability to reason about unspecified numerosities, and children younger than 5 should not.

## Method

### Participants

Fifty-two children between the ages of 3 and 5 participated. There were 21 3-year-olds (range = 3 years 0 months to 3 years 11 months; mean = 3 years 7 months; 11 boys), 17 4-year-olds (range = 4 years 1 month to 4 years 11 months; mean = 4 years 5 months; 11 boys), and 14 5-year-olds (range = 5 years 0 months to 6 years 1 month; mean = 5 years 8 months; 6 boys). They were recruited at daycare centres in Kitchener-Waterloo and nearby areas. An additional three children were excluded from the analyses for failing twice on a familiarization trial (n = 2) and object trials during the test phase (n = 1; see below).

### Procedures and stimuli

*Box Task.* The Box Task was essentially the same as the Transform Set task except that the sets children had to reason about could not be seen. The Box Task had three phases: familiarization, labeling and test. The task began with the same familiarization phase using the ball that could be made bigger and smaller (Hoberman's Sphere), followed by a labeling phase in which the experimenter laid out one of each object kind and asked the child to label them. The experimenter used the labels provided by the child during the test phase. Twelve kinds of objects were used in the study: buttons, rocks, Lego blocks, bows, sticks, pom poms, flowers, leaves, stars, beads, shells, bells.

At test, children were told that they were going to play a game with a box and the objects that they just labeled. The experimenter first showed that the box was empty, then she transferred objects into the box from an opaque cup, and said, "I'm going to put [bells] into the box". The experimenter closed the box, and asked, "Are there more [bells] in the box now?" The experimenter then performed a transformation, and asked again, "Are there more [bells] in the box now?" The purpose of repeating the same test question twice is that in piloting, I found that children had difficulty parsing the event into appropriate time points for comparison. For example, it appeared that they sometimes compared the initial state of the box (i.e., an empty state) to the post-transformation state. To scaffold children into comparing the post-transformation state to the state of the box that was immediately before transformation, I asked the test question directly before *and* after transformation. Responses to the first question were not analyzed.

Six kinds of transformation were performed: adding one object ('Plus 1'), removing one object ('Minus 1'), taking away one object and putting it back ('-A+A'), taking away one object and putting back another object that was of the same kind but of a different colour ('-A+B'), taking away three objects and putting another one back ('-3+1'), and knocking on the box. For 'Plus 1', the experimenter added an identical object to the box. For 'Minus 1', the experimenter removed an object and hid it under the table. For '–A+A', the experimenter took an object out of the box, put it in front of the child, and put it back into the box. For '–A+B', the experimenter removed an object from the box (e.g., a black rock), hid it under the table, and put in another object of the same kind but of a different colour into the box (e.g., a white rock). For '-3+1', the experimenter took out 3 objects, put them in front of the child, quickly showed him/her that the box was not empty, and added an identical object.

The purpose of adding the '–A+B' trial was to further examine the nature of children's reasoning of unspecified numerosities. This trial extended the '-A+A' trial by testing whether children understand that a transformation that results in an identity change (i.e., substitution of one individual for a different individual) does not change the numerosity of sets. The purpose of the '-3+1' trial was to examine if children's difficulty with reasoning about the effects of transformation is specific to a change of one object. Children's performance on these two trials (i.e., '-A+B' and '-3+1') also addresses the question of whether their difficulty with the '-A+A' trial in the previous experiment can be explained by the 'two-step' processing account or the 'last-action' heuristic account.

In addition to trials that required children to judge whether there were more elements ('more' trials), children were given three object trials as control trials. These trials were designed to ensure that children understood the Box Task, and involved the addition of a pen, the removal of a pen, and knocking on the box. At the beginning of the object trials, children were shown a blue pen and a red pen, and were told that they were going to play a game with the pens and the box. After showing that the box was empty, the experimenter did one of three things, and asked "Is there a red pen in the box now?" For addition, the experimenter first put a blue pen, and then a red pen into the box; for the removal of a pen, the experimenter put both the blue and the red pens into the box, and then removed the red pen; for knocking, the experimenter put the blue pen in the box, and knocked on the box. Children were allowed to make one error across the three trials. Altogether, only four children made an error; they were included in the analyses. One child made two errors and was removed from the analyses.

Nine five-year-olds completed the study in one session. The others were tested over two sessions of twelve trials each. Children who finished all trials in one session received a total of 18 ‘more’ trials and 3 object trials; those who completed in two sessions received 9 ‘more’ trials and 3 object trials in the first session, and 12 ‘more’ trials in the second session. Thus, the object trials were conducted either at the end of session 1 or at the end of the entire session. There were three trials for each transformation type. However, for children who completed over two sessions, there were 6 ‘Plus 1’ trials, with 3 ‘Plus 1’ trials for each session; the purpose was to balance the ‘yes’ and ‘no’ responses in each session. The type of transformation was pseudo-randomized such that no two consecutive trials were of the same transformation type. Children received one of two randomized orders of trials.

## Results and Discussion

### *Overall performance on Box Task*

Preliminary analyses found no order or gender effects ($p$’s > .31), so these variables were not included in subsequent analyses.

An overall proportion correct, averaging the scores across all six transformation types, was computed and was used as a dependent variable. For each post-transformation question, children received a score of 1 each time they answered correctly; the maximum score is 18. For children who received 6 ‘Plus 1’ trials, only the first 3 trials were included.

First, I analyzed if there was an effect of age. An ANOVA revealed a main effect of Age Group, $F(2,51) = 6.37$, $p = .003$, $\eta_p^2 = .20$. Tukey HSD revealed that 5-year-olds (48.8%) differed significantly from 3-year-olds (26.5%, $p = .013$), but not 4-year-olds (39.5%, $p = .33$), and that 3- and 4-year-olds did not differ from each other ($p = .14$). To better characterize children’s performance at each age, I conducted separate analyses for 3-, 4-, and 5-year-olds.

*Relationship between age and trial type*

For each age group, I first analyzed whether children's responses differed on the six different transformation types using Friedman's ANOVA with proportion of 'yes' responses as the dependent variable. Then, similar to the analyses in Experiment 3, I compared whether children are sensitive to the distinction between 'Plus 1' and other transformation types using Wilcoxon signed rank tests with a Bonferroni corrected alpha level of .01 (.05/5 comparisons for each age group). Figures 9-11 display children's proportion of 'yes' responses for each transformation type.

For 3-year-olds, responses varied across the six transformation types ($\chi 2$ (5) = 11.34, $p$ = .045; see Figure 9). However, Wilcoxon signed rank tests revealed they failed to recognize the distinction between addition and almost all of the other transformation types (see Table 6). They did, however, show some sign of distinguishing between addition and subtraction. An inspection of Figure 9 suggests an overall 'yes' bias for 3-year-olds.

For 4-year-olds, responses varied across the six transformation types ($\chi 2$ (5) = 28.12, $p < .001$; see Figure 10). Wilcoxon signed rank tests revealed that 4-year-olds were able to differentiate between addition and subtraction, and between addition and '-3+1'. The results also suggest that they were marginally able to differentiate between addition and 'Knock'. However, they did not recognize that the '-A+A' trial and the '-A+B' do not increase the numerosity of the set (see Table 6).

For 5-year-olds, responses varied across the six transformation types ($\chi 2$ (5) = 35.49, p < .001; see Figure 11). Wilcoxon signed rank tests revealed that 5-year-olds can distinguish between addition and subtraction, '-3+1', and 'Knock'. Nevertheless, they continued to show signs of difficulty in differentiating ‘–A+A’ and ‘-A+B’ from addition (see Table 6).

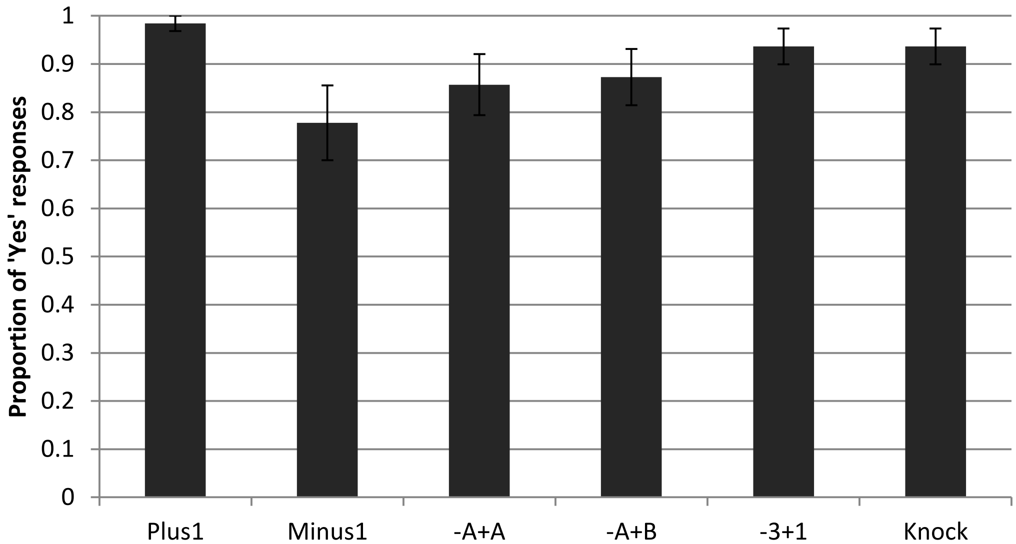


*Figure 9*. Three-year-olds’ proportion of ‘yes’ responses for hidden sets. Error bars represent standard errors of the mean.

