EXPLORING THE INTERFACE BETWEEN NUMBER WORDS AND PERCEPTUAL MAGNITUDES

by

Denitza P. Dramkin

B.Sc. (Hons.), The University of Houston, 2015

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Psychology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

August 2018

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, a thesis/dissertation entitled:

Exploring the Interface Between Number Words and Perceptual Magnitudes

submitted by Denitza P. Dramkin in partial fulfilment of the requirements for

the degree of Master of Arts

in Psychology

Examining Committee:

Darko Odic, Psychology
Supervisor

D. Geoffrey Hall, Psychology
Supervisory Committee Member

Janet F. Werker, Psychology
Supervisory Committee Member

Additional Examiner

Additional Supervisory Committee Members:

Supervisory Committee Member

Supervisory Committee Member

Abstract

As humans, we reason about quantity in at least two distinct ways—through our intuitive, approximate perception of quantity and through precise number words. With sufficient development, these two systems interface and interact, allowing us to make quick judgments with crude precision (e.g., how many items are in our shopping basket). To date, two theories have been proposed to explain the underlying mechanism of this interface between perception and language. Under the first—the *associative mapping* theory—children create item-specific associations between particular number words (e.g., "ten") and the perceptual representations that they most frequently experience. While under the second—the *structure mapping* theory—children map number words to their perceptual representations by realizing the inherent similarity in the representational structure of the two systems (e.g., both are linear dimensions where higher values represent *more/greater* amounts). Existing literature has almost exclusively focused on understanding how children create this interface in one domain of quantity (i.e., number), leaving the critical question of how children map number words to other, non-numeric domains of quantity (e.g., length, area) entirely open. This thesis explores when and how children map number words to a broader spectrum of quantities by examining their estimation abilities in number, length, and area. We find that while the perception of number, length, and area are largely independent of each other, estimation accuracy and variability are tightly linked and show a similar age of maturity, supporting the structure mapping account. These results are discussed in the broader context of how language and perception interact and change with development.

Lay Summary

As humans, we represent and reason about quantity in at least two distinct ways—through our intuitive but imprecise "sense" of quantity and through precise number words. Eventually, we learn to integrate our intuitive representations with number words. To date, two theories have been put forth to explain how this is achieved, yet the majority of the research has only explored the interface in only one domain of quantity (i.e., number), leaving open the question of how we learn to map number words to other domains (e.g., length, area). This thesis explores how and when children learn to integrate their intuitive perceptions of quantity with number words across a number, length, and area estimation task. We discuss the implications of our findings relevant to current theories in the field, including how we learn to reason about our rich perceptual world.

Preface

This thesis is an original, unpublished intellectual product of the author, D. Dramkin. The research was conducted by the author at the University of British Columbia, Centre for Cognitive Development, under the supervision of D. Odic, who was heavily involved in the research design (including the custom-made Psychtoolbox-3 scripts), data analyses, and writing process. All work and associated methods were approved by the University of British Columbia's Research Ethics Board (H14-01984: Development of Language and Quantity).

Table of Contents

Abstract ..... **iii**

Lay Summary ..... **iv**

Preface ..... **v**

Table of Contents ..... **vi**

List of Tables ..... **vii**

List of Figures ..... **viii**

Acknowledgements ..... **ix**

Dedication ..... **x**

Introduction **1**

Methods and Procedures **6**

Participants. *6*

Discrimination Task. *7*

Number Word Task. *9*

Estimation Task. *9*

Results **13**

Number Word Task. *13*

Descriptives. *13*

Age of Maturity. *18*

Correlations Within and Across Tasks. *22*

General Discussion **24**

References **30**

List of Tables

Table 1. Distribution of participants across age groups.. 14

Table 2. Discrimination performance across age... 15

Table 3. Estimation performance across age. 18

List of Figures

Figure 1. Discrimination task stimuli. 8

Figure 2. Estimation task stimuli. 10

Figure 3. Discrimination performance modeled across development and dimension. 20

Figure 4. Estimation performance modeled across development and dimension. 21

Acknowledgements

First and foremost, thank you to Darko Odic for your illuminating insight, patience, kindness, and support. I am eternally grateful to have you as my supervisor and mentor. Thank you for giving me the opportunity to learn from you. You make me a better scholar and scientist.

Thank you to my committee members, Janet Werker and Geoff Hall, for your mentorship and collaboration. Your work inspired me to apply to UBC in the first place. Thank you for enriching me with your presence and feedback.

To Carolyn Baer and Cory Bonn, thank you for being such *koala*-ty lab mates and friends. Thank you letting me learn from you and inspiring me to do better.

To my cohort and The Study Group—Miranda Sitch, Drake Levere, Kyle Gooderham, Sumeet Mutti, Gabriel Smith, and Yvette Graveline—thank you for helping me survive the perils of graduate school. Thank you for the stimulating intellectual conversations and for the silly adventures. Miranda, I am especially grateful that I have you to share this journey with.

Thank you to the research assistants that dedicated hours of their time to helping further this project—Paris Gappmayr, Michelle Indarjit, and Hillary Smith. Your hard work did not go unnoticed or unappreciated.

