An Examination of discrete and continuous quantity representations across the lifespan:

Author: Sophie Savelkouls

Boston College Electronic Thesis or Dissertation, 2019

AN EXAMINATION OF DISCRETE AND CONTINUOUS QUANTITY REPRESENTATIONS ACROSS THE LIFESPAN

Sophie Savelkouls

A dissertation

submitted to the Faculty of

the department of Psychology

in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Boston College
Morrissey College of Arts and Sciences
Graduate School

March 2019

AN EXAMINATION OF DISCRETE AND CONTINUOUS QUANTITY REPRESENTATIONS ACROSS THE LIFESPAN

Sophie Savelkouls

Advisor: Sara Cordes, Ph.D.

The format of our quantity representations is a contentious topic of study in the field of numerical cognition with researchers debating whether we use discrete (i.e. number) or continuous (e.g. area, time, volume or density) cues to make quantity judgements. It has been proposed (through the Sense of Magnitude Theory) that continuous quantities are more perceptual in nature and thus do not require the higher order cognitive processes needed to represent abstract number, making it unlikely that number is tracked in the presence of perceptual quantities. In the current dissertation, I examined claims made by the Sense of Magnitude theory by 1) investigating the accuracy with which we represent continuous quantities and the mental processes we may engage in when representing these quantities and by, 2) comparing the relative salience of discrete and continuous quantities and how this may change across development. In Project 1, I investigated the accuracy with which infants make element size discriminations and whether this ability becomes more precise with age. Project 2 examined the precision with which adults track cumulative area and uncover the process by which they do so. Lastly, Project 3 explored the relative salience of number for preschoolers by assessing their "Spontaneous Focusing on Number." Together, findings from these three projects undermine claims stating that humans at all stages of development are better at, and prefer to, attend to continuous quantities over discrete

number. Instead I propose that this dissertation suggests that humans at all stages of development are strongly attuned to number in their environment. This work not only provides insight into the way we represent quantity in our day to day lives, but it can help us understand where individual difference in mathematical achievement in school may stem from.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 2

CHAPTER 3

LIST OF FIGURES

ACKNOWLEDGEMENTS

This dissertation, and my graduate career, would not have been possible without the help of some very important people.

First of all, I would like to express my special appreciation and thanks to my advisor Sara Cordes. These past 5 years would not have been possible without your guidance, support and endless confidence that all data (even infant data) could be "saved" (as long as you run 3 follow-up studies). Thank you for having confidence in me and my data when I did not.

I also want to thank my dissertation committee: Hiram Brownell, Katherine McAuliffe and Elida Laski, for offering your time and guidance in reviewing this dissertation in all its stages.

Of course, thank you to all members of the Infant and Child Cognition Lab: grad students, postdocs, lab managers and RAs. Thank you for helping me conduct my research and providing such a warm and cooperative working environment. Karina - when we started I was worried it would be difficult to have another graduate student starting in the lab at the same time as me. However, I could not have been more wrong. Having someone to commiserate with, celebrate milestones with and complain about data with, have made these five years so much more enjoyable. Additionally, a special thank you goes to my many devoted research assistants without whom I would never have been able to collect so much data throughout the years.

None of this research would have been possible without the families and children that participated in these studies. Thank you especially to the parents that visited the lab and participated in our research. Thank you also to the lab's offsite testing partners: preschools, the Boston Children's Museum and the Living Lab at the Museum of Science in Boston.

Most of all I want to thank my two families:

My salsa family – I took my first salsa lesson the same week that I started graduate school. I didn't know then how important you all would become in keeping me motivated throughout this difficult process. Thank you for teaching me that music, dancing and lots of sequins make everything better. Sebastian – you listened to me complain too many times to count in these last few months and always managed to make me smile, I am very grateful.

My parents – without your support I never would have had the courage to come to the United States 9 years ago. You provided me with the freedom to choose my own path and showed me that just about anything is possible as long as you work hard for it. Thank you for the many late night skype calls that kept me down-to-earth. Ik hou van jullie.

INTRODUCTION

There is no doubt that human and non-human animals can represent quantities (Brannon, Lutz, & Cordes, 2006; Halberda & Feigenson, 2008d; Odic, Libertus, Feigenson, & Halberda, 2013; Xu & Spelke, 2000). However, the format of these representations and how we process these quantities at different stages of development is a much more contentious topic of study. There are a variety of dimensions of quantities we can represent – discrete quantity (i.e. number) and continuous quantities (e.g. area, time, volume or density) – and they are often strongly correlated with one another. For example, 8 cookies will not only be more numerous than 4 cookies, but 8 cookies will also take up a larger area on the plate, and the density of cookies compared to the empty space of the plate will also be greater. This inherent correlation between these different quantity dimensions means that it is difficult to determine which quantitative dimension infants, children, or adults might use when making quantity judgements, and it brings into question whether we can track these quantities independently of one another at all. Although a large amount of research has investigated our ability to represent discrete number, much less research has investigated our ability to represent continuous quantities. In order to paint a full picture of how infants, children, and adults represent quantity, it is important that we examine our abilities to represent both discrete and continuous quantities, independently of the others, as well as investigate the relative saliency of these different quantities. Given the lack of research investigating our abilities to represent continuous quantities, the current research aims not only to investigate more generally our ability to discriminate these continuous quantities, but I also examined the ways in which number can interfere with our judgements of continuous quantity, and how the relative saliency of discrete and continuous quantities may impact our attention to number in the world.

Theoretical framework

Piaget was one of the first researchers to examine children's understanding of quantities and also one of the first to suggest that number was simply too abstract of a concept for young children to grasp (Piaget, 1952, 1977). Instead Piaget suggested that children relied upon continuous quantities (also referred to as continuous extent) as a proxy for number. Since then a large body of research has continued to investigate the development of quantity understanding and the extent to which we use continuous and discrete cues to make quantity judgements. Proponents of the "Sense of Magnitude" (SoM) theory claim, similar to Piaget, that our abilities to track number are fully reliant upon an ability to track continuous quantities (Gebuis & Reynvoet, 2012b; Leibovich, Katzin, Harel, & Henik, 2017). They make this claim because they suggest that continuous cues are dependent on, and derived from, perceptual cues in nature (i.e. they are tied to a specific sensory modality) which makes them much easier to track than number which is abstract (i.e. we can track this regardless of the sensory modality). They further postulate that the abstract nature of number means that representing number requires many more higher order cognitive processes than tracking continuous quantities, and thus it is unlikely that number is tracked in the presence of perceptual quantities.

Others claim that number is not too abstract for us to represent, citing evidence that has shown that even young infants are able to discriminate number at a very early age (Lipton & Spelke, 2003; Xu & Spelke, 2000). Furthermore, they find that infants are in fact often more accurate at discriminating number compared to continuous quantities such as cumulative area or element size (Cordes & Brannon, 2011; M. E. Libertus, Starr, & Brannon, 2014).

Given these two contradictory frameworks, it is important that I briefly examine what the literature has found so far with respect to our ability to discriminate both discrete and continuous quantities.

Discrete Quantity Discrimination in Humans

Most investigations of our ability to represent number have tested infants, children or adults on their ability to discriminate two different quantities (e.g., two set sizes, two amounts of cumulative area) from one another. In particular, given the strong correlation between number and continuous quantities, to investigate whether we can discriminate number independently of other quantities, a number of sophisticated designs have been developed that control for continuous cues isolating number as the only relevant cue. For example, in the infant literature on number discrimination, habituation studies have been developed where infants are presented with a series of displays, where the number of items in the display remains constant throughout habituation but where continuous cues such as the size of the dots, and thus the cumulative surface area and contour of the items vary across habituation. Then in test, infants are shown displays with either a novel or familiar number of items, while controlling for continuous variables. Similar designs with adults have also been developed: continuous quantities are varied in such a way that it makes relying upon continuous quantities when performing numerical discriminations, a less reliable strategy. Using these designs, research has consistently shown that infants (Lipton & Spelke, 2003; Xu & Spelke, 2000), children (Halberda, Mazzocco, & Feigenson, 2008; Odic, Le Corre, & Halberda, 2015; Odic et al., 2013), and adults (Odic et al., 2013) can use number alone to make discriminations between two sets of items, suggesting that number can be tracked independently of continuous quantities.

Based on this literature, it has been determined that humans' ability to discriminate number is ratio-dependent and improves with age. Although 6-month-old infants need as much as a 1:2 ratio of change between two sets of items to notice a change in number (M. E. Libertus & Brannon, 2010; Lipton & Spelke, 2003, 2004; Xu & Spelke, 2000; Xu, Spelke, & Goddard, 2005), by 9-10 months infants are able to discriminate a 2:3 ratio (M. E. Libertus & Brannon, 2010; Lipton & Spelke, 2003, 2004; Xu & Arriaga, 2007) and by the time children reach the age of 3 they can accurately discriminate a 3:4 ratio (Halberda & Feigenson, 2008b), with adults discriminating even a 9:10 ratio (Odic et al., 2013).

Continuous Quantity Discrimination in Humans

The literature on number discrimination is quite extensive and suggests that humans can represent number independent of other continuous quantities. However, relatively less work has examined human's ability to discriminate continuous quantities, and the findings from this literature are much more ambiguous. Taking a closer look at the infant literature, a few studies have used paradigms similar to that used with number, examining whether infants can discriminate using a single continuous dimension, while controlling for all other quantity dimensions. Studies on element size have for example found that 6-month olds were able to discriminate up to a 1:2 ratio, but not a 2:3 ratio in the size of a single Elmo face (Brannon et al., 2006). Similarly, 6-month-old infants are able to detect a 1:2 ratio of difference in the duration of a single visual or auditory stimulus (Wynn & VanMarle, 2006). This suggests that infants can make continuous quantity (size or duration) discriminations at a similar level of accuracy as that of number, for single items.

However, one key difference between these studies and those of number discrimination is that these element size studies have examined quantity discrimination in the context of a *single item*, while number discrimination studies by definition present stimuli in the context of *sets*. Therefore, taking a look at the few studies that have studied continuous extent discrimination, specifically when stimuli are presented in sets, the evidence is not as clear. Studies examining element size discrimination in the context of sets have found that 7-month olds need as much as a 4-fold change to detect a difference (Cordes & Brannon, 2011). Furthermore, studies examining other continuous dimensions have found that 7-month olds successfully detect a 4-fold but not a 3-fold change in cumulative area (Cordes & Brannon, 2008), and successfully detected a 3-fold but not a 2-fold change in contour (Starr & Brannon, 2015). Thus at least with infants, the research suggests that, within the context of sets of items, infants are more precise at making numerical discriminations compared to continuous extent discriminations. Only when tracking continuous extent across a single item do infant's abilities match their number discrimination abilities. In Project 1, I extended these findings by investigating element size discriminations in the context of sets at two different ages to determine whether, like number, element area discriminations become more precise with age.

The evidence on continuous extent representations in children and adults is even sparser. Two studies did directly compare children's and adult's number and element size discriminations by presenting them either with two arrays of dots (number task) or a single irregular shape that had two different color sections (area task) and asked participants to judge which color had more (Leibovich & Henik, 2014; Odic et al., 2013). Although children and adults were more accurate and had quicker reaction times for the area task compared to the number task – similar to many studies in the infant literature – they only assessed area discrimination in the context of a single item, making a comparison to numerical discriminations involving multiple items inappropriate. Although a few studies have examined the salience of CA in the context of multiple items (Barth, 2008; Hurewitz, Gelman, & Schnitzer, 2006), very few studies have systematically compared these tracking abilities to that of number.

The Relative Salience of Discrete and Continuous Quantities

One way we can begin to answer which dimension of quantity we represent most readily is to examine the relative saliency of these different dimensions and the ways in which these different quantities might interfere with one another. A line of work with adults has investigated this question by presenting adults with interference tasks where adults are presented with arrays of dots where discrete and continuous quantity are either congruent with one another (e.g. the larger number array has a larger cumulative area) or incongruent with one another (e.g. the larger number array has a smaller cumulative area). This allows researchers to test whether number interferes with cumulative area judgements or vice versa, whether cumulative area judgements interfere with number judgements. All previous studies using this type of paradigm have examined the ways in which continuous cues may interfere with number judgements, finding that adults perform worse on incongruent trials, suggesting that continuous dimensions interfere with discrete quantity judgements (Barth, 2008; DeWind & Brannon, 2012; Gebuis &

Reynvoet, 2012a; Hurewitz, Gelman, & Schnitzer, 2006). The few studies that have examined how number interferes with continuous extent judgements (Barth, 2008; Hurewitz, Gelman, & Schnitzer, 2006) have not systematically tested how different factors such as set size (the number of items in each array) or ratio (i.e. the ratio between the number of items in each array) affected performance. This is what I investigated in Project 2 by presenting adults with a cumulative area discrimination task in which I manipulated the set size and ratio of the arrays of items. By manipulating factors such as set size, I was also able to get a better sense of the mental processes by which adults track these different types of quantities (e.g. whether adults perform any type of mental computations to represent continuous quantities like cumulative area) and how the presence of one type of quantity may affect their representation of other quantities.

Another way that we can begin to understand the relationship between our abilities to represent number and continuous extent is to investigate how salient these different dimensions are to humans. One line of work that has investigated the relative salience of quantitative information has examined the construct of "Spontaneous Focusing on Number" (SFON) which is usually defined as one's propensity to focus on number without being prompted to do so. While most tasks examining SFON measure children's tendency to use number words to describe a picture or the frequency with which children imitate the number of repetitive actions performed by an experimenter, more recent SFON tasks have specifically examined SFON when pitted against another quantitative dimension. In these match-to-sample type tasks, children can complete the task by either choosing to match based on number or another continuous dimension such as cumulative area, and greater matching using the dimension of number is considered as evidence of children being more attuned to number in their environment, at least when pitted against continuous quantities. What is still unclear from this line of research is to what extent children's SFON as measured in different types of tasks related to children's actual knowledge about number. This is what I explored in Project 3.

All in all, before we can make any claims about the ways in which humans represent quantity, we need not only investigate the accuracy with which we represent both discrete and continuous quantities, but it is also important that we investigate the relative salience of these different types of quantities, and how this may change across development. This will give us greater insight into how we inherently represent quantities and how we navigate our quantitative world.

The Proposed Dissertation

Although many studies have investigated infant's, children's, and adult's abilities to represent discrete number, relatively little work has explored our abilities to represent continuous quantities such as element area or cumulative area, specifically in the context of sets. One prominent theory in the literature, the "Sense of Magnitude" theory has claimed that our abilities to track number (an abstract quantity) are fully reliant upon an ability to track continuous quantities, which are more perceptual in nature and thus do not require the higher order cognitive processes needed to represent abstract number (Gebuis & Reynvoet, 2012b; Leibovich et al., 2017). However, before we can make any claims about the ways we represent these different quantities, it is important that we have a thorough sense of infant's, children's, and adult's acuity in representing these different quantities, specifically continuous quantities which have been examined relatively little in the literature. Additionally, it is important that we examine the relative saliency of these different dimensions and the ways in which these different quantities might interfere with one another.

Through the following three projects, I investigated human abilities to represent both discrete and continuous quantity, the mental processes they may have engaged in to represent continuous quantities, and the relative saliency of these quantities, across development. In doing this, I have also brought into question whether we truly rely on perceptual, continuous cues when tracking quantity, as the "Sense of Magnitude" theory has suggested (Gebuis & Reynvoet, 2012b; Leibovich et al., 2017). Below is a short outline of the three projects.

Project 1: Tracking the size of one item among many: element area discrimination in infancy

This project investigated the accuracy with which infants make element size discriminations. Only a few studies have examined continuous extent discriminations in infancy, and they suggest that while infants may be capable of discriminating relatively small (i.e., 2-fold ratio; (Brannon et al., 2006)) changes in element size when presented with a single item, when presented with an array of items, they are much less precise, requiring as much as a 4-fold change in order to detect change (Cordes & Brannon, 2011). I replicated and extended these previous findings by exploring whether element area discriminations, like that of number, become more precise with age. In particular, I tested whether 7- and 12-month old infants succeed in detecting a 3-fold change in element area of items within an array and determined whether continuous extent discriminations become more precise with age. I found that 7-month olds failed and 12-month old marginally succeeded at discriminating a 3-fold change in the element area of items within an array suggesting that infants are not only starting off with a lower level of element area acuity at around 6-7 months of age compared to number (i.e., infants can discriminate a 2-fold change in number at this age; Xu & Spelke, 2000), but they are also improving at a much slower rate in tracking element area than for number. These findings further undermine claims that infants are better at discriminating continuous quantities compared to number.

Project 2: The Impact of Set Size on Cumulative Area Judgements

Although we know a lot about the precision with which humans discriminate number, less is known about how precisely we are capable of tracking cumulative area, and how numerical information may or may not interfere with this ability. The aim of this project was to determine the precision with which adults track cumulative area and to uncover the process by which they do so. I presented adults with arrays of dots (of differing set sizes) and asked them to judge the relative cumulative area of the displays. Two experiments were conducted for this project, with Experiment 1 controlling for *item* density (i.e. the number of items in the display per unit of the background) and Experiment 2 controlling for *area* density (i.e. the cumulative area of the items per unit of the background). This design allowed me to investigate the following research questions: (1) Does numerical congruency matter for cumulative area judgements?; (2) How does set size impact cumulative area representations?; and finally (3) By what process do adults represent cumulative area when presented with an array of items? Similar to the number discrimination literature, I found cumulative area judgements to be ratio-dependent. More interestingly, however, I found that participants not only performed worse on trials where number was incongruent with cumulative area, but that adults performed worse as set size increased. These findings suggest that number interferes with continuous quantity judgements, suggesting that it is at least *as salient* as continuous variables, undermining claims in the literature suggesting that continuous properties are easier to represent, and more salient to adults.

Project 3: Relative Salience of Number: Preschooler's Cardinal Knowledge Relates to Spontaneous Focusing On Number for Small, but not Large, Sets

The final project of this dissertation examined the relative salience of number for preschoolers across different contexts, assessing their "Spontaneous Focusing on Number" (SFON). The main aim of this project was to examine the relation between children's number knowledge abilities and SFON, to assess how dependent SFON may be upon a child's ability to verbally encode the numbers presented. To do this, I manipulated multiple variables. First of all, given that prior studies have investigated SFON in the context of small sets exclusively, with no work exploring SFON in the context of large sets, my primary goal was to determine whether children's SFON for small and large sets similarly relate to their knowledge of number. Second of all, given that many different tasks have been used to assess SFON that have distinct task demands (verbal vs. behavioral) and stimuli (different types and quantities of other features available to the child to focus on outside of number), I examined to what extent these tasks are measuring the same underlying construct. To do this, I presented preschoolers with four distinct SFON tasks assessing their spontaneous attention to number for small (Experiment 1) and large (Experiment 2) sets of numbers. Results not only revealed no relation in SFON performance across the four distinct SFON tasks, but I found preschooler's SFON for small sets (1-4 items) to be significantly stronger than that for large sets (10-40 items). Furthermore, analyses revealed that number knowledge was only associated with SFON for small sets, but not large. Together, these findings suggest that SFON may not be a set-size independent construct, and instead may hinge upon a child's number knowledge, at least in the preschool years.

CHAPTER 1: TRACKING THE SIZE OF ONE ITEM AMONG MANY: ELEMENT AREA DISCRIMINATION IN INFANCY

Abstract

What quantitative properties infants represent when encountering a set of objects has been a topic for debate in the field of numerical cognition, with some suggesting that it should be easier to represent continuous quantities (that are perceptual in nature) than discrete number (which is not tied to any particular percept). Although we know that around 6-7 months of age, infant's representations of number are more precise than that of continuous quantities (they can discriminate a 1:2 ratio of change in number, but fail at a 1:3 ratio of change in element area), it is unclear whether continuous quantity tracking becomes more precise over the course of development similar to numerical discriminations. In particular, do continuous tracking abilities "catch up" to number discrimination at some point in infancy? Thus, in the current study, we examined 7- and 12-month-olds' element area discriminations, to determine whether like number their element area discrimination precision increases with development. We found that 7-month-olds failed to discriminate a 1:3 ratio of change in element area with only marginal success at this same ratio at 12 months of age. Therefore, even by 12 months, infant abilities to track element area are weak, and not anywhere near their numerical tracking capacities. Findings are discussed in light of claims made the Sense of Magnitude theory about how humans represent quantity.

Tracking the size of one items among many: element area discrimination in infancy

What quantitative properties do infants represent when encountering a set of objects, such as rubber ducks in their bath? One possibility is that they track number, monitoring how many individual ducks are in the bathtub. Additionally, they could track the cumulative surface area, estimating how much total yellow they see in the water. Or, they could pay attention to the size of each individual duck. Which of these quantitative properties is most salient to infants is a question that has long been debated in the literature (Piaget, 1952, 1977).

Historically, it has been suggested that number may be too abstract of a quantitative dimension for infants to track. Unlike continuous quantities which are “perceptual” in nature (i.e., dependent upon and tied to the percept), number can be tracked across multiple sensory modalities and is thought to be represented independent of perceptual qualities of the display (i.e. 3 ducks, 3 sounds, and 3 ideas are all instances of the number 3 even though each is associated with very distinct percepts). Proponents of the Sense of Magnitude (SoM) Theory have suggested that rather than tracking number (which presumably requires much higher order cognitive processes), infants are thought to rely upon perceptual, continuous quantities such as surface area or perimeter of visual items to track quantity (Henik, Leibovich, Naparstek, Diesendruck, & Rubinsten, 2011; Mix, Levine, & Huttenlocher, 1997).

However, we now know this isn’t the case. A myriad of studies have examined infant number discrimination abilities, finding that preverbal infants *can* track number even as newborns (Coubart, Izard, Spelke, Marie, & Streri, 2014; Izard, Sann, Spelke, & Streri, 2009), and that they do so even when all other continuous quantitative cues are controlled (Xu & Spelke, 2000; Xu et al., 2005). On the other hand, although a substantial amount of research has examined infant numerical abilities, much less is known about the developmental trajectory of infant abilities to track continuous quantities. The few studies that have examined this question have found that around 6-7 months of age, infants are relatively poor discriminators of cumulative area (CA) and element area (EA) compared to their numerical discrimination abilities (Brannon, Abbott, & Lutz, 2004; Cordes & Brannon, 2008, 2011; M. E. Libertus et al., 2014). However, this previous work has focused on continuous quantity tracking only in 6-7 month olds. Although we know that numerical discrimination abilities become more precise over the course of development, it is unknown whether continuous quantity tracking increases in parallel, and in particular, whether at some point in infancy continuous tracking abilities may "catch up" to number discrimination in infancy. Thus, in the current study, we examined 7-month olds' (to replicate prior work) and 12-month-olds' (to extend this work to an older age range) EA discriminations, to determine whether like number their EA discrimination precision increases with development.

One recurring problem in the literature examining infant quantitative discrimination abilities is that because discrete (i.e. number) and continuous (e.g. cumulative area, density, element area) quantitative properties are highly correlated with one another, it is difficult to determine which quantitative dimension infants attend to. This has led researchers to design paradigms that isolate one quantitative dimension (e.g. number) while controlling for the others (e.g. CA, EA or density). For example, to assess infant number discrimination abilities independent of continuous variables, habituation studies have presented infants with displays where the *number* of items remains identical across habituation trials but the element area (and thus cumulative area and contour length) of the items varies across displays, making these continuous variables unreliable cues for tracking (e.g., Xu & Spelke, 2000). The use of this type of paradigm has led to the discovery that while 6-month-old infants are able to make numerical discriminations of a 1:2 ratio (e.g. 8 vs. 16 items), they fail at a 2:3 ratio (e.g. 8 vs. 12 items; Lipton & Spelke, 2004; Wynn & VanMarle, 2006; Xu & Spelke, 2000; Xu, Spelke, & Goddard, 2005), suggesting that the ease with which infants compare two sets of items depends on the ratio between them. Furthermore, by 9 months of age, infants succeed at detecting a 2:3 ratio change in number (Wood & Spelke, 2005; Xu & Arriaga, 2007), and by 3 or 4 years, children successfully discriminate a 3:4 ratio (Odic et al., 2013), indicating that with age and experience, children are able to discriminate smaller ratios of change in number.

Despite the evidence showing that infants can use number alone to track quantities, some theorists continue to believe that in actuality, infants, children and adults rely on perceptual continuous quantities such as EA or CA when estimating or tracking number (Gebuis & Reynvoet, 2012a, 2012b; Leibovich et al., 2017; Mix, Huttenlocher, & Levine, 2002).These claims come from research revealing that numerical estimates (in children and adults) are impacted by changes in the continuous perceptual qualities of a display, such that e.g., an array of 10 items may be judged to have more items if the individual items within the array are smaller in size (Ginsburg & Nicholls, 1988; Tokita & Ishiguchi, 2010, 2013). Although it is undisputed that continuous quantities may bias numerical estimates, it is debated whether these biases reflect an inherent reliance upon continuous cues (instead of number; Leibovich et al., 2017) or instead an inability to inhibit attention to irrelevant quantitative information when tracking number (Cantrell, Boyer, Cordes, & Smith, 2015; Cantrell & Smith, 2013). This is where the infant work becomes relevant.

If it was the case that numerical abilities were dependent upon tracking of continuous quantities, we would expect infant abilities to discriminate area (and other continuous quantities) to be comparable to, or even more precise than, that of number. That is, given that infants can discriminate a 1:2 ratio of change in number at 6 months, we would expect this same age group to also be able to discriminate 1:2 ratio of change in area. Note, however, given claims that infants rely upon continuous quantities when tracking number – in the context of sets of objects – it is important to assess infant abilities to track continuous quantities using comparable stimuli – that is, in the context of *sets of objects.* Only a handful of studies have examined the acuity of infant discriminations of continuous properties in the context of sets, and the ones that have suggest that infants are relatively poor discriminators of continuous extent in numerical contexts. For example, when CA is the sole relevant cue for discrimination (i.e., number and other continuous cues are controlled for), 6-month olds have been found to require as much as a 4-fold change in CA to detect a change (Brannon et al., 2004; Cordes & Brannon, 2008; M. E. Libertus et al., 2014). This low precision is remarkable given that infants of this age are able to track a 2-fold change in number. Thus, 6-month old infants are able to track number with significantly greater precision than changes in CA, contradicting claims of a Sense of Magnitude Theory (Leibovich et al., 2017).

However, it may be that CA discriminations are particularly difficult because representing *cumulative* area requires tracking surface area across multiple sets of objects, a process which could potentially require the engagement of a computation process (e.g. summing across all the items in the array to get a total estimate). If that is the case, one would imagine that tracking the size of an individual object within the context of an array of objects should be a significantly simpler process since it should not involve any computational processing – simply the representation of area of a single item. While it is true that research has revealed that 6-7-month-old infants can discriminate the area of a single item (presented in isolation) with comparable precision to that of their number tracking abilities (i.e., they can detect a 1:2 ratio change in EA, but not a 2:3 ratio – identical patterns as that found with number in this age group; Brannon, Lutz, & Cordes, 2006; see Feigenson, 2005 for a review) – research involving sets of objects have provided less support for SoM. Only one study to date has done this, and unlike the findings with a single element, Cordes and Brannon (2011) found that 7-month-olds failed to discriminate 1:3 ratio of EA when presented in the context of an array of items. In fact, infants needed as much as a 1:4 ratio of change to detect a difference. This finding is particularly surprising given that (in theory) the task demands should be identical whether infants are asked to track the area of a single element when it is presented singly, or in the context of an array of items; to succeed infants need only pay attention to the size of one item in the set. In practice, however, it seems that EA discriminations prove significantly more challenging for infants in the context of a set of items.

Together, the few studies that have examined continuous extent discriminations in infancy suggest that they are remarkably poor at using area alone to track quantity. Notably, however, since these studies have only tested infants around 6-7 months of age, it is unclear whether infants show a similar developmental trajectory in their area discriminations compared to number. If, as proponents of the Sense of Magnitude Theory suggest, continuous quantities are easier to discriminate than number, we would not only expect infants to discriminate EA with similar (or better) precision than number, but we would expect a parallel increase in acuity with age, similar to what we see in number. In terms of number, we know that infants can discriminate a 1:2 ratio of change at 6 months and a 2:3 ratio at 9 months. In terms of element area, previous studies suggest that they fail to discriminate a 1:3 ratio at 7 months, yet it is still unclear whether they improve in their area tracking abilities over the infancy period, and whether their area tracking abilities may eventually "catch up" to that of their numerical tracking abilities. In the current study, we examined the developmental trajectory of infant EA discriminations to determine whether infants improve in their EA discriminations with age. To do this, we tested 7- and 12-month-old infants on an EA discrimination task where both age groups were presented with a 1:3 ratio of change in area – a ratio of change 7 month olds have previously been shown to fail to detect. Replicating prior findings, we expected 7-month-olds to again fail to notice a 1:3 ratio of change (Cordes & Brannon, 2011); however, if infant area discriminations – like that of number – improve over the course of development and even potentially catch up to their numerical tracking abilities – we predicted that 12 month olds should succeed in detecting a 1:3 ratio of change.

Methods

Participants

Twenty 7-month olds (*M* = 6m29d, *Range* = 6m16d – 7m13d) and twenty 12-month olds (*M* = 12m4d, *Range* = 11m17d – 12m24d) participated in this study. An additional ten 7-month olds were excluded due to: parental or sibling interference (N=2), technical problems (N=2) or not finishing at least 4 out of 6 test trials (N=6). Furthermore, an additional eight 12-month olds were excluded for parental or sibling interference (N=2) or not finishing at least 4 out of 6 test trials (N=6).