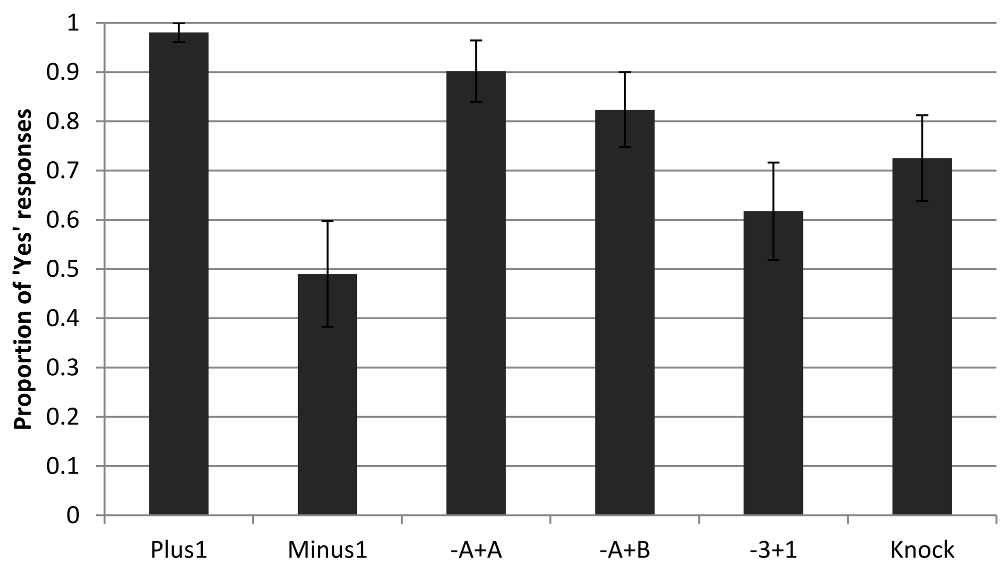


*Figure 10*. Four-year-olds’ proportion of ‘yes’ responses for hidden sets. Error bars represent standard errors of the mean.

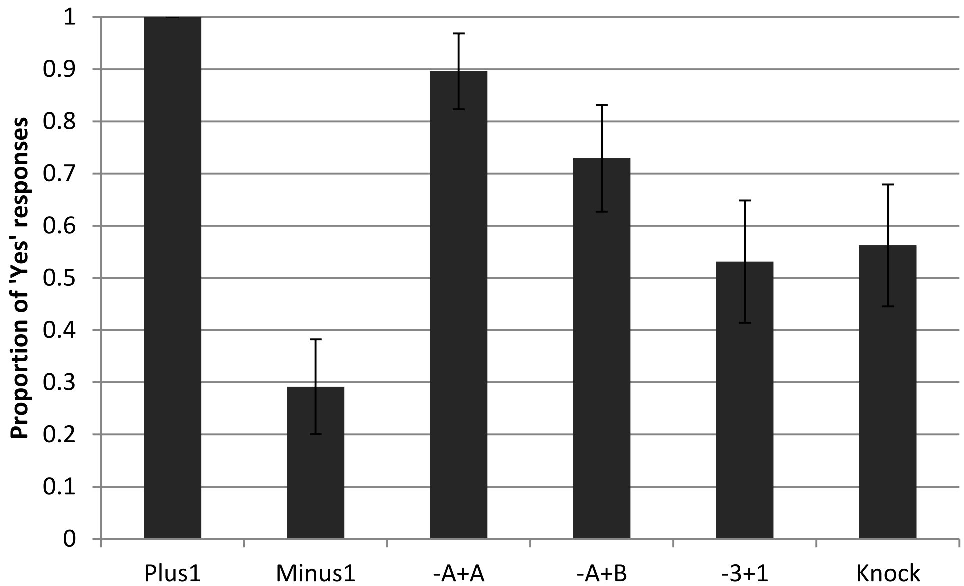


*Figure 11*. Five-year-olds' proportion of 'yes' responses for hidden sets. Error bars represent standard errors of the mean.

| | **3-year-olds** | **4-year-olds** | **5-year-olds** |
|---|---|---|---|
| **Plus 1 – Minus 1** | -2.34 (.017)** | -2.99 (.003)*** | -3.38 (.001)*** |
| **Plus 1 – Minus A Plus A** | -1.90 (.058) | -1.34 (.18) | -1.34 (.18) |
| **Plus 1 – Minus A Plus B** | -1.80 (.072) | -2.07 (.038) | -2.23 (.026) |
| **Plus 1 – Minus 3 Plus 1** | -.96 (.34) | -2.72 (.007)*** | -2.75 (.006)*** |
| **Plus 1 – Knock** | -96 (.34) | -2.36 (.018)** | -2.60 (.009)*** |

*Table 6.* Z-scores (p-values) from Wilcoxon signed rank tests for 3-, 4-, and 5-year-olds. *** significant difference with Bonferroni correction (adjusted alpha = .01). ** marginally significant difference with Bonferroni correction.

Consistent with previous experiments, Experiment 4 showed that 3-year-olds likely lack the ability to reason about unspecified numerosities, and this ability emerges at 4 years of age. It is important to note that 4-year-olds in this experiment distinguish between addition and '-3+1' but continue to have difficulty with reasoning about '-A+A' and '-A+B'. This provides evidence that children's difficulty with reasoning about the effect of '-A+A' in the previous experiment cannot be explained by the fact that '-A+A' involves two steps, because '-3+1' also involves two steps and children are succeeding at differentiating between 'Plus 1' and '-3+1' by the age of 4. These results also suggest that children are likely not relying on a 'last-action' heuristic. Given that the last action of '-A+A', '-A+B', and '-3+1' all involve the addition of an individual, if children are relying on this heuristic, they should respond that there are more elements when tested on each of these three types of transformations. But children respond that '-3+1' does not increase the numerosity of a set, and this rules out the possibility that children simply recruit a 'last-action' heuristic.

Nevertheless, the current experiment also showed that 5-year-olds have trouble recognizing that the subtraction and addition of the same individual and the substitution of an individual do not affect the numerosity. This was somewhat inconsistent with findings from Experiment 3, which showed that by 5, children can reason about unspecified numerosities, and specifically, that 5-year-olds understand how adding and removing the same individual does not change the numerosity of visible sets. There are two possible explanations for these conflicting findings.

First, children's difficulty with the '–A+A' and '–A+B' trials may be explained by an unwillingness to answer 'no' to the question, "Are there more [bells] in the box now?" when the numerosity of a set remains the same. For example, a handful of children responded "the same" to the test question about ‘more’ on ‘–A+A’ and ‘–A+B’ trials, but these same children had no difficulty answering yes/no for addition and subtraction. Moreover, this difficulty may be driven by the pragmatics of the task. For example, when asked whether there were more elements in the set, children had to generate a list of possible responses – i.e., more, less, the same number. In the case of visible sets, it may be easier for children to generate all three responses but in the case of hidden sets, ‘less’ may be a better and more salient response than ‘the same number.’ In other words, children may have difficulty generating ‘the same number’ as a possible alternative response to a question about ‘more’. This may explain why children have no difficulty with the ‘-A+A’ trial in Experiment 3 but exhibit difficulty in the current experiment. Children who were presented with hidden sets may only be able to generate ‘less’ as a possible alternative. This may explain why children have particular difficulty answering ‘no’ on trials in which the numerosity remains unchanged (i.e., ‘-A+A’, ‘-A+B’). [13]

Second, it is possible that children interpret ‘more’ non-comparatively – i.e., as if it means ‘some’. The evidence for the ‘more = some’ interpretation comes from children’s justifications. For example, when the experimenter took away one object or knocked on the box, some children answered that the box contained more elements because “there’s some.” This suggests that in the context of the Box Task, they may not always be interpreting ‘more’ comparatively. However, children do appear to interpret ‘more’ comparatively for some types of transformation (e.g., ‘Plus

---

[13] This difficulty may also be related to the semantics of adjectives. For example, pairs of adjectives such as tall/short are different from pairs of adjectives such as dead/alive. A person who is *not* tall does not mean that the person is short, but a person who is not dead does mean that the person is alive. This difference in meaning highlights a distinction that is documented in the semantics literature: tall/short are gradable adjectives and dead/alive are absolute adjectives. It is possible that children are interpreting ‘more’ as if it is an absolute adjective, when in fact it behaves more similarly to gradable adjectives, and thus, they refuse to say no to the question about ‘more’ unless the set contains fewer items.

1', 'Minus 1', '-3+1'), suggesting that the "more = some" interpretation cannot account for all of the data. To probe the robustness of findings from previous experiments, and determine whether 5-year-olds can represent unspecified numerosities, I conducted Experiment 5 with the objective of removing some of the task demands children experienced in Experiment 4.

## Experiment 5: Children's reasoning about numerosity changes with quantified sets of hidden objects

Results from Experiment 4 provide inconclusive evidence with respect to children's representation of unspecified numerosities, and one possibility was that children were not clear on the numerical aspect of the task. To address this concern, in Experiment 5, I labelled the hidden sets with a numerical expression (e.g., "there are 29 bells in the box") and used the numerical expression in a 'more than' construction (e.g., "are there more than 29 bells in the box now?"). The objective of introducing these changes was to highlight the task of numerical comparison that was required of the child. Lastly, given that 5-year-olds may have difficulty with reasoning about numerosity change with hidden sets, I also tested 6-year-olds in this experiment.

### Method

#### Participants

Forty-one children between the ages of 3 and 6 participated. There were 13 3-year-olds (range = 3 years 2 months to 3 years 11 months; mean = 3 years 7 months; 6 boys), 10 4-year-olds (range = 4 years 4 months to 4 years 11 months; mean = 4 years 8 months; 5 boys), 9 5-year-olds (range = 5 years 0 months to 5 years 10 months; mean = 5 years 4 months; 4 boys), and 9 6-year-olds (range = 6 years 0 months to 6 years 11 months; mean = 6 years 4 months; 4 boys). They were recruited in daycare centres and kindergartens in Kitchener-Waterloo and nearby areas.

Procedures and stimuli

*Box Task (Number Word Trials)*. The Box Task (Number Word Trials) is almost identical to Experiment 4 with the exception that the experimenter used a number word in the introduction and in the test question, and the test question was only asked once after transformation.

After the familiarization phase (with Hoberman's Sphere) and the labeling phase, all children first received three training trials in which the experimenter used a small number word – 'two' – to label a set of objects. For the training trials and the subsequent test trials, the experimenter first showed that the box was empty, and then s/he said "I'm going to put [2] [bells] into the box. There are [2] [bells] in the box." The experimenter performed the transformation, and asked "Are there more than [2] [bells] in the box now?"

After the three training trials, the experimenter immediately began the test trials. The only difference between the training trials and the test trials was the number words that were used. At test, different number words were used depending on the child's age: 'nine' was used for three-year-olds and four-year-olds, 'nineteen' was used for five-year-olds, and 'twenty-nine' was used for six-year-olds. The transformations were identical to Experiment 4: adding one object, removing one object, taking away one object and putting it back ('-A+A'), taking away one object and putting back another object of the same kind but of a different colour ('-A+B'), taking out three objects and putting another one back ('-3+1'), knocking on the box.