To Crystal D. Tran and Audrey Utti, thank you for believing in me and encouraging me to believe in myself. Were it not for you, I would not have pursued this passion of mine. Thank you for inspiring me and for your continued mentorship and friendship.

Thank you to the participants and their families for taking the time to help further this research, and to the Department of Psychology for their resources and financial support.

To the countless friends who offered words of encouragement, advice, and help: thank you. I am forever grateful for your support.

Thank you to Derek James for providing endless hugs, tea lattes, and emotional support.

To Asen Angelov, thank you for inspiring me with your worth ethic. I am proud to call you my older brother.

To my family—my parents especially—thank you for giving me the freedom to pursue my passion and motivation to continue. Mom, I am especially grateful for your wisdom; I am all that I am because of you. обичам те.

Dedication

Everything that I do is for my bubble of joy, my light, my little brother—

Clinton Alexander Miles.

Introduction

As adults, we represent and reason about quantity in at least two ways. The first is through utilizing our intuitive, automatic—albeit imprecise—perceptual representations of quantity. For instance, without counting or measuring, we can instantly judge which set of dots is more numerous, which line is longer, or which blob is bigger (see Figure 1). These intuitive quantity representations also permit a variety of mathematical operations to be carried over them, including addition, subtraction, and division (Dehaene, 2001, 2009, 2011; Bonny & Lourenco, 2013), providing us with quick and effortless ways to think and reason about quantity. Moreover, we share these perceptual representations with most—if not all—non-human animals (Howard, Avarguès-Weber, Garcia, Greentree, & Dyer, 2018; Lucon-Xiccato, Gatto, & Bisazza, 2018; Jordan & Brannon, 2006; Piffer, Petrazzini, & Agrillo, 2013; Platt & Johnson, 1971; Premack & Woodruff, 1978) and have them readily available from birth onward (Izard, Sann, Spelke, & Streri, 2009; Dehaene, 2001; Feigenson, Dehaene, & Spelke, 2004). The cost of such a readily available system is its imprecision: how well we discriminate between two quantities is dependent on their ratio, a constraint known as Weber's law (see Odic & Starr, 2018; Cantlon, 2018). As a result, our ability to judge quantity is limited by perceptual noise, or imprecision, and the closer two quantities appear to be, the harder it is to tell them apart.

The second way of representing and reasoning about quantity is through number words, most often through a slowly learned counting routine (Sarnecka & Carey, 2008; Wynn, 1992; Benoit, Lehalle, & Jouen, 2004; Wiese, 2007; Le Corre & Carey, 2007). Number words allow us to represent quantity precisely and discretely. Thus, a key advantage of reasoning about quantity in this way is that it is ratio-independent: while a plate of 100 cookies is perceptually near-indistinguishable from a plate of 101, verbally or symbolically distinguishing between these two amounts (e.g., “one-hundred” vs. “one-hundred-and-one”) becomes as easy as visually distinguishing between much larger differences (e.g., ratio of 2.0, or 200 vs. 100 cookies). The cost to using such a precise system, however, is the difficulty in mastering it: no non-human animal has access to precise number representations (Beran, Rumbaugh, & Savage-Rumbaugh, 1998; Boysen & Capaldi, 2014; Tomonaga, 2008), children in Western cultures take a very long time to master the counting routine (Le Corre & Carey, 2007), and some human cultures that lack number words, such as the Amazonian Pirahã tribe, apparently lack access to these representations altogether (Frank, Everett, Fedorenko, & Gibson, 2008; Holden, 2004; Pica, Lemer, Izard, & Dehaene, 2004).

Although children’s intuitive sense of quantity and the acquisition of exact number words have often been studied in isolation, these two systems also interface and interact starting from around age five onward, allowing children and adults to fluidly convert their intuitive representations into exact ones and vice-versa (Le Corre & Carey, 2007; Odic, Le Corre, & Halberda, 2015). For example, a typical preschooler shown a display of dots on a screen for just a few seconds—too quick to count—can attach a precise number word to their perception of number (e.g., estimating that there are “ten” dots on the screen), and if given a precise number word (e.g., “eight”), young children can in turn produce an approximate representation of that amount by tapping on the table that number of times (Odic et al., 2015). The interface between these two systems—the perceptual and the linguistic—has been theorized to be critical in children’s higher-level concepts of basic mathematical operations (Libertus, Feigenson, & Halberda, 2013; Libertus, Odic, & Halberda, 2012; Bonny & Lourenco, 2013; Libertus, Feigenson, & Halberda, 2011), fractions (Schneider & Siegler, 2010; Siegler, Fazio, Bailey, & Zhou, 2013; Siegler, Thompson, & Schneider, 2011), and currency (Marques & Dehaene, 2004).

In this thesis, I explore the mechanisms that support the interface between children's intuitive representations of number, length, and area and exact number words. While the existing literature has extensively focused on understanding how children map number words to their intuitive (or "gut") sense of number—the Approximate Number System (ANS)—it has left the question of how children map number words to other, non-numeric domains of quantity entirely open. Yet, as I discuss below, understanding when and how children map number words to a broader spectrum of quantity representations is key to our understanding of development not only because we, as adults, have access to precise categories of length, size, etc. (e.g., inches, squared meters, etc.), but also because any similarities and differences in the acquisition of this interface across dimensions can inform theories of how children learn to map perceptual representations to language and vice-versa in the first place.