Apparatus

Stimuli were presented on one 19'' inch computer monitor mounted on a black wall in a dimly lit room. Infants sat in a highchair or on their parent's lap facing the monitor, approximately 20" from the screen, with the video camera recording the infant's looking towards the monitor. For the purposes of online coding, a recording of the infant's face was presented on a large TV monitor in a separate experimental room. Using a gamepad, the online coder recorded the infant's eye gaze throughout the duration of the study. For the purposes of offline coding, the recording of the infant's face as well as a recording of the stimuli the infant saw was multiplexed and recorded using digital recording software.

Design

All infants were habituated to homogeneous sets of dots that were identical in EA. Across trials, the EA of the dots was kept constant, but the number of dots in each display varied (and thus CA of the arrays varied from trial-to-trial too). Following habituation, infants were presented with six test trials that alternated between two types: 1) novel EA trials, where the EA of the dots involved a 3-fold increase (or decrease, dependent upon condition) compared to habituation or, 2) familiar EA trials, where the EA of the dots was identical to those in habituation. The number of items in the two arrays differed from the average number of items in habituation by an equal proportion – making number an irrelevant cue for discrimination. Moreover, the CA of the two test arrays were equated, making CA an irrelevant cue as well. Whether participants' first test trial was novel or familiar was counterbalanced across participants. Additionally, as in Cordes & Brannon, 2011, for the novel test trials, half of participants were presented with trials where the novel arrays involved a 3-fold decrease in element area (Test Condition A), while the other half of participants were presented with novel arrays that involved a 3-fold increase in element area (Test Condition B).

Procedure

Infants were first presented with a short attractor video to orient them to the computer monitor. Once they looked at the attractor video for at least 2 consecutive seconds, they began the habituation portion of the experiment. During habituation, each display was presented until either 1) the infants had looked a minimum of .5 seconds to the display and then looked away for 2 consecutive seconds or 2) the infants had looked to the display for a maximum of 60 seconds. Infants moved to test once they met the habituation criteria, such that looking during the last three trials had declined 50% compared to the first three habituation trials in which the infant had looked for a total of at least 12 seconds (as per Cordes & Brannon, 2009, 2011; Xu & Spelke, 2000). Infants were presented with a minimum of 6 habituation trials and a maximum of 16 trials. If infants had still not met the habituation criteria after the maximum 16 habituation trials, they moved onto test anyway. There were a total of 6 different habituation stimuli that were presented in blocks such that each of the 6 stimuli were presented once before the set repeated a second time. The order of the trials within each block was randomized.

In test, infants were shown 6 test trials, alternating between novel and familiar test trials. Identical to habituation, the stimuli remained on the monitor until the infant looked to the display for a minimum of .5 seconds and then looked away for 2 consecutive seconds or if the infant looked to the display for a total of 60 seconds.

Stimuli

All dot arrays were green dots on a white background, created with Adobe Illustrator CS5 Software (See Figure 1.1 for stimuli). In habituation, each dot had an element area of $3cm^2$, with the number of dots used for each display spaced logarithmically between 6-30 dots, averaging out to 15.67 dots per display and thus with an average CA of $47cm^2$. In test, the number of dots in the display (8 or 24 dots) was approximately equidistant from the number of dots in habituation. Thus, for participants in Test Condition A (where the novel test display involved a 3-fold *decrease* in EA), the familiar test trials contained 8 dots of $3cm^2$ and the novel displays contain 24 dots of $1cm^2$, resulting in a CA of both test displays being $24cm^2$. For participants in Test Condition B (where the novel test display involved a 3-fold increase in EA), the familiar test trials contained 24 dots that were each $3cm^2$ and the novel displays contain 8 dots that were each $9cm^2$, resulting in a CA of $72cm^2$ for both types of test displays. The background display size was identical across habituation and test, making the density of the arrays in test also approximately equidistant from the average density in habituation.

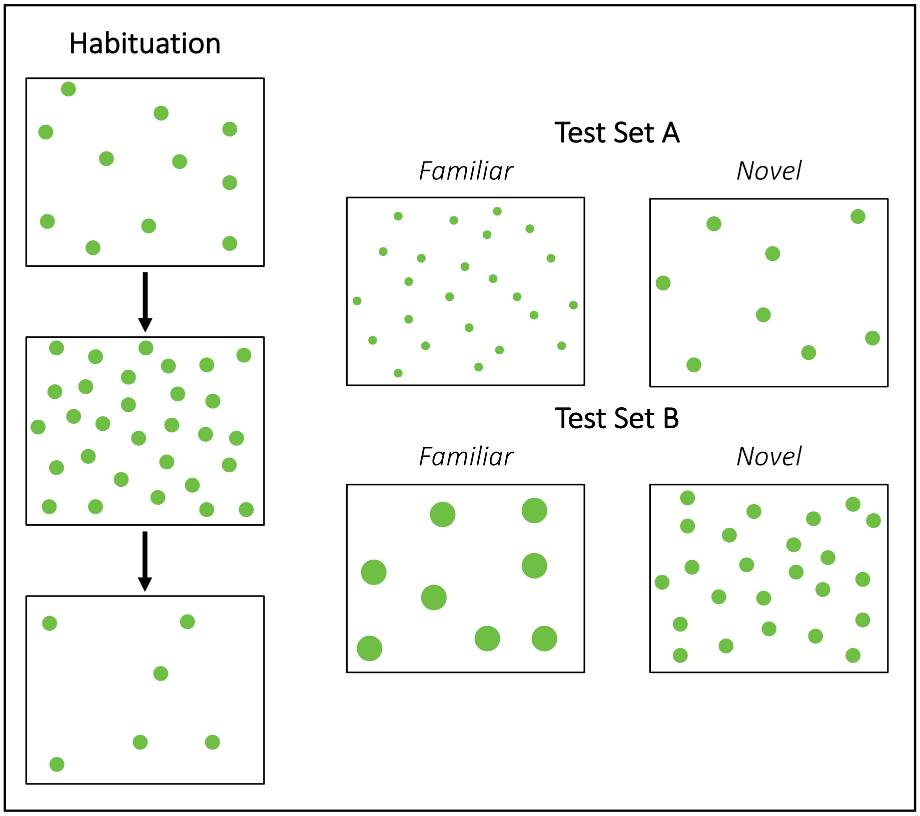

Figure 1.1 *Sample stimuli*
Infants were habituated to arrays that were constant in EA, but that varied in number and CA. In test, infants were presented with displays that were the same EA as in habituation (familiar displays) and displays that changed in EA by a 3-fold change. Test Set A *decreased* in EA 3-fold, Test Set B *increased* in EA 3-fold.

Data Processing & Coding

All videos were coded online by one experienced coder blind to the experimental condition, and offline by a second blind coder using Preferential Looking Coder (K. Libertus, 2008), a program that codes each frame (100ms) of the video. If reliability between these coders was less than 85% overall, or if any individual test trial reliability was less than 85%, a third coder recoded the video offline. Reliability was then calculated between all three coders, and the coder with the highest reliability with the other codes was used for data analysis. Using this coding system, reliability between coders was found to be 94.7%.

Looking times that were longer than 3 standard deviations away from the mean were treated as outliers and were replaced with the next longest looking time that was within 3 standard deviations (as per Cordes & Brannon, 2009, 2011).

Results

7-month olds

A paired samples t-test revealed that infants significantly reduced their looking from the first three (M = 10.62, SE = 1.39) to last three habituation trials (M = 6.57, SE = .89; $t(19) = 2.63$, $p = .02$). Seven babies met the habituation criterion.

A 2 x 2 repeated measures ANOVA testing the within-subjects factor of Test Trial Type (Novel or Familiar) and the between-subjects' factor of Test Condition (Decrease in Size vs Increase in Size) found no main effect of either Test Trial Type ($F(1,18) = .08$, $p = .78$, $\eta_p^2 = .004$) or Test Condition ($F(1,18) = .28$, $p < .59$, $\eta_p^2 = .02$), nor any interaction ($F(1,18) = .44$, $p = .52$, $\eta_p^2 = .02$; See Figure 1.2). Only 9 out of 20 infants looked longer to the novel test trials, which was not significantly different from chance ($p = .75$, Binomial statistic). Replicating prior research Cordes & Brannon (2011) our findings reveal that 7-month olds failed to detect a 3-fold change in EA in the context of sets.

12-month olds

We again ran a paired samples t-test and found a significant reduction in looking from the first three (M = 6.72, SE = .55) to last three habituation trials (M = 3.86, SE = .49; $t(19) = 4.11$, $p < .001$). Eleven babies met the habituation criteria.

A 2 x 2 repeated measures ANOVA testing the within-subjects factor of Test Trial Type (Novel or Familiar) and the between-subjects' factor of Test Condition (Decrease in Size vs Increase in Size) found a marginal main effect of Test Trial Type ($F(1,18) = 2.95$, $p = .10$, $\eta_p^2 = .14$) and a significant effect of Test Condition ($F(1,18) = 190.99$, $p < .001$, $\eta_p^2 = .91$) such that participants had longer looking times ($M = 6.38$, SE = .54) overall in the Increasing condition (when the items in test were 3 times larger than those in habituation) compared to the decreasing condition ($M = 4.26$, SE = .54; when the items in test were 3 times smaller than those in habituation). There was no significant Test Trial Type by Test Condition interaction ($F(1,18) = .47$, $p = .50$, $\eta_p^2 = .03$).

Importantly, 15 out of 20 infants looked longer to the novel test trials ($p = .02$, Binomial statistic) which was significantly above chance. Furthermore, when we looked at the first pair of test trials, we did find that infants looked significantly longer to the first novel ($M = 7.49$, SE = 1.05) compared to the first familiar test trial ($M = 5.15$, SE = .68; $t(19) = 2.07$, $p = .053$, $d = .46$) although this was not the case for the second or third pair of test trials (p's > 29). This suggests that although infants as a group only revealed a marginal preference in looking towards the novel test trial, when examining the first test trials alone infants did show a stronger preference for the novel test trials. Additionally, individually a significant majority of infants looked longer to the novel test trials.

Although we do not have strong evidence for 12-month olds successfully discriminating a 3-fold change in EA, this ability does seem to have improved marginally with age.

Combined Analysis

Next, we conduct an omnibus ANOVA testing the within-subjects factor of Test Trial Type (novel or familiar) and the between-subjects factor of Age Group (7 or 12 months). Similar to our findings with each age group separately, there was no main effect of Test Trial Type ($F(1,18) = 1.86$, $p = .18$, $\eta_p^2 = .05$), nor was there a main effect of Age Group ($F(1,18) = .28$, $p = .60$, $\eta_p^2 = .01$). However, there was a (barely) marginally significant interaction between these two variables ($F(1,18) = 2.72$, $p = .11$, $\eta_p^2 = .07$).

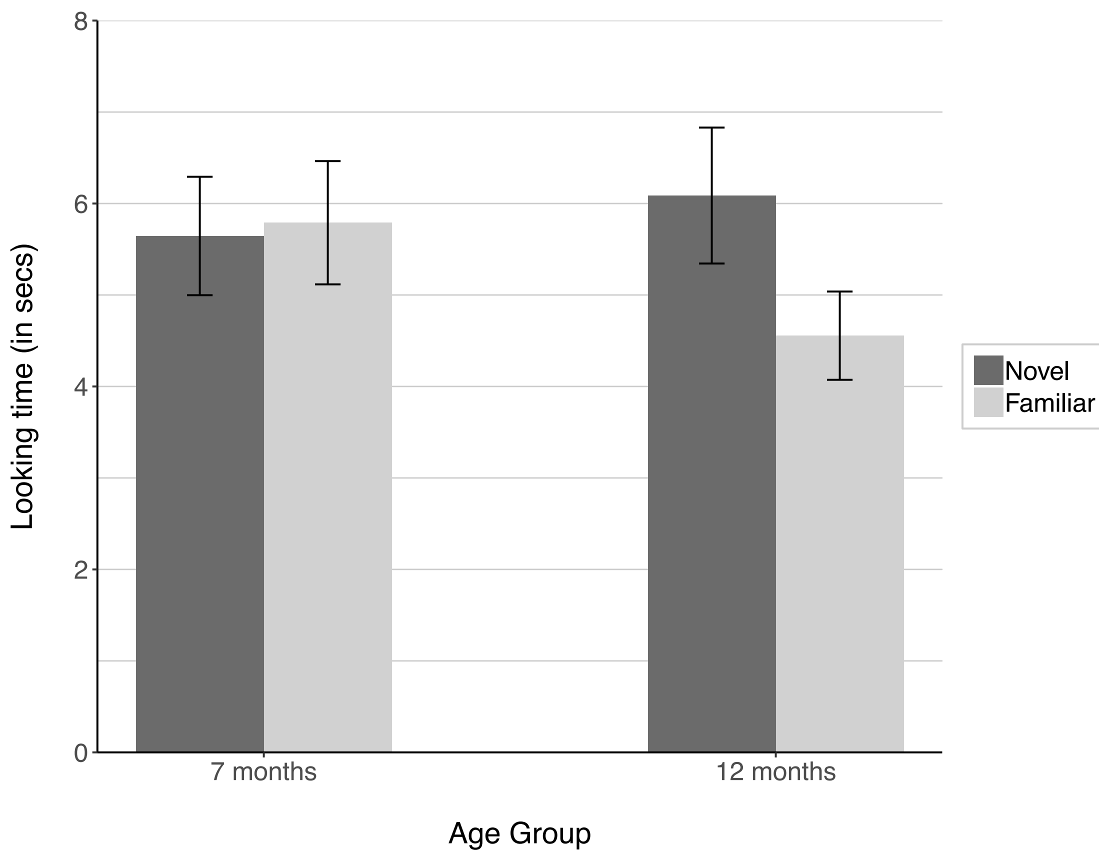

Figure 1.2. *Mean looking times for the novel and familiar test trials for 7- and 12-month olds.* Error bars represent standard error.

Discussion

Although a myriad of studies have examined infant abilities to make numerical discriminations, less is known about the developmental trajectory of infant abilities to track continuous quantities such as cumulative or element area. This is particularly relevant given that proponents of the Sense of Magnitude (SoM) Theory suggest that infants rely on perceptual, continuous quantities to track quantities (Gebuis & Reynvoet, 2012a; Henik et al., 2011; Leibovich et al., 2017; Mix et al., 1997) and, as such, should be better at discriminating continuous quantities than number. If so, we would not only expect infants to be able to discriminate smaller ratios of change for EA than that of number, but they should also demonstrate improvements in their EA discrimination abilities during infancy, at the very least at the same pace as that of number.

In the current study, we provided the first test of infant EA discrimination abilities at two different developmental time points: 7 and 12 months. We found that at 7-months, infants failed to discriminate a 1:3 ratio of change in EA, replicating previous findings (Cordes & Brannon, 2011). At 12 months, however, the data area somewhat mixed, revealing marginal success in detecting a 3-fold change in EA. While the findings are not robust, the implications are. Importantly, it is clear that even by 12 months of age, infant abilities to track EA are weak, and not anywhere near their numerical tracking capacities. Prior research suggests that infants at this age are able to discriminate a 2:3 ratio change in number (or possibly even finer ratios, Lipton & Spelke, 2003), suggesting that they are almost 200% more precise at discriminating changes in number compared to EA. Furthermore, this means that infants not only start off with a lower level of area tracking acuity in early infancy (around 6-7 months of age), but their precision in area tracking improves at a much slower rate than that of number over the course of the infant period.

All in all, these findings further undermine claims made by the SoM theory positing that infants rely upon perceptual quantities, such as EA, instead of number to track quantities (Gebuis & Reynvoet, 2012a; Henik et al., 2011; Leibovich et al., 2017; Mix et al., 1997). Not only does previous research suggest that infants *can* make numerical discriminations independent of continuous cues (Lipton & Spelke, 2003; Xu & Spelke, 2000; Xu et al., 2005), our research (together with others: Brannon et al., 2004; Cordes & Brannon, 2009, 2011; Starr & Brannon, 2015) reveals that infants are remarkably poor trackers of continuous quantities in the context of numerical sets. As such, it seems highly unlikely that infants would resort to tracking continuous quantities over number when presented with a set of objects.

How do we interpret the finding that infants are able to discriminate EA with a higher level of precision when presented with a single item (Brannon et al., 2006), compared to when presented with an array of items? In theory, the process by which children complete these tasks should be the same; to succeed infants need only pay attention to the size of one item, whether or not those items are presented in the context of an array of items or as a single item. In practice, however, EA discriminations become much more difficult in the context of a set of items. Why might this be the case? One possibility is that the presence of multiple items in an array encourages infants to pay attention to all items, and this spreading of attention across items takes away from the level of attention they pay to any single item, decreasing their precision in representing EA. Another possibility is that the mere presence of multiple items encourages infants to focus on quantitative features that pertain to the set as a whole, such as the number of items in the set or their CA (although this seems unlikely given previous evidence showing that infants are also poor discriminators of CA; Brannon et al., 2004; Cordes & Brannon, 2008). However, since infants are not given any reliable cues for number or CA during the habituation trials, they fail to form a coherent representation of items in habituation which leads to failure to dishabituate to the novel display in test. That is, changes in number across the sets may, in fact, *detract* from EA discriminations in this context. Future work would benefit from addressing to what extent the size of the set (whether presented with a single item or an array of items) may affect how infants represent quantity.

The current study provides further evidence in the longstanding debate as to which quantitative dimensions are most salient to infants: number or continuous quantities. Not only did we replicate previous findings showing that infants start off with a lower level of EA acuity at 6-7 months compared to that of number (failing at a 1:3 ratio for EA but succeeding at a 1:2 ratio for number; Xu & Spelke, 2000), we also demonstrated that infants improve at a slower rate for EA discriminations compared to that of number. Our findings, along with others from the previous literature, continue to undermine claims made by SoM theory, instead providing strong support for the idea that infants may, in fact, rely upon number when tracking quantity.

CHAPTER 2: THE IMPACT OF SET SIZE ON CUMULATIVE AREA JUDGEMENTS

Abstract

The ability to track number has long been considered more difficult than tracking continuous quantities. Evidence for this claim comes from work revealing that continuous properties (specifically cumulative area) influence numerical judgements, such that adults perform worse on numerical tasks when cumulative area is incongruent with number. If true, then continuous extent tracking abilities should be relatively precise, and unimpeded by numerical features. However, few studies have directly examined this hypothesis by characterizing adult abilities to discriminate arrays on the basis of continuous quantities. The aim of the present study was to determine the precision with which adults track cumulative area and to uncover the process by which they do so. We presented adults with arrays of dots (of differing set sizes) and asked them to judge the relative cumulative area of the displays. Similar to the number discrimination literature, we found cumulative area judgements to be ratio-dependent. More interestingly, however, we found that participants not only performed worse on trials where number was incongruent with cumulative area, but that adults performed worse as set size increased. These findings suggest that number interferes with continuous quantity judgements, suggesting that it is at least *as salient* as continuous variables, undermining claims in the literature suggesting that continuous properties are easier to represent, and more salient to adults.

The Impact of Set Size on Cumulative Area Judgements

Representing quantity is an important skill for human and non-human animals alike. Whether you are a human deciding just how many apples you will need to make your favorite apple pie, or a mosquitofish deciding where in the ocean you can find the highest density of zooplankton, the ability to represent approximate quantities is important for day-to-day life. However, which quantities we rely upon for these important decisions has been a topic of debate (Gebuis & Reynvoet, 2012a; Leibovich et al., 2017; Savelkouls & Cordes, 2017). Human and non-human animals can represent discrete quantity (i.e. number; Humans: e.g., Halberda & Feigenson, 2008; Non-human animals: e.g., Brannon & Terrace, 1998; Meck & Church, 1983) but they can also represent continuous quantities (also referred to as continuous extent[1]) such as area, volume, length or density (Humans: e.g., Brannon, Lutz, & Cordes, 2006; Odic, 2018; Non-human animals: e.g., Boysen, Berntson, & Mukobi, 2001). Furthermore, these discrete and continuous quantities are strongly correlated with one another: e.g., 10 apples are not only more numerous than 5 apples, but their cumulative volume, weight, surface area, and density are also greater. This naturally strong correlation between discrete and continuous variables has led researchers to question the extent to which we track these quantities independently of each other.

A majority of the research investigating humans' quantitative abilities has focused on our ability to represent discrete number. While substantial research has supported the idea that infants, children, and adults are remarkably good at representing number (Halberda & Feigenson, 2008a; Odic et al., 2015; Xu & Spelke, 2000), not everyone

[1] Note: For the purposes of this paper, "continuous quantity" will exclusively refer to visual quantities, not e.g., time.

agrees. Proponents of the “Sense of Magnitude” (SoM) theory take a neo-Piagetian approach to number representation, suggesting that our abilities to track number are fully reliant upon an ability to track continuous quantities (Gebuis & Reynvoet, 2012b; Leibovich et al., 2017). The premise of this argument is that continuous quantities such as element area (EA), cumulative area (total area of all items in an array; CA) or density (the ratio of the number/area of items and the size of the display) are derived from and dependent upon the perceptual qualities of the display, and thus are significantly easier to track than number. In contrast, number is considered to be an abstract quantity –it can be tracked using many different sensory modalities (vision, sound and even touch), and even compared across modalities (e.g. it is possible to compare the number of voices heard to the number of people seen). As such, the ability to track number has been considered to involve much higher order cognitive processes than tracking continuous quantities, making it unlikely that number is tracked in the presence of other perceptual quantities.

As a direct test of these claims, researchers have investigated whether we can track number independent of continuous properties. Although substantial work reveals that continuous properties can bias numerical judgements (Gebuis & Reynvoet, 2012a; Hurewitz et al., 2006; Leibovich, Henik, & Salti, 2015), researchers have successfully developed paradigms that systematically control for continuous properties that typically correlate with number, providing strong evidence that humans are capable of tracking number independent of continuous perceptual variables (Halberda & Feigenson, 2008b; Lipton & Spelke, 2003; Odic et al., 2013; Starr, Libertus, & Brannon, 2013a; Xu & Spelke, 2000; Xu et al., 2005). However, less work has explored the converse; that is, how well can we track continuous properties independent of number? Are continuous properties easier to track? Notably, as posited by the SoM theory, numerical abilities would only be dependent upon continuous extent tracking if and only if continuous extent representations are more precise and more salient than numerical ones. If our abilities to track continuous properties are less refined than our abilities to track number, then it is unlikely that we would rely upon less precise continuous representations. Although a myriad of studies have examined number discrimination abilities in the context of competing continuous extent information (Barth, 2008; DeWind & Brannon, 2012; Gebuis & Reynvoet, 2012a; Hurewitz, Gelman, & Schnitzer, 2006), little research has directly examined our abilities to discriminate arrays on the basis of continuous quantity in the context of competing numerical information. The aim of the present study was to examine adult abilities to discriminate CA across arrays with differing numerical information. We had two goals: (1) to determine the precision with which adults track CA across various set sizes and (2) to investigate if and how CA discriminations are influenced by numerical information. That is, we aimed to understand the process by which we track continuous quantities – are CA representations dependent upon number? This latter question is of theoretical importance because it can speak to how our representations of number and continuous quantities may be related.

How well do adults discriminate CA?

Most studies that have examined adult CA tracking abilities have used discrimination or numerical Stroop type tasks. In these types of tasks, on some trials continuous properties and number are incongruent with one another (e.g. the array with a greater number of items has a smaller CA) and on other trials, number and continuous properties are congruent (e.g. the array with the greater number of items also has a greater CA). Although explicit instructions are to judge the relative numerosity of the displays (i.e., not to attend to cumulative area), multiple studies have found that adults consistently perform worse on incongruent compared to congruent trials (Barth, 2008; DeWind & Brannon, 2012; Gebuis & Reynvoet, 2012a; Hurewitz, Gelman, & Schnitzer, 2006). This has been taken as evidence that even when given explicit instructions to pay attention to number, adults *automatically* process continuous properties of the set (even when irrelevant to the task). These findings have provided the basis for claims that continuous perceptual properties are more readily and precisely tracked than number.

However, is it true that adults track CA with relatively greater precision than that of number? If numerical judgements were fully dependent upon continuous properties, then one would expect our ability to track continuous quantities to be more refined than that of number. That is, humans should be at least as good at discriminating arrays based upon continuous quantities as they are at discriminating arrays based upon number. However, very few studies have specifically examined adult abilities to discriminate continuous properties. A handful of studies have compared area and numerical tracking abilities in infants, children, and adults reporting similar, or even more precise, abilities to discriminate the area of a single item relative to their abilities to discriminate number (Brannon et al., 2006; Leibovich & Henik, 2014; Odic et al., 2013). Yet, critically, the only way to address claims (such as those made by SoM theory) that continuous extent is more readily tracked over number is to examine the precision of CA representations in the context of sets where numerical information is available. Importantly, these studies have found performance costs for incongruent trials such that CA judgements are less accurate when number is incongruent with CA (Barth, 2008; Hurewitz et al., 2006)[2] . That is, number may be just as salient as number in the context of sets[3].

There is some evidence to suggest that continuous quantity discriminations may be less precise than numerical discriminations, at least in human infants. Studies that have investigated 6-7-month-olds' abilities to track the size of individual objects (element area, or EA) when presented in the context of an array of items, or the CA of an array of items have found that infants needed as much as a 1:4 ratio of change in both EA and CA to detect a change. This 1:4 ratio of change is notably greater than the 1:2 ratio of change necessary to detect changes in number (Brannon et al., 2004; Cordes & Brannon, 2008, 2011). In sum, the infant literature suggests that infants are better at tracking number than continuous dimensions in the context of sets.

However, research with infants presents some limitations that make it difficult to fully uncover the extent of human CA representations. First of all, infant measures of quantity discrimination are always implicit (i.e. one cannot give infants any instructions as to what they should do in a task), thus one has little control over what they are doing in a task. It is impossible to be explicit about task demands to the infant, confounding salience of a dimension with precision. That is, it is conceivable that infants are

---

[2] Barth (2008) ran additional models of the data and concluded that this decreased performance in incongruent trials was not due to interference between the two dimensions of number, but instead could be explained by the fact that participants underestimated individual element areas resulting in more difficult discriminations. Importantly, we highlight that these findings are consistent with claims that area judgements are more difficult than numerical ones.

[3] Note Hurewitz et al. (2006) reported CA interfered with numerical judgements more so than number interfered with CA judgements. Importantly, however, the authors did not systematically match the ratio of change across the two dimensions, such that numerical differences across displays may have been significantly smaller (and thus less salient) than CA differences. As such, it is inappropriate to judge relative salience of the two quantities.

incredibly precise trackers of CA, but simply less likely to focus on this dimension in general. Secondly, infant's limited attention abilities mean that researchers can only present infants with a few trials, which limits the types of questions researchers can answer. Questions regarding whether changes in number may influence CA tracking abilities and/or how precisely we can track small changes in CA are not addressable with this population. Moreover, CA tracking abilities may change over the course of development, likely due to changes in maturation, other domain-general abilities (i.e., working memory, attention), and visual acuity, making it important to investigate CA tracking in adult populations in addition to younger populations.

Processes involved in CA representation

So how do we represent CA when presented with an array of items? The previous literature has presented us with two mutually exclusive theoretical possibilities.

One possibility, which we will refer to as the 'Direct Perception' hypothesis, is consistent with assumptions of the Sense of Magnitude (SoM) theory. According to the Direct Perception hypothesis we are able to track surface area directly from the perception – that is, we directly perceive exactly how much area is covered just as readily as we notice the color or luminance of the items. Importantly, this direct abstraction does not require extensive cognitive processing, such as summing surface area across individual items, and thus does not require individuating items in the array. As such, the number of items in the display – that is the number of items over which CA is tracked – is irrelevant to CA tracking and thus set size should have no impact on the precision of CA acuity. Support for this hypothesis comes from studies revealing similar infant CA discrimination abilities for small sets (2-3 items) as that of large sets (10-15 items)[4], suggesting that CA acuity is unaffected by set size (Cordes & Brannon, 2008). This Direct Perception Hypothesis underlies many neo-Piagetian claims (e.g. Gebuis & Reynvoet, 2012b; SoM: Leibovich et al., 2017; Mix, Huttenlocher, & Levine, 2002) positing the direct perception of continuous extent quantities, thus making them easier to track than abstract quantities, like number.

On the other hand, we propose an alternative possibility, known as the 'Computation' hypothesis (see also Barth, 2008). Rather than representing CA directly, according to the Computation hypothesis, we may track CA by representing the surface area of individual items within an array (likely through direct perception of the surface area of individual items) and then summing across these representations (e.g. adding representations of individual areas together). Because prior research suggests that mental summation is not a completely noiseless process (Cordes et al., 2007), each addition process contributes noise to the representation. Thus, precision in the representation of CA should decrease as the number of elements in the display increases. Unlike the Direct Perception hypothesis, under the Computation hypothesis CA acuity should be affected by the number of items in the display, with worse acuity as the set size increases. The Computation hypothesis is supported by prior research revealing infants are significantly better (i.e., more precise) at tracking the area of a single item, compared to tracking the CA of multiple items (Brannon, Abbot, & Lutz, 2004; Brannon, Lutz, & Cordes, 2006;

[4] This study relied upon a standard habituation looking-time paradigm, revealing that infants were capable of discriminating a 4-fold, but not a 3-fold, change in CA, across exclusively small and exclusively large sets. Because infant habituation techniques do not lend themselves to fine-grained assessments of discrimination abilities, it is not possible to determine whether discrimination capabilities may have varied somewhat as a function of set-size between the 4-fold and 3-fold changes.

Cordes & Brannon, 2008). Additionally, Barth (2008) compared quantitative models of adult CA judgements and determined that a summation account provided the best explanation for the data.