Results and Discussion

*Overall performance on Box Task (Number Word Trials)*

Preliminary analyses found no order or gender effects ($p$'s > .64), so these variables were not included in subsequent analyses.

First, I analyzed children's performance on the training trials involving small number words by age group. Table 7 shows that as children grow older, they perform better on reasoning about sets that were labeled with small number words such as ‘two’. Binomial tests showed that the proportion of 6-year-olds who responded correctly on at least 2 out of 3 trials for sets labelled with ’two’ differed significantly from 3-year-olds, $p$ = .02, and marginally from 4-year-olds, $p$ = .07, but did not differ from 5-year-olds, $p$ = .13. [14]

| | **3-year-olds** | **4-year-olds** | **5-year-olds** | **6-year-olds** |
|---|---|---|---|---|
| **At least 2 out of 3 correct** | 7 | 7 | 7 | 9 |
| **N** | 13 | 10 | 9 | 9 |

*Table 7.* Number of children who answered at least 2 out of 3 questions correctly on trials labelled with ‘two’.

For the test trials involving larger number words (i.e., nine, nineteen, twenty-nine), children received a score of 1 for each transformation that they answered correctly. An overall proportion correct on test trials averaging the scores across all transformation types was computed for each child and was used as a dependent variable. ANOVA revealed a main effect of Age Group, $F(3,37) = 6.47$, $p = .001$. Tukey HSD revealed that 6-year-olds (87.7%) differed significantly from 3- and 4-year-olds ($p$*’s* < .042), and were marginally different from 5-year-olds ($p = .071$). To better characterize children’s performance at each age, I conducted separate analyses for 3-, 4-, 5-, and 6-year-olds.

[14] Arguably, children who fail to answer correctly on at least 2 trials on ‘two’ may have difficulty with comprehending the test question, and should thus be excluded from the analyses. However, due to my relatively small sample sizes in each age group, all children were included in the analyses.

*Relationship between age and trial type*

For each age group, I first analyzed whether children's responses differed on the six different transformation types using Friedman's ANOVA with proportion of 'yes' responses as the dependent variable. Then, similar to the analyses in Experiment 4, I compared whether children are sensitive to the distinction between 'Plus 1' and the other transformation types using Wilcoxon signed rank tests with a Bonferroni corrected alpha level of .01 (.05/5 comparisons for each age group). Figures 12-15 display children's proportion of 'yes' responses for each transformation type.

For 3-year-olds, responses did not vary across the six transformation types ($\chi 2$ (5) = 5.74, $p$ = .33; see Figure 12). Wilcoxon signed rank tests revealed that the use of a quantified expression did not help 3-year-olds in reasoning about the effects of transformation for hidden sets (see Table 8). This may not come as a surprise given that half of the 3-year-olds appear to have difficulty even with training trials on 'two', which is a numeral that is likely within their range of known number word knowledge.

For 4-year-olds, responses varied across the six transformation types ($\chi 2$ (5) = 10.9, $p$ = .007; see Figure 13). The use of a quantified expression revealed a relatively similar pattern of results for the 4-year-olds compared to Experiment 4 (see Table 8). They continued to show difficulty recognizing that the addition and subtraction of the same individual does not increase the numerosity of a set. However, 4-year-olds in the current experiment were able to recognize that the substitution of an individual does not increase the numerosity of a set, while failing to notice that knocking on the box was numerically-irrelevant.

For 5-year-olds, responses did not vary across the six transformation types ($\chi 2$ (5) = 8.3, $p$ = .14). This is somewhat surprising given that an inspection of Figure 14 suggests that 5-year-olds show some understanding of the effects of transformation such as addition, subtraction, and '-3+1' (see Table 8).

For 6-year-olds, responses varied across the six transformation types ($\chi 2$ (5) = 33.2, *p* < .001; see Figure 15). Wilcoxon signed rank tests revealed that they were able to differentiate between all the transformation types (see Table 8).

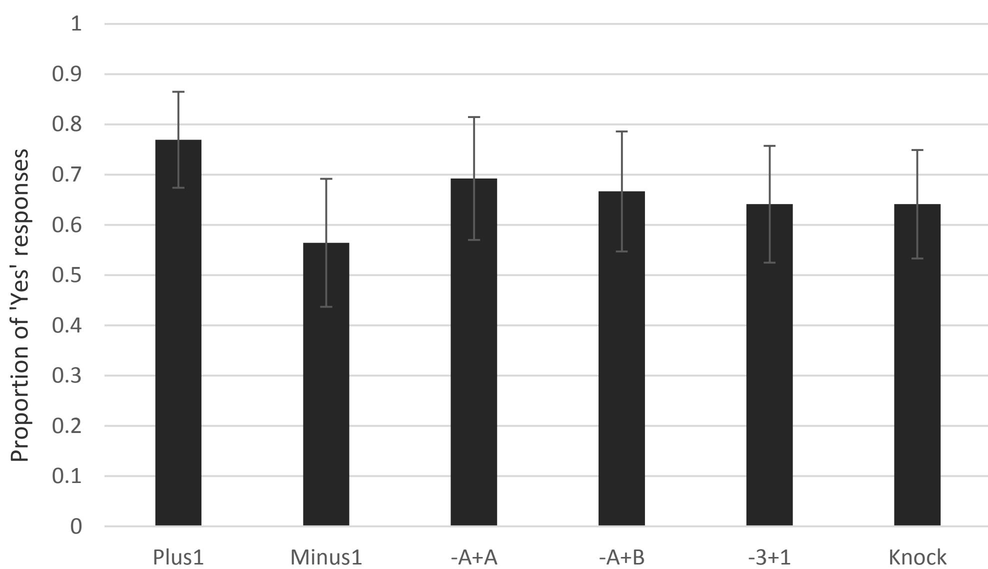


*Figure 12*. Three-year-olds' proportion of 'yes' responses for hidden sets labelled with a number word. Error bars represent standard errors of the mean.

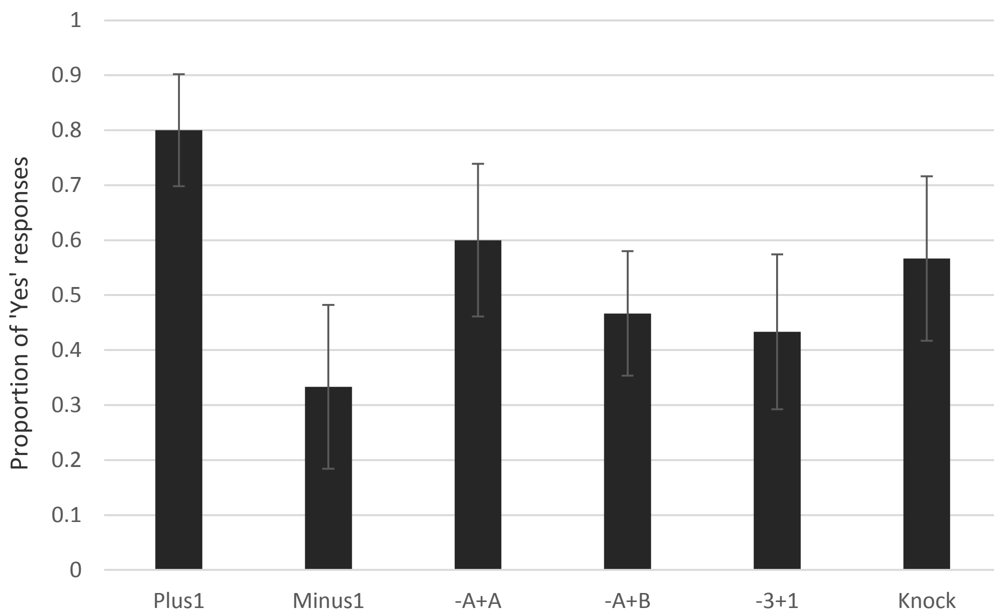


*Figure 13*. Four-year-olds' proportion of 'yes' responses for hidden sets labelled with a number word. Error bars represent standard errors of the mean.

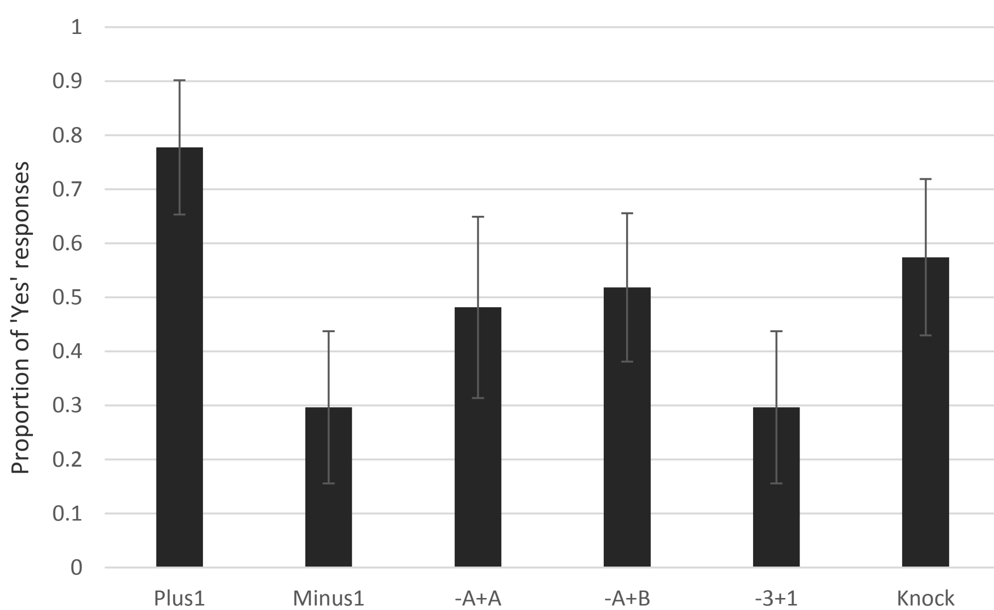


*Figure 14*. Five-year-olds' proportion of 'yes' responses for hidden sets labelled with a number word. Error bars represent standard errors of the mean.