At present, two mechanisms have been put forth to explain how children acquire the interface between number words and the ANS (Sullivan & Barner, 2010, 2011, 2013). Under the *associative mapping* theory, children create item-specific associations between particular number words (e.g., "ten") and the perceptual representations that they most frequently experience. As a result, given that we all experience smaller numbers more often than larger numbers (Piantadosi, 2016), we should expect that children acquire the mapping between words like "one", "two", and "three" well before acquiring the mapping for words like "eighty-six". Consistent with this theory, children learn number words in a slow, graded fashion, first associating "one" with exactly *1* item, then "two" with exactly *2* items, then "three" with *3* items, etc. until they eventually understand the principal of cardinality (Le Corre & Carey, 2007). Furthermore, work by Sullivan and Barner (2010) has shown that adults are unwilling to adjust their estimation behaviour in response to feedback or outright experimenter instructions, if the range of numbers being viewed/estimated is below 20, which is exactly in the range where we would expect associative mapping to form.

Yet, while the associative mapping theory accounts for how children first map number words to small quantities, it does not explain how children eventually learn to reason about larger quantities (Sullivan & Barner, 2011). The alternative—the *structure mapping* theory—posits that children map number words to the ANS by realizing the inherent similarity in the representational structure of the two systems: both are linear dimensions where higher values represent *more/greater* amounts. Thus, rather than creating item-by-item mappings, children may instead generate links between number words and the ANS based on analogy, proportional reasoning, or an understanding of ordinality between the two systems (Sullivan & Barner, 2010, 2013; Barth, Starr, & Sullivan, 2009). In support of this account, studies have shown that participants' entire range of verbal estimates can be impacted by mislabeling visually presented stimuli (e.g., calling a set of 30 items, "twenty-five") or misinforming participants about the highest numerosity shown (e.g., being told that there will be at most "seventy-five" dots despite there actually being 350; Sullivan & Barner, 2010). However, this account still fails to explain why the initial process of matching number words to the ANS occurs in such a graded fashion, requiring item-by-item pairings (Le Corre & Carey, 2007; Sarnecka & Carey, 2008).

To gain further insights into the mechanisms that support the relationship between perceptual quantity representations and precise number words, we can compare and contrast how children form this mapping for dimensions beyond just number, including length and area. If the interface is formed through associative mapping, then we should expect that children's acquisition of the interface in one dimension (e.g., number) has no bearing on the acquisition of another dimension (e.g., length), as children's associative experiences between these dimensions should be independent. In other words, under this account, we should find that the interface between number words and our perceptual representations of number, length, and area emerge separately and independently of each other.

If, on the other hand, children acquire the interface through structure mapping, then the very moment that a child has realized the common ordinal structure between number words and the ANS, they should be able to extend it to other, non-numeric dimensions, as well (e.g., length, area). Therefore, under this account, we predict that children should be able to map between number words and their intuitive representations of number, length, and area at roughly the same time and using a shared cognitive mechanism.

Previous work has shown that although our intuitive sense of number, length, and area share the same psychophysical signatures (e.g., ratio-dependence, consistent with Weber's law), children show independent precision and development across the three dimensions (Odic, 2018; Odic, Libertus, Feigenson, & Halberda, 2013). For example, while both number and area acuity steadily improve throughout childhood, children exhibit adult-like precision in area discrimination much earlier than do they do in number (Odic, 2018; Odic et al., 2013). Moreover, how well children and adults distinguish between quantities in one domain (e.g., number) has not been shown to relate to their abilities to distinguish quantity in another (e.g., length, area, time, density; Odic, 2018; Odic et al., 2013; Cordes & Brannon, 2011, 2008), leading some researchers to propose that dimensions like number, length, and area are represented by domain-specific perceptual abilities (Burr & Ross, 2008; Ross & Burr, 2010); for alternative accounts see Dakin, Tibber, Greenwood, and Morgan (2011), Gebuis and Reynvoet (2012), Gebuis and Van Der Smagt (2011), or Leibovich, Katzin, Harel, and Henik (2017).

In the present experiment, we test 5- to 12-year-olds and adults on two tasks—a number/length/area *discrimination* task, and a number/length/area *estimation* task—to measure individual and developmental differences in the precision of each participant's perceptual representations of quantity (i.e., discrimination) and the quality of the interface between these perceptual representations and precise number words (i.e., estimation). In turn, we examine whether children's developing interface between number words and the representations of number, length, and area emerges independently and in piecemeal, as predicted by the associative mapping theory, or together and as a single shared ability, as predicted by the structure mapping theory. In addition, to make sure that any commonalities in the interface between number words and the perceptual representations of number, length, and area are emerging primarily because of the interface itself, we also control individual and developmental differences in these perceptual representations themselves through the discrimination task.

Methods and Procedures

Participants.