The Current Study

Importantly, no studies have directly examined the effect of set size on cumulative area tracking abilities in adults. With supporting evidence for both the Direct Perception and Computation hypotheses, it is still unclear how we track CA and how these abilities may hinge upon numerical information in the display. One key distinction between the two hypotheses is in the role that number plays in CA discriminations. While the Direct Perception hypothesis assumes that number should have no effect of CA acuity, the Computation hypothesis predicts less CA acuity with increasing set size. In the current study, we presented adults with a CA discrimination task in which they judged which of two arrays of items had a greater CA. We manipulated 4 variables: the CA Ratio (the ratio between the CAs of two displays – a way of varying the relative difficulty of the comparison to provide a means of assessing CA acuity), Set Size (how many items were in each display), the Number Ratio (the ratio between the number of items in the two displays) and Congruency (whether the display with the larger number of dots had the smaller or greater CA). By manipulating these 4 variables, we explored the following questions:

(1) *Does numerical congruency matter for CA judgements?* If (according to the SoM theory) CA is relatively more salient than number, then CA judgements should be unaffected by whether or not numerical information is consistent or inconsistent with CA magnitude. On the other hand, some prior work suggests that numerical information is

automatically processed, even in the context of a CA judgement task (Barth, 2008; Hurewitz et al., 2006), suggesting both quantities may be similarly salient. If so, then congruent numerical information should promote CA discrimination performance and/or incongruent numerical information should hinder CA discrimination performance. We aim to both explore whether numerical congruency has an impact and then, by comparing performance to a number neutral condition (in which both displays have an equal number of items), we will determine whether numerical congruency either facilitates and/or hinders CA judgements.

(2) *Do numerical differences across the displays impact our CA tracking abilities?* That is, do participants perform worse on incongruent trials when the difference in number between the two arrays (i.e., the ratio between the number of dots in each array) is more salient than when it is less salient (i.e. the ratio between the two set sizes is closer to 1)? Conversely, might bigger numerical differences on congruent trials support CA judgements? To investigate this question, we presented 3 distinct Number Ratios (1, 1.33 and 1.5) across trials to explore how numerical changes may matter for CA judgements.

(3) *How does set size impact CA acuity?* To investigate whether larger sets produce less accurate CA representations (as predicted by the Computation hypothesis), we presented participants with arrays containing exclusively small (2-4 dots), medium (6-9 dots), or large (9-15) sets. Furthermore, since evidence suggests infants can track the area of a single item with greater precision than the CA of multiple items (Brannon, Lutz, & Cordes, 2006; Cordes & Brannon, 2008), we also presented participants with single item size comparisons to assess whether a similar pattern is found later in development.

Across two experiments, we addressed these research questions by asking adult participants to rapidly judge which of two simultaneously presented visual arrays had the larger CA. Importantly, number was irrelevant to our task demands and thus should have had no influence on performance. In both experiments, we manipulated the numerical and CA ratio, set size and numerical congruency.

Experiment 1

Given that few studies have examined adult CA judgements, Experiment 1 examined the effect of Set Size on adults' CA discrimination performance. Participants completed a discrimination task in which they were asked to choose which of two arrays of blue dots had a greater CA.

Methods

Participants

Seventy-eight Boston College students participated in our study in exchange for cash or course credit (66 female, *M* = 18.89 years, *Range* = 18-26 years). Informed consent was obtained from all participants.

Procedure

Each participant completed the study on a computer with a 22'' monitor. Participants were first presented with an instruction screen that informed them that on each trial, they should choose the display of dots with the "greater amount of blue, therefore the greater cumulative area of blue." Each trial consisted of two side-by-side displays of blue dots and participants made a forced choice judgement about the presented pair of displays by pressing the left or right arrow key on the keyboard. They were first presented with a minimum of four practice trials; only once they had responded correctly on three of the four practice trials did they move onto the test trials (all participants moved to the test trials after the first set of four practice trials). The practice trials were designed to be very easy for the participant, with the CA Ratio changing 3-fold across the two displays, and the number of items (set size) varying from 2-12 dots (this was identical in range to the test trials). Next, participants received 190 test trials, with a break every 50 trials (3 breaks total). Participants were encouraged to look away from the computer and talk to the experimenter during the break. The order of the trials, as well as which display was presented on the left or right side of the screen was randomized for each participant.

Across trials we manipulated the following variables: CA Ratio (1.15, 1.33, 1.45, 1.6, or 1.9), Number Ratio (1, 1.33 or 1.5), Congruency (Congruent, Incongruent), and Set Size (Small (2-4), Medium (6-9) or Large (9-15) sets) to determine how these factors influenced participants' CA judgements. Participants experienced a total of 180 trials involving arrays of multiple items: 5 CA ratios x 3 Number Ratios x 2 Congruency x 3 Set Sizes x 2 trials. In addition, intermixed amongst the multiple item trials were 10 single item trials: 2 trials x 5 CA ratios (See Table 2.1 for breakdown of trials). Importantly, however, trials involving the Number Ratio 1 were neutral trials since they were neither congruent nor incongruent. Lastly, for each CA Ratio, we included two trials that were "single" trials in which each display contained only one dot. These trials would allow us to compare participants' area discriminations involving a single item compared to CA discriminations involving multiple items (trials with small, medium and large set sizes).

Number Ratio	1, 1.33 or 1.5						Single Trials
Set Size	Small		Medium		Large		
Congruency	Cong.	Incong.	Cong.	Incong.	Cong.	Incong.	

Table 2.1. *Experiment 1: Breakdown of the number of stimuli per variable manipulated.* Participants were tested on 5 CA Ratios (1.15, 1.33, 1.45, 1.6, or 1.9). Then for each CA Ratio, trials were broken down as illustrated above for the variables of Number Ratio, Set Size and Congruency. This resulted in total of 180 multiple items trials (5 CA ratios x 3 Number Ratios x 2 Congruency x 3 Set Sizes x 2 trials = 180 trials). Additionally, participants were tested on 2 single trials per CA Ratio, for a total of 10 single trials. In total, participants therefore were presented with 190 trials (180 multiple items trials + 10 single trials).

Upon completing the test trials, participants were asked two questions. The first question was open-ended and asked participants to briefly describe the strategies they used to make their decisions in the task. We included this question as a general gauge of the approaches people took when comparing cumulative areas. The second question was multiple-choice and asked participants which of the following choices best described what cues they used to make their decision(s). Participants chose one from a list of five possible answers: (1) I squished all the dots together for each image and used that as a cue, (2) I used the number of dots in each image as a cue, (3) I used the density of dots in each image as a cue, (4) I estimated the average size of the dots on each image and used that as a cue, (5) I tried to find the smallest/biggest dot in each image and used that as a cue. We included these options because we predicted that these would encompass the most commonly used strategies[5].

Stimuli

Stimuli were created using Adobe Illustrator (See Figure 2.1 for stimuli). For our CA values, we generated a list of 12 random CA values between 20-45 cm^2 (this range was deemed reasonable for our display size, ensuring that each individual dot would not become so small that they would be hard to see, or so big that they would not fit within the stimulus background) and used these 12 values for each Number and CA Ratio. These random numbers were then multiplied by the appropriate CA Ratio to determine the CA of the comparison display. Thus, across all CA Ratios tested, the CA values ranged between 23-83.6 cm^2.

To ensure that participants would not be able to use the size of individual dots as a cue for discrimination, the dots in each array were heterogeneous in size. The individual dot sizes were randomly chosen to fall within 35% of the average element area (as per Lidz, Pietroski, Halberda, & Hunter, 2011). We also controlled for item density (which we here defined as the number of items per given display size) so that the two arrays being compared had identical densities, although across trials densities did vary from .009 - .03 items/cm^2. Thus, the item background ranged from 226.7 cm^2 to 460cm^2.

[5] Participants' responses on these two questions was not related to their actual performance on the task and therefore these questions have not been further analyzed in our results section.

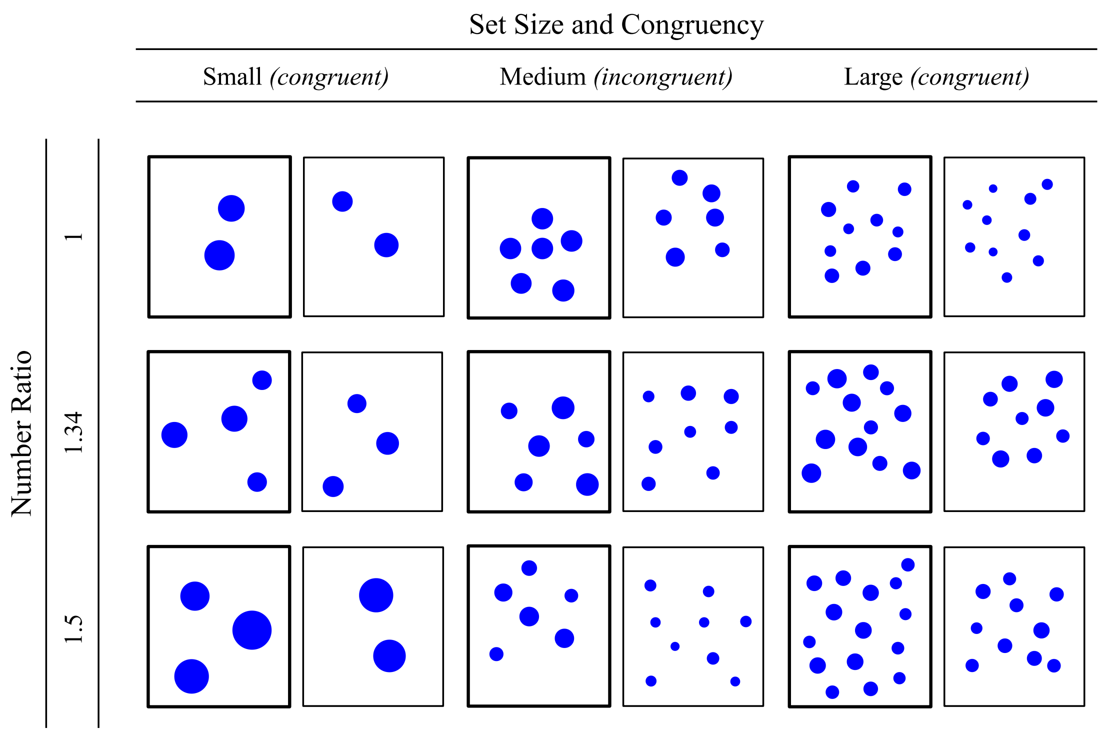

Figure 2.1. *Example stimuli pairs for the 1.6 CA Ratio.*
Stimuli are broken down by Number Ratio, Set Size, and Congruency. For each stimuli pair, the image on the left with the darker border has the larger CA.

Data Processing

Our primary dependent measure involved accuracy on the task. Participants whose performance fell 3 standard deviations above or below the mean for their overall percentage correct on the task were excluded from any analyses (N=1). We also calculated Weber fractions for each participant using only the data from the trials with sets greater than 1. Weber fractions (w) are defined as the smallest change between two quantities that can be reliably be detected. We estimated w using a psychophysical model using Gaussian random variables as has been done in previous research (Halberda & Feigenson, 2008b; Izard, Pica, Spelke, & Dehaene, 2008; Moyer & Bayer, 1976). In short, we inputted each participants' accuracy on the four hardest CA Ratios (1.15, 1.3, 1.45 and 1.6) and manipulated a single free parameter *w* until we found a Weber curve that best fitted the data and that minimized error (Halberda & Feigenson, 2008b).

Results & Discussion

On average, participants performed well on our task with 85.76% accuracy (*Range* = 60.53-98.42%). The average Weber fraction across all trials with multiple items was $w = .23$, (*Range* = .07-1.2; SD = .16, SE = .018). We compared this Weber fraction to the Weber fraction previously reported by Odic et al. (2013) with adults on a number discrimination task ($w = .13$, SE = .02, SD = .057, N = 8). Since the two samples varied so widely in their standard deviations and the number of participants tested, we conducted an independent samples t-test assuming unequal variances using the Welch-Satterthwaite procedure for unequal variances. This revealed that the weber fraction for our CA discrimination task was significantly higher than that reported for the numerical discriminations, $t(16.22) = 3.35$, $p < .01$, suggesting that CA acuity in our study was significantly worse than prior reports of numerical acuity (a higher weber fraction indicates lower acuity).

Does numerical congruency matter for CA judgements?

To explore our first research question, we specifically examined accuracy on those trials involving arrays with more than one item (excluding single item trials) in which the number of items differed across arrays (i.e., where numerical ratio did not equal one). We ran a 5 (CA Ratio: 1.15, 1.3, 1.45, 1.6, 1.9) x 2 (Number Ratio:1.33 or 1.5) x 2 (Congruency: Congruent vs. Incongruent) repeated measures ANOVA on data from those trials. Not surprisingly, results revealed a main effect of CA ratio ($F(4, 304) = 192.92$, $p < .001$, $\eta_p^2 = .72$) such that performance improved as CA ratio got larger. A paired samples t-test between all 5 levels revealed that performance on all CA ratios was significantly different from one another (p's < .001). Analyses also revealed a main effect of congruency ($F(1,76) = 19.06$, $p < .001$, $\eta_p^2 = .54$), such that participants performed significantly better on congruent ($M = 93.55\%$) compared to incongruent trials ($M = 71.30\%$), in line with previous research (Barth, 2008; Hurewitz et al., 2006) and in contradiction of predictions of the SoM Theory. Results also revealed a significant CA ratio x Congruency interaction (See Figure 2.2), $F(4,304) = 28.40$, $p < .001$, $\eta_p^2 = .27$. Paired samples t-tests revealed that regardless of CA ratio, participants performed significantly better on the congruent compared to incongruent trials (p's < .001); however, this difference became smaller as the CA ratio became easier (i.e., further from 1). That is, not surprisingly, numerical congruency had a greater impact on more difficult

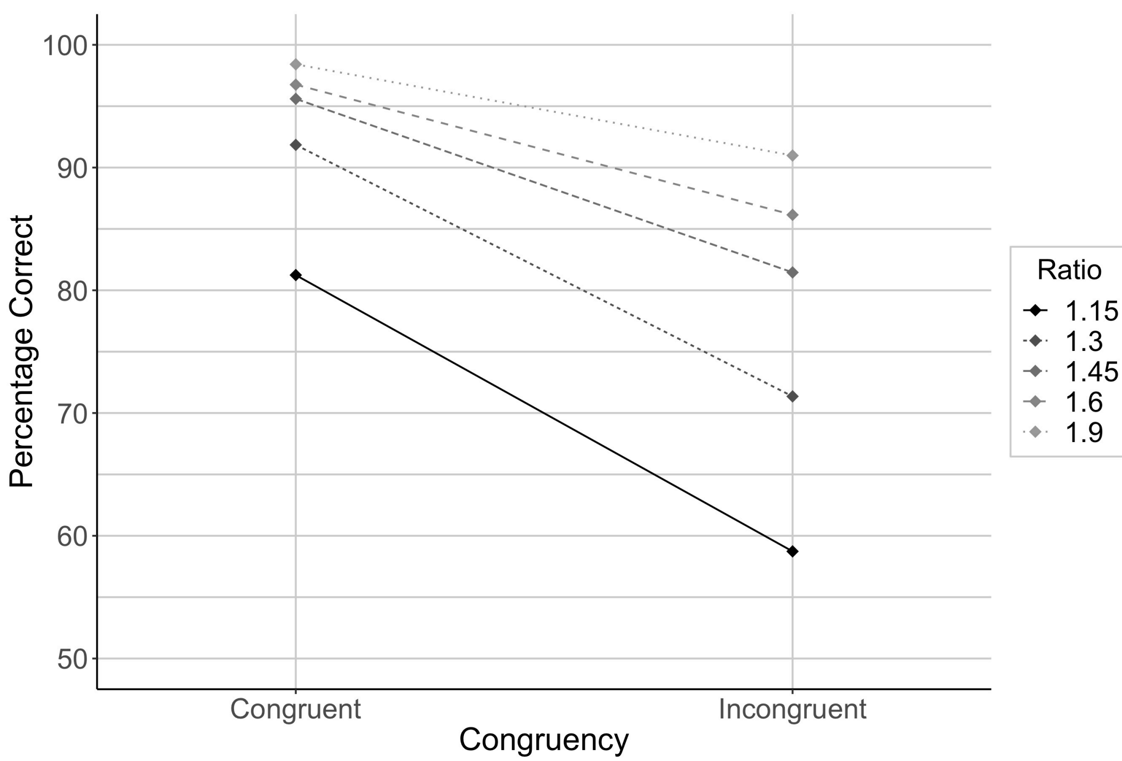

Figure 2.2. *Experiment 1: Percent Correct as a function of CA Ratio (1.15, 1.3, 1.45, 1.6 and 1.9) and Congruency (Congruent and Incongruent).*

CA discriminations. Lastly, analyses revealed a main effect of Number Ratio ($F(1,76) = 20.68$, $p < .001$, $\eta_p^2 = .21$) which was qualified by a significant Number Ratio x Congruency interaction ($F(1,76) = 16.03$, $p < .001$, $\eta_p^2 = .17$). Although adults performed better on congruent compared to incongruent trials for both numerical ratios (1.33 ratio: $t(76) = 8.26$, $p < .001$, $d = 1.40$; 1.5 ratio: $t(76) = 9.73$, $p < .001$, $d = 1.67$), the impact of congruency was greater when the numerical ratio between the two arrays was greater (that is, number mattered more when the numerical difference was more salient; 1.33 numerical ratio: $M_{difference} = 19.61\%$, $t(76) = 4.00$, $p < .001$, $d = .24$; 1.5 numerical ratio: $M_{difference} = 24.89\%$; $t(76) = 4.00$, $p < .001$, $d = .24$; see Figure 2.3). Furthermore, although there was no difference in performance between the two Number Ratios for congruent trials ($M_{difference} = 0.52\%$; $t(76) = .81$, $p = .41$, $d = .13$) there was for

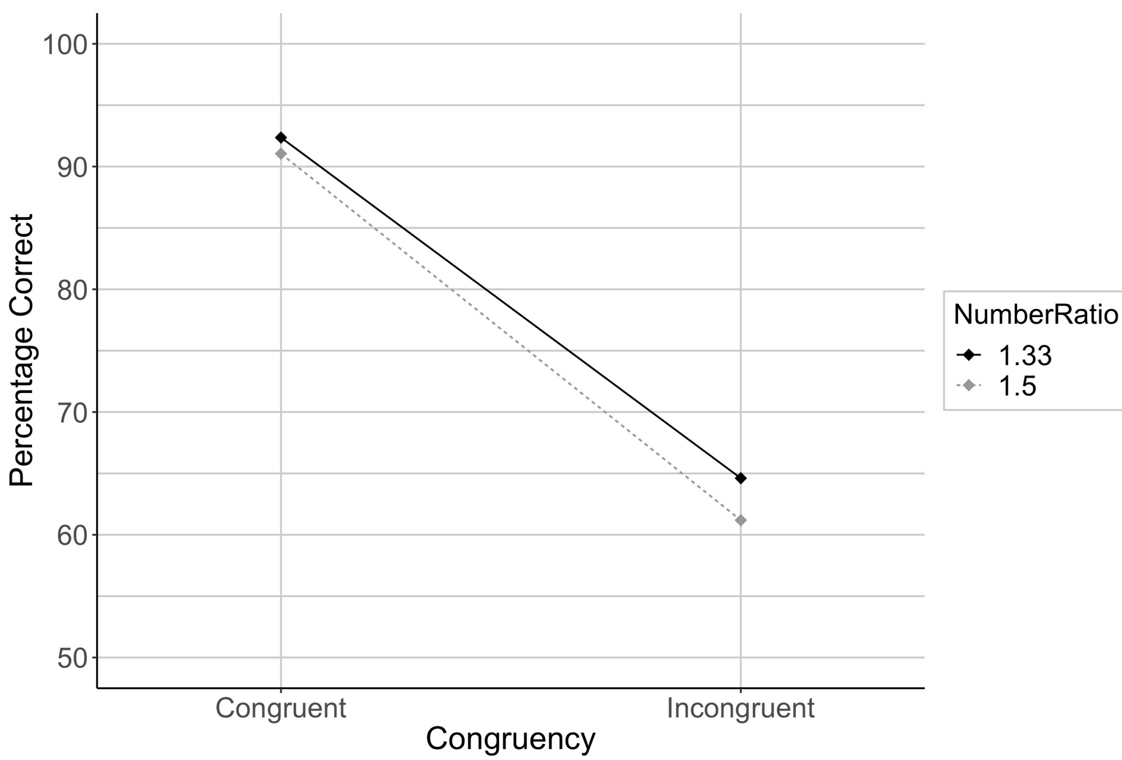

Figure 2.3. *Experiment 1: Percent Correct as a function of Number Ratio (1.33 and 1.5) and Congruency (Congruent and Incongruent).*

incongruent trials ($M_{difference}$ = 5.80%, $t(76) = 4.85$, $p < .001$, $d = .32$).

Do numerical differences across the displays impact our CA tracking abilities?

Next, we examined the impact of numerical congruency. That is, relative to neutral trials (trials where the number of items was the same in both displays i.e., Number Ratio 1) did numerical congruency promote performance, did numerical incongruency detrimentally impact performance, or both? We performed a repeated measures ANOVA comparing performance across all three types of trials (neutral, congruent, incongruent; collapsing across all numerical and CA ratios). The main effect of congruency was significant, $F(2, 152) = 90.86$, $p < .001$, $\eta^2 = .55$. Paired samples t-tests revealed that performance on the incongruent trials (71.30%) was significantly worse than on the

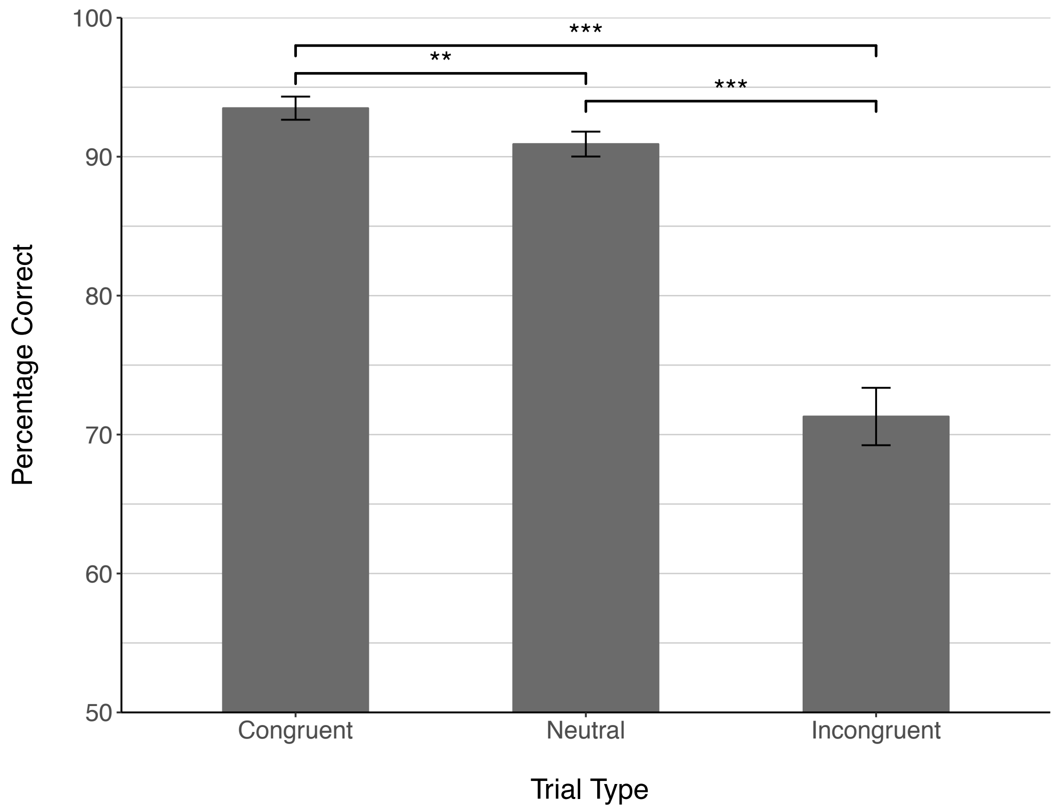

Figure 2.4. *Experiment 1: Percent Correct as a function of Number Ratio (1.33 and 1.5) and Congruency (Congruent and Incongruent).* Error bars represent standard error.

neutral trials (90.91%, $t(76) = 10.93$, $p < .001$, $d = 1.36$) and that performance on the congruent trials (93.55%) was significantly better than the neutral trials ($t(76) = 2.80$, $p < .01$, $d = .053$, See Figure 2.4). Thus, conflicting numerical information (i.e., incongruency) detrimentally impacted performance relative to neutral trials, but consistent numerical information (i.e. congruency) also facilitated performance relative to neutral trials. Although incongruent numerical information appeared to detrimentally impact performance significantly more so than congruent numerical information benefited performance, it should be noted that the high level of performance on congruent trials may have led to ceiling effects in performance, limiting the extent to which performance could benefit from congruent numerical information.

How does set size impact CA acuity?

To investigate how differing set sizes impacted CA discrimination performance, we ran a repeated measures ANOVA comparing performance across the four set sizes (single item, small set, medium set, large set). Importantly because single item trials necessarily were number neutral (i.e., a comparison of one item to one item cannot involve congruent or incongruent trials), we limited this analysis to only number neutral trials (those trials where number was identical in both arrays i.e., Number Ratio 1). Analyses revealed a significant effect of set size ($F(3,228) = 8.05$, $p < .001$, $\eta^2 = .10$), driven by significantly better performance on the Single trials ($M = 94.36$) compared to Small ($M = 91.22$), Medium ($M = 90.26$), or Large sets ($M = 89.81$, p's $< .01$, d's $> .95$). Although performance tended to decrease as a function of increasing set size, the difference in performance across Small, Medium, and Large set sizes did not reach significance (p's $> .15$; See Figure 2.5).

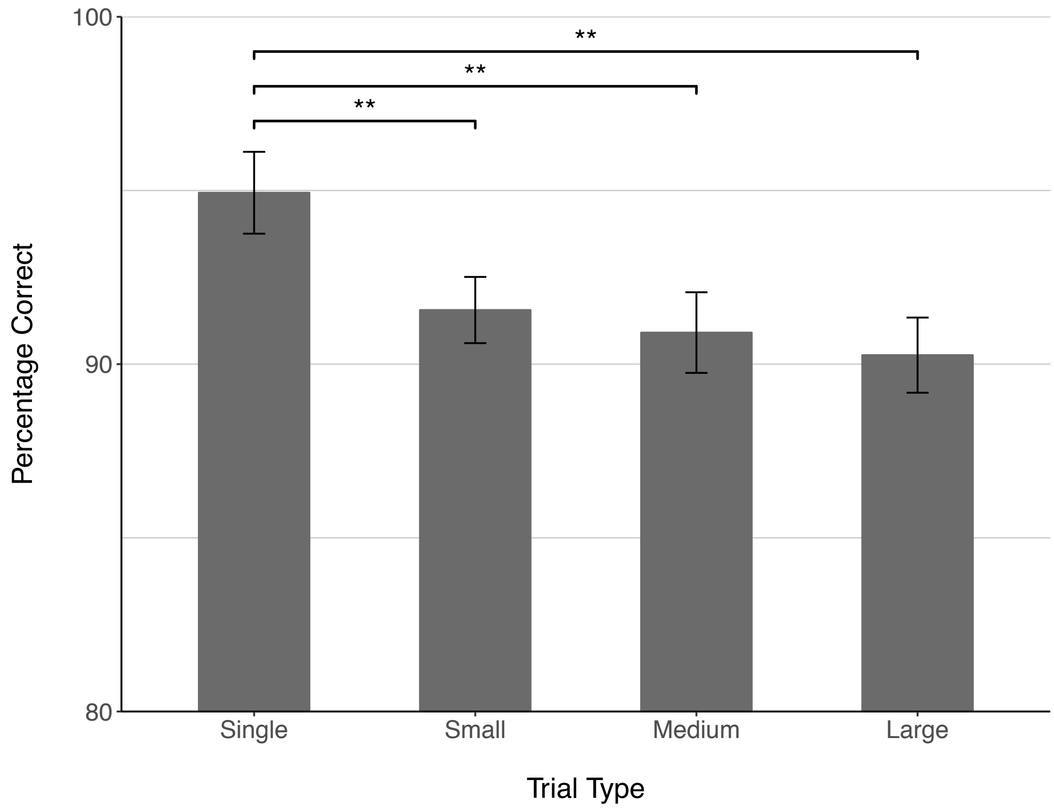

Figure 2.5. *Experiment 1*: *Percent Correct as a function of Set Size (Single, Small, Medium and Large).* Error bars represent standard error.
* $p < .05$, ** $p < .01$, *** $p < .001$

Because it is conceivable that the processes involved in tracking area in our single item trials may have been distinct from those involving more than one item (i.e., direct perception of area of a single item versus a potential computation process for tracking CA of a group of objects), we performed one additional analysis to explicitly compare performance as a function of set size for only those trials involving arrays of items. We calculated the slope relating performance to small, medium, and large set sizes (dummy coding small as 1, medium as 2, and large as 3), examining only performance on the neutral trials, with a Number Ratio 1. We found a negative slope of -.006 that did not differ significantly from 0, $t(77)=1.38$, $p=.17$, $d = -.15$[6]. This suggests that although performance decreased as set size increased, this trend was not significant.