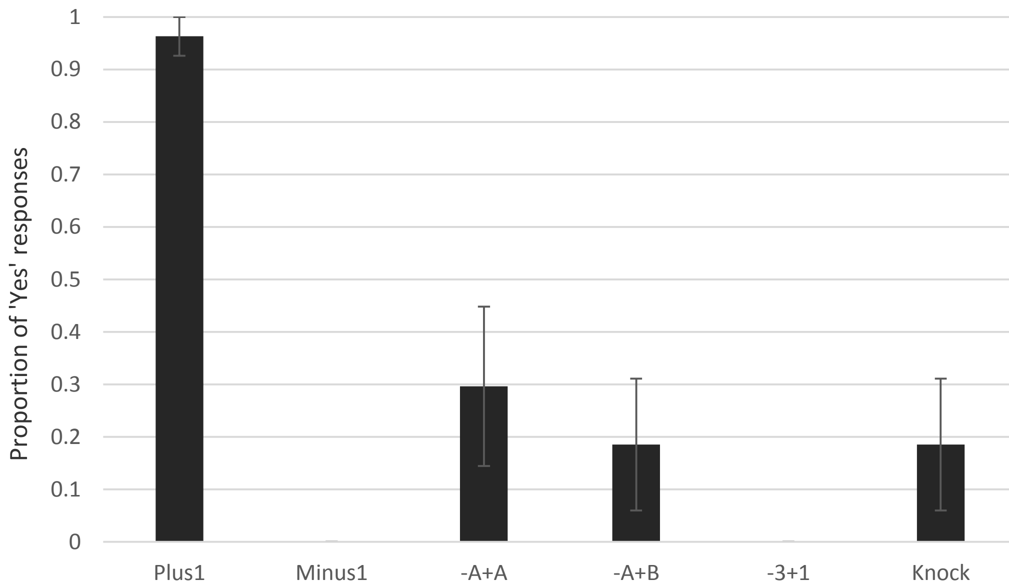


*Figure 15*. Six-year-olds' proportion of 'yes' responses for hidden sets labelled with a number word. Error bars represent standard errors of the mean.

| | 3-year-olds | 4-year-olds | 5-year-olds | 6-year-olds |
|---|---|---|---|---|
| **Plus 1 – Minus 1** | 1.83 (.068) | -2.38 (.017)** | -2.01 (.044) | -2.89 (.004)*** |
| **Plus 1 – Minus A Plus A** | -1.29 (.20) | -1.79 (.74) | -.97 (.34) | -2.46 (.014)** |
| **Plus 1 – Minus A Plus B** | -.95 (.34) | -2.38 (.017)** | -1.38 (.17) | -2.24 (.008)*** |
| **Plus 1 – Minus 3 Plus 1** | -1.59 (.11) | -2.37 (.018)** | -2.21 (.027) | -2.89 (.004)*** |
| **Plus 1 – Knock** | -1.84 (.066) | -1.78 (.075) | -1.02 (.31) | -2.64 (.008)** |

*Table 8*. Z-scores (p-values) from Wilcoxon signed rank tests for 3-, 4-, 5-, and 6-year-olds. *** significant difference with Bonferroni correction (adjusted alpha = .01). ** marginally significant difference with Bonferroni correction.

Results from Experiment 5 suggest that children do not reason about unspecified numerosities until 6 years. This is somewhat inconsistent with the results of Experiments 3 and 4, showing that by around 5 years of age, children can reason about the effects of transformation algebraically. In fact, an inspection of the graph suggests that performance of 5-year-olds improved in the current experiment compared to Experiment 4. The lack of significant results from the Wilcoxon signed rank tests could be due to the relatively small sample sizes and insufficient power to detect a significant difference between the paired trial types.

Regardless of the pattern of results, one may argue that Experiment 5 did not necessarily tap into children's representation of unspecified numerosities. Indeed, it is possible that 4-year-olds know the meaning of 'nine' and can recruit particular facts about 'nine' to reason about addition and subtraction. The same is true for 5 and 6-year-olds who may know the meaning of 'nineteen' and 'twenty-nine'. Nevertheless, it provides evidence that labelling hidden sets of objects with number words increases their ability to reason about numerosities, and children's pattern of responses is similar to those of Experiments 3 and 4, where no numerosity was given.

## Chapter Three Discussion

Three experiments explored the development of children's capacity to reason about unspecified numerosities – i.e., that the addition of one object necessarily increases the numerosity of a set regardless of set size. In each of the experiments, children were asked to reason about how different types of transformation changed the numerosity of a set of objects. Specifically, children were asked whether there were more elements in the set after the transformation. To assess when children begin to show a capacity to reason about unspecified numerosities, set size was manipulated such that children could not recruit representations of particular numbers (Experiment

3) and sets of objects were presented out of view of children, requiring them to reason about unspecified numerosities to conclude whether or not the transformation affected the numerosity of hidden sets (Experiment 4).

In Experiment 3, children were asked to reason about how the addition of one element, subtraction of one element, the combined operation of addition and subtraction, and moving one object affected the numerosity of a visible set. Results from this experiment showed an interesting developmental trajectory: 3-year-olds were able to reason about the effect of transformations on particular numerosities, but were unable to reason about unspecified numerosities; starting at the age of 4, children understood whether addition, subtraction, and rearranging objects increase the numerosity of sets; and by 5, they also correctly reasoned that the addition of one object and subtraction of one object cancels each other out and thus does not affect numerosity.

To provide stronger evidence for children's ability to reason about unspecified numerosities, children were asked to reason about the effects of transformations on the numerosity of hidden sets in Experiment 4. This completely eliminates perceptual access to the sets. Although the overall analysis revealed that 5-year-olds had difficulty reasoning about the effects of transformations, trial type analyses suggested a pattern that was mostly consistent with the findings from Experiment 3. Specifically, 5-year-olds demonstrated an understanding of whether addition, subtraction, and non-numerical transformation such as knocking on the box (similar to rearranging objects in the previous experiment) affected the numerosity of a set. However, they had trouble differentiating between addition and the subtraction and addition of the same individual ('-A+A') and the substitution of an individual ('-A+B').

Thus, Experiment 4 provided somewhat inconclusive evidence regarding when children possess the ability to reason about unspecified numerosities. In Experiment 5, I examined this conflicting finding and probed children's understanding of the effects of transformation on hidden sets by modifying the paradigm in Experiment 4. Specifically, hidden sets were labelled with a numerical expression to highlight the task of numerical comparison that was required of children. Due to the poor overall performance of 5-year-olds in Experiment 4, 6-year-olds were also recruited for this experiment. Results showed a pattern of results that is somewhat similar to Experiment 4: 4-year-olds are able to distinguish between addition and subtraction, and between addition and '-3+1'. Although the transformation types analyses were not significant for the 5-year-olds, this is likely due to insufficient sample size, and the pattern of results (see Figure 14) suggested that they were able to distinguish between some of the transformation types. By 6 years, there is strong evidence that children are capable of reasoning about whether all of the numerical and non-numerical transformations increase numerosity, and this provides evidence that they have the capacity to reason about unspecified numerosities.

Overall, 3-year-olds fail to reason about unspecified numerosities, as suggested by their poor performance on reasoning about visible sets and hidden sets across all three experiments. Four-year-olds show emergent understanding that the addition of one object increases the numerosity of a set, but the subtraction of one object and the combined operation of 'Minus 3' and 'Plus 1' does not. However, the finding that 4-year-olds fail to distinguish between addition and the non-numerical transformations (i.e., 'Move 1' and 'Knock') may raise the question of whether 4-year-olds truly demonstrate an ability to reason about number in the case of addition and subtraction. I will return to this point later in this section. Finally, both 5- and 6-year-olds demonstrate the ability to reason about unspecified numerosities. Although 5-year-olds have difficulty in reasoning that '–A+A' and '–A+B' do not increase numerosity, this may be due to a difficulty in negating a question about 'more' when the numerosity of the set remains unchanged (i.e., a bias to respond 'no' only when the set contains fewer objects). Moreover, given 4-year-olds' emergent understanding of addition and subtraction, 5-year-olds' success on reasoning about the effect of '–A+A' in visible sets, and 6-year-olds' success on reasoning about the effects of all transformation types, there is reason to suggest that by 5, children have the ability to reason about unspecified numerosities.

*An alternative explanation: quantification by continuous cues and not number?*

Although the current studies were designed to test children's understanding of transformations on the *numerosity* of a set, an alternative strategy is to perform comparison on the basis of continuous extent of the set. There are at least two ways to quantify a set of objects by continuous extent in the current experiments. First, children can quantify on the basis of the total mass of the objects. For example, when one Lego block was added to a set of Lego blocks in Experiment 3, not only was the number of Lego blocks increased, but the amount of "Lego block stuff" (or the amount of plastic) also increased. Similar arguments can be made for Experiments 4 and 5 (e.g, when a Lego block went into the box, both the number of blocks and the amount of 'Lego block stuff' increased). While the number of objects and their total mass are confounded in the study, there are two important reasons to interpret children's performance as being based on reasoning about number. First, it has been argued that the construal of an entity may influence the kind of quantification people perform on an object or a set of objects. Prasada and colleagues have argued that humans have an intuitive sense of how entities are to be construed (Prasada, Ferenz, & Haskell, 2002). For example, whether we construe an entity as an object kind or a substance kind is dependent upon its structure (e.g., shape, regularity of structure). If an entity is construed as an object kind (e.g., honey-dipper), then its structure is perceived to be non-arbitrary, whereas if an entity is construed as a substance kind (e.g., wooden stuff), then its structure is perceived to be arbitrary. To illustrate, when a honey-dipper is construed as a honey-dipper, it has the structure it does *because* it is a honey-dipper; the structure is not accidental. However, when we construe an entity as a piece of wood, we do not think that it has the structure it does because it is wood; its structure is more likely to be accidental. This arbitrariness in structure affects the process of quantification: object construals call for quantification by number, whereas substance construals call for quantification by continuous cues (Prasada et al., 2002). In the current experiment, almost all of the stimuli were object kinds. The only exception might be rocks. However, the rocks used in the current study all have a regular shape (i.e., they are round like pebbles), and thus, following Prasada's proposal, it is likely that the rocks are still construed of as objects.