A total of 90 children (47 males) between the ages of 5- to 12-years-old (M = 8;9) participated in the study, ranging from 5;1 (5 years, 1 month) to 12;11 years (12 years, 11 months). We chose this age range, as some work suggests that children begin establishing the interface between their intuitive representation of quantity and number words from age five onwards (Le Corre & Carey, 2007; Barth et al., 2009). An additional 14 children were tested, but failed to complete both tasks in full due to unwillingness to continue.

Children were recruited from the Greater Metro-Vancouver area in British Columbia, Canada through the Early Development Research Group database at the University of British

Columbia (UBC). All participants were individually tested at the Centre for Cognitive Development at UBC in a quiet room on a 13” MacBook Air running custom-made Psychtoolbox-3 scripts. Consent was obtained from the parent or legal guardian present during the time of the study, and experimenters received verbal assent before each child’s participation. Following the study, participants were rewarded with a small prize (e.g., toy, t-shirt, or book) and a certificate of appreciation.

In order to generate more predictive models of the age of maturity and development across domains, 14 college-age adults (2 males; ranging from 19;6 – 34;6 years, *M* = 24;1) were also tested on identical Discrimination and Estimation tasks and were rewarded with university course credit and/or a certificate of appreciation for their time. An additional adult was recruited, but eliminated from analysis for failure to complete the Estimation Task, due to technical issues. This left us with a final sample of 104 participants (5- to 12-year-olds and adults).

Discrimination Task.

Participants were first presented with three dimensions: number (i.e., “which side has more dots”), length (i.e., “which line is longer?”), and area (i.e., “which blob is bigger?”). The number stimuli were spatially separated collections of yellow dots on the left and blue dots on the right side of the screen; the line stimuli were two randomly oriented lines, with a yellow line on the left and a blue line on the right side on the screen; the area stimuli were two amorphous blobs, with a yellow blob on the left and a blue blob on the right side on the screen (Figure 1). Each side had a colour-matched cartoon character (Spongebob for yellow and a Smurf for blue), and participants were asked to indicate their answer verbally or by pointing. All responses were entered in on the computer (via pressing F for yellow/Spongebob or J for blue/Smuf) by the experimenter to minimize the influence of motor control on the results. Adult participants were allowed to press the computer keys themselves during this task.

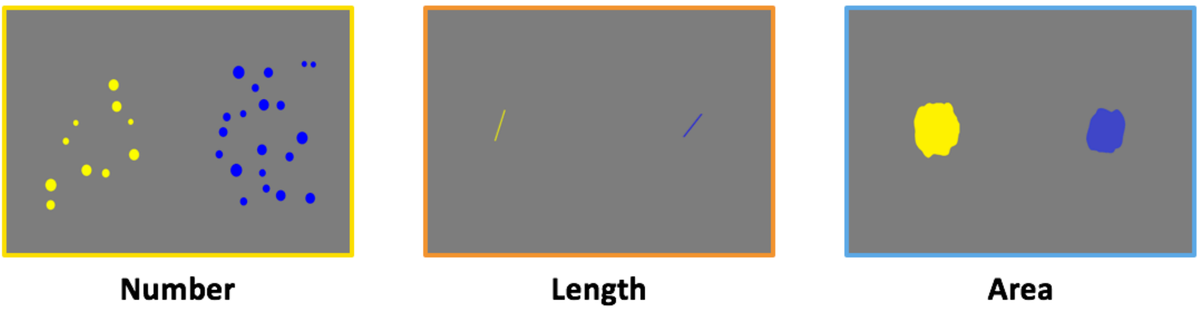

Figure 1. Discrimination task stimuli (without the cartoon characters), in which participants had to judge which side was more numerous (e.g., which side has more dots, which line is longer, which blob is bigger).

The stimuli were shown on the screen for only 500 milliseconds (ms) to prevent the participants from counting (Cordes, Gallistel, Gelman, & Latham, 2007). The task began with 6 easy practice trials—2 in each dimension—to ensure that participants understood the task. Subsequently, participants completed 192 trials (64 per dimension) in an intermixed order, thus eliminating any potential order effects. To alter difficulty, each trial varied in one of 4 ratios which were identical across the three dimensions: 2.0 (e.g., 20 vs. 10 dots), 1.50 (e.g., a 12 vs. an 18 cm line), 1.20 (e.g., a 120 mm^2 vs. a 100 mm^2 blob), and 1.07. Participants were given feedback based on their performance that was either positive (i.e., a female computer voice saying, “Good job!” or “Great!”) or negative (i.e., “Oh no, that’s not right!”), and the experimenter would encourage the participant to continue. Children and adults took between 10 – 12 minutes to complete the task.

For each of the three dimensions—number, length, area—the dependent variable was accuracy (i.e., percent correct across all trials) and the Weber fraction (*w*), which roughly indexes the hardest ratio an individual can discriminate at around 75% accuracy (i.e., a higher *w* corresponds to greater imprecision in number representations and poorer discrimination performance; Halberda & Feigenson, 2008; Lidz, Pietroski, Halberda, & Hunter, 2011; Odic, 2018; Pica et al., 2004). As discussed in detail in Results, we used a standard and frequently used psychophysical model to estimate each participant's *w* separately from their guessing rate (*g*).

Number Word Task.