Altogether, results from Experiment 1 suggest that CA discriminations in adults are ratio-dependent, hindered and possibly facilitated by numerical congruency, but less impacted by set size. Furthermore, we found that although adults are significantly better at discriminating the size of a single item compared to discriminating CA across multiple items, the size of the actual set (e.g. whether there were 2-3 items or 7-8 items) did not have an effect on CA performance, consistent with the Direct Perception Hypothesis.

Experiment 2

The goal of Experiment 2 was the replicate findings of Experiment 1 with some small changes. Firstly, in Experiment 1, we controlled for *item* density by holding constant the number of items in the display per unit background area. Although this controlled for density of the items within the display, it did not allow us to rule out that participants relied upon the relative amount of white background within the display. In Experiment 2, we controlled for *area* density by holding constant the relevant surface area (to be tracked) per unit background to ensure that density did not drive our pattern of results. Moreover, in Experiment 2, we eliminated the 1.9 CA Ratio since adults performed at ceiling on this ratio, thus Experiment 2 tested adults on only 4 CA Ratios: 1.15, 1.3, 1.45 and 1.6.

[6] Even when we excluded the easiest 1.9 ratio where participants were performing at ceiling, we found a negative slope of -.005 that did not differ significantly from 0, $t(76) = .87$, $p = .39$, $d = -.10$.

Methods

The methods of Experiment 2 were identical to Experiment 1 except for the following:

Participants

A total of 54 undergraduate students from Boston College participated in our study in exchange for cash or course credit (41 female, *M* = 19.65 years, *Range* = 18-26 years). Since the effect sizes we obtained in Experiment 1 were larger than expected, we reduced our sample size in this Experiment. All participants provided informed consent.

Procedure

Experiment 2 included the following variables and their levels: CA Ratio (1.15, 1.33, 1.45, 1.6), Number Ratio (1, 1.33 or 1.5), Congruency (Congruent, Incongruent), and Set Size (Small, Medium, Large). Similar to Experiment 1, we continued to have 38 trials per CA Ratio (this includes 2 single item trials per CA Ratio), leading to a total of 152 unique trials. To increase the precision in our measurement, we presented participants with the 152 unique trials 3 times over the course of the experiment (yielding 456 trials total). The trials were organized in blocks such that a participant was presented with all 152 unique trials before the trials would be repeated, with an unlimited break every 100 trials (5 breaks total). Otherwise, procedures were identical to Experiment 1.

Stimuli

The only changes made to the stimuli was that we now controlled for density by dividing the CA by the size of the display. The density was identical across the two arrays to be compared in each trial, although across trials densities did vary from .08-.15

(CA/background area in cm^2). Thus, the area of our background ranged from 193.3 – 300 cm^2.

Results & Discussion

Consistent with the fact that we dropped the easiest CA ratio in this experiment, performance was significantly less accurate here compared to Experiment 1 ($t(116) = 2.19$, $p = .03$, $d = .40$), with an average of 80.19% correct (Range = 50.00-96.05%). Since we suspected that this was due to the exclusion of the easiest 1.9 CA Ratio, we performed a second independent samples t-test this time comparing performance on Experiment 1 and 2 excluding the easiest 1.9 CA Ratio in Experiment 1 as well (M = 83.47%, Range = 58.55-98.02%) and the difference in performance was no longer statistically significant ($t(116) = 1.03$, $p = .31$, $d = .19$), suggesting that lower performance in this Experiment was due to the fact that we eliminated our easiest CA Ratio.

The average weber fraction across participants was $w = .32$, (*Range* = .09-2.1, SD =.36, SE=.05), this was marginally worse than in Experiment 1, ($t(127) = 1.90$, $p = .06$, $d = .32$). We again using the Welch-Satterthwaite procedure for unequal variances to compare our weber fraction to that which has been previously reported by Odic et al. (2013) and found a significant difference in performance, $t(54.96) = 3.43$, $p < .01$, suggesting that CA acuity in our study was significantly worse than prior reports of numerical acuity.

Does numerical congruency matter for CA judgements?

As in Experiment 1, we examined accuracy on trials involving arrays of multiple items in which number was either congruent or incongruent. A 4 (CA Ratio: 1.15, 1.3, 1.45, 1.6) x 2 (Number Ratio: 1.33 or 1.5) x 2 (Congruency: Congruent vs. Incongruent) repeated measures ANOVA again revealed a main effect of CA ratio, $F(3, 159) = 167.88$, $p < .001$, $\eta_p^2 = .76$. Performance on all 4 CA ratios was significantly different from one another (p's<.001), with performance improving as the CA Ratio increased, replicating our previous finding that cumulative area discriminations adhere to Weber's Law. In contradiction of predictions of the SoM Theory, we found a significant main effect of Congruency, ($F(1,53) = 60.62$, $p<.001$, $\eta_p^2 = .53$) such that participants performed significantly better on congruent ($M = 91.70\%$) compared to incongruent ($M = 62.90\%$) trials. We also again found a CA ratio x Congruency interaction, $F(3,159) = 24.89$, $p < .001$, $\eta_p^2 = .32$. Paired samples t-tests revealed that regardless of CA ratio, participants performed significantly better on the congruent compared to the incongruent trials;

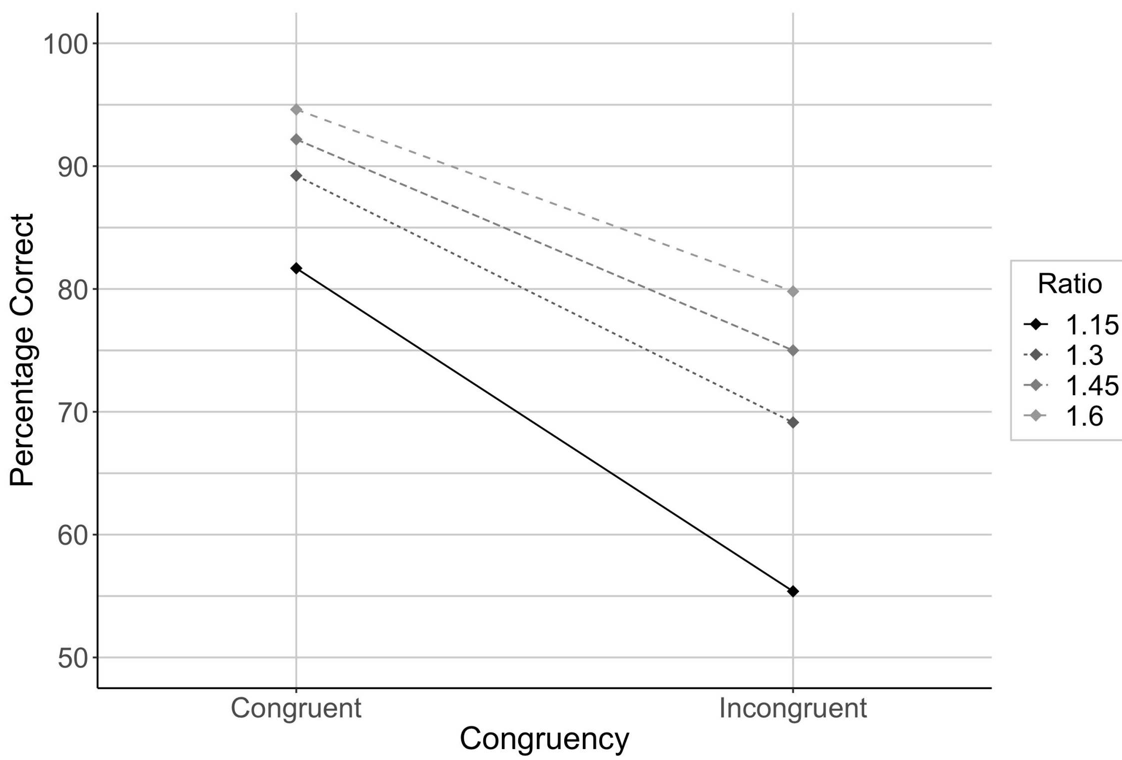

Figure 2.6. *Experiment 2 Percent Correct as a function of CA Ratio (1.15, 1.3, 1.45, 1.6 and 1.9) and Congruency (Congruent and Incongruent).*

however, again the impact of congruency lessened as the CA ratio increased (i.e., became easier; *p*'s < .001; See Figure 2.6).

Lastly, we replicated our main effect of Number Ratio ($F(1,53)=14.19$, $p<.001$, $\eta_p^2 = .21$) with participants performing better on the 1.33 Ratio ($M = 78.49\%$) compared to the 1.5 Ratio ($M = 76.12\%$; See Figure 2.7). However, unlike Experiment 1, we did not find a Number Ratio x Congruency interaction, $F(1,53) = 2.33$, $p = .13$, $\eta_p^2 = .04$ (although the pattern of results was identical across experiments).

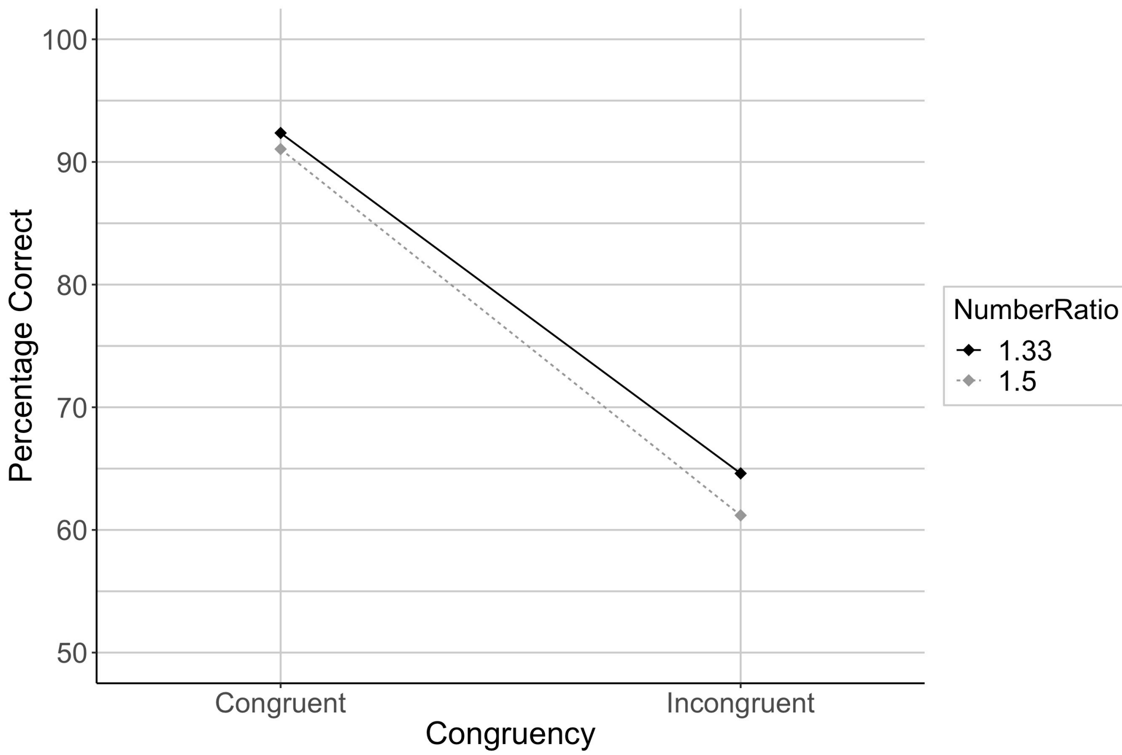

Figure 2.7. *Experiment 2: Percent Correct as a function of the Number Ratio (1.33 and 1.5) and Congruency (Congruent and Incongruent).*

Do numerical differences across the displays impact our CA tracking abilities?

As in Experiment 1, we examined to what extent congruency or incongruency between number and CA was helping or hurting performance in this task. Therefore, we ran a repeated measures ANOVA comparing performance on the number-neutral trials (trials of the Number Ratio 1) with performance of congruent trials and incongruent trials (again, collapsing all analyses across the 1.33 and 1.5 Number Ratio and across all CA ratios). Once again the ANOVA was significant, $F(2, 106) = 61.29$, $p < .001$, $\eta_p^2 = .54$. Paired samples t-tests revealed that performance on congruent trials (91.72%) was significantly better than the neutral trials (84.28%, $t(53) = 4.82$, $p < .001$, $d = 1.49$), performance on the incongruent trials (62.90%) was significantly worse than the neutral trials (84.28%, $t(53) = 8.85$, $p<.001$, $d = 1.07$). Once again, this suggests that competing numerical information interferes with CA judgements, and consistent numerical information also facilitates CA judgements (See Figure 2.8).

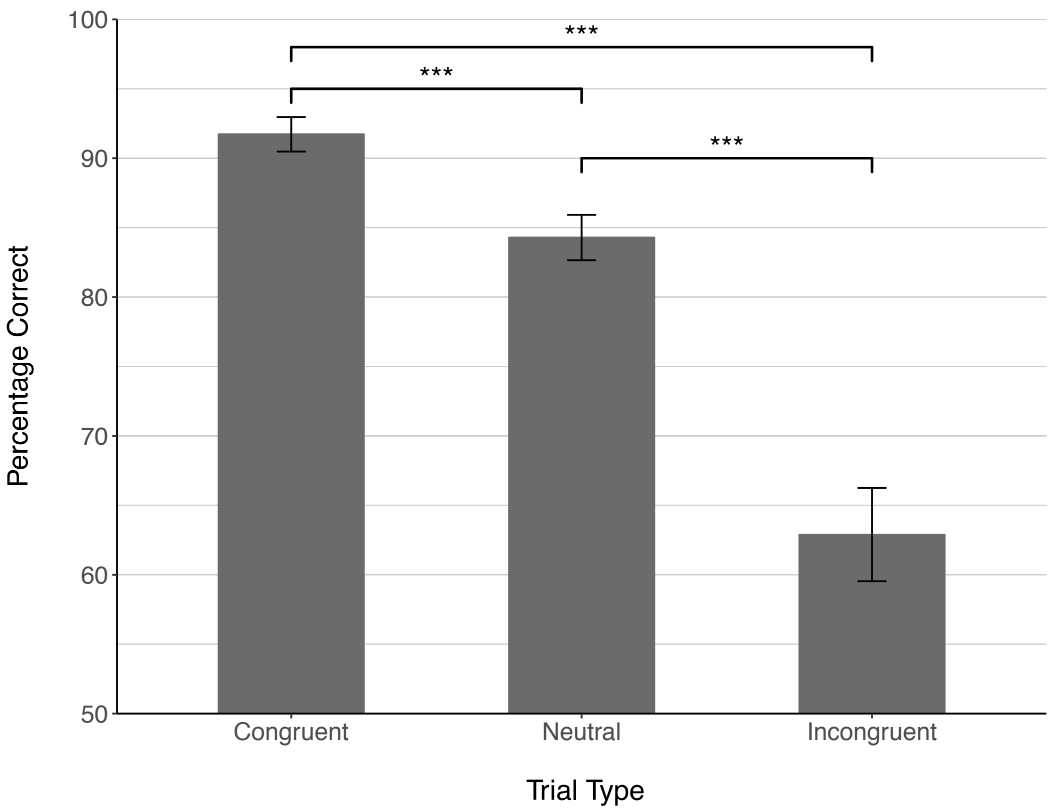

Figure 2.8. *Experiment 1: Percent Correct as a function of Congruency (Congruent, Neutral and Congruent).* Error bars represent standard error.
† $p < .10$, * $p< .05$, ** $p< .01$, *** $p<.001$

How does set size impact CA acuity?

To examine the effect of set size on CA discriminations, we again looked at those trials where number was held constant within trials, but differed across trials. That is, we compared performance on Single trials to Small, Medium and Large trials involving the 1 Number Ratio. A within-subjects ANOVA again revealed a significant effect of set size ($F(3,159) = 20.55$, $p < .001$, $\eta_p^2 = .28$), and follow-up paired samples t-tests revealed this was once again driven by the Single trials ($M = 90.35\%$) where participants performed significantly better than Small ($M = 86.84\%$), Medium ($M = 82.94\%$) or Large trials ($M = 83.06\%$; p's < .001). In contrast to Experiment 1, however, we found participants performed significantly better on Small compared to Medium and Large trials (p's < .001; See Figure 2.9)

Next, to specifically address the impact of set size on performance on trials involving sets, we calculated the slope for performance on small, medium, and large set sizes. Here we examined only performance on the neutral trials, with a Number Ratio 1, and found a negative slope (slope = -.019) that differed significantly from 0 ($t(53) = 3.54$, $p < .001$, $d = -.48$). This suggests that there was a steady decrease in performance as set size increased.

Results from Experiment 2 replicate our findings in Experiment 1. Not only did we replicate previous findings revealing that CA discriminations abide by Weber's law, congruency between number and CA plays an important role in discrimination performance. Once again, our results suggest that number is more likely to interfere with CA judgements than to assist them. Moreover, Set Size mattered in that adults were significantly better at discriminating the size of single items compared to the CA of multiple items, and the size of the actual set (e.g. whether there were 2-3 items or 7-8 items) did have some (although not a robust) effect on CA performance.

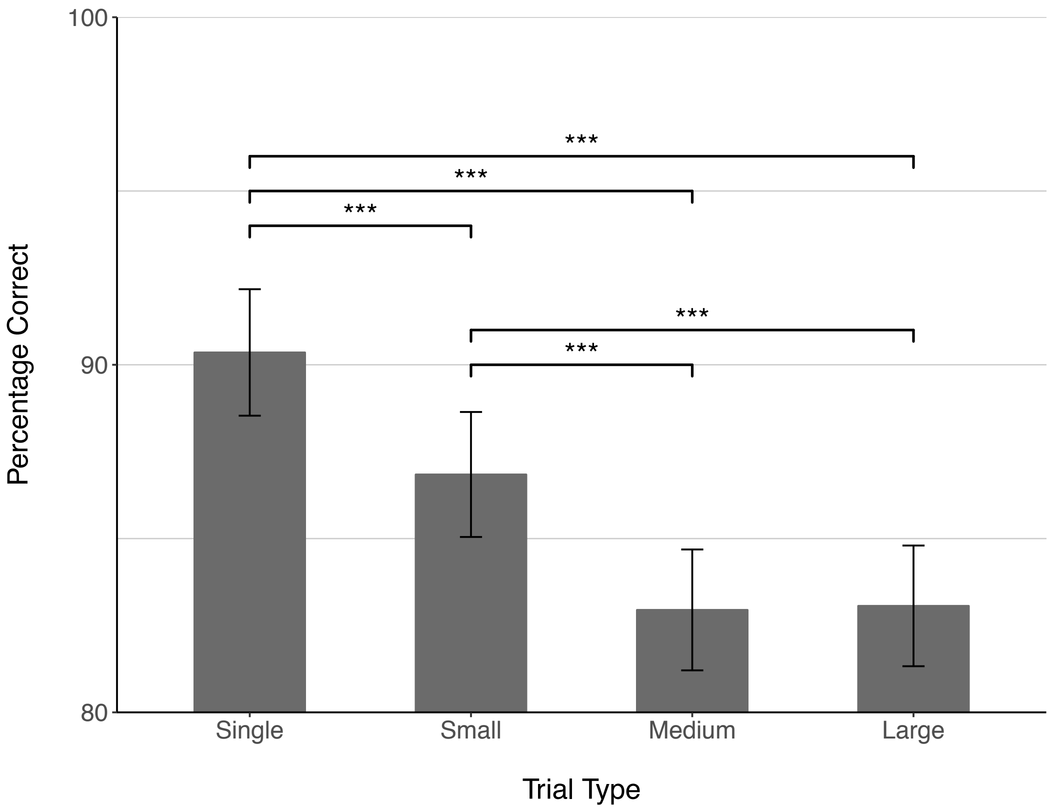

Figure 2.9. *Experiment 2*: *Percent Correct as a function of Set Size (Single, Small, Medium and Large).* Error bars represent standard error.
* *p*< .05*,* ** *p*< .01, *** *p*<.001

Combined Analysis

We combined data from Experiment 1 and 2 and compared the average weber fraction across participants across Experiments, which was w = .27, (*Range* = .07-2.1), to weber fractions previous reported by Odic et al. (2013) for adult numerical discrimination tasks (using the Welch-Satterthwaite procedure for unequal variances) and once again found a significant difference in performance, $t(23.17) = 4.09$, $p < .001$, with participants performing significantly worse on the CA task.

Given that our results regarding the effect of set size on performance were inconclusive (we found a significant effect of set size in Experiment 2 but not 1), we combined our data from Experiments 1 and 2 to run another slope analysis (excluding data from the 1.9 ratio in Experiment 1), again looking only at the neutral trials with Number Ratio 1. We found a negative slope of -.011, which was significantly different from 0, $t(130) = 2.59$, $p = .011$, $d = -.21$ suggesting that overall, participants did show increasingly worse performance as set size increased.

General Discussion

The aim of this study was to investigate if and how CA discriminations are influenced by numerical information in adults. Although many previous studies have investigated adult abilities to discriminate discrete quantity (i.e. number discrimination; Halberda & Feigenson, 2008c; Odic et al., 2013), very few studies have examined adults' performance on discrimination tasks that involve continuous properties, such as CA. This is a particularly interesting question given the claims made by proponents of the SoM that continuous quantities should be more easily represented than number because unlike number which is abstract, continuous quantities are perceptual in nature and are not tied to any specific sensory modality (Gebuis & Reynvoet, 2012b; Leibovich et al., 2017). As a direct test of these claims, many studies have investigated whether we can track number independent of continuous properties, finding that even human infants can do so (Halberda & Feigenson, 2008b; Lipton & Spelke, 2003; Odic et al., 2013; Starr et al., 2013a; Xu & Spelke, 2000; Xu et al., 2005). However, less work has explored the converse; that is, how well can we track continuous properties independent of number?

Our first aim was to examine whether congruency between number and area played a role in adults' discrimination performance. If CA is significantly more salient and easy to represent than number, then CA acuity should be more precise than that of number. This did not appear to be the case. In addition to replicating previous findings suggesting that CA discriminations were ratio-dependent (Halberda & Feigenson, 2008a; Odic et al., 2013), we also found that the average weber fraction associated with CA discriminations was not significantly lower than those previously reported for number discrimination (Odic et al., 2013), contradicting any claims that adults are *better* at discriminating continuous quantities compared to number (Leibovich & Henik, 2014). Moreover, if CA is more salient and easy to represent than number, then numerical information should be less likely to interfere with CA representations than vice versa. Again, this did not appear to be the case. Across two experiments, we found that participants overwhelmingly performed better on congruent trials (e.g. when the arrays with the larger number of dots also have a greater CA) compared to incongruent trials (e.g. when the larger number array has a smaller CA), replicating previous findings (Barth, 2008; DeWind & Brannon, 2012; Gebuis & Reynvoet, 2012a; Hurewitz, Gelman, & Schnitzer, 2006). Not surprisingly, numerical information had a greater impact on adult performance for the most difficult CA judgements. That is, we found that numerical congruency had a larger effect for harder CA discriminations and for trials where the ratio in number between the two arrays was larger.

Moreover, our study expanded upon previous research by exploring whether numerical congruency facilitated CA judgements, whether numerical incongruency hindered CA judgements, or both, specifically when compared to neutral trials where the number of items in both arrays is identical (i.e. trials where the ratio of number was 1). We found both to be the case: incongruency between Number and CA significantly hurt performance *and* congruency between these two variables boosted performance, suggesting that adults *can* and *will* use all available quantity information in making quantitative judgements, whether or not this information is helpful or hurtful.

The final, and most important goal of this study was to understand the process by which adults represent CA when presented with an array of items. In particular, we compared two possible hypotheses. On the one hand, the 'Direct Perception' hypothesis assumed that we extract how much surface area we see in the display without any reliance upon individuating the items in the array, which is consistent with SoM theory (Gebuis & Reynvoet, 2012b; SoM: Leibovich et al., 2017; Mix, Huttenlocher, & Levine, 2002). On the other hand, the 'Computation' hypothesis proposed that adults track the sizes of each individual item within the array and perform a summation process to arrive at an estimate of the CA of the array (as proposed by Barth, 2008). To distinguish between these two accounts, we investigated how set size affected performance on this task. Assuming that this summation process contributes error to the representation (Cordes et al., 2007), the 'Computation' hypothesis would predict a decrease in performance as the number of items in the array increased since each additional item adds to the error in the summation or computation process. The 'Direct Perception' hypothesis would not predict a relationship between CA acuity and set size. Indeed, we found that adults performed significantly more accurately on trials where they were presented with single items (trials that required them to make element area comparisons) compared to trials with small (2-4 dots), medium (6-9 dots), or large (9-15) sets. Furthermore, although analyses were mixed across experiments, slope analyses combining data from Experiments 1 and 2 demonstrated that as set size increased, accuracy in making CA judgements decreased. This provides support for the Computation Hypothesis, suggesting that when making CA judgements, adults represent the surface area of individual items within an array and sum across these representations to gain a representation of the array's CA. These findings are in line with previous findings by Barth (2008), whose computational model suggested that a computational account of CA representation provided the best explanation for the data.

Furthermore, our finding that adults find it easier to discriminate single items compared to sets of items suggests that distinct processes may be at play when tracking the area of an individual item versus tracking the area of an array. In the former case, other continuous quantities – such as the diameter of the item – may serve as a reasonable cue for discrimination and thus area may not even be tracked under these circumstances. In the case of an array of items, though, it is unlikely that successful discrimination can take place without tracking the area of the items within the array. Moreover, the fact that differences in acuity persist in judging the area of a single item versus an array of items emphasizes that if we do want to make comparisons between adult abilities to represent discrete and continuous properties, or make claims about the saliency of continuous variables in the context of numerical stimuli, it is important that we test both in the same context – that is, within arrays of objects. Given that tests of numerical discrimination by definition require adults to make estimates or computations across multiple items, one should similarly examine the representation of continuous dimensions in the context of sets. In fact, evidence with infants supports this: Brannon et al. (2006) found that while 7 month olds infants were to discriminate a 1:2 ratio in EA of a single item, when presented with arrays of items this same age group needed as much as a 1:4 change to detect a difference in the EA (Cordes & Brannon, 2011). Although infants were asked to do the same exact type of task, the mere presence of multiple items detracted from infant's abilities to represent EA.

In conclusion, our results are in stark contrast with claims of a SoM theory. In particular, our results suggest that CA discriminations – in the context of multiple items – are not more precise than that of numerical discriminations, and may in fact be even marginally *less* precise. Moreover, though it was completely irrelevant to the task demands and thus not a reliable cue for tracking, we find that number is automatically processed in the context of CA judgements suggesting that number is at least as salient as this continuous quantity. Importantly, by examining the effect of set size on these discriminations, we found support for the Computation hypothesis, suggesting that adults represent CA by summing across individual items with an array. Future work will be needed to determine whether the computation mechanisms that adults are engaging in when making these discriminations are present early on in development or whether they are learned over time. That is, it would be interesting to investigate whether infants and children show similar set size signatures in their CA judgements as adults.

CHAPTER 3: RELATIVE SALIENCE OF NUMBER: PRESCHOOLER'S NUMBER KNOWLEDGE RELATES TO SPONTANEOUS FOCUSING ON NUMBER FOR SMALL, BUT NOT LARGE, SETS

Abstract

Much research has examined the reciprocal relationship between a child's spontaneous focus on number (SFON) in the preschool years and later mathematical achievement. However, many different tasks have been used in the literature to assess SFON that have distinct task demands (verbal vs. behavioral) and stimuli (different types and quantities of other features available to the child to focus on outside of number) and it is unclear to what extent these tasks are measuring the same underlying construct. Moreover, prior studies have investigated SFON in the context of small sets exclusively, but no work has explored whether individual differences exist in SFON for large sets and whether these differences relate to math ability. In the current study, preschoolers were presented four distinct SFON tasks assessing their spontaneous attention to number for small (Experiment 1) and large (Experiment 2) sets of number. Results revealed no relation in SFON performance across the four distinct SFON tasks. Moreover, in contrast to predictions of a single construct of SFON, preschooler's SFON for small sets (1-4 items) was significantly stronger than that for large sets (10-40 items) and analyses revealed that number knowledge was only associated with SFON for small sets, but not large. Together, findings suggest that SFON may not be a set-size independent construct, and instead may hinge upon a child's number knowledge, at least in the preschool years. The role of number language and how it relates to children's SFON are discussed.

Relative Salience of Number: Preschooler's Number Knowledge Relates to Spontaneous Focusing on Number for Small, but not Large, Sets

There are documented individual differences in children's tendencies to pay attention to number in their natural environment, with some children naturally attending to number in their day-to-day lives more than others (Hannula-Sormunen & Lehtinen, 2005; Hannula et al., 2007). A child's propensity to focus on number without any prompting has been called Spontaneous Focusing on Number (SFON; Hannula & Lehtinen, 2005). Much research has examined the reciprocal relationship between SFON and math skills; finding not only that individual differences in SFON in the preschool years predict later long-term measures of math achievement (Hannula-Sormunen, Lehtinen, & Räsänen, 2015; Hannula-Sormunen, Lepola, & Lehtinen, 2010; McMullen, Hannula-Sormunen, & Lehtinen, 2015), but also that early number skills acquired in the preschool years (such as counting abilities) relate to SFON a few years later (Hannula-Sormunen & Lehtinen, 2005).