In addition, in the current experiments, nouns were used with count syntax (e.g., "Does Winnie The Pooh have more bells now?"), thus signifying quantification by number. Previous studies have shown that children are sensitive to such syntactic cues when asked to make 'more' judgments (Barner & Snedeker, 2005; Gathercole, 1985; see also Gordon, 1985). For example, when 4-year-olds were asked questions such as "Who has more shoes/toothpaste?", where one alternative contained a greater number of objects but less overall mass (e.g., six tiny shoes, six tiny portions of toothpaste) and the other alternative contained fewer objects but with greater overall mass (e.g., two giant shoes, two giant portions of toothpaste), they responded that six tiny shoes were more shoes than two giant shoes, and that two giant portions of toothpaste were more toothpaste than six tiny portions. Importantly, this pattern of finding has also been demonstrated for nouns that can be used flexibly with mass syntax or count syntax. For example, when asked "Who has more *rocks*?" children responded on the basis of the number of rocks, but when asked "Who has more *rock*?" they responded on the basis of total mass (Barner & Snedeker, 2005). These results provide strong evidence that children are sensitive to syntactic cues when making 'more' judgments. Given that only count nouns were used in the current study, there is reason to suggest that children are responding to the question about 'more' on the basis of number.

However, this argument leaves open one question: children in Piaget's conservation of number task responded on the basis of length even though natural object kinds were used (e.g., beads). This raises the question of whether the use of count syntax or the construal of object kind necessitates quantification by number. I speculate that in the Piagetian task, the conflict between perceptual cues and the numerosity of the set is too distracting and makes a comparison of numerosity difficult. One piece of evidence suggests that this might be the case. Using Piaget's conservation of number task, Winer (1978) found that children perform better on the conservation of number task when shown small sets (two, three) than large sets (five, six). One explanation is that children know the numerosity of small sets, thus making the perceptual conflict less distracting. If this explanation is true, then it is possible that when perceptual cues are completely removed (e.g., children who were shown hidden sets of objects in the current study), children are more likely to attend to the syntactic cue and to the nature of the object kind in a quantity judgment task. Moreover, even when shown visible sets, it can be argued that the only trial that induces a perceptual conflict is the non-numerical transformation in which one object is moved. Thus, there is reason to believe that children in the current study did attend to the syntactic cue and construed the stimuli as object kinds, causing them to reason about the *numerosity* of the sets.

A second alternative to number is contour (Experiment 3 only). For example, when one Lego block was added to a set of Lego blocks, both the number of objects and the contour of the set of objects increased. If children are adopting a 'contour strategy', then one may expect them to deploy it for both small and large sets in Experiment 3, and their performance on small sets should be similar to that on large sets. However, this is not what I found. As shown in Experiment 3, children across all age groups were consistently better on comparing small sets than large sets. One can argue that the 'contour strategy' is only applied for reasoning about large sets, and that children are recruiting parallel individuation when shown small sets. If this is the case, children should be more likely to respond that the set contains 'more' in the case of the 'Move 1' trial than in the 'Plus 1' trial because the contour has increased more in the former than in the latter case. However, as shown in Figures 6-8 across the 3-, 4-, and 5-year-olds, the proportion of 'yes' responses was numerically higher for 'Plus 1' trials than for 'Move 1' trials, which is not predicted on the 'contour strategy' account.

*Relating to previous research: Reasoning about unspecified numerosities and number word acquisition*

As noted in the Introduction, much previous research on children's understanding of number words is motivated by the assumption that children's understanding of numerical symbols may reveal the nature of their numerical knowledge. The present experiments adopted a different strategy in examining children's numerical knowledge. In particular, children were asked to reason about changes in numerosity, and not how number words change their application in the context of numerical and non-numerical transformations.

Previous studies on number word meanings and the current set of experiments reveal a similar pattern of results on children's developing numerical knowledge. First, Experiment 3 showed that by 5, children understand how rearranging objects, addition, and subtraction affect the numerosity of sets. A similar pattern was also revealed in trial type analyses in Experiment 4. This converges with findings from Lipton and Spelke (2006), who showed that 5-year-olds understand that the number word that labels a large set of objects changes when objects are removed from or added to the set, but the number word remains the same when objects are simply rearranged. Importantly, the current experiments are the first to show that even before children turn 5, they have the capacity to reason about some aspects of unspecified numerosities.

Second, Experiment 3 also showed that by 3, children are capable of reasoning about how rearranging objects, addition, and subtraction affect the numerosity of small sets. This is consistent with studies on children's understanding of small number words and their sensitivity to numerosity changes in the small number range (e.g., Gelman, 1977; Bullock & Gelman, 1977; Condry & Spelke, 2008; Schaffer, Eggleston, & Scott, 1974).

In addition, across the three experiments, it has been shown that 3-year-olds fail to correctly reason about whether addition and subtraction increase the numerosity of large sets. This finding has implications for the conflicting findings on children's understanding of number word application (Sarnecka & Gelman, 2004; Condry & Spelke, 2008; Brooks et al., 2013). In previous studies, when presented with one set of objects and two number words, 3-year-olds demonstrate understanding of the application of number words. They apply a different number word when a numerical transformation is performed, and the original number word when a non-numerical transformation is performed. However, when shown two sets of objects, they are at chance. Using experiments on novel objects that were similar in design to the experiments on number word application, Brooks et al. (2013) argue that children's success on the one-set task reflects sensitivity to pragmatic cues provided by use of the alternative label, and not genuine knowledge about the application of number words. Specifically, they argue that children first interpret number words as quantifiers such as 'some' and 'all'. Thus, they understand that number words denote numerosities, and they may also know that a change in numerosity can affect the application of number words. However, they do not necessarily know that a change of *one* element is sufficient to change the application of number words. Importantly, in the one-set task, children can succeed by making a pragmatic inference. When given an alternative number word that directly contrasts with the original label, the experimenter is inviting the child participants to think that the new label is relevant to the change in numerosity. Therefore, on Brooks et al.'s account, understanding that adding one element changes the application of number words does not necessarily reflect knowledge of the semantics of number words, but can instead be explained by the use of pragmatic inference. This contrasts with Sarnecka and Gelman's view, which argues that the knowledge that a change of one element is sufficient to change number word application *is* part of the semantics of number words. Nevertheless, on both accounts, children at the age of 3 should understand that number words are about numerosities, and that the application of number words changes in the context of numerical transformation but not in the non-numerical transformation. Yet the current studies found that 3-year-olds fail to identify whether addition and subtraction affect the numerosity of sets. What may explain this discrepant finding?

One possibility is that the current experiments included the '-A+A' trial, which was not investigated in previous experiments that used the one-set paradigm. Perhaps this explains why 3-year-olds in previous studies demonstrate some knowledge of numerical transformation, but those in the current studies do not (at least not with large sets). To evaluate this possibility, I removed the '-A+A' trial and analyzed children's performance against chance in Experiment 3. Post-hoc analyses showed that 3-year-old children still performed at chance for large number trials ($M$ = 54.4% vs. $M$ = 46.1% when '-A+A' was included; $p$'s = .49, one-sample t-tests). Therefore, the inclusion of this trial type does not seem to be responsible for the discrepancy.

Another possibility is that it may be harder for children to respond to questions about more/less (as in the current experiments) than about same/different (as in previous studies on number word meanings). Future studies can investigate how children's reasoning about 'more' develops with their corresponding understanding of number word meanings.

*Relating to previous research: Interference of continuous cues in tracking numerosity changes*

Not surprisingly, the current results on trials in which a spatial transformation was performed (e.g., moving 1 object) in Experiment 3 replicate Piaget's (1952) finding that children are often influenced by non-numerical cues when asked to track numerosity. Specifically, in the classic Piagetian conservation of number paradigm, children are presented with two rows of objects that are in one-to-one correspondence. In some contexts, the experimenter then moves one row of objects to make it longer and asks whether there is still the same number of objects. Children younger than 6 often respond that the longer row has 'more' objects, attending to non-numerical cues such as length in response to a question about the numerosity of a set. In the current experiments, 3- and 4-year-old children attended to non-numerical cues and responded that a set contained more elements after one object was moved in the context of large sets. Nevertheless, by 5, children became less inclined to respond that the numerosity had increased under non-numerical transformation. Thus the present studies find earlier competence than Piaget did, suggesting that the tendency to rely on non-numerical cues in numerical tasks diminishes by around 5 years of age.

*Relating to previous research: Young children's interpretation of the relationship between set identity and the numerosity of sets*

My findings also extend a recent study on children's use of one-to-one correspondence cues to track changes in numerosity (Izard, Streri, & Spelke, 2014). They found that children can use one-to-one correspondence to track small sets (2 or 3) but not large sets (5 or 6) of items. In this study, 2 1/2–year-old children who had not learned number word meanings beyond 'four' were given one-to-one correspondence cues to track sets of items. They were presented with a set of finger puppets that were sitting on branches of a toy tree; the puppets and the branches either corresponded one-to-one or they did not. This set up made clear when there was a difference of one puppet (e.g., 5 puppets with 6 branches). Then, the experimenter placed the puppets in an opaque box, and acted out stories that resulted in a transformation of the set (e.g., addition of one puppet, subtraction of one puppet, removing one puppet and returning it to the set, substituting one puppet with an identical-looking puppet). Children were then allowed to search in the box after the transformation, and search times were used as a dependent variable. For example, if children understand the effect of addition, then they should understand that when 1 puppet is added to a set of 5 puppets, it results in a set of 6 puppets. Thus, after returning 5 puppets to the tree, they should keep searching for the 6$^{th}$ puppet. Izard and colleagues found that for large sets, children could not compute the effects of adding one or subtracting one from the initial set on the correspondence between branches and puppets. However, they were successful when the transformation involved the addition and removal of the same puppet, searching longer for the 6$^{th}$ puppet when the set was expected to contain 6 puppets. Under a similar transformation, one in which a puppet was replaced with an identical-looking puppet, children failed to search longer for the 6$^{th}$ puppet.

Findings from this study are similar to the current experiments in two ways. First, consistent with Izard et al.'s study, I found that 3-year-olds in Experiment 3 failed to distinguish the addition of one element and subtraction of one element when shown large sets of objects. Second, similar to Izard et al., I also found that when shown small sets, children are capable of reasoning about the effects of addition and subtraction on the numerosity of sets. However, my findings diverge in the case of the addition and removal of the same individual item or puppet. Unlike Izard et al., both 3- and 4-year-olds in the current experiments incorrectly reasoned that when an individual was removed and returned back to the set (i.e., -A+A), the numerosity of the set changed. Thus, findings from the current experiments reveal a more protracted development of understanding the effects of removing and returning the same individual on the numerosity of a set.