Because children can dramatically vary in their knowledge of the verbal count sequence, children were given a short Number Word Task following completion of the Discrimination Task and before the start of the Estimation Task, in which they were asked to count to at least 25, a target that previous standardized testing shows should be expected for 5-year-old children (Ginsburg & Baroody, 2003). Experimenters recorded whether each child reached the target number during counting or the highest number counted to without error, such as skipping a number or double counting. Additionally, children were asked to determine which of two numbers was greater (i.e., "12 vs. 15," "18 vs. 12," "20 vs. 10") to ensure that they had a basic understanding of cardinality. Responses were recorded for participants' failure or success on the latter task, which served as a control to ensure that they understood number words and cardinality, necessary for estimation.

Estimation Task.

In the Estimation Task, participants were asked to verbally assign a number word to the amount of stimuli they saw across number, length, and area trials. In order to control for experience that they may have had with particular units (e.g., cm, mm^2, etc.), which could influence their estimation performance, all participants were provided with novel units. These novel units were introduced one at a time, across 6 training phases, with 2 consecutive training phases per dimension. For length, participants were introduced to a "blicket," which appeared on the screen as a single orange line segment (Figure 2). They were then told a story about how the "*blickets* live in a house," which subsequently appeared on screen, and were then asked to estimate "how many *blickets*" they thought were in that house. In order to demonstrate how the units worked, once participants gave an estimate, they were shown the correct number of *blickets*, which was always either 2, 3, or 4. Following the two training trials for blickets, participants were shown 2 training trials for a unit in area estimation (i.e., a blue amorphous blob called a *modi*), and then 2 training trails for units for number estimation (i.e., a yellow dot called a *toma*), which followed the same pattern of training noted earlier (Figure 2).

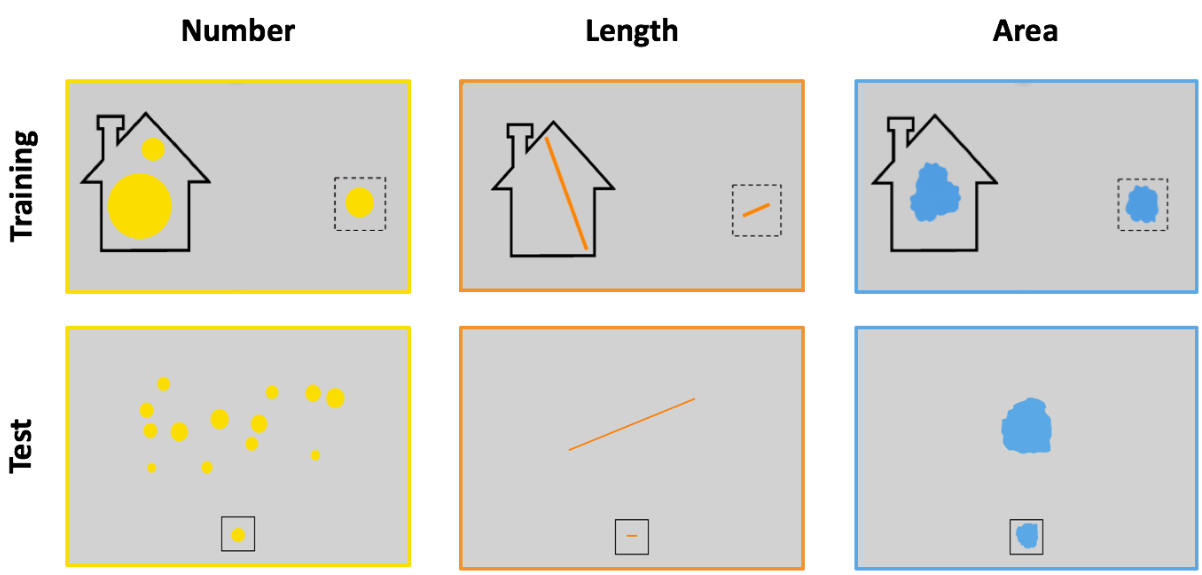

Figure 2. Estimation task stimuli for training and testing, in which participants had to judge how many items were displayed (e.g., how many *tomas*, how many *blickets*, how many *modies*). The top panels illustrate the training stimuli during which children were acquainted with the units in the context of objects living in a house. The bottom shows the actual testing stimuli.

Subsequently, participants completed 96 test trials (32 per dimension) in an intermixed order. On each trial, they would see the unit—either a blue blob (i.e., *modi*), an orange line (i.e., *blicket*), or a yellow dot (i.e., *toma*)—and, when ready, they would see a larger amorphous blob, a larger orange line, or a collection of yellow dots rapidly appear and disappear. The stimuli would stay on the screen for 500 ms, too quick to count (Cordes et al., 2007). Each participant would then be asked to estimate, "how many *modies/blickets/tomas*" they saw, with the use of the quantifier (i.e., how *many*) and plural syntax (e.g., mod*ies*) guiding them towards providing us with a number word. The experimenter would then write down the verbal estimate provided on a response sheet.