However, the limits of SFON are still unclear. That is, is SFON a general attribute pertaining to attention to any numerical information in the environment, or might it be set size- or task- specific? In the current study, we investigated two different research questions: (1) How does children's SFON differ as a function of the task used to assess it? Multiple tasks have been used to assess SFON, some that are verbally-based and others that are purely behavioral (i.e. not reliant upon language), yet little is known about how SFON may differ as a function of these tasks. There is some evidence to suggest that a child's SFON may vary substantially across distinct SFON tasks (Batchelor, Inglis, & Gilmore, 2015), however this study did not explore all SFON tasks in the literature, choosing instead to focus on the two most prominent SFON tasks. A systematic comparison in performance on different SFON measures can therefore help elucidate how task demands may affect a child's likelihood of focusing on number spontaneously. We also asked: (2) How does set size affect preschooler's SFON? To date, investigations of SFON in preschoolers have mostly been limited to small (<5 items) sets of items which young children are more likely to be able to enumerate and count compared to larger sets. Do children spontaneously attend to number even when they may be unable to accurately track the exact number of items present? Or might evidence of SFON be dependent upon their enumeration abilities? Getting an answer to these questions is important in furthering our understanding of the construct of SFON and the nature of its relationship with other numerical abilities.

Spontaneous Focus on Number

In the literature, an important distinction is made between knowing specific numerical skills (e.g., how to count), and knowing that these skills are relevant to the task at hand (e.g., that counting might be a relevant strategy for a particular task; Hannula-Sormunen & Lehtinen, 2005). Although the majority of research in this domain tends to focus on the former (what mathematical knowledge do children possess and how can we teach it to them?), recent studies have also begun to focus on the latter (when do children realize that number might be an important dimension to focus on and/or relevant to the task?). SFON has been described as an attentional process that precedes (and is distinct from) enumeration (i.e., counting, identifying cardinality) with the idea being that SFON may predispose children to realize that they *should* individuate and enumerate objects (Hannula-Sormunen & Lehtinen, 2005). These ideas are supported by research revealing that SFON around ages 3-4 predicts children's enumeration and counting skills at ages 5-6 (Hannula-Sormunen & Lehtinen, 2005; Hannula et al., 2007) and rational number knowledge at age 12 (McMullen et al., 2015). Moreover, SFON at age 6 is positively correlated with arithmetic skills but not reading skills two years later, suggesting that SFON is truly domain-specific (Hannula-Sormunen et al., 2010) and not simply a proxy for domain-general capacities such as IQ or working memory.

One limitation of this research, however, is that SFON has only been tested with small sets (<5 items). In theory, SFON should pertain to attention to all numerical information, and should not be dependent upon an ability to enumerate the number of items, but instead reflect the recognition that number is a relevant dimension to attend. Since tests of SFON in preschoolers have only involved sets that they are also able to enumerate (i.e., count and identify the cardinality), it is unclear whether the ability to enumerate is a component of, or a necessary precursor to, demonstrating SFON for this age group. If SFON is truly a generalized attention to number, then it should not depend on the size of the sets involved nor on the child's ability to enumerate the sets. That is, children should demonstrate similar levels of SFON for large sets as they do for small sets.

The distinction between testing children on small and large sets is particularly important given what we know about children's developing number knowledge. Even before children master the count procedure – that is, before they acquire a full understanding of cardinality and the meaning of *all* number words – they can have an understanding of the meaning of *some* of the number words (Le Corre & Carey, 2007). The process of learning to count is a lengthy process that progresses through a series of stages over a period of 1-2 years. Research suggests that between the ages of 2.5 – 4, children start off as “subset knowers,” meaning that they have an understanding of the meaning of a subset of numbers but have yet to grasp the cardinal principle more generally (i.e., that the last number word used in a count list is the cardinality of the set). Subset knowers go through a step-wise process where they learn the meaning of the number word “one”, then “two” and so on. Not until after children learn the meaning of “four” do children acquire the cardinal principle – that the last number word used in a count refers to the cardinality of the set – and begin to understand the purpose of counting (thus becoming “cardinal-principle knowers” or CP-knowers; Wynn, 1992). At this point, it is expected that they now understand the meaning of all number words within their count list. This distinction between subset knowers and cardinal principle knowers is an important one because it represents qualitative differences in children’s behavior. For example, cardinal principle knowers more consistently recognize that counting is an effective strategy, whereas subset-knowers are less likely to spontaneously count in the face of a numerical task (Gordon, Chernyak, & Cordes, submitted; Le Corre & Carey, 2007; Posid & Cordes, 2018; Wynn, 1992).

Given that subset-knowers are able to accurately identify the cardinality of a small set, but not a large one, it is important to determine how this cardinal understanding may impact the likelihood of the child demonstrating SFON. That is, do subset-knowers demonstrate SFON at similar levels for small and large sets? Do we see differences between subset-knowers and CP-knowers? If, as Hannula-Sormunen and collegues (2010) claim, SFON is truly a domain-specific attentional phenomenon that is distinct from children’s actual numerical abilities, one would not only expect similar levels of

SFON for small and large sets, but (similar to previous findings) one would expect SFON for both small and large sets to correlate with number knowledge. To our knowledge only one study has tested SFON with both small and large sets. Cantlon et al. (2010) tested children on a match-to-sample task where children could either match stimuli using number or cumulative surface area, presenting them with sets ranging from 1-12, and they found that regardless of set size, children chose the number match over the cumulative surface area match at above chance levels. Although they found that children performed better on any comparison that involved set size one, because their trials had small and large sets intermixed, they were not able to address whether the degree of SFON differed as a function of set size. Furthermore, since they did not test children's number knowledge, they were not able to address whether their SFON differed as function of the child's number knowledge (i.e., subset- versus CP-knower).

Current Measures of SFON

A secondary research question concerned reconciling different measures of SFON that have been used in prior literature. Three types of tasks have traditionally been used to measure SFON: A) Imitation tasks, B) Picture tasks, and C) Choice tasks. Although all three have been used across studies, it is unknown whether these tasks assess the same underlying SFON construct.

In Imitation tasks, children as young as 3 years are shown a series of repetitive actions by an experimenter and are asked to imitate the experimenter (Hannula-Sormunen & Lehtinen, 2005; Hannula-Sormunen et al., 2015, 2007). For example, Hannula and Lehtinen (2005) showed participants a mailbox and letters of two different colors. The experimenter would put a certain number of letters of each color in the mailbox and the child was asked to "do exactly like I just did" without any mention of number. The dependent variable in this task is whether or not the child imitates the number of actions the experimenter undertook, or alternatively whether they use number words while doing their imitation. Thus, this task relies very little upon a child's linguistic capacities and instead requires a purely nonverbal behavioral response.

Picture tasks are newer in the literature (Batchelor et al., 2015) and have typically been used with slightly older children (4-6 year olds) compared to the Imitation task. In the Picture task, children are presented with a picture containing a complex scene including sets of items and are asked to describe what they see in the picture. As in other SFON tasks, the child is not told about the nature of the task or what the experimenter is looking for in their description. In this case, the dependent variable is whether or not the child uses any number or quantity words in their descriptions. Importantly, both of these tasks hinge upon the ability to track exact number either verbally or behaviorally. Thus, any measure of SFON obtained from these measures necessarily correlates with an ability to encode exact number, something thought to be dependent upon number word learning.

Choice tasks have more recently been designed as another way to measure SFON (Cantlon, Safford, & Brannon, 2010; Chan & Mazzocco, 2017). The Choice task involves an ambiguous match-to-sample game where children are asked to select the picture that "best matches" a sample picture (typically involving an array of items). On critical trials, one of the choice pictures matches the sample picture on the dimension of number, while the other matches the sample on another quantitative dimension such as Cumulative Area (CA; Cantlon, Safford, & Brannon, 2010), color, or shape (Chan & Mazzocco, 2017). By directly pitting number against other dimensions, these critical trials provide a measure of children's SFON, by measuring their relative preference for number over this other dimension. Cantlon et al. (2010) found that 3-5 year old children spontaneously focused on number over CA at a rate higher than chance alone, a finding that has been replicated with English and Japanese populations (see also Cantrell, Kuwabara, & Smith, 2015). Chan and Mazzocco (2017) found that the degree to which children matched based upon number depended on which other dimension number was pitted against. For example, 4-5 year olds were much more likely to pick the number match when the other available choices involved relatively low-salience features (e.g., pattern and orientation) compared to highly salient features (e.g., color and shape).

There are some important differences between these three SFON tasks that call into question whether they measure the same underlying construct. Although the Picture task relies upon a verbal response, both the Imitation and Choice tasks are behavioral measures of SFON. The verbal requirements of the Picture task may prevent some children with limited communication abilities from being able to demonstrate an attention to number and even furthermore it may require a level of comfort with number words that children in the preschool years (when this task has been used) simply do not have. This may explain why previous research has found that although performance on both the Picture and Imitation tasks predict arithmetic skills years later (Picture Task: Batchelor et al., 2015; Imitation Task: Hannula-Sormunen et al., 2010), performance on the two tasks do not correlate with one another (Batchelor et al., 2015; Rathé, Torbeyns, Hannula-Sormunen, & Verschaffel, 2016). This suggests that these tasks may tap into distinct aspects of SFON (verbal vs. behavioral) or alternatively, the verbal demands of the picture task may mask individual differences in underlying attention to number that children are unable to fully express verbally.

Furthermore, all three tasks vary in terms of the quantity and type of other features (outside of number) available to the child. While the Choice task presents children with a limited number of other dimensions to attend to over number (in that it pits one or two dimensions against number while controlling for other features), the Picture and Imitation tasks place few, if any, limits on the dimensions that a child may deem relevant for responding. For example, in the Imitation task, children could attend to the color or orientation of the cards placed in the mailbox, the facial expressions of the experimenter, or many other possible features. In the Picture task, anything in the picture is fair game. The Choice task, on the other hand, was specifically designed to measure the relative salience of number when pitted against only one or two other dimensions (e.g., number pitted against cumulative area or ratio), and as such it is allows for a more systematic interpretation of SFON. Therefore, in addition to systematically measuring SFON, it allows researchers to measure just how salient number is when compared to other quantities such as cumulative area (Cantlon et al., 2010).

Importantly, unlike the other two SFON tasks, the Choice task is the only task that allows for an assessment of SFON for large sets. Whereas the imitation task would be too cumbersome if children were expected to imitate as many as 5-10 actions, and the picture task requires a specific enumeration of a large set of objects which may be taxing for the child, the choice task allows children the opportunity to match based upon an approximate estimate of the number of items within the set.

Thus, across two experiments, we administered three different SFON tasks to 2.5-5 year old children: the Picture Task (Batchelor et al., 2015), the Mailbox Imitation task (Hannula-Sormunen & Lehtinen, 2005) and an Area Choice Task in which number was pitted against Cumulative Area. Following these SFON tasks, children were administered the Give-N task - a standard measure of children's number knowledge and cardinal understanding (Wynn, 1992). To explore SFON for small and large sets, we manipulated the size of the sets presented in the Area Choice task across experiments, such that children were presented with exclusively small sets (<4 items) in Experiment 1, and exclusively large sets (10-40 items) in Experiment 2. Since the Picture and Imitation tasks do not lend themselves to testing SFON for large sets, we manipulated set size only on the Area Choice task.

Our first aim was to examine the relationship between our four SFON measures to see 1) whether they were correlated with one another and 2) which measure of SFON correlated most strongly with Number Knowledge. Our second aim was to compare preschoolers' SFON for small and large sets by examining whether there were similar levels of SFON when presented with small (Experiment 1) and large (Experiment 2) sets in the Area Choice task. Given that previous research has shown a strong relationship between SFON (when tested with small sets) in the preschool years and number knowledge, we assessed whether children's SFON for small and large sets similarly relate to their knowledge of number. If as has been suggested by Hannula-Sormunen and collegues (2010), SFON is a generalized attention to number (i.e., regardless of what number) then we should expect 1) similar levels of SFON for small (Experiment 1) and large sets (Experiment 2) and, 2) a child's actual knowledge of number (i.e. whether they are a Subset- or CP- knower) should be similarly related to performance on the small and large SFON tasks. On the other hand, if number knowledge does play a role in children's level of SFON, and these two are not as distinct as has been suggested in the literature, we may find 1) children to show greater SFON for small compared to large sets in the Area Choice task, and, 2) children's number knowledge should only relate to SFON in cases in which the child is able to enumerate the sets involved. That is, CP knowers, who are able to count and have knowledge of number words for small and large sets, may show SFON for both small and large sets, but Subset knowers may only demonstrate SFON for small sets.

Experiment 1

Methods

Participants

Participants were 118 preschoolers (Range 2.5-5.1 year olds; Mean age = 3.65, SD = .65, 73 Female). An additional 6 participants were excluded for experimenter error (n = 3) or for only completing a single task or less (n = 3). Of our 118 participants, 66 participants completed all four of our tasks with the rest of our participants completing at least 2 or more of our tasks (See Table 3.1 for a breakdown of participants included in each task). Given that a substantial number of participants were not able to complete all tasks, many of our analyses will look at a subset of participants.

Participants were recruited from the Greater Boston Area and either participated in lab or at their preschool or after school program. Of the 56% of our sample that provided demographic information, 86% of families identified as Caucasian, 5% as

Asian, and 9% as biracial. Furthermore, 98% of mothers and 83% of fathers responded as having completed a Bachelor's Degree or higher.

	Included	Excluded	Exclusion reasons	Average Age Included	Average Age Excluded
Picture Task	106	12	(1) N = 10 (2) N = 2	3.66 (2.5 - 5.1)	3.60 (2.5 - 4.8)
Area Choice Task	105	13	(1) N = 4 (2) N = 9	3.69 (2.5 – 5.1)	3.37 (2.5 - 4.3)
Mailbox Task	99	19	(1) N = 17 (2) N = 2	3.68 (2.5 - 5.1)	3.49 (2.5 - 4.5)
Give-N	97	21	(1) N = 12 (2) N = 9	3.66 (2.5 - 4.9)	3.60 (2.5 - 5.1)

Table 3.1. *Participant Exclusions Experiment 1.*
Exclusions were divided into two categories: (1) experimenter errors or errors with equipment (video cameras, computers etc.) or (2) a failure on the participant's part to finish the task (for the Area Choice task, participants needed to complete more than 2/3 of Probe trials to be included, for all other tasks participants needed to complete all trials to be included.

Design

Participants completed four different tasks in the following order: Picture Task, Area Choice Task, Mailbox Imitation Task, and the Give-N Task[7]. The Give-N task was presented last because it is a measure of number knowledge and we did not want to cue in the participants that we were assessing number until after all the SFON tasks were administered.

[7] Participants completed a fourth SFON task, the Proportion Choice Task, that was identical to the Area Choice task, except that the alternative (non-number) feature for participants to match on was the proportion of red to blue items in the display. This task was completed between the Mailbox Imitation Task and the Give-N Task. Since participants performed very poorly on this task – only 30 participants performed above chance on the standard trials – we will not include this task in the rest of this dissertation.

Tasks & Procedure

Picture Task. *Adapted from Batchelor, Inglis, & Gilmore (2015).* Children were presented with three cartoon pictures taken from children's books, presented on laminated cardboard (21 x 21 cm). The pictures were chosen because they were fairly simple pictures that clearly contained small sets of items, in a numerical range (1-4 items) that children could attend to and have a verbal label for. The pictures also contained many different colors, shapes, and animal characters, providing many other options to label and talk about, other than number.

Children were given three opportunities to talk about each of three pictures, in a set order. The experimenter introduced the task by saying: "This game is all about pictures. I am going to show you a picture and I want you to tell me everything you can see in the picture. Are you ready?" The researcher then put the first picture in front of the child and asked "What do you see in this picture?" When the child was finished talking, the experimenter prompted the child twice more saying: "Great! What else do you see?" After the final, third prompt, the researcher moved onto the next picture. Children were given 3 prompts to talk for each picture (unlike the original version, which provided only a single opportunity; Batchelor et al., 2015) because pilot testing revealed the additional prompting helped children overcome their initial shyness and reluctance to talk. If, after any of the prompts, the child said they did not see anything else in the picture, the experimenter moved onto the next picture.

Area Choice Task. *Adapted from Cantlon, Safford, & Brannon (2010).* The aim of this tasks was to measure which of two quantitative dimensions, number vs. cumulative area (CA), children would spontaneously use in a delayed match-to-sample task on a tablet. Children were first shown a single sample stimulus in the center of the tablet screen and were told: “I want you to look at this picture very carefully and when you are done remembering the picture, I want you to touch it.” If the child seemed reluctant to touch the tablet on the first few trials, the experimenter would prompt the child again or touch the tablet for them if they indicated they were done remembering the picture. Next, participants were shown two options and were asked: “Which picture best matches the one you just saw? This picture, or this one?” The experimenter would point to or circle each picture, to make it explicit what the two options were. The task was self-paced, meaning participants could choose when to move on from the sample stimulus and take as much time as needed to make their choice. Instructions were only repeated on the first few trials or whenever participants became distracted and needed re-prompting. At no point were children given any explicit instructions on how they should match the pictures.

This task included two types of trials. In Standard trials, one of the two choice stimuli matched the sample stimulus on *both* dimensions (i.e., number and CA; the correct match), while the other choice stimulus did not match on either of the two dimensions (incorrect match). In Standard trials, children were rewarded only for choosing the ‘correct match’, in this case a positive auditory and visual stimulus played on the tablet. Choosing the ‘incorrect match’ resulted in a red ‘x’ appearing on the screen with no auditory stimulus. In Probe trials, one of the choice stimuli matched the sample in terms of number, but not on CA (number match), while the other stimulus matched the sample in terms of CA, but not on number (area match). In Probe trials, participants were rewarded regardless of their choice. See Figure 3.1 for an example of Standard and Probe trials.

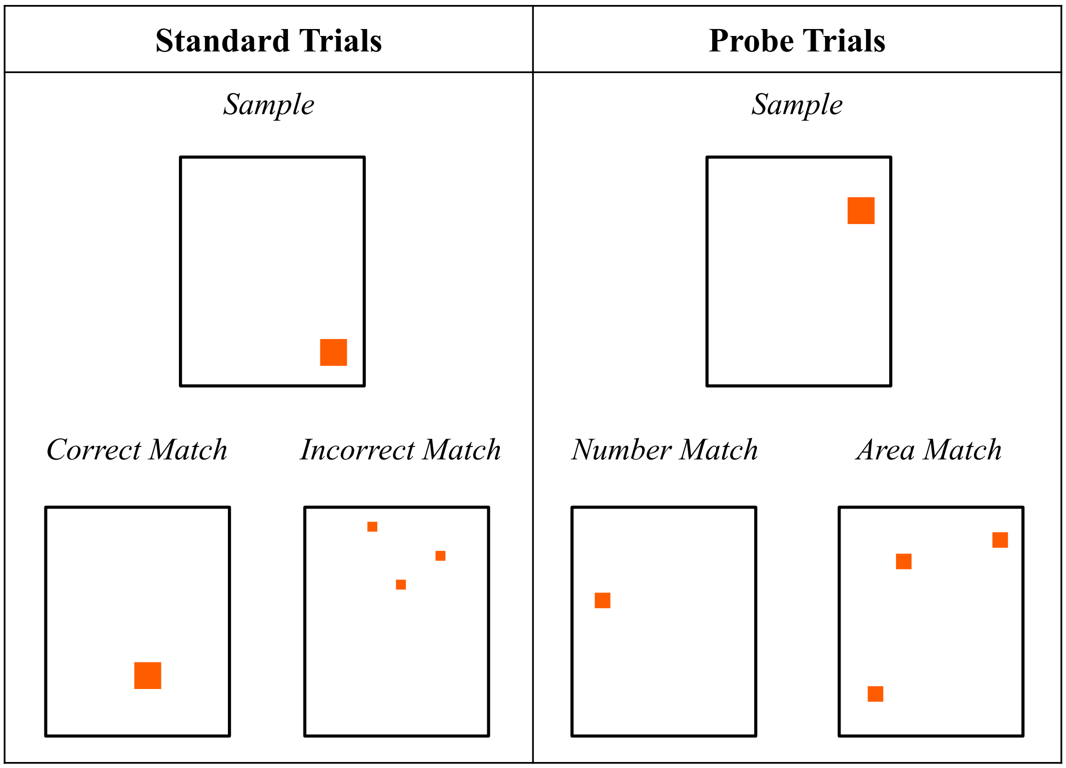

Figure 3.1. *Stimuli from the Area Choice task from Experiment 1.*
For all trials, participants first saw a sample picture, followed by two choice pictures. For Standard Trials (left pane), the Correct choice matched the sample on both Number *and* Area and an Incorrect choice matched the sample on *neither* Number *or* Area. For Probe Trials (right pane), the Number Choice matched the Sample on Number *not* Area, and the Area Choice matched the sample on Area *not* Number.

To familiarize participants with the task, they were first shown 6 Standard trials (referred to as 'Practice trials'), followed by 12 Test trials (a randomized mix of 6 new Standard trials and 6 Probe trials). To keep participants motivated, a short (16 secs) attractor video played halfway through the task. Thus, the entire task consisted of 18 trials (6 practice + 12 test), with an attractor video after the first nine trials.

Stimuli consisted of orange squares randomly placed on a white background (17 x 11 cm) with the element size of the squares homogenous within displays. For *Standard trials*, the number of items within the sample stimulus consisted of the numbers 1 through 4 and the CA of the sample items were chosen from one of three possible cumulative areas (4800, 8800 and 12800 pixels^2), ensuring that each CA was paired with equal frequency with each number. For the choice stimuli, the correct match had the same exact number of items and CA (and thus the same element area) as the sample stimulus, with only the placement of the squares in the display differing between the sample and the correct match. To determine the number and CA of the incorrect match we first of all ensured that we had all pairwise combinations of the numbers 1 through 4 between the sample and incorrect match stimulus (e.g., if the sample had 1 item, the incorrect match had 2, 3 or 4 items with equal frequency). This meant that the number ratio between the sample and incorrect match stimulus ranged from 0.25 to 0.75. Then to determine the CA of the incorrect matches, we took all element sizes generated from the sample stimuli, and randomly assigned those across the incorrect matches. In doing so, we made sure that 1) the size of items for the incorrect match were not identical to those of the sample/correct match and 2) the CAs for the incorrect match stimuli were larger than the CAs for the sample/correct match stimuli for half of trials and smaller for the other half. We decided to choose our incorrect match CA in this way because this would allow for the CA ratios between the sample and incorrect choice stimuli to be similar in range (Range: 0.38 to 0.73) to the number ratio range of .25-.75.

For *Probe Trials,* the sample stimuli were identical in number and CA as the sample stimuli of the *Standard* trials, with only the placement of the squares in the display differing. The Area Choice stimulus was identical in CA to the sample stimulus, and the Number Choice stimulus was identical in number to the sample stimulus. Then, to determine the number of items in the Area Choice stimulus, and the CA of the Number Choice stimulus, we made sure that the ratio of difference between the number of items in the sample and area match stimuli was identical to the ratio of difference between the CA in the sample and number match stimuli. For example, if the number ratio between the sample (e.g., 1 item) and the area match stimuli (e.g., 2 items) was .5, then the CA ratio between the sample (e.g., 4800 pixels^2) and the number choice stimuli (e.g., 9600 pixels^2) was also .5. For half of the probe trials, the CA of the number choice was greater than the sample, while the other half was smaller than the sample (by either multiplying or dividing by the given ratio).

Mailbox (Imitation) Task. *Adapted from Hannula & Lehtinen (2005).* Materials included a small mailbox placed in front of the child and 15 yellow laminated letters spread out across the table. Given that we were working with a younger age group than previous research using this task, we did some pilot testing and chose to simplify the task in several ways. First, children were presented with envelopes all of a single color, rather than two different colored sets. We also chose to test children only on very small quantities, 1 on the first trial and 2 on both the second and third trial. Lastly, rather than placing the letters in a pile, we fanned out all the letters because pilot testing revealed that our younger age group had trouble picking up an individual letter when placed in a pile, and often accidentally picked up multiple letters without realizing.

The experimenter introduced this task by pointing to the mailbox and the letters, saying, “This is my mailbox and these are my letters. In this game, I want you to watch very carefully what I do, and then I want you to do exactly like I did.” The experimenter then proceeded by picking up one letter and putting it in the mailbox, then asking: “Now can you do exactly like I did and tell me when you are done?” If the child stopped putting letters in the mailbox, but did not tell the researcher that they were done, the experimenter would ask “Are you done doing exactly like I did?” waiting for the child’s confirmation before proceeding. For both the second and third trial, the experimenter put 2 letters in the mailbox, each with a separate motion, repeating the exact same instructions. To avoid children putting all letters in the mailbox on the first or second trial, leaving no letters for the next trial(s), the experimenter always interrupted the child after they had put three letters in (a clear indication they were not imitating number) and told the participant “Ok, now it’s my turn!” If a child attempted to put envelopes in during the experimenter’s demonstration, the experimenter would stop the child and tell them “Wait, it is still my turn, it will be your turn next.”

Give-N Task. *Adapted from Wynn (1992)*. Children were introduced to a pond (a small blue basket) and 20 small yellow rubber ducks and were told that the ducks like to go into the pond. The experimenter started by showing one duck jumping into the pond, and then after removing the duck, asked the child “Can you put *one* duck into the pond?” Once the child was done putting ducks into the pond, the experimenter verified “Is that *one* duck?” If the child said yes, the experimenter went onto the next trial; if the child said no, they were given an opportunity to fix what they had done until they were happy that there was *one* duck in the pond. If the child correctly put one duck into the pond, the experimenter asked for larger set sizes, with the number of ducks requested, N, increasing from 1-6. Using a titration method, each time the child successfully put N ducks into the basket, they were asked for N + 1 ducks, but if the child failed on N ducks, they were then asked for N – 1. To reduce the number of trials children had to perform, the experimenter skipped the set sizes of 2 and 5 ducks when going up the titration ladder. However, if the child failed to correctly place 3 or 6 ducks into the pond, then the experimenter asked for 2 or 5 ducks respectively. The task ended when the child: 1) succeeded in correctly placing N ducks into the pond twice and failed on N + 1 twice or 2) succeeded twice on the N = 6 trial.

Data Processing & Coding

Picture Task. This task was transcribed using Computerized Language Analysis (CLAN), which was made available through the Child Language Data Exchange System (CHILDES; MacWhinney, 2000). We used the CHAT (Codes for the Human Analysis of Transcripts) transcription format at the utterance level. For each of the three pictures (trials), we used CLAN software to perform a frequency count of any number or quantity-related words found in the transcripts that were said by the participants. The number and quantity words that we searched for included: the number words 1-10, many, more, less, little, lot, count, big, and small. As per Batchelor et al., (2015), on each trial children received a score of 1 if they used any number or quantity words (regardless of how many) and a score of 0 if they did not. Therefore, children could get a maximum score of 3 on this task. Twenty percent of participants were transcribed by a second coder, frequency analyses of quantity word use were done on these transcripts, and then these participants were also given a score 0-3. The level of consistency between both coders’ scores were calculated using linear weighted Kappa, which resulted in a Kappa score of .81.

Area Choice Task. Performance on the Practice trials was not analyzed since participants were still attempting to understand the rules of the task. Although performance on Standard trials was used to measure participants’ degree of understanding of the aim of the task, these trials were not a measure of SFON since participants could use number and/or area as a cue for matching. Thus, our only measure of SFON in this task was the proportion of number matches on Probe Trials with higher scores on the Probe Trials reflecting a greater tendency to match on number rather than on the other quantitative dimension (i.e., CA).

Mailbox Task. As per Hannula and Lehtinen (2005), children were considered to have spontaneously focused on number for any trial if participants met any of the following requirements: a) they put the same number of letters in the mailbox as the experimenter, b) their utterances included number words - regardless of whether they were the correct number words - (e.g., “I am putting in two at the same time”) or quantity more generally (e.g., “How many did you put in?”), and/or c) they used gestures/fingers to denote numbers. Thus, children are given credit for displaying the correct numerical behavior (requirement a) and/or a numerical/quantitative verbal/gestural response (requirements b-c). Scores for this task were binary such that children scored 1 on a trial if they demonstrated any or all of the above measures of SFON, and a score of 0 if they did not. Therefore, children could get a maximum score of 3 on this task. Twenty percent of participants were coded by a second coder and reliability between the two coders was calculated using linear weighted Kappa, which resulted in a Kappa score of .80.

Give-N. The child's Give-N score was the highest number of ducks they successfully put into the basket[8].

Results

Individual SFON Measures

Picture Task. Overall, very few quantity words were produced in this task. On average, participants used number words on .97 of the three trials (See Figure 3.2).

Mailbox Task. Participants focused on number more on this task, with 80.8% of participants imitating the number of actions or using number words on at least one trial

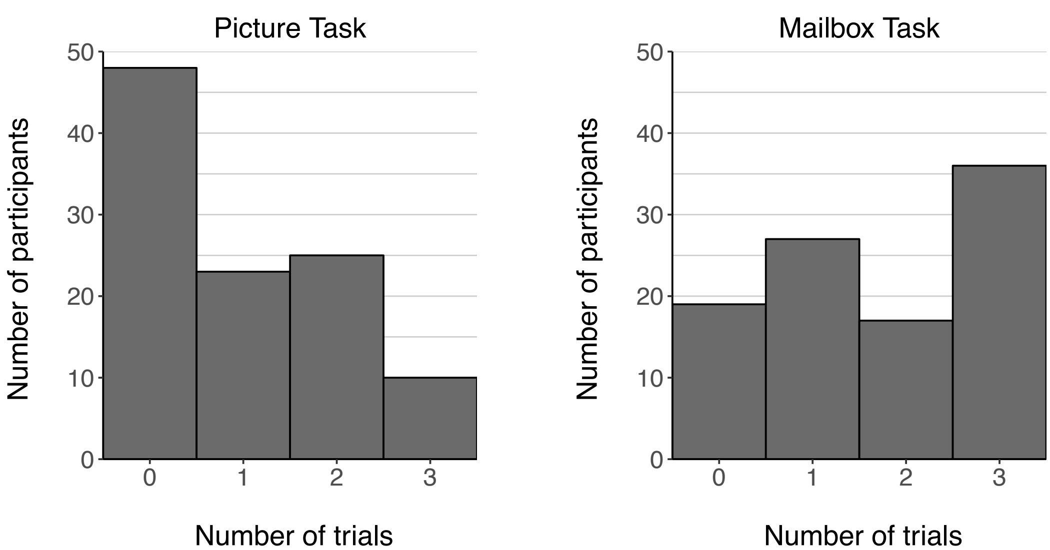

Figure 3.2 *Histograms depicting the number of trials on which participants used quantity words (Picture Task, left) or imitated number (Mailbox Task, right) in Experiment 1.*

[8] About three quarters of our participants were tested on an extended version of Give-N in which they were asked for *N* from 1-10. For consistency, we will only be reporting participants score as 1-6. For those participants that received the extended version, we gave them a score of 6, if 1) they correctly placed 6 ducks into the pond twice or 2) if they correctly placed both 6 and 7 ducks into the pond.