A difference in methodology might explain this discrepant finding. To succeed on the current experiments, children have to understand how removing an object and returning it to the set affects its numerosity. However, in Izard et al.'s study, children could succeed on the analogous transformation without thinking about its effect on numerosity. As argued by Izard and colleagues, it may be that young children (or non-CP-knowers) did not interpret the one-to-one correspondence between branches and puppets as determining numerical equality between them, but rather interpreted each of the branches as referring to one individual. In other words, they may not have interpreted the presence or absence of one-to-one correspondence as indicating that "the number of branches is equal/not equal to the number of puppets", but rather interpreted the situation as "each branch has one puppet" (see Sarnecka & Gelman, 2004; Sarnecka & Wright, 2013, for evidence that non-CP-knowers fail to use one-to-one correspondence cues to infer the numerosity of sets). Therefore, children in Izard et al. may have succeeded on the transformation that is equivalent to the '-A+A' trial in the current study because it may be easier to reason about the effect of '–A+A' at the level of individuals than at the level of its effect on the numerosity of sets.

*Conclusions about the nature of representations of unspecified numerosities*

The pattern of results reported in this Chapter raises important questions about the nature of representations of unspecified numerosities. The data clearly indicate a developmental trend in children's reasoning about how transformations to a set affect numerosity. But what is developing?

On one view, children are born with the capacity to represent unspecified numerosities. It is important to note that hypothesizing that an ability is innate does not mean infants possess that ability from birth (e.g., puberty is an innate biological mechanism that does not occur until our early teenage years). On the nativist view, children do not acquire representations of unspecified numerosities because the representation of particular numerosities presupposes one to represent unspecified numerosities, and thus, it does not make sense to hypothesize that children construct representations of unspecified numerosities. Rather, as children grow older, they develop an ability to *apply* reasoning principles that operate over representations of unspecified numerosities, and the principles that operate over such representations require exposure to sets of objects and numerical language to fully develop. This may explain why there was an effect of block in Experiment 3: children who were tested on small number trials first performed better than those who were tested on large number trials first. Children may need to engage in a numerical environment before recruiting representations of unspecified numerosities. On this account, what is changing over the early developmental years is how the numerical principles such as addition and subtraction can be applied to representations of unspecified numerosities. This explains why children's understanding of the effects of transformation on numerosity develops gradually: they first learn that addition but not subtraction increases the numerosity of a set, followed by an understanding of how the two numerical operations cancel each other out. Therefore, what develops is not the representations of unspecified numerosities per se but the ability to reason numerically.

On the other hand, it can also be argued that children construct representations of unspecified numerosities as they grow older. One possibility is to construct representations of unspecified numerosities from representations of particular numbers. As the current findings suggest, children are capable of reasoning about small sets (1-3) before reasoning about large sets (14-16) and hidden sets. On the constructivist account, children first notice how numerical principles apply to representations of particular numerosities, and then make an induction of these principles over all possible numbers. Specifically, they know that (1) small sets of objects have a particular numerosity, and that (2) there is an order for small sets, i.e., a set of three is more than a set of two, and a set of two is more than a set of one. It has also been shown that even infants can perform addition over small sets of individuals (Wynn, 1992b). Thus, the analogies children observe in the small number range may be powerful enough to allow them to make a crucial induction: any sets of objects will have a numerosity x, which will be different from x + 1.

*Conclusions*

Since Piaget, the bulk of research on children's developing understanding of number focuses on how children represent and reason about representations of particular numerosities. Studies on children's understanding of numerical knowledge that applies to all possible numbers are rare (e.g., Evans, 1983; Hartnett, 1991). The present findings shed light on both the extent and limits of children's representations of unspecified numerosities. Across three experiments, I found that by 5, children demonstrate knowledge about how numerical and non-numerical transformations affect the numerosity of a set. Although children younger than 5 fail to reason entirely correctly about unspecified numerosities, by 4, they begin to show a general understanding of addition and subtraction.

In sum, by investigating how children reason about the effects of numerical and non-numerical transformations in the context of numerical comparison, we gain insight into children's representation of numerical knowledge. The present studies contribute to our understanding of children's developing numerical reasoning by providing converging evidence that children develop the ability to reason about unspecified numerosities at around the age of 5. Future studies should investigate how the different aspects of numerical knowledge such as the one-to-one correspondence principle and the successor principle develop with children's ability to reason about unspecified numerosities. This will provide important insights into the origins and development of the concept of the natural numbers.

# Chapter Four: General Conclusions

Using numerical comparison as a case study, this dissertation explored the limits and extent of representations of number through the study of the development of representations of particular numerosities and unspecified numerosities.

I began by asking how rich the innate non-verbal representations of sets are in representing number. A growing body of literature has shown evidence for two distinct non-verbal representations of sets of objects – the approximate number system (ANS) and an object tracking system termed parallel individuation (e.g., Dehaene, 1997; Feigenson, Dehaene, & Spelke, 2004; Xu, 2003; Carey, 2009). Some have argued that ANS is our *only* innate non-verbal system for representing numerical relations, and there is much evidence to suggest that the ANS can represent numerical relations early in infancy (e.g., Brannon, 2002; Suanda, Tompson, & Brannon, 2008; McCrink & Wynn, 2004). However, there is little consensus as to whether parallel individuation can *also* represent numerical relations between two sets of objects. This motivates the research reported in Chapter 2. Using a numerical comparison task, children between the ages of 2 ½ and 4 ½ were shown two sets of objects and were asked which set contained more elements. In two experiments, I found that young preschoolers (2 ½ to 3 ½) were more accurate on comparing small sets (1 to 3, and likely, 4) than large sets (6 or more) even though the ratios of the numerosities were the same in the small and large numerosity comparisons. Importantly, I found the same pattern of results in children who had not learned any number word meanings. These findings suggest that preschoolers use distinct systems to compare small and large numerosities: parallel individuation for small sets and the ANS for large sets. These findings make an important contribution to the field of numerical cognition: I provided the first evidence that in addition to the ANS, parallel individuation, a non-verbal representational system for representing objects, can be used to support the computation of numerical comparison of two distinct sets prior to the acquisition of numerical language.

The third chapter investigated the development of another type of representation of number – namely, representations of unspecified numerosities. Investigations into the concept of the natural numbers must not only account for representations of particular numerosities, but they must also explain how we represent unspecified numerosities. As adults, we are capable of determining whether statements such as '$x + 1 > 1$' are true without knowing the exact value of x. The ability to represent unspecified numerosities is an important aspect of numerical knowledge that is fundamental not only in the field of mathematics, but also in other fields such as physics and computer science. But when and how does the representation of unspecified numerosities develop? In Chapter 3, I examined when children are able to reason about the effects of different types of transformation on the numerosity of a set, as a way to explore the development of the representation of unspecified numerosities. In three experiments, children between the ages of 3 and 5 observed as the experimenter performed a transformation on a set of visible or hidden objects, and were asked whether there were more elements in the set. The experiments were designed to examine when children begin to understand that '$x + 1 > x$' regardless of the value of x. I found that the ability to reason about unspecified numerosities emerges at around age 4, and is fully in place by around 5-6 years of age. These findings make three important contributions: (1) I provided evidence that 5-year-olds show the capacity to reason about unspecified numerosities in a task that does not require them to interpret number word meanings; (2) I showed that children younger than 5, and in particular, 4-year-olds, have some limited abilities to reason about representations of unspecified numerosities; (3) Contrary to the view that studies on number word application may underestimate children's numerical knowledge, I used a task that did not employ number words and found converging evidence that is consistent with previous research on children's knowledge of *unknown* number words (Lipton & Spelke, 2006).

*The representation of unspecified numerosities: What is developing?*

The findings from Chapter 3 suggest that 3-year-olds fail to reason about unspecified numerosities. This ability may emerge at around age 4, and is fully in place by age 5. What is developing between the ages of 3 and 6 that may account for these developmental changes? One possibility is the experience of schooling. At the time of testing, both 4- and 5-year-olds spent half of each week at a daycare (or at home) and the other half at an elementary school with a junior and senior kindergarten program. Thus, it is possible that being in a school environment enhances children's numerical reasoning abilities, perhaps via structured activities. Nevertheless, given that 4- and 5-year-olds experience the same amount of schooling but demonstrated a different level of understanding of numerical transformations in my experiments, it seems unlikely that schooling alone can explain the developmental differences. A second possibility is that children's ability to ignore perceptual cues develops between the years of 4 and 5. In particular, given that 4-year-olds failed to distinguish between 'Move 1' and 'Plus 1' in Experiment 3, their poor performance compared to 5-year-olds could stem from perceptual interference. However, in Experiments 4 and 5 where perceptual cues were completely removed, 4-year-olds still demonstrated difficulty with numerical reasoning. A third possibility is that pragmatic abilities develop between the ages of 4 and 5. In addition to the non-numerical transformations ('Move 1', 'Knock'), 4-year-olds also had difficulty with '-A+A' across all experiments. This raises the possibility that 4-year-olds' poorer performance was due to their developing pragmatic abilities. For example, as mentioned earlier, 4-year-olds may not be able to spontaneously generate 'same number' as a possible response alternative to a question about 'more'. They may interpret a question about 'more' as if 'more' is an absolute adjective – i.e., when a set does not contain 'more', it contains 'less'. I speculate that this particular difficulty about interpreting 'more' may only arise in the context of questions when children have to generate possible response alternatives. These three possibilities are not mutually exclusive. It is likely that there could be multiple sources of difficulty for young children's failure to reason about number.

In addition, while I argue that the development of children's numerical reasoning abilities develops gradually, it is possible that the change occurring when children turn 5 is more sudden. For example, given that 4-year-olds failed to distinguish between addition and the non-numerical transformations ('Move 1', 'Knock'), it leaves open the question of whether they are truly engaging in numerical reasoning in the context of addition and subtraction. As noted above, the use of real object kinds and count syntax should make it salient that the task is about discrete quantification – quantification on the basis of the number of objects. Moreover, as mentioned in Chapter 3, both Sarnecka and Gelman (2004) and Brooks and colleagues (2013) showed that by age 2 ½ to 3, children understand the type of transformation that licenses a change in number word, and the type of transformation that does not license a change in number word. This provides some evidence that the changes I found in the current set of experiments are more quantitative in nature and due to other aspects of the tasks. Future studies could investigate this in more detail by including a trial in which continuous cues are controlled (e.g., removing three tiny bells and adding 1 giant bell that has a larger volume than the combined volume of 3 tiny bells).