The correct answer on each trial was always either 5 (e.g., a line five times the length of the standard), 8 (e.g., 8 dots), 13 (e.g., a blob thirteen times larger than the standard), or 21 across each dimension. Participants were not provided with any direct feedback during the test trials, but were encouraged to continue (e.g., "Alright, let's do another one!", "You're doing well! Let's do some more."). If they gave nonsense numbers, such as a combination of two or more numbers (e.g., "eleventy-four") or numbers that exceeded the realm of possibility (e.g., 1 million/billion/trillion), participants were prompted to provide their "next best guess" or asked if they were "sure." Additionally, if participants responded with "one" or "a big one," they were reminded of the rules of the game (e.g., "Remember, how the *blickets* cuddled together in the house? How many *blickets* do you think were cuddling there?"), and then asked to provide their "next best guess." For the participants who declined to alter their responses, their estimates were recorded on the response sheet as is, and the experimenter proceeded to the next trial. Children and adults took 15 – 20 minutes to complete this task.

Consistent with prior research, the dependent variable was—for each dimension— (1) the estimation accuracy and (2) estimation variability (see Odic et al., 2015). For each participant, we plotted their verbal estimates against the true/objective stimulus; in ideal circumstances, an individual with a perfect mapping between number words and their intuitive sense of quantity should (on average) always say “eight” for the collection of 8 dots, or an 8 cm line, etc. As a result, the linear slope—or beta in a standard linear regression—between their responses and the objective quantity is the index of how accurate one’s mapping is, with a slope of 1.0 indicating perfect mapping, a slope of 0 indicating no mapping at all, and a slope between those two values indicating some degree of under-estimation. To better evaluate the degree of over- and under-estimation from a perfect mapping (i.e., a slope of 1.0), we took the absolute differences in the beta values from 1.0 for each dimension, yielding our measure of *estimation accuracy* (i.e., how closely what participants reported across various quantities matched the objective quantities shown).

In addition, we calculated each participant’s *estimation variability*, corresponding to the sigma value in a standard linear regression (i.e., the average standard deviation of their estimates divided by the objective magnitude presented, with values closer to 0 indicating more and more precise mapping). Examining the variability of the estimates, thus, allowed to us to interpret how participants’ estimates were distributed around the objective quantities.

By using both measures—estimation accuracy (i.e., difference in slope from 1.0) and estimation variability (i.e., standard deviation divided by the true quantity shown)—we indexed how well each participant estimated across number, length, and area. Relying on both measures was key, as in theory participants can demonstrate similar performance when examining estimation accuracy, but differ dramatically in their estimation variability. Consider a child who, when shown 15 dots over multiple trials, estimates “fourteen”, “fifteen”, and “sixteen”, against a child who estimates “ten”, “fifteen”, and “twenty-five”; while both these children have an average estimate of 15, the latter child has much worse precision than the former. Thus, we rely on both measures to more robustly index estimation abilities across dimensions (as we show in the Results, these two measures were additionally not correlated, further supporting the notion that they index two separate aspects of performance).

Results

Number Word Task.

Only 4 (M = 6;7) children failed to answer some portion of the Number Word Task correctly, with one child failing to count correctly beyond 21 and three children answering one of the three number comparison questions incorrectly. Additionally, one child was not given the Number Word Task due to experimenter error. Because of overall excellent performance we saw no reason to exclude any participants from the Estimation Task analyses based on these results.

Descriptives.

Next, we examine overall performance in each of the three dimensions across Discrimination and Estimation tasks to explore participants' intuitive perceptual representations and, more importantly, the interface between these representations and precise number words. To analyze this data, we grouped participants into the nearest age-group (Table 1), forming 8 groups for children ranging from 5- to 12-years of age; adult participants were all categorized as being 24-years-old (i.e., the mean group age), though for the purposes of these analyses, the binning method is irrelevant as we are treating age categorically.

Age Group	n	Mean Age	SD of Age	Min Age	Max Age
5	10	5.53	0.26	5.15	5.90
6	10	6.53	0.33	6.10	6.90
7	10	7.51	0.24	7.13	7.86
8	15	8.44	0.24	8.01	8.96
9	21	9.51	0.29	9.01	9.92
10	10	10.47	0.28	10.13	10.94
11	10	11.33	0.34	11.01	11.95
12	4	12.57	0.40	12.18	12.95
24	14	23.80	4.37	19.56	34.54

Table 1. Distribution of participants across age groups.

In Discrimination, a 3 (Dimension: Number, Length, Area) x 4 (Ratio: 2.0, 1.5, 1.2, 1.07) x 9 (Age group: 5, 6, 7, 8, 9, 10, 11, 12, 24) Greenhouse-Geisser corrected Mixed-Measures ANOVA with accuracy as the dependent variable (DV) showed a main effect of Dimension ($F(1.66, 158.12) = 78.74$; $p < .001$, $\eta_p^2 = .45$), Ratio ($F(2.28, 216.63) = 409.77$ $p < .001$, $\eta_p^2 = .81$), Age ($F(8, 95) = 9.57$; $p < .001$, $\eta_p^2 = .45$), and a significant Dimension x Ratio interaction ($F(4.23, 401.94) = 13.73$; $p < .001$, $\eta_p^2 = .13$), but no significant Dimension x Age interaction ($F(13.32, 187.60) = 1.03$; $p = .42$ $\eta_p^2 = .08$), nor a Ratio x Age interaction ($F(18.24, 216.63) = 1.19$; $p = .27$, $\eta_p^2 = .09$), nor a Dimension x Ratio x Age interaction ($F(33.85, 401.94) = 1.03$; $p = .41$, $\eta_p^2 = .08$). As can be seen in Table 2, these significant effects broadly replicate previous work in the literature (e.g., Odic, 2018): all three dimensions show strong ratio effects (consistent with Weber's law), participants performed the worst on the Number discrimination trials and best on the Area and Length trials, and performance on all three dimensions improved with age (see also Figure 3). We also found significant correlations between each of the three tasks and Age, even when we treated it continuously.