(See Figure 3.2). On average, participants matched on number on 1.71 of the three trials. The majority of numerical responses involved correctly imitating the number of actions (84% of numerical responses were due to a correct imitation of the experimenter's actions) suggesting higher evidence of SFON on the Mailbox task was due to the fact that children were able to give a behavioral response, rather than relying on their verbal skills alone (for literature on how emerging knowledge is expressed through behavior before speech see (Goldin-Meadow & Breckinridge Church, 1986; Hamamouche, Chernyak, & Cordes, under review).

Area Choice Task. Children performed significantly above chance on Standard Trials (65.5%; $t(104) = 6.05$, $p < .001$) and on Probe Trials children chose the 'number match' significantly more often than chance (65.1%; $t(104) = 6.76$, $p < .001$). Thus, when number was pitted against area for small sets, children were more likely to attend to numerical information, relative to area. Children's performance on both the Standard and Probe trials correlated positively with age (Standard: $r = .41$, $p < .001$; Probe: $r = .29$, $p < .01$; See Table 3.2 for all correlations of Experiment 1).

Relations between SFON tasks

Aligning with other work, there were no significant correlations in performance on any of our SFON tasks when controlling for age (p's > .20; See Table 3.2 for all correlations of Experiment 2). However, to determine whether the response mode (verbal vs. behavioral) could explain why we did not find a correlation between the different tasks, we separated those trials in the Mailbox task in which children gave a behavioral vs. a verbal response, and correlated those to the other SFON tasks. Not only did verbal responses in the Mailing Task correlate positively with the Picture task ($r = 0.21$, $p <.05$),

Variable	Age	2	3	4	5
1. Picture Task	. 01 (*N*=106)	-.02 (*N*=94)	.04 (*N*=94)	-.11 (*N*=93)	-.04 (*N*=87)
2. Area Task– Standard	.41*** (*N*=105)		.46*** (*N*=105)	.13 (*N*=89)	.25* (*N*=89)
3. Area Task – Probe	.28** (*N*=105)			.04 (*N*=89)	.27* (*N*=89)
4. Mailbox Task	.28*** (*N*=99)				.003 (*N*=83)
5. Give-N	.72*** (N=97)				

Table 3.2. *Correlation Matrix Experiment 1.*
The second column with the "Age" heading lists the correlations of each of our tasks with age. The rest of the table are pairwise partial correlations between our difference tasks when controlling for age.
*p< .05, **p<.01*** p<.005

behavioral responses in the Mailing task correlated with the Area Choice task for Standard (r = 0.22, p <.05), but not Probe trials (r = 0.11, p <.29). This suggests that response mode played a significant role in determining how a child performed on each of the particular tasks.

The Relation Between Number Knowledge and SFON Measures

Children's performance on Give-N was quite variable (M = 3.65, Range: 0-6) and as expected, was highly correlated with age (r = .72, p < .001). Therefore, all analyses examining the relation between our SFON tasks and Give-N controlled for age (in months). Surprisingly, Give-N performance was only positively correlated with performance on one SFON task - the Standard (r = .27, p < .05) and Probe trials (r = .29, p < .01) of the Area Choice task when controlling for age. None of the other SFON measures correlated with Give-N performance (p's > 0.3). Follow-up analyses explored whether the relation between Give-N performance and Probe Trial performance on the Area Choice task held across knower-levels, or if it was specifically driven by the distinction between Subset and CP- knowers. To answer this question, we divided our participants in 2 groups: Subset knowers (0-5 knowers) and CP knowers. First of all, independent samples t-tests found that both Subset (57.8%; $t(51) = 2.46$, $p = .02$, $d = .32$) and CP knowers (76.5%; $t(36) = 7.98$, $p < .001$, $d = 1.29$) performed above chance on probe trials, yet an ANCOVA comparing performance between Subset and CP knowers with Age as a covariate did find CP knowers to be performing significantly better than Subset knowers on Probe trials ($F(2, 86) = 10.26$, $p < .01$, $\eta_p^2 =.11$).

Next, we determined to what extent performance on the Probe trials held across all Subset knower levels (i.e. 0-5 knowers). A step-wise regression model with Age and Give-N predicting Probe trial performance was not significant (R^2=.06, $F(2, 49) = 1.01$, p=.37). The fact that the regression model did not hold for all knower-levels, but that subset knowers did perform above chance on these probe trials, suggests that it is specifically participants' knowledge of the cardinal principle that related to how strongly children paid attention to number in our task.

Discussion

Overall, findings from Experiment 1 revealed significant variability across our individual measures of SFON in preschoolers. In particular, whereas the Picture task revealed near floor performance – likely due to the verbal requirements of the task – preschoolers were significantly more likely than chance to match number on both the Mailbox task and the Area Choice task. Thus, adding to other work in this domain, results of Experiment 1 reveal no consistency in SFON performance across distinct SFON tasks.

The findings that verbal SFON on the Mailbox task correlated with SFON in the verbal Picture task, and that behavioral performance on the Mailbox task correlated with behavioral performance on the standard trials of the Area Choice task, suggest that response mode played a large role in determining the extent of SFON that children demonstrated across these three tasks. Otherwise, we saw no correlations between tasks that had different response modes (i.e. no correlations between the Picture and Area Task). These various SFON tasks, that have all traditionally been used to measure SFON therefore may measure distinct aspects of a child's cognition (i.e. fluency with language), either separate from or in addition to measuring SFON.

Interestingly, we also found that despite prior reports relating SFON in these different tasks with counting and later math ability, our findings revealed that the only SFON measure that related to our measure of a precursor to formal math - Give-N performance - was the Area Choice Task. In particular, children's mastery of the cardinal principle – not their number knowledge per se - significantly related to their tendency to demonstrate an attention to number in the Choice task. Although, it is not clear why only one of the tasks – the Area Choice Task - correlated with children's cardinal knowledge, given that this task may have been the least ambiguous of them all, it seems possible that the Area Choice task was simply the most straightforward task in measuring SFON.

Experiment 2

All of our SFON tasks in Experiment 1 presented children with sets of items in the small number range (< 4 items), replicating prior SFON studies where children were likely able to easily quantify the set sizes present. In Experiment 2, we examined SFON for large sets of items, specifically exploring whether this positive correlation between cardinal knowledge and SFON is exclusive to small sets or holds across all set sizes. If SFON is a truly general numerical construct, then we should see comparable levels of SFON for large sets as we do for small sets. Moreover, cardinal knowledge should also correlate with SFON for large sets – just as it does for small sets.

In Experiment 2 we presented children with large sets of items (10-40) in the Area Choice Task. Since measuring SFON via the Picture and Mailbox tasks is less conducive to large sets (doing so would be cumbersome and taxing for the child), we continued to present participants with small sets for these two SFON tasks.

Methods

Participants

Participants were 103 2.5 – 5.1 year olds (Mean age = 3.70, SD = .78, 52 Females). An additional 5 participants were excluded for experimenter error (n=1), parental interference (n=1), or failure to complete more than one task (n = 3). Of our final sample of 103 participants, 74 participants completed all 5 tasks, with the rest completing a subset of the tasks (See Table 3.3 for a breakdown of participants included in each task). Of the 45% of participants that provided us with demographic information, 74% identified as Caucasian, 8% as Asian, and 8% as biracial. Furthermore, 93% of mothers and 91% of fathers completed a Bachelor's Degree or higher.

Design

Participants completed four tasks in the following order: Picture Task (identical to Experiment 1), Area Choice Task (same structure as Experiment 1 with new stimuli), Mailbox Task (identical to Experiment 1), and Give-N (identical to Experiment 1)[9].

	Included	Excluded	Exclusion reasons	Average Age Included	Average Age Excluded
Picture Task	92	11	(1) N = 6 (2) N = 3 (3) N = 2	3.77 (2.5 – 5.1)	3.12 (2.5 – 4.1)
Area Choice Task	96	7	(1) N = 1 (2) N = 6	3.72 (2.5 – 5.1)	3.51 (2.5 – 5.0)
Mailbox Task	100	3	(1) N = 2 (2) N = 1	3.70 (2.5 – 5.1)	4.04 (2.7 – 5.1)
Give-N	96	7	(1) N = 2 (2) N = 5	3.75 (2.5 – 5.1)	3.21 (2.5 – 4.1)

Table 3.3 *Participant Exclusions Experiment 2*.
Exclusions were divided into three categories: (1) experimenter errors or errors with equipment (video cameras, computers etc.), (2) a failure on the participant's part to finish the task (for the Area Choice task, participants needed to complete more than 2/3 of Probe trials to be included, for all other tasks participants needed to complete all trials to be included) or (3) interference from the parent (e.g. telling the child how to

Tasks & Procedure

All tasks were identical to Experiment 1 except for stimuli in the Area Choice Task:

Area Choice Task. The procedures for this task were identical to Experiment 1, but the set sizes involved were larger and stimuli involved sets of purple squares (See Figure 3.3). To match the parameters to that of Experiment 1, for *Standard trials*, the number of items within the sample stimulus were increased 10-fold from Experiment 1

[9] Since we found that participants performed very poorly on the Proportion Choice Task in Experiment 1, we made some modifications on this task for Experiment 2 in an attempt to make it easier, as well as including a second Proportional Reasoning Task. However, even with these modifications, participants continued to performed below chance on both these tasks and therefore we have excluded both tasks from further analyses.

and consisted of the numbers 10, 20, 30 and 40 and the CAs for the sample stimuli were chosen from the following three possible CAs: 7200, 13200 and 19200 pixels. The CA's were larger than Experiment 1 to account for the fact that the number of items increased and we wanted to make sure no individual item was so small that they would become difficult to individuate by the participant. Our design manipulations for the correct and incorrect match stimuli were identical to Experiment 1: the correct match stimulus had the same exact number of items and CA as the sample stimulus and for the incorrect match, the number ratio between the sample and incorrect match stimulus still ranged from 0.25 to 0.75 and the CA ratios ranged between 0.38 to 0.73. *Probe trials* were designed in the exact same way as Experiment 1 but using the new set sizes and CAs.

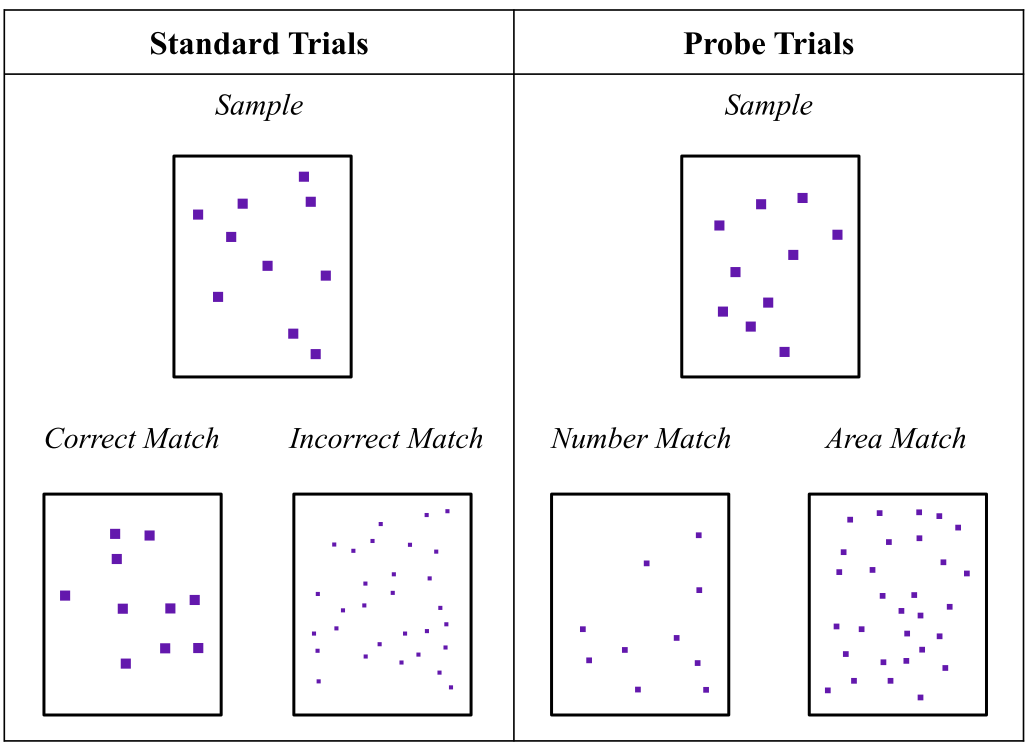

Figure 3.3. *Stimuli from the Number vs. Area task from Experiment 2.*
The procedure was identical to the Number vs. Area task from Experiment 1. The stimuli were changed such that the number of items in each display in Experiment 1 were multiplied by 10, thus creating large sets (ranging from 10-40 items).

Data Processing & Coding

All tasks were scored identically to Experiment 1. For the Picture and Mailbox task, data from 20% of participants were transcribed and/or recoded by a second coder and reliability between the two coders was calculated using linear weighted Kappa which resulted in Kappa values of .93 and .87 for the Picture and Mailbox task respectively.

Results

Measures of SFON

Picture Task. Similar to Experiment 1, there were very few number words used on this task, with participants using number words on only .79 of the three trials on average (See Figure 3.4). However, unlike Experiment 1, we did find a significant correlation between number word usage on this task and age (r = .34, p < .001; See Table 3.4 for all correlations for Experiment 2).

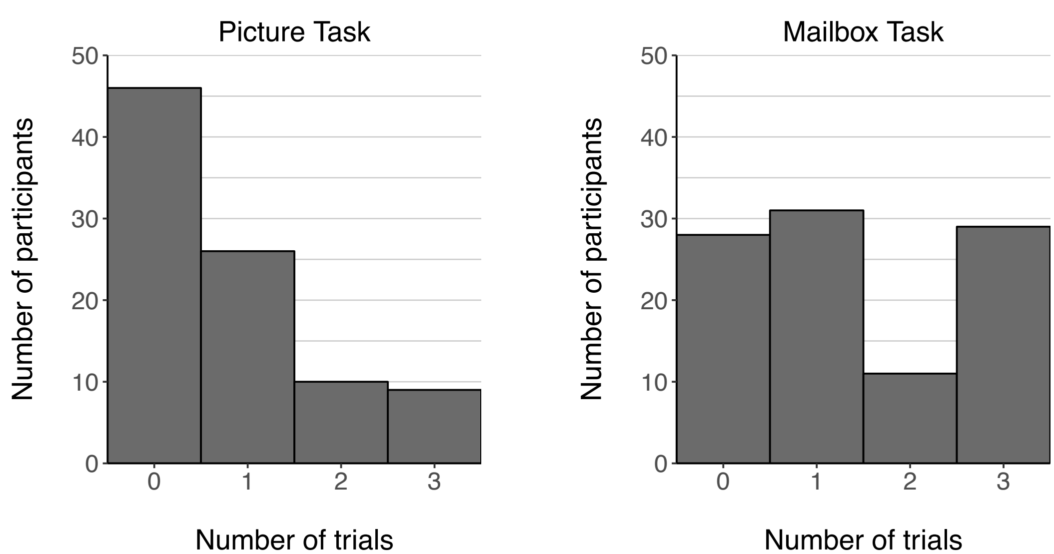

Figure 3.4. *Histograms depicting the number of trials on which participants used quantity words (Picture Task, left) or imitated number (Mailbox Task, right) in Experiment 2.*

Mailbox Task. SFON on the Mailbox task was again higher than that of the Picture task, with 72% of our participants imitating number on one or more trials (See Figure 3.4). On average, participants matched on number on 1.41 of the 3 trials. Similar to Experiment 1, it appears that better performance on the Mailbox task was driven by children's behavioral nonverbal responses (93.9% of SFON scores were behavioral, not verbal). Similar to the Picture task, here too performance was correlated with age (r = .21, p = .03).

Area Choice Task. Even in the context of large sets, participants performed significantly above chance on Standard trials (57.8%; $t(95) = 2.83$, $p < .01$, $d = .22$). However, participants were not more likely than chance to select the number match on Probe trials (50.2%; $t(95) = .09$, $p = .93$, $d = .009$). Therefore, although children were able to make a match when both area and number were confounded in Standard trials,

Variable	Age	2	3	4	5
1. Picture Task	.34*** (*N*=92)	.22* (*N*=85)	.12 (*N*=85)	-.05 (*N*=89)	.16 (*N*=87)
2. Area Task–Standard	.39*** (*N*=95)		.09 (*N*=95)	-.11 (*N*=92)	.08 (*N*=90)
3. Area Task – Probe	.08 (N=95)			.08 (*N*=92)	.06 (*N*=91)
4. Mailbox Task	.21* (*N*=100)				.08 (*N*=93)
5. Give-N	.74*** (*N*=96)				

Table 3.4. *Correlation Matrix Experiment 2.*
The second column with the "Age" heading lists the correlations of each of our tasks with age. The rest of the table are pairwise partial correlations between our difference tasks when controlling for age.
*p< .05, **p<.01*** p<.005

when number was pitted against area in the Probe trials, number did not continue to be a salient cue for matching. This suggests that for large sets participants no longer demonstrated a clear reliance on number and instead showed no preference for either number or area. This could either mean that some participants matched on number and others matched on CA – in which case we would expect to see a bimodal distribution of results – or, alternatively, that participants responded randomly and did not have a consistent strategy – in which case we would expect an approximately normal distribution. A histogram of the number of children choosing number at different levels (see Figure 3.5) was found to be fairly normally distributed suggesting that most children randomly selected their responses on Probe Trials, with no clear strategy. Despite the

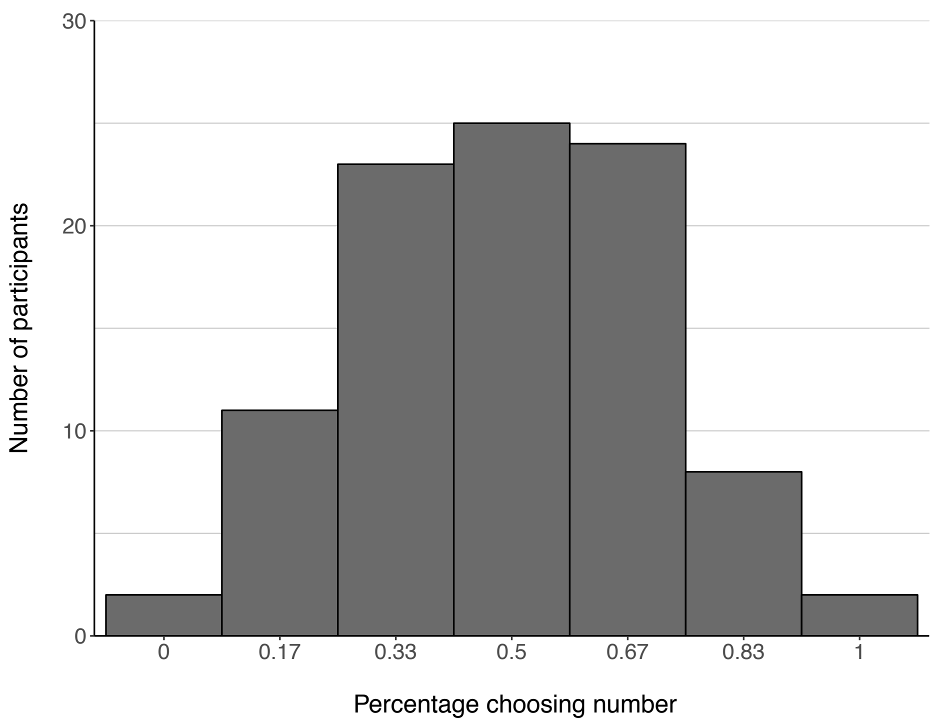

Figure 3.5. *Histogram depicting the number of participants choosing number on the. Area Choice task in Experiment 2.*

poor performance on Probe trials we did find age to correlate positively with performance on the Standard but not the Probe trials (Standard: $r = .39$, $p < .001$; Probe: $r = .08$, $p = .44$).

Relations between SFON tasks

There were no significant correlations between any of our SFON tasks when controlling for age (p's >.05; see Table 3.4). As in Experiment 1, we separated those trials in the Mailbox task in which children gave a behavioral vs. a verbal response to determine the extent to which task demands may have played a role in performance on these SFON tasks. Again, verbal responses in the Mailing Task correlated positively with performance on the Picture task ($r = 0.26$, $p <.05$), however the behavioral responses did not correlate with any of the SFON tasks (*p's* > .25).

Number Knowledge and SFON

Performance on our Give-N task was again highly variable ($M = 4.13$, Range: 0-6) and was again strongly correlated with age ($r = .74$, $p < .001$). In contrast to Experiment 1, we found no correlation between any of our SFON tasks and number knowledge when age was controlled for (p's > .10). This suggests that while number knowledge correlated with children's spontaneous focusing on number in the Are Choice Task when sets were small, this was not the case with large sets.

However, although the relationship between Give-N performance and SFON performance does not hold across all knower-levels, it is possible that we do see a different pattern of performance based on whether participants were Subset or CP-knowers. Breaking our participants down into Subset knowers (0-5 knowers) and CP knowers, we found neither Subset (46.8%; $t(49)=1.08$, $p =.29$, $d = -.15$) or CP knowers

(54.3%, $t(39)=1.23$, $p=.23$, $d=.19$) matched based on number above chance, nor did Subset and CP knowers differ significantly in their performance on Probe trials ($t(88)=1.65$, $p=.10$, $d=.33$) .

Discussion

Mirroring findings of Experiment 1, our four measures of SFON revealed no consistency within participants. Although we did find that verbal responses in the Mailing Task correlated positively with performance on the Picture task, none of our other SFON measures were correlated with one another. Again, these findings suggest that these four tasks measure distinct aspects of a child's cognition, and not a singular pure construct of SFON. These findings raise questions regarding the validity of these measures in assessing children's SFON since tasks demands in particular seem to play a role in the level of SFON preschoolers demonstrate.

In contrast to Experiment 1, when presented with large sets in the Area Choice Task of Experiment 2, children 2.5 – 5 years of age did not focus on number over cumulative area more than chance. Given that children, of all knower levels, should have easily been able to discriminate between the large numbers we presented them (preschool-aged children have been shown to discriminate a 3:4 of change, the hardest ratio tested here; Odic, Libertus, Feigenson, & Halberda, 2013), this finding suggests that a lack of focusing on number cannot be explained by an inability to discriminate the numerosities presented. Instead these findings lead us to conclude that when sets are large, number becomes less salient to children with less number knowledge. Notably, although children did not select the numerical match at above chance levels, they also did not select the area match at above chance levels either. Thus, it is not the case that children found area to be more salient in the context of large sets, but simply that number was not a salient dimension.

Furthermore, we also did not see a relationship between children's preference for number in the context of large sets (Area Choice Task) and their cardinal knowledge. Combined with findings from Experiment 1, this suggests the possibility that SFON may not necessarily be an independent construct of numerical attention in general, but instead may reflect a child's ability to attend to numbers that s/he can quantify in the world around her/him. When tested only on small sets, those children with greater number knowledge focused on number. However, when tested on large sets – sets that went beyond the scope of children's number knowledge – we no longer saw this relationship between children's knowledge of number words and their SFON. A direct comparison between our findings in Experiment 1 and 2, may be able to tell us more about this different pattern of results for Small and Large sets.

Combined Analyses Experiment 1 & 2

Before analyzing the effect of set size (i.e. Experiment) on SFON, we wanted to verify that there was no significant difference between participants in either Experiment in age or Give-N performance and found this not to be the case (*p*'s >.5). We ran an ANCOVA with Age as a covariate comparing performance on the Standard trials of the Area Choice Tasks in Experiments 1 and 2 and found a significant effect of Experiment ($F(1, 197) = 5.51$, $p = .02$, $\eta = .03$), with participants performing significantly better on Standard trials when presented with small (65.5%) compared to large (57.8%) sets. A similar analysis on probe trial performance similarly revealed that participants were significantly more likely to select the number match on probe trials when presented with small (65.1%) compared to large sets (50.2%; $F(1, 197) = 22.58$, $p < .001$, $\eta_p^2 = .10$)[10].

Given that previous studies found that children perform better on trials where the sample stimulus contained only a single item (1-item arrays; Cantlon et al., 2010), we wanted to verify that our previous findings showing better performance on small sets was not driven by those trials that included 1 item in the sample. Thus, we reran the above analyses excluding trials that included 1 item in the sample for Experiment 1 and those matched trials in Experiment 2 that included 10 items in the sample. An ANCOVA with Age as a covariate found no significant difference in performance on small (62.0%) and large sets (57.0%) for standard trials ($F(2, 197) = 1.80$, $p = .18$, $\eta_p^2 = .01$), although participants continued to be significantly more likely to select a number match on probe trials when presented with small (66.3%) compared to large sets (52.3%; $F(2, 195) = 17.65$, $p < .001$, $\eta_p^2 = .08$)[11]. Thus, performance on the Standard trials of the Area Choice Task of Experiment 1 may have been slightly boosted by performance on trials involving a single item – likely because those trials included the largest ratio difference of 1 vs 4, making it easiest for children to notice numerical changes. However, importantly, Probe trial performance was unaffected by the exclusion of single item trials, suggesting that the pattern of greater numerical matching in Experiment 1 compared to Experiment 2 was not driven by single-item trials, but instead by the distinction in set sizes.

Next, we ran a regression to test to what extent performance on the Probe trials was dependent on Age, Set Size (Experiment 1 vs Experiment 2), and Number Knowledge (Give-N; see Table 3.5). Age, the two predictors (Experiment and Give-N)

---

[10] These findings held when Give-N was used as a covariate instead of Age, *p*'s <.01

[11] This pattern of results held when Give-N performance was used as a covariate.

and their interaction was entered into a step-wise regression model and the model was significant (R^2=.18, $F(4, 174) = 9.31$, $p < .001$). We found that better number knowledge was associated with better performance on the probe trials (b = .045, SE_b =.014. ß = 38, $p < .01$) and there was also a significant interaction (b = -.05, SE_b =.02. ß = -37, $p < .05$) consistent with our finding that the effect of Number Knowledge on Probe trial performance was dependent on Set Size (i.e. Experiment). In particular, when presented with small sets (Experiment 1) each increase in Knower Level resulted in a .05 increase in performance on Probe trials. We reversed our dummy coding of the Experiment

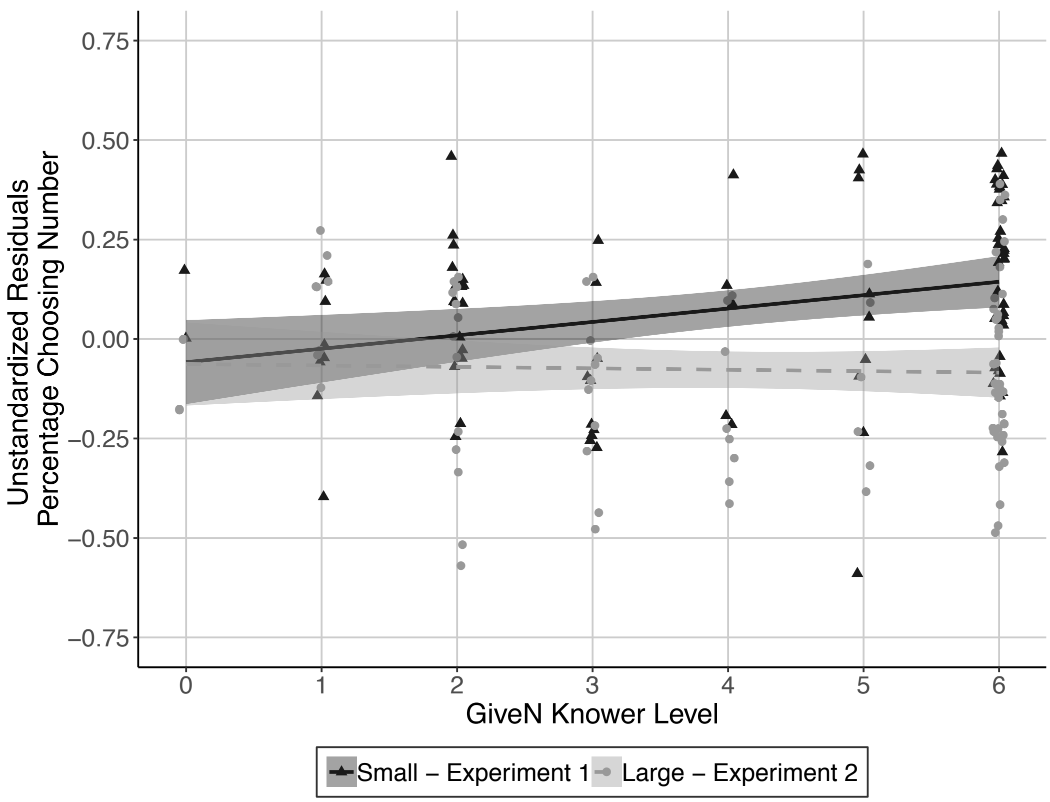

Figure 3.6. *Scatterplot displaying the relationship between participants' Number Knowledge (measured through Give-N task) and their percentage choosing the number match on the Number vs. Area Task (using unstandardized residuals), when controlling for age on Experiments 1 and 2.*
Data points have been jittered to reduce overplotting.

variable to determine whether Number Knowledge was still significant for large Sets (Experiment 2), and found this was not the case (b = .01, SE_b =.02. ß = .09, p = .48). This confirms previous findings indicating that Number knowledge only had an effect on Probe trial performance, when the Probe trials included small sets (See Figure 3.6)[12].