*Implications and convergent evidence for number word learning*

Although the present dissertation focuses on non-verbal representations of numerical relations, it has implications for children's acquisition of verbal numerical knowledge, and specifically, number word meanings. In Chapter 2, I found evidence that parallel individuation can support numerical computations, despite the lack of an explicit representation of number in the system. Importantly, I also found that children who had not acquired any number word meanings can likely recruit parallel individuation to compare small numerosities. This provides some evidence that parallel individuation could support the acquisition of early number word meanings, and may explain why children acquire number words in sequence. Using the Give-A-Number task as discussed in Chapter 2, it has been consistently shown that children learn the first few number words in order: they first learn the meaning of 'one', then 'two', 'three', and sometimes 'four', before inducing the cardinal principle – i.e., that the last word in a counted set refers to the numerosity of the set (e.g., Wynn, 1990, 1992a; Le Corre et al., 2006; Barner, Libenson, Cheung, & Takasaki, 2009; Sarnecka et al., 2007; Le Corre, Li, Huang, Jia, & Carey, under review; Sarnecka & Lee, 2009; Lee & Sarnecka, 2010, 2011). While the source of delay regarding children's acquisition of early number words remains an open question (e.g., Le Corre & Carey, 2007; Wagner & Johnson, 2011), the findings from Chapter 2 provide some support for the hypothesis that children's meanings for small number words may first be mapped onto representations of individual objects before they make the cardinal principle induction.

Investigations of the development of unspecified numerosities reported in Chapter 3 provide convergent evidence for number word learning. One motivation for the experiments in that chapter was to examine children's generalized numerical knowledge without recruiting their number word knowledge. Although knowledge about number word meanings reflects children's understanding of numerical concepts, it could also underestimate their ability. In particular, their difficulty in previous studies on number word application (e.g., Condry & Spelke, 2008) may have reflected a lack of knowledge about how number words represent numerical concepts. Thus, it is crucial to investigate children's numerical understanding in the absence of number words. Surprisingly, in such a task, I did not find an earlier emergence of numerical reasoning abilities. Results from Chapter 3 show that children are capable of reasoning about general number by age 5, the same age at which Lipton and Spelke (2006) found evidence for children's understanding of number word application in the context of unknown number words. This converging result tentatively suggests that 5-year-old children begin to reason about unspecified numerosities at the same time that they understand that there are numbers outside of their count list. Future studies can adopt a within-subject design to explore the relationship between these pieces of numerical knowledge that are closely related to each other.

*Numerical comparison on two distinct sets and on one set of objects*

The two chapters used numerical comparison as a case study to examine the nature of representations of number in children. At first glance, it may appear that results from the two chapters contrast with each other. On one hand, as shown in Chapter 2, children begin to be able to compare small numerosities as young as 2 ½ years of age, and large numerosities starting at around 3 ½ years of age. On the other hand, as shown in Chapter 3, their ability to compare a set of objects before and after transformation does not emerge for large sets until age 4 or 5. What may explain such discrepant findings?

One important factor that warrants discussion is the methodological differences. In Chapter 2, children were presented with two distinct, static sets, whereas children in Chapter 3 were shown a set that was transformed. Thus, although children across all experiments reported in this dissertation were asked to make 'more' judgments, the way the stimuli were presented may require different computations of numerical comparison. In particular, children who participated in the experiments in Chapter 3 were required to reason about the effects of transformation in the context of numerical comparison. Thus, the task facing children in these experiments is computationally more complex than for those who participated in experiments in Chapter 2. For example, in Chapter 3, when shown hidden sets, children have to generate a representation of the set before the transformation, store it in memory, and compare it to a representation that they generate after the transformation. Arguably, the task may be simpler in the case of visible sets. However, children are still required to store the pre-transformed set in memory in order to compare it to the post-transformed set. In contrast, children who participated in the experiments in Chapter 2 were asked to compare two sets of objects that were both in sight, and children did not have to update their representations of these sets or store them in memory. These differences in the process of numerical comparison may explain why children demonstrate earlier competency in comparing two distinct sets of objects than in reasoning about transformation effects on a single set of objects. Another possibility for the discrepant results is that the representation of unspecified numerosities is constructed much later. The origin of the representation of unspecified numerosities remains an open question that will likely generate a fruitful area for research.

*Conclusions*

In summary, this dissertation expanded the scope of evidence regarding the nature and development of representations of number in two important ways. First, I documented evidence for rich innate representations of sets. In addition to the ANS, parallel individuation – a system for representing individual objects – has been shown to support numerical comparison in young children, suggesting that it is one of the developmental roots for representing numerical relations.

Second, I found that children's ability to reason about unspecified numerosities develops gradually, and appears relatively late in development, not until 5 years of age. Together, these results suggest that the numerical reasoning principles that operate over representations of unspecified numerosities may develop later than computations that operate over representations of particular numerosities.

## References


Abreu-Mendoza, R.A., Soto-Alba, E.E., & Arias-Trejo, N. (2013). Area vs. density: influence of visual variables and cardinality knowledge in early number comparison. *Frontiers in Psychology*, *4*, 1-10.

Barner, D., & Snedeker, J. (2005). Quantity judgments and individuation: Evidence that mass nouns count. *Cognition*, *97*, 41-66.

Barner, D., & Snedeker, J. (2006). Children's early understanding of mass-count syntax: individuation, lexical content, and the number asymmetry hypothesis. *Language Learning and Development*, *2*, 163-194.

Barner, D., Libenson, A., Cheung, P., & Takasaki, M. (2009). Cross-linguistic relations between quantifiers and numerals in language acquisition: Evidence from Japanese. Journal of Experimental Child Psychology, 103, 421-440.

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

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

Barth, H., La Mont, K., Lipton, J., Dehaene, S., Kanwisher, N., and Spelke, E. (2006). Nonsymbolic arithmetic in adults and young children. *Cognition*, *98*, 199-222.

Brannon, E.M. (2002). The development of ordinal numerical knowledge in infancy. *Cognition*, *83*, 223-240.

Brannon, E. M., & Van de Walle, G. A. (2001). The development of ordinal numerical competence in young children. *Cognitive Psychology*, *43*, 53-81.

Brooks, N., & Audet, J., & Barner, D. (2013). Pragmatic inference, not semantic competence, guides 3-year-olds' interpretation of unknown number words. *Developmental Psychology*.

Bullock, M., & Gelman, R. (1977). Numerical reasoning in young children: The ordering principle. *Child Development, 48,* 427-434.

Cantlon, J.F., Fink, R., Safford, K. E. & Brannon, E. M. (2007). Heterogeneity impairs numerical matching but not numerical ordering in preschool children. *Developmental Science*, *10*, 431-440.

Cantlon, J.F., Safford, K.E., & Brannon, E. (2010). Spontaneous analog number representations in three-year-old children. *Developmental Science, 13*, 289-297.

Carey, S. (2009). The origin of concepts. New York: Oxford University Press.

Chi, M., & Klahr, D. (1975). Span and rate of apprehension in children and adults. *Journal of Experimental Child Psychology, 19*, 434-439.

Choo, H. & Franconeri, S. L. (2013). Enumeration of small sets violates Weber's Law. *Psychonomic Bulletin & Review*, *21*, 93-99.

Clearfield, M. W. & Mix, K. S. (1999). Number versus contour length in infants' discrimination of small visual sets. *Psychological Science*, *10*, 408-411.

Clearfield, M.W., & Mix, K.S. (2001). Amount versus number: Infants' use of area and contour length to discriminate small sets. *Journal of Cognition and Development*, *2*, 243-260.

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

Cordes, S., Gelman, R., Gallistel, C. R., & Whalen, J. (2001). Variability signatures distinguish verbal from nonverbal counting for both large and small numbers. *Psychonomic Bulletin*

*& Review*, 8, 698-707

Coubart, A., Izard, V., Spelke, E.S., Marie, J., & Streri, A. (2014). Dissociation between s small and large numerosities in newborn infants. *Developmental Science, 17*, 11-22.

Dehaene, S. (1997). *The number sense*. New York: Oxford University Press.

Evans, D.S. (1983). Understanding zero and infinity in the early school years (Unpublished doctoral dissertation). University of Pennsylvania.

Feigenson, L. (2005). A double dissociation in infants' representation of object arrays. *Cognition, 95*, B37-B48.

Feigenson, L. & Carey, S. (2003). Tracking individuals via object-files: Evidence from infants' manual search. *Developmental Science, 6,* 568-584.

Feigenson, L., & Carey, S. (2005). On the limits of infants' quantification of small object arrays. *Cognition, 97*, 295-313.

Feigenson, L., Carey, S., & Hauser, M. (2002a). The representations underlying infants' choice of more: Object-files versus analog magnitudes. *Psychological Science, 13*, 150-156.

Feigenson, L., Carey, S., & Spelke, E.S. (2002b). Infants' discrimination of number vs. continuous extent. *Cognitive Psychology, 44*, 33-66.

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

Feigenson, L. & Halberda, J. (2004). Infants chunk object arrays into sets of individuals. *Cognition, 91*, 173–190.

Frank, M.C., & Barner, D. (2011). Representing exact number visually using mental abacus. *Journal of Experimental Psychology: General, 141*, 134-149.

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

Gathercole, V. (1985). More and more and more about more. *Journal of Experimental Child Psychology, 40*, 73-104.

Gallistel, C. R. (1989) Animal cognition: the representation of space, time and number. *Annual Review of Psychology*, 40, 155-189.

Gallistel, C. R., & Gelman, R. (2000). Non-verbal numerical cognition: From reals to integers. *Trends in Cognitive Sciences, 4*, 59-65.

Gallistel, C. R. (2007). Commentary on Le Corre & Carey. *Cognition, 105*, 439-445.

Gelman, R. (1977). How children reason about small numbers. In N. Castellan, D. B. Pisoni and G. Potts (Eds.), *Cognitive Theory*. Hillsdale, NJ: Erlbaum, Vol. 2, (8), 219-238.