Discrimination Task										
	Dimension	**Mean**	**SD**	**Correlations with Age**	**Growth Modelling**					
					a	***k***	***m***	***b***	**Age of Maturity**
Accuracy	*Number*	78.82	9.28	.38**	50	85.02	4.05	0.35	16.15
	Length	89.57	6.72	.37**	50	92.32	1.99	0.44	12.05
	Area	88.31	6.03	.42**	50	89.74	4.54	1.70	7.11
w	*Number*	0.20	0.15	-.26**	0.12	1	2.45	-0.51	11.21
	Length	0.09	0.07	-.30**	0.06	1	-0.19	-0.45	9.88
	Area	0.09	0.07	-.34**	0.07	1	1.75	-0.67	8.50
									** $p < .01$

Table 2. Discrimination performance across age. Note that correlations between Age and *w* are Spearman *rho* values, due to the non-normal distribution of *w* values. As described in the Results section, *a* corresponds to the bottom asymptote (the minimum value with development), *k* to the top asymptote (the maximum value with development), *m* to the age at which development has reached roughly halfway, and *b* to the rate of development. Because lower *w* values are better, growth rates for this dependent variable are negative.

Because each dimension obeyed Weber's law (i.e., was ratio-dependent) we fit a standard psychophysical model to estimate each participant's perceptual precision for each of the three dimensions (i.e., their Weber fraction, *w*; Odic, 2018; Halberda & Feigenson, 2008; Lidz et al., 2011; Pica et al., 2004):

$$Accuracy = (1-g) * \Phi\left(\frac{Ratio - 1}{w\sqrt{1 + Ratio^2}}\right) * \frac{g}{2}$$

where Φ is the Gaussian cumulative distribution function, *w* is each participant's Weber fraction, and *g* is each participant's guess/lapse rate (the consistent percentage of trials on which they took a random guess). This model assumes the underlying representations are distributed along a continuum of Gaussian/Normally distributed random variables, consistent with prior theories of intuitive quantity representations (Lidz et al., 2011). Because each representation is distributed across a continuum, two values will naturally overlap in representation, leading to confusion. As a result, the smaller the ratio is between two quantities, the greater overlap there will be of their Gaussian representations, and thus the greater difficulty participants will have with being able to discriminate between them. Therefore, lower *w* values indicate better performance. To estimate each participant's *w,* we fit their accuracy data over ratios to the above equation through R's *mle2* function.

A 3 (Dimension: Number, Length, Area) x 9 (Age group: 5, 6, 7, 8, 9, 10, 11, 12, 24) Greenhouse-Geisser corrected Mixed-Measures ANOVA, with Weber fraction (*w*) values as the DV, revealed a main effect of Dimension ($F(1.42, 135.06) = 34.66$; $p < .001$, $\eta_p^2 = .27$), a main effect of Age ($F(8, 95) = 5.47$, $p < .001$, $\eta_p^2 = .32$), but no Dimension x Age interaction ($F(11.37, 135.06) = 0.66$; $p = .78$, $\eta_p^2 = .05$). As shown in Table 2, we found average *w* values that replicate previously established patterns in the literature (e.g., Odic, 2018; Halberda & Feigenson, 2008; Halberda, Ly, Wilmer, Naiman, & Germine, 2012): Number *w* was significantly worse than that of Area and Length, but in the correct range given previous findings in these age groups, and there were general improvements with age even when treated continuously.

Next, we examined Estimation Task performance, focusing on both estimation accuracy and estimation variability as dependent variables. To account for instances in which participants may have taken a random guess, we excluded estimates that were more than 3 standard deviations away from each participant's mean guess and below 2. We find that, with the exception of Area Estimation, there are no correlations between estimation accuracy and estimation variability, even when controlling for age and even when excluding adults: Area ($rho_{adults} = -.38$, $n = 104$, $p < .001$; $rho_{noadults} = -.41$, $n = 90$, $p < .001$), Number ($rho_{adults} = .14$, $n = 104$, $p = .14$; $rho_{noadults} = .031$, $n = 90$, $p = .77$), Length ($rho_{adults} = -.13$, $n = 104$, $p = .18$;

$rho_{noadults}$ = -.18, n = 90, p = .09). In other words, these two signatures of estimation performance index separate abilities, at least in the case of Length and Number. Because of this, we report all estimation accuracy vs. variability analyses separately below.