[12] A third experiment was conducted with adults, testing them on the Pictures Task and the Small and Large version of the Area Choice task, along with a few measures of math abilities. This Experiment generally showed high levels of SFON on all tasks, although again a significantly greater likelihood of selecting the number match for small, compared to large, sets. See Supplementary materials for a full report of Experiment 3.

Variable	B	Std. Error	ß	t	p	R^2	R^2 change
Step 1						.02	.02
Age	.05	.03	.14	1.84	.07		
Step 2						.16	.14
Age	-.01	.03	-.04	.34	.74		
Experiment	-.16	.03	-.33	4.70	<.001		
Give-N	.03	.01	.25	2.42	<.05		
Step 3						.18	.02
Age	-.004	.03	-.01	.12	.90		
Experiment	-.01	.08	-.03	.18	.86		
Give-N	.05	.01	.38	3.19	<.01		
Experiment x Give-N	-.05	.02	-.37	2.11	<.05		

Table 3.5. *Regression with Probe Trials as our dependent measure. Experiment 1 (Small sets) was coded as 0, and Experiment 2 (Large sets) was coded as 1.*
*p< .05, **p<.01*** p<.005

General Discussion

The aim of the current study was to examine the relationship between four different SFON measures that have been used in the literature to determine 1) whether these measures were correlated with one another and 2) which measure of SFON correlated most strongly with Number Knowledge. Secondly, we were interested in comparing preschooler's SFON for small and large sets by examining whether there were similar levels of SFON when presented with small (Experiment 1) and large (Experiment 2) sets in the Area Choice task. Given the evidence showing that SFON, tested in the preschool years, is predictive of later long-term measures of math achievement (Hannula-Sormunen et al., 2015, 2010; McMullen et al., 2015), it is important that we get a better understanding of what the limits of SFON are to further our understanding of its relationship with other numerical and mathematical abilities.

Our first research question pertained to how performance on different SFON tasks may be related, allowing us to measure how context and task demands affects children's tendencies to pay attention to number. We tested children on three different tasks (Imitation Task, Picture Task, and Number vs. Area Choice Task) and found that none of our SFON tasks correlated with one another (controlling for age). The only exception was that when breaking down performance on the Mailbox task down by whether children had given a verbal or behavioral response, verbal responses alone correlated with Picture task performance in both Experiment 1 and 2, and the behavioral responses correlated with the Area Choice task in Experiment 1 only. These findings therefore replicate to some extent previous research showing that performance on the Picture and Imitation tasks as a whole do not correlate with one another (Batchelor et al., 2015; Rathé et al., 2016).

Instead, these findings suggest that the difference in response modes of the three tasks may play a role in the level of SFON children show. Although the Picture Task relied entirely on verbal expression of numerical information, the Mailbox task allowed children to express their attention to number either verbally or through behavioral imitation. Our finding that only verbal responses, but not behavioral responses, in the Mailing task correlated with Picture task performance, and that verbal responses made up a very small percentage of numerical responses in the Mailing task (in Experiments 1 and 2 respectively, 84% and 93.9% of numerical responses were behavioral, not verbal), suggests that the verbal nature of the Picture Task may hinder children's expression of SFON. It seems likely that being able to talk spontaneously about number is a skill that develops after the ability to imitate or act using number (as in the Mailbox or Choice Task). This is in line with other literature on children's use of gesture for example, showing that while children may not verbally be able to express any emerging knowledge or skills that they are learning, they may be able to express it using a behavioral mode like gesture (Goldin-Meadow & Breckinridge Church, 1986). In fact, when adult participants (who are significantly more verbal than children) were asked to perform the same SFON tasks (see Supplementary materials for Experiment 3, our adult participants), we found adults were significantly more likely to talk about number than our child participants, suggesting that the verbal limitations of the task may have played a role in the low numerical performance. As such, on top of measuring SFON, the different SFON tasks may also capture individual differences in other basic cognitive skills related to response mode. One way to investigate this possibility would be to include a test of children's verbal abilities in future studies using the Picture task to determine to what extent children's verbal proficiency may impact their ability to display SFON in this task.

Another, non-mutually exclusive possibility for why participants performed so differently on these different tasks is that the different SFON measures may not measure the same underlying SFON construct. When we take a look at the three existing measures of SFON and what they have in common and where they differ, it becomes clear that there seem to be some discrepancies in what they are measuring. For example, it is unclear what role accuracy plays in SFON. In the Mailbox task (apart from the rare case where children used number words during the task) the measure of SFON is also an accuracy measure. Children who may have been attending to number, but fail to accurately imitate the correct number of actions, would not be considered to be engaging in SFON. While this is also the case in our Choice task, it was less of an issue because the particular numbers presented were expected to be within the range of values that children can track and compare using basic estimation abilities (Halberda & Feigenson, 2008). On the contrary, in the Picture task, regardless of whether children's number word use refers to the correct amount, as long as they use number words, they are given credit for engaging in SFON. Although, this was not an issue in our study (since participants attended less to number overall on the Picture task), this may be masked by the high verbal demands of this task.

Our second research question concerned how set size may impact the likelihood of a child demonstrating SFON. Prior research has only tested SFON with small sets (<4 items), sets that preschoolers are typically able to enumerate and have the number words for, making it difficult to determine whether SFON is a general numerical construct, or specific to enumerable numerosities. If, as Hannula-Sormunen and collegues (2010) claim, SFON is general attention to discrete numerical information, then we should expect two things: (1) SFON should vary very little across set sizes and (2) SFON for all set sizes, small and large alike, should relate to number knowledge. To explore these possibilities, we presented children with an Area Choice Task involving small sets (Experiment 1) and large sets (Experiment 2), allowing us to compare levels of SFON across set sizes. In contrast to predictions of a single construct of SFON, preschooler's SFON (as measured by probe trial performance on the Area Choice task) for small sets (1-4 items) was significantly greater than that for large sets (10-40 items). Moreover, regression analyses revealed that number knowledge was only associated with SFON for small sets, but not large.

These findings suggest that children's number knowledge may play an integral role in SFON, at least in the preschool years. Support for this idea comes from research suggesting that language plays an important role in solidifying certain concepts, and even remembering them across time. For example, in the domain of color, participants perform better at color discrimination tasks (Winawer et al., 2007) and have better memory for colors (Uchikawa & Shinoda, 1996) when the colors they are tested on have distinct linguistic labels (e.g., shades of green vs. blue), compared to when they are part of the same linguistic category (e.g., shades of blue). Similarly, in the domain of number, societies such as the Amazonian Pirahã tribe that speak a language that does not have words to represent exact numerical quantities show deficits in remembering the exact cardinality of large sets, suggesting that language for number may be particular useful in terms of memory and attention for number concepts (Frank, Everett, Fedorenko, & Gibson, 2008).

Relating these findings to what we know about SFON, it is possible that in the preschool years when children are learning number words, these words allow for the encoding of numerical information in their memory, and this improved memory for, and awareness of, number may be the primary driver of individual differences in SFON at this age. Thus it is possible that at least in the preschool years, SFON when assessed may be in part a reflection of a child's ability to encode number exactly – that is, SFON may be better described as a proxy for children's enumeration abilities - but not their spontaneous focusing on *any* numerical information (i.e., children's representation of large sets). Furthermore, this could mean that findings showing that SFON in the early preschool years (Hannula-Sormunen & Lehtinen, 2005) predicts later math ability, may be accounted for by the fact that number knowledge predicts later math ability (e.g., Libertus, Feigenson, & Halberda, 2011). Future studies should explore whether the relationship between SFON and later math abilities holds when controlling for differences in number knowledge.

Importantly however, our findings cannot make any claims about the relationship between SFON and enumeration in older children. Since, by the age of 5-6, children typically have already mastered the cardinal principle, individual differences in SFON for these older children may still reflect a true spontaneous attention to number and/or a mastery of verbal counting. Given that we only tested 2.5 – 5-year-olds, an age at which children are in the process of learning the meaning of number words and the cardinality principle, we cannot make claims regarding SFON for older children.

One surprising finding was that we did not replicate previous research showing a relationship between children's performance on the Imitation task (i.e., the Mailbox Task) and their number knowledge (Hannula-Sormunen & Lehtinen, 2005; Hannula et al., 2007). We believe this can be explained by a difference in our measures of number knowledge. Hannula-Sormunen and Lehtinen (2005) measured cardinality using the "caterpillar task" which presents children with caterpillars with a different number of legs, and they are then asked to bring "just enough" socks for all the legs. This task in many ways resembles an imitation or choice task, since children are first presented with the quantity (e.g., number of legs) that they then need to imitate (e.g., number of socks). On the other hand, our measure of cardinality was the more commonly used Give-N task where children are verbally instructed to put a certain number of ducks into a pond, however this task cannot be solved through imitation or matching. It is therefore possible that the structural similarity between the caterpillar task and the SFON imitation task could explain why Hannula-Sormunen and Lehtinen (2005) found a relationship between number knowledge and the Imitation task, while we did not.

In light of our findings, how should we interpret past research showing that SFON relates to children's later arithmetic and math achievement (Hannula-Sormunen & Lehtinen, 2005; Hannula-Sormunen et al., 2010)? Although we do not doubt that the ability to attend to number plays an important role in children's numerical development, given that preschool-aged children only demonstrate SFON for small sets once they have reliably learned how to track number via the counting process, it is clear that acquiring a symbolic system (i.e., language) that encodes number plays a very important role in what children pay attention to. In fact, we propose that in preschool, SFON is not a truly independent construct from cardinal knowledge and enumeration and that instead it is a reflection of an ability to encode exact number (i.e., small sets). In that case, the relationship between SFON and later math ability may be driven by the strong correlation between number knowledge and math ability. Furthermore, the fact that task demands played such an important role in whether or not children demonstrated SFON, further supports our hypothesis that SFON as has been tested in the current literature, does not seem to be a distinct construct from number knowledge in the preschool years.

In breaking down how different measures of SFON relate to children's number knowledge, we have gained a better understanding of some of the informal and self-initiated practices that young children engage in with respect to number and math. We hope that this in turn can inspire educators to develop tools and practices that continue to promote this spontaneous interest in number and math later in children's academic career.

Supplementary Materials

Our findings from Experiments 1 and 2 that children are more likely to focus on number when presented with small sets compared to large sets, suggest that children may pay attention to number when presented with numbers that they know and have number words for. If that is the case, then one would expect adults, who are able to enumerate and have number words for both small and large sets, to show SFON equally for small and large sets. Therefore, in Experiment 3, we tested adult participants on both the small and large Area Choice task, so we could determine whether adult participants would show differential SFON based on set size. We also tested our adults on the SFON Picture Task. In Experiments 1 and 2, we saw no correlation between our Picture task and any of the other SFON tasks. Since it is possible that this lack of correlation is due to the fact that the Picture task is a verbal task, and therefore is more demanding on children who have limited vocabularies, we wanted to compare whether adults would not just show higher levels of SFON, but also whether we would see a correlation between adult performance on the Picture task and the choice tasks. Lastly, we included two math measures for adults to complete. Since previous studies have found that SFON in children is predictive of their later math abilities, we wanted to investigate whether this relationship may also hold in adults.

Methods

Participants

Twenty-five undergraduate students (Mean age = 19.13, SD = 1.04, 20 Females) participated in this study. An apriori power analysis with G*power (Faul, Erdfelder, Buchner, & Lang, 2009) indicated that we would need 20 subjects to have 80% power for detecting a difference in performance between the small and large Number vs. Area task, with the traditional .05 statistical significance criteria and an effect size of .66 (the same effect we found comparing Experiment 1 and 2 Probe trials with our preschoolers). All participants completed all tasks, except one participant who did not complete the picture task due to experimenter error.

Design

Participants completed seven tasks: Synonym Task (a distractor task), Picture Task (identical to Experiment 1 and 2), Small Area Choice Task (identical to Experiment 1), Large Area Choice Task (identical to Experiment 2), Subjective Numeracy Task, Math Fluency Task, and Applied Problems. The tasks were completed in the above order except that the order of the two Area Choice Tasks (Small vs. Large) was counterbalanced across participants.

Measures

All tasks were identical to Experiments 1 and 2 except for the following:

Synonym Task. This task was used as a distractor task. Since some participants may have been familiar with the focus of research in our lab (numerical cognition), participants were first given a non-numerical task to make it seem as though number was not the primary focus of the study.

The synonym task is subtest from the Oral Vocabulary section of the Woodcock Johnson Test of Achievement (Woodcock, McGrew, & Mather, 2001). During this task, participants were presented with 9 words and were asked to read the word out loud and then to provide a synonym for that word. Since this was a distractor task, performance on this task was not recorded.

Picture Task. We presented adults with the same 3 pictures, in the same order, as Experiments 1 and 2. However, since pilot testing revealed that adults did not need more than one prompt to describe the pictures in great detail, we changed the instructions for adult participants so that they were given a time limit of 30 seconds to describe everything they saw on the picture. We recorded both whether participants used number words on each of the three pictures as well as the quantity of number words that were used.

Area Choice Tasks. Participants were presented with both the Small version of the task **(**identical to the Choice task of Experiment 1) and the Large version (identical to the Choice task of Experiment 2) separately. Since the task order was counterbalanced across participants, adults were given instructions only for the first Area Choice task that they completed and these instructions were identical to the instructions given to children in Experiments 1 and 2. For the second task, participants were told that they would "do another matching task that is a little different".

Subjective Numeracy Scale (SNS). To measure participants' level of confidence with mathematical tasks as well as their preference for numerical information over prose information, we administered the SNS, a measure that has been validated and found to correlate well with objective measure of mathematical achievement (Fagerlin et al., 2007). This scale presents participants with 8 questions, with the first four questions measuring how able and comfortable they are with everyday scenarios that involve number or mathematical calculations (e.g. "How good are you with calculating a 15% tip?") and the second set of four questions measures participants' preference for information presented numerically or with prose ("When people tell you the chance of something happening, do you prefer that they use words ("it rarely happens") or numbers ("there's a 1% chance")). Participants responded on a 1 – 6 scale where low scores (1-3) signified less comfort with numerical representations while larger scores (4-6) suggest that participants preferred numerical information. A separate score for participants' ability/comfort and preference for number is also calculated.

Math Fluency Task. This was another subtest of the Woodcock Johnson Test of Achievement (Woodcock et al., 2001) and it measured participant's abilities to solve simple arithmetic problems involving addition, subtraction and multiplication. The subjects were given a double-sided piece of paper with 160 questions and were given 3 minutes to complete as many problems as possible. Participants' accuracy on the problems as well as the total amount of problems completed was recorded.

Applied Problems. We modified this from the Applied Problems subtest of the Woodcock Johnson Test of Achievement (Woodcock et al., 2001), a subtest that is designed to measure participants' ability to solve complex word problems. The problems are organized in order of difficulty, with the first few questions designed to be challenging for children but easy for adults. Therefore, since pilot testing revealed that adult participants were performing at ceiling on the first set of questions, to reduce the number of questions participants would have to answer, we chose to only test participants on the 8 last questions of the test (the 8 most difficult questions). Participants were instructed to complete all problems and were told that they would be timed. In our analyses, we measured performance both by looking at percent correct as well as the time taken to complete the questions. If participants had not completed all problems after 15 minutes, they were stopped. The questions included real life scenarios that required mathematical computations to complete. An example question is: "*Doug works part time at a music store. As an employee, he received a discount of 10 percent on all purchases he makes. How much does he have to pay for a CD that sells for six dollars and ninety cents?"*

Data Processing & Coding

Since we were interested in examining more generally how adult math performance correlated with SFON, but not how any particular math task (Math Fluency or Applied Problems) related to SFON, we created a combined Math Score by transforming raw scores on the two tasks into z-scores, and then averaging across those the z-scores. We used this combined Math score in further analyses

Results and Discussion

Measures of SFON

Picture Task. Of the 23 participants that completed this task, only 2 participants did not use number words on all three trials. Given these ceiling effects, it was more useful to examine the average quantity of number words used . On average, summed across the three trials, participants used 14.74 number words (Range: 6 – 24).

Area Choice Tasks. For the Standard trials, all participants got 100% of trials correct on the Small Choice Task and performed slightly less well, but still significantly above chance on the Large Choice Task (93.1%; $t(24) = 15.25$, $p < .001$). However, performance was significantly better on the Small compared to the Large Choice task ($t(23) = 2.46$, $p < .02$). Furthermore, for the Probe trials, participants matched on number significantly above chance for both the Small (98.6%; $t(24) = 50.61$, $p < .001$) and Large Choice Task (73.6%; $t(24) = 6.82$, $p < .001$), but once again participants were more likely to select the numerical match on the Small compared to the Large Choice Task ($t(23) = 6.66$, $p < .001$).

Relations between our three SFON Measures. None of our measures of SFON were significantly correlated with one another (*p*'s >.13)

Measures of Mathematical Ability

SNS. The maximum score on the SNS was 40 points, indicating a high level of comfort with numerical information in everyday situations. On average participants scored 26.56 (Range:14-33), which suggests a slight preference for numerical over non-numerical information in everyday situations.

Math Fluency. Of the 160 problems, on average participants attempted 133.08 problems (Range: 96 - 160) in the 3-minute allotted time period, with 4 participants completing all 160 problems. When only taking into account the number of problems participants attempted, we see that the average performance was very high (95.49%).

Applied Problems. Participants completed all problems and on average got 52.08% of problems correct (*Range*: 25.00 – 75.00%). In terms of time, participants on average took 8mins 28secs to complete the 8 problems.

Relationship between SFON and Mathematical Ability

There were no significant correlations between our two Choice Tasks and our combined Math score. There was however a negative relationship between the quantity of number words used in the Picture Task and adult Math performance ($r = -.44$ $p = .03$). No other significant correlations were found.

Overall, we find that performance on neither the small or large choice tasks correlated with adult math scores. Given our findings in Experiment 1 and 2 that children showed greater SFON when they had more number knowledge (higher Give-N score), we expected that since adults have knowledge of both small and large sets, that we would see a positive correlation between both the small and large choice task and math ability. However, given that adults were at ceiling in the Small Choice task and performing very highly on Large Choice task, this may explain why we did not find any correlations here.

IMPLICATIONS AND FUTURE DIRECTIONS

The ability to represent quantities is extremely important for our day-to-day functioning. We use our ability to track quantities to judge how much food to buy at the grocery store, to notice how fast we are driving, or to determine whether we have enough time to finish our favorite TV show. Although many of these actions have become automatic to us as adults, how do young infants and children represent quantities in the world and how salient are different quantities to them? Furthermore, can adults track number alone or must they rely upon cues from continuous quantities as well? These questions have been of interest to some of the first developmental psychologists such as Piaget, but to this day continue to attract researchers in the field of numerical cognition.

One reason the nature of our quantitative representations has interested so many researchers is that number is the only quantity we learn about extensively through formal training in school, and children's abilities to represent discrete number at an early age is predictive of later math achievement (Geary, 2011; Jordan, Kaplan, Locuniak, & Ramineni, 2007; Starr, Libertus, & Brannon, 2013b). Furthermore, math achievement has been shown to be one of the most important predictors of later academic success (Duncan et al., 2007; Romano, Babchishin, Pagani, & Kohen, 2010), yet many children fall behind in math achievement in school (Byrnes & Wasik, 2009; Chatterji, 2005; Cross, Woods, & Schweingruber, 2009). To develop tools that can help children struggling in math, it is important not only that we understand how we represent quantity at an early age, but also how these representations of quantity may change throughout development.

Aside from discrete number which children are exposed to extensively through formal education, there are a variety of continuous quantities that we can also represent such as size, cumulative area, time, or density, for example. Furthermore, there is naturally strong correlation between discrete and continuous variables such that a basket of 10 apples will not only be more numerous than 5 apples, but will also have a larger cumulative surface area and will also be more densely packed in a basket. Due to the fact that these quantities are so highly correlated with one another, researchers have questioned the extent to which we track these quantities independently of one another. In particular, proponents of the neo-piagetian "Sense of Magnitude" theory claim that because continuous quantities are "perceptual" in nature (they can be represented directly from the percept) they must be much easier to track than number, a quantitative dimension thought to be tracked independent of perceptual qualities of the display (i.e. number is abstract in nature; Gebuis & Reynvoet, 2012; Leibovich, Katzin, Harel, & Henik, 2017). Due to this, Sense of Magnitude theorists suggest that young infants who have yet to be formally taught about number, should only be able to discriminate quantities using continues cues because continuous quantities do not require higher order cognitive processes needed to represent abstract number. Then, once infants develop an understanding of the correlation between discrete and continuous quantities, often after formal education, children and adults learn how to disentangle number from continuous quantities. Importantly though, even though older children and adults *can* use number to represent quantity, they will only do this as a last resort. Thus, two important implications of this theory are that 1) infants, children and adults should represent number with less accuracy than continuous cues and 2) when presented with both discrete and continuous quantities, number should be less salient.

The aim of this dissertation was to investigate these two implications of the Sense of Magnitude theory as a way to determine whether this theory accurately models how humans represent quantity at different stages of development. To do this, I presented three separate projects, each examining a different stage in development. In Project 1, I investigated infant abilities to discriminate continuous quantities (specifically element area) and compared these abilities to previous work on infant number discrimination abilities. If continuous quantities are easier to represent than number in infancy, then we would not only expect infants to discriminate element area with similar (or better) precision than number, but we would expect a parallel increase in acuity with development, similar to what we see in number. Not only did we replicate previous findings showing that 7 month olds cannot discriminate a 1:3 ratio of change in element area (when this same age group has successfully discriminate a 1:2 ratio in number; Xu & Spelke, 2000) by 12 months of age our infants still failed to discriminate this large ratio of change. Thus, not only are infants less accurate at representing element area compared to number, but even by 12 months of age, infant element area tracking capacities are not anywhere near that of number.

Next, in Project 2 we continued to investigate human abilities to represent continuous quantities by investigating the extent to which discrete quantity may interfere with adult's judgements of continuous quantity. If continuous quantities are truly easier to represent and more salient than number, then conflicting numerical information should not affect adult abilities to make cumulative area discriminations. In two experiments, we found adult cumulative area discriminations to be similar to that of number – a finding in contrast to Sense of Magnitude theorists positing that continuous extent discriminations should be relatively more precise than that of number. Moreover, numerical information interfered with cumulative area judgements such that as the number of items in the display increased, adults performed progressively worse. Together, these findings suggest that number is at least *as salient* as continuous variables and thus, together with our findings from Project 1, we have found no evidence to suggest that continuous quantities are easier to represent than number.

Lastly, in Project 3, I addressed the second implication of Sense of Magnitude theory which is that continuous quantities should be more salient than number. To examine this question of the relative salience of different quantities, I assessed preschoolers on Spontaneous Focusing on Number (SFON) tasks, which examine children's tendencies to spontaneously focus on number without being prompted. The previous literature on SFON has shown that number is very salient to preschool aged children, with levels of SFON correlating with later math achievement (Hannula-Sormunen & Lehtinen, 2005; Hannula-Sormunen et al., 2010), however few SFON investigations have pitted number against another quantitative dimension. In two experiments, we found that when number was pitted against cumulative area, preschool aged children found number to be more salient than cumulative area but only when presented with small sets of items (1-4 items; Experiment 1). Thus, again contradicting claims of SoM, we find cumulative area to be less salient than number to young preschoolers, at least for small sets. Notably, when presented with larger arrays (10-40 items; Experiment 2) children showed no preference for either number or cumulative area suggesting that – at least in the preschool period – SFON may simply reflect a child's ability to enumerate particular set sizes. Relating these findings back to Sense of Magnitude theory, Project 3 found no evidence to suggest that continuous quantities are more salient to children than number. When number was directly pitted against cumulative area, preschool aged children did not show a preference for cumulative area; in fact, when presented with small sets, children showed a strong preference for number.

Building on previous work that has investigated how humans at different stages of development represent quantity, these three projects demonstrate that discrete number is represented more accurately, and is more salient to infants, children and adults than are most continuous quantities. These findings, along with previous findings in the numerical cognition literature, undermine claims made by proponents of the Sense of Magnitude theory stating that humans at all stages of development are better at, and prefer to attend to continuous quantities. Instead our findings align more closely with the idea that we are born with a "sense of number" (Dehaene, 1997; Feigenson, Dehaene, & Spelke, 2004). Evidence for this comes from studies that have tested infants (even newborns, Izard, Sann, Spelke, & Streri, 2009) on their ability to discriminate number across visual (Xu & Spelke, 2000; Xu et al., 2005) or auditory displays (Lipton & Spelke, 2003), to match number across modalities (Coubart et al., 2014; Izard et al., 2009), and to distinguish increases in number from decreases in number (Brannon, 2002; de Hevia & Spelke, 2010).

An earlier theory proposed by Spelke, in fact suggests that number is one of the five core knowledge systems that we are endowed with from birth (Kinzler & Spelke, 2007; Spelke, 2000, 2003, 2017). Spelke proposed that all humans, regardless of culture, are equipped with five core systems that are innate and evolutionary ancient, that are the basis for any other knowledge we may gain throughout our experiences in the world. Under this view, our ability to discriminate number from an early age is due to the fact that we are born with a core knowledge system for number (an approximate number system) that is present across cultures and species. Then, as we are exposed to the natural language in our culture, our approximate sense of number develops into a more exact sense of number. Importantly though, the extent to which we develop an understanding of exact number is fully dependent on our culture and its natural language. That is, while the core systems are not unique to humans, the development of an exact numerical representation is, and our species is able to develop this ability specifically because we are the only species that have a developed language faculty.

Evolutionary Perspective: why have we developed a sense of quantity

Why might humans have evolved the ability to track quantity? Taking a look at the animal literature may be particularly useful here as it suggests that there are at least two important benefits of having a sense of quantity. The first is that it allows human and non-human animals to maximize food intake that is crucial for survival. Studies that have tested animals on their ability to choose the larger of two sets of foods items have shown that dogs (Ward & Smuts, 2007), coyotes (Baker, Shivik, & Jordan, 2011) and orangutans (Call, 2000) reliably choose the set that is largest in both number *and* cumulative area. Furthermore, studies that have tried to determine which quantitative cue non-human animals use when making quantitative discriminations in food related contexts have found that although many animal species can use number alone to make food comparisons, they are often biased by other continuous cues that may be present, such as the size of the largest individual item in a set (Chimpanzees: Beran & Rumbaugh, 2001; Guppies: Lucon-Xiccato, Miletto Petrazzini, Agrillo, & Bisazza, 2015) or even the amount of movement of the prey (Salamanders: Krusche, Uller, & Dicke, 2010).

A second evolutionary important reason for having developed a sense of quantity is for the purpose of avoiding predators or hunting for prey. Angelfish will choose the larger of two shoals of conspecifics to join because this is more likely to provide safety against predators (Gómez-Laplaza & Gerlai, 2011). Female African Lions are less likely to approach and contest another social group when they hear calls of three intruding lions nearby, as opposed to only one, suggesting that they use number to make assessments about the safety of their approach (McComb, Packer, & Pusey, 1994). Similarly, male chimpanzees will make their decision on whether to approach based on their relative numbers compared to an intruding group of chimpanzees (Wilson, Hauser, & Wrangham, 2001). Lastly, counting even plays a role in the egg-laying practices of American coots (Lyon, 2003). American coots engage in brood parasitism, the practice whereby female birds lay some of their eggs in the nests of conspecifics. In protection of their own brood, research suggests that coots will count their eggs, discounting 'parasitic' eggs, as a way to make decisions about how many future eggs to lay. Thus, the ability to visually keep track of the number or surface area of the eggs in the nest is crucial for decisions related to clutch sizes, which in turn has important consequences for the fitness of the American coot. Therefore, it is clear that for social human and non-human animals, representing quantity is crucial for survival. Not only does it allow animals to keep track of the relative sizes of different social groups, it allows them to make efficient choices when hunting for prey, all to ensure the survival of their species. Importantly, in the context of tracking predator/prey relationships, continuous quantities are significantly less important than relative number (i.e., the number of approaching lions is more important than the duration of their calls), perhaps suggesting that social relationships may underlie the evolution of discrete number tracking abilities in humans.

The role of education in our preference for number

Even if Spelke's core knowledge theory is correct and our ability to represent number (at least approximate number) is innate, it is likely that the emphasis on number in most formal education systems further drives the preference for number later on in development. In fact, my finding in Project 3 that children's knowledge of number was correlated with the children's preference for number over cumulative area suggests that children's preference for and knowledge of number go hand in hand. Furthermore given that a myriad of studies continue to find a strong relationship between children's number understanding or counting skills in the preschool years and their later math achievement (Aunio & Niemivirta, 2010; Jordan et al., 2007; Jordan, Kaplan, Ramineni, & Locuniak, 2009), it is not surprising that such a strong emphasis has been put on developing counting skills in the United States common core standards for example (Common Core Standards initiative, 2010).