Gelman, R., & Gallistel, C.R. (1978). The child's understanding of number. Cambridge, MA: Harvard University Press

Gilmore, C., Attridge, N., Clayton, S., Cragg, L., Johnson, S., Marlow, N., Simms, V., & Inglis, M. (2013). Individual differences in inhibitory control, not non-verbal number acuity, correlate with mathematics achievement. *PLOS ONE, 8(6),* e67374.

Gordon, P. (1985). Evaluating the semantic categories hypothesis: The case of the count/mass distinction. *Cognition, 20*, 209-242.

Halberda, J., & Feigenson, L. (2008). Developmental change in the acuity of the "Number Sense": The approximate number system in 3-, 4-, 5-, 6-year-olds and adults. *Developmental Psychology, 44,* 1457-1465.

Halberda, J., Mazzocco, M., & Feigenson L. (2008). Individual differences in nonverbal number acuity correlate with maths achievement. *Nature, 455*, 665-669.

Hartnett, P. (1991). The development of mathematical insight: from one, two, three to fininty. (Unpublished doctoral dissertation). University of Pennsylvania.

Hauser, M.D., Carey, S., & Hauser, L.B. (2000). Spontaneous number representation in semi-free-ranging rhesus monkeys. *Proceedings of the Royal Society, London 267*, 829-833.

Hyde, D. C., & Spelke, E. S. (2009). All numbers are not equal: an electrophysiological investigation of small and large number representations. Journal of Cognitive Neuroscience, 21, 1039-1053.

Hyde, D. C., & Spelke, E. S. (2011). Neural signatures of number processing in human infants: Evidence for two core systems underlying numerical cognition. *Developmental Science, 14*, 360-371.

Jevons, W. S. (1871). The power of numerical discrimination. *Nature*, *3*, 363-372.

Izard, V., Streri, A., & Spelke, E.S. (2014). Toward exact number: Young children use one-to-one correspondence to measure set identity but not numerical equality. *Cognitive Psychology, 72*, 27-53.

Le Corre, M., & Carey, S. (2007). One, two, three, four, nothing more: An investigation of the conceptual sources of the verbal counting principles. *Cognition, 105*, 395-438.

Le Corre, M., & Carey, S. (2008). Why the verbal counting principles are constructed out of representations of small sets of individuals: A reply to Gallistel. *Cognition, 107*, 650-662.

Le Corre, M. (2014). Children acquire the later-greater principle after the cardinal principle. *British Journal of Developmental Psychology, 32*, 163-177.

Le Corre, M., Van de Walle, G., Brannon, E., & Carey, S. (2006). Re-visiting the competence/performance debate in the acquisition of the counting principles. *Cognitive Psychology, 52*, 130-169.

Le Corre, M., Li, P., Huang, B., Gia, J, & Carey, S. (under review). Language and number: A comparison of young Mandarin and English Learners.

Lee, M. D., & Sarnecka, B.W. (2010). A model of knower-level behavior in number-concept development. *Cognitive Science, 34*, 51-67.

Lee, M. D., & Sarnecka, B.W. (2011). Number-knower levels in young children: Insights from Bayesian modeling. *Cognition*, *120,* 391-402.

Leslie, A. M., Gelman, R., Gallistel, C. R. (2008). The generative basis of natural number concepts. *Trends in Cognitive Sciences, 12*, 213-218.

Leslie, A. M., Gallistel, C.R., Gelman, R. (2007). Where integers come from. In Carruthers, P., Laurence, S., Stich, S., (Eds.) The innate mind, Vol 3: Foundations the future, New York: Oxford University Press, 109-138.

Lipton, J.S., & Spelke, E.S. (2006). Preschool children master the logic of number word meanings. *Cognition*, 1-10.

Libertus, M., Feigenson, L., & Halberda, J. (2011). Preschool acuity of the approximate number system correlates with school math ability. *Developmental Science, 14*, 1292-1300.

Lipton, J. S., & Spelke, E. S. (2003). Origins of number sense: Large number discrimination in human infants. *Psychological Science*, *14*, 396-401.

Luck, S.J., & Vogel, E.K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390*, 279-281.

Massachusetts, Department of Elementary and Secondary Education. (2011). *Massachusetts Curriculum Framework for Mathematics: Grades Pre-Kindergarten to 12*. Retrieved from http://www.doe.mass.edu/frameworks/current.html.

Meck, W. H., & Church, R. M. (1983). A mode control model of counting and timing processes. *Journal of Experimental Psychology: Animal Behaviour and Process, 9*, 320-334.

McCrink, K. & Wynn, K. (2004) Large-number addition and subtraction by 9-month-old infants. *Psychological Science, 15,* 776-781.

McCrink, K. & Wynn, K. (2007) Ratio abstraction by 6-month-old infants. *Psychological Science, 18*, 740-746.

Mix K.S., Huttenlocher, J., & Levine, S.D. (1996). Do preschool children recognize auditory-visual numerical correspondences? *Child Development, 67*, 1592-1608.

Moher, M., Tuerk, A.S., & Feigenson L. (2012). Seven-month-old infants chunk items in working memory. *Journal of Experimental Child Psychology, 112*, 361-377.

Moyer, R. S., & Landauer, T. K. (1967). Time required for judgment of numerical inequality. *Nature*, *215*, 665-670.

Odic, D., Libertus, M., Feigenson, L., & Halberda, J. (2013). Developmental change in the acuity of approximating area and approximating number. *Developmental Psychology, 49,* 1103-1112.

Odic, D., Pietroski, P., Hunter T., Lidz, J., & Halberda, J. (2012). Young children's understanding of "more" and discrimination of number and surface area. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39,* 451-461.

Ontario, Ministry of Education. (2006). *The Kindergarten Program Revised.* Retrieved from http://www.edu.gov.on.ca/eng/curriculum/elementary/kindercurrb.pdf.

Piaget, J. (1952). *The child's conception of number.* London : Routledge & Kegan Paul, Ltd.

Prasada, S., Ferenz, K., & Haskell, T. (2002). Conceiving of entities as objects and stuff. *Cognition, 83*, 141-165.

Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision, 3*, 179–197.

Revkin, S. K., Piazza, M., Izard, V., Cohen, L., & Dehaene, S. (2008). Does subitizing reflect numerical estimation? *Psychological Science, 19,* 607-614.

Rosenberg, R. D., & Feigenson, L. (2013). Infants hierarchically organize memory representations. *Developmental Science, 16*, 610-621.

Ross-Sheehy, S., Oakes, L. M., & Luck, S. J. (2003). The development of visual short-term memory capacity in infants. *Child Development, 74,* 1807-1822.

Rousselle, L., & Noël, M. P. (2008). Basic numerical skills in children with mathematics learning disabilities: A comparison of symbolic vs. non-symbolic number magnitude processing. *Cognition, 102*, 361–395.

Rousselle, L., Palmers, E., & Noël, M. P. (2004). Magnitude comparison in preschoolers: What counts? Influence of perceptual variables. *Journal of Experimental Child Psychology, 87,* 57–84.

Rips, L.J., Bloomfield, A., & Asmuth, J. (2008). From numerical concepts to concepts of number. *Behavioural and Brain Sciences, 31,* 623-642.

Sarnecka, B.W., & Gelman, S.A. (2004). Six does not just mean a lot: Preschoolers see number words as specific. *Cognition, 92,* 329-352.

Sarnecka, B. W., & Lee, M. D. (2009). Levels of number knowledge in early childhood. *Journal of Experimental Child Psychology, 103,* 325-337.

Sarnecka, B.W., Kamenskaya, V.G., Yamana, Y., Ogura, T., & Yudovina, J.B. From grammatical number to exact numbers: Early meanings of “one," “two,” and “three” in English, Russian, and Japanese. *Cognitive Psychology, 55,* 136-168.

Sarnecka, B. W., & Wright, C. E. (2013). The idea of an exact number: Children’s understanding of cardinality and equinumerosity. *Cognitive Science. 37*, 1493-1506.

Schaeffer, B., Eggleston , V. H., & Scott , J. L. (1974). Number development in young children.

*Cognitive Psychology, 6*, 357-379.

Shapiro, S. (2000). *Thinking about mathematics: The philosophy of mathematics*, Oxford University Press, Oxford, UK

Sophian, C. (1987). Early developments in children's use of counting to solve quantitative problems. *Cognition and Instruction, 4,* 61-90.

Starkey, P., & Cooper, R.G. (1980). Perception of numbers by human infants. *Science, 210,* 1033-1035.

Suanda, S.H., Tompson, W., & Brannon, E. (2008). Changes in the ability to detect ordinal numerical relationshps between 9 and 11 months of age. *Infancy, 13,* 308-337.

Trick, L., & Pylyshyn, Z. W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychological Review, 101*, 80 - 102.

Wagner, J. B., & Johnson, S. C. (2011). An association between understanding cardinality and analog magnitude representations in preschoolers. *Cognition, 119*, 10-22.

Whalen, J., Gallistel, C. R., & Gelman R.. (1999). Non-verbal counting in humans: The psychophysics of number representation. Psychological Science, 130-137.

Winer, G.A. (1974). Conservation of different quantities among preschool children. *Child Development, 45*, 839 – 842.

Wood, J. N., & Spelke, E. S. (2005). Infants' enumeration of actions: Numerical discrimination and its signature limits. *Developmental Science, 8(*2), 173-181.

Wynn, K. (1990). Children's understanding of counting. *Cognition, 36*, 155-193.

Wynn, K. (1992a). Child’s acquisition of the number words and the counting system. Cognitive Psychology, 24, 220-251.

Wynn, K. (1992b). Addition and subtraction by human infants. *Nature, 358*, 749-750.

Uller, C., Carey, S., Huntley-Fenner, G., & Klatt, L. (1999). What representations might underlie infant numerical knowledge? *Cognitive Development*, *14*, 1-36.

Xu, F. (2003) Numerosity discrimination in infants: Evidence for two systems of representations. *Cognition, 89*, B15-B25

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

Xu, F., Spelke, E., & Goddard, S. (2005). Number sense in human infants. *Developmental Science, 8*, 88-101.

Xu, F., & Arriaga, R.I. (2007). Number discrimination in 10-month-old infants. *British Journal of Developmental Psychology, 25*, 103-108.