Unfortunately, attention to number early in development can be at the expense of other important quantitative information. For example, recent work has revealed children's overt attention to number may compromise their abilities to track non-symbolic proportional information – a skill necessary for the later acquisition of formal fractions (Boyer & Levine, 2015; Boyer, Levine, & Huttenlocher, 2008; Hurst & Cordes, 2018; Jeong, Levine, & Huttenlocher, 2007). For example, in non-symbolic proportional matching tasks (e.g. the “Wally-Bear Juice Task” in which children have to find the cup that has the same proportion of juice/water as a target cup), children are more successfully at finding the correct match when the non-symbolic fraction (e.g. 4/6) is depicted continuously (e.g. the cup has 2/3 juice and 1/3 water) versus discreetly (e.g. the cup would have 4 juice pieces and 2 water pieces; Boyer & Levine, 2015; Boyer et al., 2008; Hurst & Cordes, 2018). This discrepancy in performance has been attributed to the fact that children are biased towards number in discrete situations (they will count the pieces) at the expense of proportional information. Although various hypotheses have been put forth as to why this bias may exist (e.g. over-emphasis of number in early education or under-emphasis of proportional information), it is clear that this strong preference for number is not always helpful. It seems therefore that a combination of an innate system for number along with our educational system that strongly emphasizes number, even to the detriment of developing other quantitative skills, explains why humans, from infants to adults, show such a strong preference for number over other quantitative information.

Future Directions: The origin of individual differences in number sense

The current dissertation has highlighted just how attuned humans of all ages are to number in their environment. Yet, the reality is that there are great disparities in math proficiency in children even as early as first grade. This begs the question at what point in a child’s formal or informal education, this gap in number and math understanding emerges. Only one study has directly examined the relationship between infant’s early number sense and how this relates to their mathematical achievement years later, finding that number sense at 6 months predicted math achievement at 3 years, even when controlling for general intelligence (Starr et al., 2013b). What this study was not able to address, however, is what led to these individual differences at 6 months. Similarly, research in the domain of SFON suggests that individual differences in SFON in kindergarten was also predictive of mathematical achievement in second grade, although again it is not clear what the source of these early individual differences are (Hannula-Sormunen, 2014; Hannula-Sormunen et al., 2010). One possibility, that we addressed in Project 3, is that SFON in the preschool years, may really be a proxy for children's enumeration abilities, and thus perhaps the individual differences we see in SFON are really due to individual differences in children's counting abilities.

Future work would benefit from examining the source and onset of these individual differences, especially given that the ability to discriminate number is present even in newborns (Izard et al., 2009). One way to examine this question would be to track individual differences in number discrimination throughout infancy. Although we know that infants generally follow a steady developmental trajectory of improving in their number discrimination such that newborns can discriminate a 1:3 ratio, 6-month-olds discriminate a 1:2 ratio and 9-month-olds discriminate a 2:3 ratio (Lipton & Spelke, 2004; Xu & Spelke, 2000), it is unclear whether on an individual basis, infants start off with the same number sense at birth, and if infants that fall behind at some point in infancy, continue to stay behind in their number discrimination abilities a few months later. On the other hand, another non-mutually exclusive possibility is that the difficulties we see in math achievement in grade school develop in early childhood when children are being exposed to early number skills such as counting in the home and school environment. In fact, a myriad of studies have shown that there is a large discrepancy in the quantity and quality of math talk in the home and at school (Elliott, Braham, & Libertus, 2017; Gunderson & Levine, 2011; Klibanoff, Levine, Huttenlocher, Vasilyeva, & Hedges, 2006; Rowe, Levine, Suriyakham, Gunderson, & Huttenlocher, 2010). Another possibility is that differences in the exposure to number and counting skills in the home and school already has an effect on young infant's development of a number sense, and that this delay in reaching number discrimination milestones, affects their later math achievement. As such, it is not clear from the current research *when* these individual differences first occur which makes implementing successful interventions difficult.

Conclusion

In conclusion, this dissertation examined infant, child, and adult abilities to represent continuous quantities and the relative saliency of these continuous quantities compared to number. Not only did I demonstrate that infants and adults are poor at discriminating continuous quantities, adult representations of continuous quantities were affected by the presence of numerical information, suggesting that number is at least as salient as continuous quantity. Lastly, when pitted against each other, preschool aged children preferentially attended to number over continuous quantities. By examining these questions across development, and in combination with previous findings in the literature, I have shown not only that humans are not very accurate at representing continuous quantities but also that they do not find those perceptual cues to be particularly salient in their environment, undermining claims made by the prominent Sense of Magnitude Theory (Gebuis & Reynvoet, 2012a; Leibovich et al., 2017). By investigating these questions, I have provided insight into the way we represent and use different dimensions of quantity in our day to day lives, helping us understand how individual differences in the salience of these early representations may affect the way children and adults perform on tests of mathematical achievement in later life.

REFERENCES

Aunio, P., & Niemivirta, M. (2010). Predicting children's mathematical performance in grade one by early numeracy. *Learning and Individual Differences*, *20*(5), 427–435. http://doi.org/10.1016/j.lindif.2010.06.003

Baker, J. M., Shivik, J., & Jordan, K. E. (2011). Tracking of food quantity by coyotes (Canis latrans). *Behavioural Processes*, *88*(2), 72–75. http://doi.org/10.1016/j.beproc.2011.08.006

Barth, H. C. (2008). Judgements of discrete and continuous quantity: An illusory Stroop effect. *Cognition*, *109*(2), 251–266. http://doi.org/10.1016/j.cognition.2008.09.002

Batchelor, S., Inglis, M., & Gilmore, C. (2015). Spontaneous focusing on numerosity and the arithmetic advantage. *Learning and Instruction*, *40*, 79–88. http://doi.org/10.1016/j.learninstruc.2015.09.005

Beran, M. J., & Rumbaugh, D. M. (2001). "Constructive" enumeration by chimpanzees (Pan troglodytes)on a computerized task. *Animal Cognition*, *4*(2), 81–89. http://doi.org/10.1007/s100710100098

Boyer, T. W., & Levine, S. C. (2015). Prompting children to reason proportionally: Processing discrete units as continuous amounts. *Developmental Psychology*, *51*(5), 615–620. http://doi.org/10.1037/a0039010

Boyer, T. W., Levine, S. C., & Huttenlocher, J. (2008). Development of Proportional

Reasoning: Where Young Children Go Wrong. *Developmental Psychology*, *44*(5), 1478–1490. http://doi.org/10.1037/a0013110

Boysen, S. T., Berntson, G. G., & Mukobi, K. L. (2001). Size matters: Impact of item size and quantity on array choice by chimpanzees (Pan troglodytes). *Journal of Comparative Psychology*, *115*(1), 106–110. http://doi.org/10.1037/0735-7036.115.1.106

Brannon, E. M. (2002). The development of ordinal numerical knowledge in infancy. *Cognition*, *83*(3), 223–240. http://doi.org/10.1016/S0010-0277(02)00005-7

Brannon, E. M., Abbott, S., & Lutz, D. J. (2004). Number bias for the discrimination of large visual sets in infancy. *Cognition*, *93*(2), B59-68. http://doi.org/10.1016/j.cognition.2004.01.004

Brannon, E. M., Lutz, D., & Cordes, S. (2006). The development of area discrimination and its implications for number representation in infancy. *Developmental Science*, *9*(6). http://doi.org/10.1111/j.1467-7687.2006.00530.x

Brannon, E. M., & Terrace, H. S. (1998). Ordering of the numerosities 1 to 9 by monkeys. *Science*, *282*(5389), 746–749. http://doi.org/10.1126/science.282.5389.746

Brannon, E. M., Wusthoff, C. J., Gallistel, C. R., & Gibbon, J. (2001). *NUMERICAL SUBTRACTION IN THE PIGEON: Evidence for a Linear Subjective Number Scale* (Vol. 12). Retrieved from https://cpb-us-w2.wpmucdn.com/web.sas.upenn.edu/dist/2/51/files/2016/10/Psychological-

Science-2001-Brannon-238-43-1yoq60e.pdf

Byrnes, J. P., & Wasik, B. A. (2009). Factors predictive of mathematics achievement in kindergarten, first and third grades: An opportunity-propensity analysis. *Contemporary Educational Psychology*, *34*(2), 167–183. http://doi.org/10.1016/j.cedpsych.2009.01.002

Call, J. (2000). Estimating and operating on discrete quantities in orangutans (Pongo pygmaeus). *Journal of Comparative Psychology*, *114*(2), 136–147. http://doi.org/10.1037/0735-7036.114.2.136

Cantlon, J. F., & Brannon, E. M. (2006). Shared system for ordering small and large numbers in monkeys and humans. *Psychological Science*, *17*(5), 401–406. http://doi.org/10.1111/j.1467-9280.2006.01719.x

Cantlon, J. F., Safford, K. E., & Brannon, E. M. (2010). Spontaneous analog number representations in 3-year-old children. *Developmental Science*, *13*(2), 289–297. http://doi.org/10.1111/j.1467-7687.2009.00887.x

Cantrell, L., Boyer, T. W., Cordes, S., & Smith, L. B. (2015). Signal clarity: an account of the variability in infant quantity discrimination tasks. *Developmental Science*, *18*(6), 877–893. http://doi.org/10.1111/desc.12283

Cantrell, L., Kuwabara, M., & Smith, L. B. (2015). Set size and culture influence children's attention to number. *Journal of Experimental Child Psychology*, *131*, 19–37. http://doi.org/10.1016/j.jecp.2014.10.010

Cantrell, L., & Smith, L. B. (2013). Open questions and a proposal: A critical review of the evidence on infant numerical abilities. *Cognition*, *128*(3), 331–352. http://doi.org/10.1016/j.cognition.2013.04.008

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

Chatterji, M. (2005). Achievement gaps and correlates of early mathematics achievement: Evidence from the ECLS K-first grade sample. *Education Policy Analysis Archives*, *13*. Retrieved from https://eric.ed.gov/?id=EJ846750

Cordes, S., & Brannon, E. M. (2008). The difficulties of representing continuous extent in infancy: Using number is just easier. *Child Development*, *79*(2), 476–489. http://doi.org/10.1111/j.1467-8624.2007.01137.x

Cordes, S., & Brannon, E. M. (2009). The relative salience of discrete and continuous quantity in young infants. *Developmental Science*, *12*(3), 453–463. http://doi.org/10.1111/j.1467-7687.2008.00781.x

Cordes, S., & Brannon, E. M. (2011). Attending to one of many: When infants are surprisingly poor at discriminating an item's size. *Frontiers in Psychology*, *2*(APR), 65. http://doi.org/10.3389/fpsyg.2011.00065

Coubart, A., Izard, V., Spelke, E. S., Marie, J., & Streri, A. (2014). Dissociation between small and large numerosities in newborn infants. *Developmental Science*, *17*(1), 11–22. http://doi.org/10.1111/desc.12108

Cross, C. T., Woods, T. A., & Schweingruber, H. (2009). *Mathematics Learning in Early Childhood: Paths Toward Excellence and Equity*. *SciencesNew York*.

de Hevia, M. D., & Spelke, E. S. (2010). Number-Space Mapping in Human Infants. *Psychological Science*, *21*(5), 653–660. http://doi.org/10.1177/0956797610366091

Dehaene, S. (1997). Babies who count. In *The Number Sense: How the Mind Creates Mathematics* (pp. 41–63).

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., … Japel, C. (2007). School readiness and later achievement. *Developmental Psychology*, *43*(6), 1428–1446. http://doi.org/10.1037/0012-1649.43.6.1428

Elliott, L., Braham, E. J., & Libertus, M. E. (2017). Understanding sources of individual variability in parents’ number talk with young children. *Journal of Experimental Child Psychology*, *159*, 1–15. http://doi.org/10.1016/j.jecp.2017.01.011

Fagerlin, A., Zikmund-Fisher, B. J., Ubel, P. A., Jankovic, A., Derry, H. A., & Smith, D. M. (2007). Measuring numeracy without a math test: Development of the subjective numeracy scale. *Medical Decision Making*, *27*(5), 672–680. http://doi.org/10.1177/0272989X07304449

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, (41), 1149–1160.

Feigenson, L. (2005). A double-dissociation in infants’ representations of object arrays.

Cognition, *95*(3), B37–B48. http://doi.org/10.1016/j.cognition.2004.07.006

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*(7), 307–314. http://doi.org/10.1016/j.tics.2004.05.002

Frank, M. C., Everett, D. L., Fedorenko, E., & Gibson, E. (2008). Number as a cognitive technology: Evidence from Pirahã language and cognition. *Cognition*, *108*(3), 819–824. http://doi.org/10.1016/j.cognition.2008.04.007

Geary, D. C. (2011). Cognitive predictors of achievement growth in mathematics: A 5-year longitudinal study. *Developmental Psychology*, *47*(6), 1539–1552. http://doi.org/10.1037/a0025510

Gebuis, T., & Reynvoet, B. (2012a). The interplay between nonsymbolic number and its continuous visual properties. *Journal of Experimental Psychology: General*, *141*(4), 642–648. http://doi.org/10.1037/a0026218

Gebuis, T., & Reynvoet, B. (2012b). The role of visual information in numerosity estimation. *PLoS ONE*, *7*(5). http://doi.org/10.1371/journal.pone.0037426

Ginsburg, N., & Nicholls, A. (1988). Perceived numerosity as a function of item size. *Perceptual and Motor Skills*, *67*(2), 656–658. http://doi.org/10.2466/pms.1988.67.2.656

Goldin-Meadow, S., & Breckinridge Church, R. (1986). The mismatch between gesture and speech as an index of transitional knowledge. *Cognition*, *23*, 43–71. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.715.4570&rep=rep1&type=pdf

Gómez-Laplaza, L. M., & Gerlai, R. (2011). Can angelfish (Pterophyllum scalare) count? Discrimination between different shoal sizes follows Weber's law. *Animal Cognition*, *14*(1), 1–9. http://doi.org/10.1007/s10071-010-0337-6

Gordon, R., Chernyak, N., & Cordes, S. (n.d.). Children's spontaneous use of gesture in a numerical task.

Gunderson, E. A., & Levine, S. C. (2011). Some types of parent number talk count more than others: Relations between parents' input and children's cardinal-number knowledge. *Developmental Science*, *14*(5), 1021–1032. http://doi.org/10.1111/j.1467-7687.2011.01050.x

Halberda, J., & Feigenson, L. (2008a). Change in the acuity of the "Number Sense": The approximate number system in 3-, 4-, 5. *Developmental Psychology*. Retrieved from http://psycnet.apa.org/journals/dev/44/5/1457.html

Halberda, J., & Feigenson, L. (2008b). Developmental change in the acuity of the "Number Sense": The Approximate Number System in 3-, 4-, 5-, and 6-year-olds and adults. *Developmental Psychology*, *44*(5), 1457–65. http://doi.org/10.1037/a0012682

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

http://doi.org/10.1037/a0012682

Halberda, J., & Feigenson, L. (2008d). Set representations required for the acquisition of the “natural number” concept. *Behavioral and Brain Sciences*, *31*(6), 655–656. http://doi.org/10.1017/S0140525X08005566

Halberda, J., Mazzocco, M. M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature*, *455*(October), 665–668. http://doi.org/10.1038/nature07246

Hamamouche, K. A., Chernyak, N., & Cordes, S. (n.d.). Sharing scenarios facilitate division performance in preschoolers. *Submitted*.

Hannula-Sormunen, M. M. (2014). Spontaneous focusing on numerosity and its relation to counting and arithmetic. *The Oxford Handbook of Numerical Cognition*, *1*(September), 275–290. http://doi.org/10.1093/oxfordhb/9780199642342.013.018

Hannula-Sormunen, M. M., & Lehtinen, E. (2005). Spontaneous focusing on numerosity and mathematical skills of young children. *Learning and Instruction*, *15*(3), 237–256. http://doi.org/10.1016/j.learninstruc.2005.04.005

Hannula-Sormunen, M. M., Lehtinen, E., & Räsänen, P. (2015). Preschool Children’s Spontaneous Focusing on Numerosity, Subitizing, and Counting Skills as Predictors of Their Mathematical Performance Seven Years Later at School. *Mathematical Thinking and Learning*, *17*(2–3), 155–177. http://doi.org/10.1080/10986065.2015.1016814

Hannula-Sormunen, M. M., Lepola, J., & Lehtinen, E. (2010). Spontaneous focusing on numerosity as a domain-specific predictor of arithmetical skills. *Journal of Experimental Child Psychology*, *107*(4), 394–406. http://doi.org/10.1016/j.jecp.2010.06.004

Hannula, M. M., Rasanen, P., Lehtinen, E., Hannula-Sormunen, M. M., Rasanen, P., & Lehtinen, E. (2007). Development of Counting Skills: Role of Spontaneous Focusing on Numerosity and Subitizing-Based Enumeration. *Mathematical Thinking and Learning*, *9*(1), 51–57. http://doi.org/10.1207/s15327833mtl0901_4

Henik, A., Leibovich, T., Naparstek, S., Diesendruck, L., & Rubinsten, O. (2011). Quantities, amounts, and the numerical core system. *Frontiers in Human Neuroscience*, *5*, 186. http://doi.org/10.3389/fnhum.2011.00186

Hurewitz, F., Gelman, R., & Schnitzer, B. (2006). Sometimes area counts more than number. *Proceedings of the National Academy of Sciences*, *103*(51), 19599–19604. http://doi.org/10.1073/pnas.0609485103

Hurst, M. A., & Cordes, S. (2018). Attending to relations: Proportional reasoning in 3-to 6-year-old children. *Developmental Psychology*, *54*(3), 428–439. http://doi.org/10.1037/dev0000440

Izard, V., Pica, P., Spelke, E. S., & Dehaene, S. (2008). Exact equality and successor function: Two key concepts on the path towards understanding exact numbers. *Philosophical Psychology*, *21*(4), 491–505. http://doi.org/10.1080/09515080802285354

Izard, V., Sann, C., Spelke, E. S., & Streri, A. (2009). Newborn infants perceive abstract numbers. *Proceedings of the National Academy of Sciences*, *106*(25), 10382–10385. http://doi.org/10.1073/pnas.0812142106

Jeong, Y., Levine, S. C., & Huttenlocher, J. (2007). The Development of Proportional Reasoning: Effect of Continuous Versus Discrete Quantities. *Journal of Cognition and Development*, *8*(2), 237–256. http://doi.org/10.1080/15248370701202471

Jordan, N. C., Kaplan, D., Locuniak, M. N., & Ramineni, C. (2007). Predicting First-Grade Math Achievement from Developmental Number Sense Trajectories. *Learning Disabilities Research & Practice*, *22*(1), 36–46. http://doi.org/10.1111/j.1540-5826.2007.00229.x

Jordan, N. C., Kaplan, D., Ramineni, C., & Locuniak, M. N. (2009). Early Math Matters: Kindergarten Number Competence and Later Mathematics Outcomes. *Developmental Psychology*, *45*(3), 850–867. http://doi.org/10.1037/a0014939

Kinzler, K. D., & Spelke, E. S. (2007). Core systems in human cognition. *Progress in Brain Research*. http://doi.org/10.1016/S0079-6123(07)64014-X

Klibanoff, R. S., Levine, S. C., Huttenlocher, J., Vasilyeva, M., & Hedges, L. V. (2006). Preschool children's mathematical knowledge: The effect of teacher "math talk." *Developmental Psychology*, *42*(1), 59–69. http://doi.org/10.1037/0012-1649.42.1.59

Krusche, P., Uller, C., & Dicke, U. (2010). Quantity discrimination in salamanders. *Journal of Experimental Biology*. http://doi.org/10.1242/jeb.039297

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*(2), 395–438. http://doi.org/10.1016/j.cognition.2006.10.005

Leibovich, T., & Henik, A. (2014). Comparing performance in discrete and continuous comparison tasks. *Quarterly Journal of Experimental Psychology*, *67*(5), 899–917. http://doi.org/10.1080/17470218.2013.837940

Leibovich, T., Henik, A., & Salti, M. (2015). Numerosity processing is context driven even in the subitizing range: An fMRI study. *Neuropsychologia*, *77*, 137–147. http://doi.org/10.1016/j.neuropsychologia.2015.08.016

Leibovich, T., Katzin, N., Harel, M., & Henik, A. (2017). From ’sense of number’to ’sense of magnitude’–The role of continuous magnitudes in numerical cognition. *Behavioral and Brain*, *40*. http://doi.org/10.1017/S0140525X16000960

Libertus, K. (2008). Preferential Looking Coder.

Libertus, M. E., & Brannon, E. M. (2010). Stable individual differences in number discrimination in infancy. *Developmental Science*, *13*(6), 900–6. http://doi.org/10.1111/j.1467-7687.2009.00948.x

Libertus, M. E., Feigenson, L., & Halberda, J. (2011). Preschool acuity of the approximate number system correlates with school math ability. *Developmental Science*, *14*(6), 1292–1300. http://doi.org/10.1111/j.1467-7687.2011.01080.x

Libertus, M. E., Starr, A., & Brannon, E. M. (2014). Number trumps area for 7-month-old infants. *Developmental Psychology*, *50*(1), 108–112. http://doi.org/10.1037/a0032986

Lidz, J., Pietroski, P., Halberda, J., & Hunter, T. (2011). Interface transparency and the psychosemantics of most. *Natural Language Semantics*, *19*(3), 227–256. http://doi.org/10.1007/s11050-010-9062-6

Lipton, J. S., & Spelke, E. S. (2003). Origins of Number Sense: Large-number discrimination in Human Infants. *Psychological Science*, *14*(5), 396–401. http://doi.org/10.1111/1467-9280.01453

Lipton, J. S., & Spelke, E. S. (2004). Discrimination of Large and Small Numerosities by Human Infants. *Infancy*, *5*(3), 271–290. http://doi.org/10.1207/s15327078in0503_2

Lucon-Xiccato, T., Miletto Petrazzini, M. E., Agrillo, C., & Bisazza, A. (2015). Guppies discriminate between two quantities of food items but prioritize item size over total amount. *Animal Behaviour*, *107*, 183–191. http://doi.org/10.1016/j.anbehav.2015.06.019

Lyon, B. E. (2003). Egg recognition and counting reduce costs of avian conspecific brood parasitism. *Nature*, *422*(6931), 495–499. http://doi.org/10.1038/nature01505

McComb, K., Packer, C., & Pusey, A. (1994). Roaring and numerical assessment in contests between groups of female lions, Panthera leo. *Animal Behaviour*, *47*(2), 379–387. http://doi.org/10.1006/anbe.1994.1052

McMullen, J., Hannula-Sormunen, M. M., & Lehtinen, E. (2015). Preschool spontaneous focusing on numerosity predicts rational number conceptual knowledge 6 years later. *ZDM - Mathematics Education*, *47*(5), 813–824. http://doi.org/10.1007/s11858-015-0669-4

Meck, W. H., & Church, R. M. (1983). A mode control model of counting and timing processes. *Journal of Experimental Psychology. Animal Behavior Processes*, *9*(3), 320–334. http://doi.org/10.1037/0097-7403.9.3.320

Mix, K. S., Huttenlocher, J., & Levine, S. C. (2002). Math without words: Quantitative development in infancy and early childhood. New York: Oxford University Press.

Mix, K. S., Levine, S. C., & Huttenlocher, J. (1997). Numerical abstraction in infants: another look. *Developmental Psychology*, *33*(3), 423–428. http://doi.org/10.1037/0012-1649.33.3.423

Moyer, R. S., & Bayer, R. H. (1976). Mental comparison and the symbolic distance effect. *Cognitive Psychology*, *8*(2), 228–246. http://doi.org/10.1016/0010-0285(76)90025-6

Odic, D. (2018). Children's intuitive sense of number develops independently of their perception of area, density, length, and time. *Developmental Science*, *21*(2), e12533. http://doi.org/10.1111/desc.12533

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

Odic, D., Libertus, M. E., Feigenson, L., & Halberda, J. (2013). Developmental change in the acuity of approximate number and area representations. *Developmental Psychology*, *49*(6), 1103–1112. http://doi.org/10.1037/a0029472

Piaget, J. (1952). *The child's conception of number*. Routledge & Paul. Retrieved from https://books.google.com/books/about/The_Child_s_Conception_of_Number.html?id=1ET995R5VHEC

Piaget, J. (1977). *The essential Piaget*. (H. E. Gruber & J. J. Vonèche, Eds.). New York: Basic Books.

Posid, T., & Cordes, S. (2018). How high can you count? Probing the limits of children's counting. *Developmental Psychology*, *54*(5), 875–889. http://doi.org/10.1037/dev0000469

Practices, N. G. A. C. for B., & Council of Chief State School Officers. (2010). *Common Core State Standards for mathematics: Kindergarten introduction.*

Rathé, S., Torbeyns, J., Hannula-Sormunen, M. M., & Verschaffel, L. (2016). Kindergartners' Spontaneous Focusing on Numerosity in Relation to Their Number-Related Utterances During Numerical Picture Book Reading. *Mathematical Thinking and Learning*, *18*(2), 125–141. http://doi.org/10.1080/10986065.2016.1148531

Romano, E., Babchishin, L., Pagani, L. S., & Kohen, D. (2010). School readiness and later achievement: Replication and extension using a nationwide Canadian Survey. *Developmental Psychology*, *46*(5), 995–1007. http://doi.org/10.1037/a0018880

Rowe, M. L., Levine, S. C., Suriyakham, L. W., Gunderson, E. A., & Huttenlocher, J. (2010). What counts in the development of young children's number knowledge? *Developmental Psychology*, *46*(5), 1309–1319. http://doi.org/10.1037/a0019671

Savelkouls, S., & Cordes, S. (2017). Numerical intuitions in infancy: Give credit where credit is due. *Behavioral and Brain Sciences*, *40*, e189. http://doi.org/10.1017/S0140525X16002272

Spelke, E. S. (2000). Core knowledge. *The American Psychologist*, *55*(11), 1233–43. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11280937

Spelke, E. S. (2003). What Makes Us Smart? Core Knowledge and Natural Language. *Language in Mind: Advances in the Study of Language and Thought*. Retrieved from https://philpapers.org/rec/SPEWMU

Spelke, E. S. (2017). Core Knowledge, Language, and Number. *Language Learning and Development*, *13*(2), 147–170. http://doi.org/10.1080/15475441.2016.1263572

Starr, A., & Brannon, E. M. (2015). Evidence against continuous variables driving numerical discrimination in infancy. *Frontiers in Psychology*, *6*, 923. http://doi.org/10.3389/fpsyg.2015.00923

Starr, A., Libertus, M. E., & Brannon, E. M. (2013a). Infants show ratio-dependent number discrimination regardless of set size. *Infancy*, *18*(6), 927–941. http://doi.org/10.1111/infa.12008

Starr, A., Libertus, M. E., & Brannon, E. M. (2013b). Number sense in infancy predicts mathematical abilities in childhood. *Proceedings of the National Academy of Sciences*, *110*(45), 18116–18120. http://doi.org/10.1073/pnas.1302751110

Tokita, M., & Ishiguchi, A. (2010). How might the discrepancy in the effects of perceptual variables on numerosity judgement be reconciled? *Attention, Perception, and Psychophysics*, *72*(7), 1839–1853. http://doi.org/10.3758/APP.72.7.1839

Tokita, M., & Ishiguchi, A. (2013). Effects of perceptual variables on numerosity comparison in 5-6-year-olds and adults. *Frontiers in Psychology*, *4*(JUL), 431. http://doi.org/10.3389/fpsyg.2013.00431

Uchikawa, K., & Shinoda, H. (1996). Influence of basic color categories on color memory discrimination. *Color Research and Application*, *21*(6), 430–439. http://doi.org/10.1002/(SICI)1520-6378(199612)21:6<430::AID-COL5>3.0.CO;2-X

Ward, C., & Smuts, B. B. (2007). Quantity-based judgements in the domestic dog (Canis lupus familiaris). *Animal Cognition*, *10*(1), 71–80. http://doi.org/10.1007/s10071-006-0042-7

Wilson, M. L., Hauser, M. D., & Wrangham, R. W. (2001). Does participation in intergroup conflict depend on numerical assessment, range location, or rank for wild chimpanzees? *Animal Behaviour*, *61*(6), 1203–1216. http://doi.org/10.1006/anbe.2000.1706

Winawer, J., Witthoft, N., Frank, M. C., Wu, L., Wade, A. R., & Boroditsky, L. (2007). Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Sciences*, *104*(19), 7780–7785.

http://doi.org/10.1073/pnas.0701644104

Wood, J. N., & Spelke, E. S. (2005). Infants' enumeration of actions: Numerical discrimination and its signature limits. *Developmental Science*, *8*(2), 173–181. http://doi.org/10.1111/j.1467-7687.2005.00404.x

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III Tests of Achievement*. *Test*.

Wynn, K. (1992). Children's acquisition of the number words and the counting system. *Cognitive Psychology*, *24*(2), 220–251. http://doi.org/10.1016/0010-0285(92)90008-P

Wynn, K., & VanMarle, K. (2006). Six-month-old infants use analog magnitudes to represent duration. *Developmental Science*, *9*(5). http://doi.org/10.1111/j.1467-7687.2006.00508.x

Xu, F., & Arriaga, R. I. (2007). Number discrimination in 10-month-old infants. *British Journal of Developmental Psychology*, *25*(1), 103–108. http://doi.org/10.1348/026151005X90704

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition*, *74*(1), 1–11. http://doi.org/10.1016/S0010-0277(99)00066-9

Xu, F., Spelke, E. S., & Goddard, S. (2005). Number sense in human infants. *Developmental Science*, *8*, 88–101. http://doi.org/10.1111/j.1467-7687.2005.00395.x