A 3 (Dimension: Number, Length, Area) x 9 (Age group: 5, 6, 7, 8, 9, 10, 11, 12, 24) Greenhouse-Geisser corrected Mixed-Measures ANOVA with estimation accuracy as the DV revealed a main effect of Dimension ($F(1.99, 189.90) = 95.73$; $p < .001$, $\eta_p^2 = .50$), a main effect of Age ($F(8, 95) = 3.21$; $p = .003$, $\eta_p^2 = .21$), but no significant Dimension x Age interaction ($F(15.99, 189.90) = 1.29$; $p = .21$, $\eta_p^2 = .10$). As can be seen in Table 3, in contrast to the Discrimination Task, participants had the *best* estimation accuracy on the Number estimation trials and the worst estimation accuracy on the Area estimation trials, with Length in-between. Additionally, across all three dimensions estimation accuracy improved (i.e., approached 0) with age (see also Figure 4).

In addition, a 3 (Dimension: Number, Length, Area) x 9 (Age group: 5, 6, 7, 8, 9, 10, 11, 12, 24) Greenhouse-Geisser corrected Mixed-Level ANOVA with estimation variability as the DV revealed a main effect of Dimension ($F(1.79, 169.67) = 69.84$; $p < .001$, $\eta_p^2 = .42$), a main effect of Age ($F(8, 95) = 6.66$; $p < .001$, $\eta_p^2 = .36$), but no significant Dimension x Age interaction ($F(14.29, 169.67) = 0.74$; $p = .73$, $\eta_p^2 = .06$). As can be seen in Figure 4, in the case of estimation accuracy, we also found the best performance (i.e., lowest variability) for Number estimation trials, and the worst performance for Area estimation trials, with Length in the middle, when examining estimation variability (Table 3).

Together, these results stand in strong contrast to the Discrimination results: while children and adults generally struggled the most with Number discrimination, they were the least accurate and least precise for Area and Length estimation. However, these results do not yet answer the question of whether the interface emerges and peaks at the same time for the three dimensions, or if it is built in a piecemeal fashion. To understand that, we next calculate the approximate age of maturity for each dimension, followed by examining correlations in estimation performance across the three dimensions.

Estimation Task										
	Dimension	**Mean**	**SD**	**Correlations with Age**	**Growth Modelling**				
					a	***k***	***m***	***b***	**Age of Maturity**
Accuracy	*Number*	0.24	0.19	-.19^	0.21	1	4.07	-1.1	6.52
	Length	0.35	0.21	-.36**	0.20	1	2.80	-0.24	14.08
	Area	0.58	0.24	-.32**	0.64	0.42	11.72	-0.57	17.06
Variability	*Number*	0.17	0.10	-.60**	0.12	1	2.45	-0.51	11.21
	Length	0.27	0.11	-.53**	0.18	1	1.55	-0.32	15.28
	Area	0.31	0.13	-.41**	0.23	1	1.51	-0.32	15.04

** $p < .01$
^ $p = .06$

Table 3. Estimation performance across age. Note that all correlations with Age are Spearman *rho* values because both estimation accuracy and variability are non-normally distributed. As described in the Results section, *a* corresponds to the bottom asymptote (the minimum value with development), *k* to the top asymptote (the maximum value with development), *m* to the age at which development has reached roughly halfway, and *b* to the rate of development. Because lower estimation accuracy and variability values are better, growth rates for these dependent variables are negative.

Age of Maturity.

To better understand precisely when each dimension reaches its developmental plateau, we modelled the discrimination and estimation data through a standard logistic growth model (e.g., Kersey, Braham, Csumitta, Libertus, & Cantlon, 2018; Odic, 2018; Marceau, Ram, Houts, Grimm, & Susman, 2011; Ram & Grimm, 2015),

$$Accuracy\ (acc) = a + \left(\frac{k - a}{1 + \mathrm{e}^{-b(age - m)}}\right)$$

where *a* is the bottom asymptote, *k* is the top asymptote, *b* is the rate of growth, and *m* is the midpoint or age between onset and end of development (see Table 2 and Table 3). In the case of discrimination accuracy, the top asymptote (*k*) represents the end-point of development, whereas in the case of estimation accuracy and estimation variability the bottom asymptote (*a*) represents the end-point of development, since lower values are indicative of better performance. Although this model has four free parameters, in practice we would fix one (either *k* or *a*) depending on whether growth grows towards a positive asymptote (e.g., accuracy) or towards a negative asymptote (e.g., estimation accuracy and variability). We estimated these parameters by fitting the above equation to the age group level, through R's *mle2* function.

As shown in Figure 3 and 4, we found that the logistic growth model was an excellent fit to the data across dimensions and tasks (Table 2 and 3)[2]. Given the fit parameters, we can estimate the approximate age of maturity by checking at what age the model estimates that the data is within 5% of the adult-like peak level of performance (see Odic, 2018).

[2] The one exception to this was the model's fit for area estimation accuracy, where we found issues with getting the model to converge. This may be a result of area slopes showing bimodal patterns of development, as can be seen in Figure 4. An alternative possibility is that we captured different types of strategies across development used to estimate in area. This could also reflect differences in how the task was understood. Thus, we interpret our results lightly for this dimension.

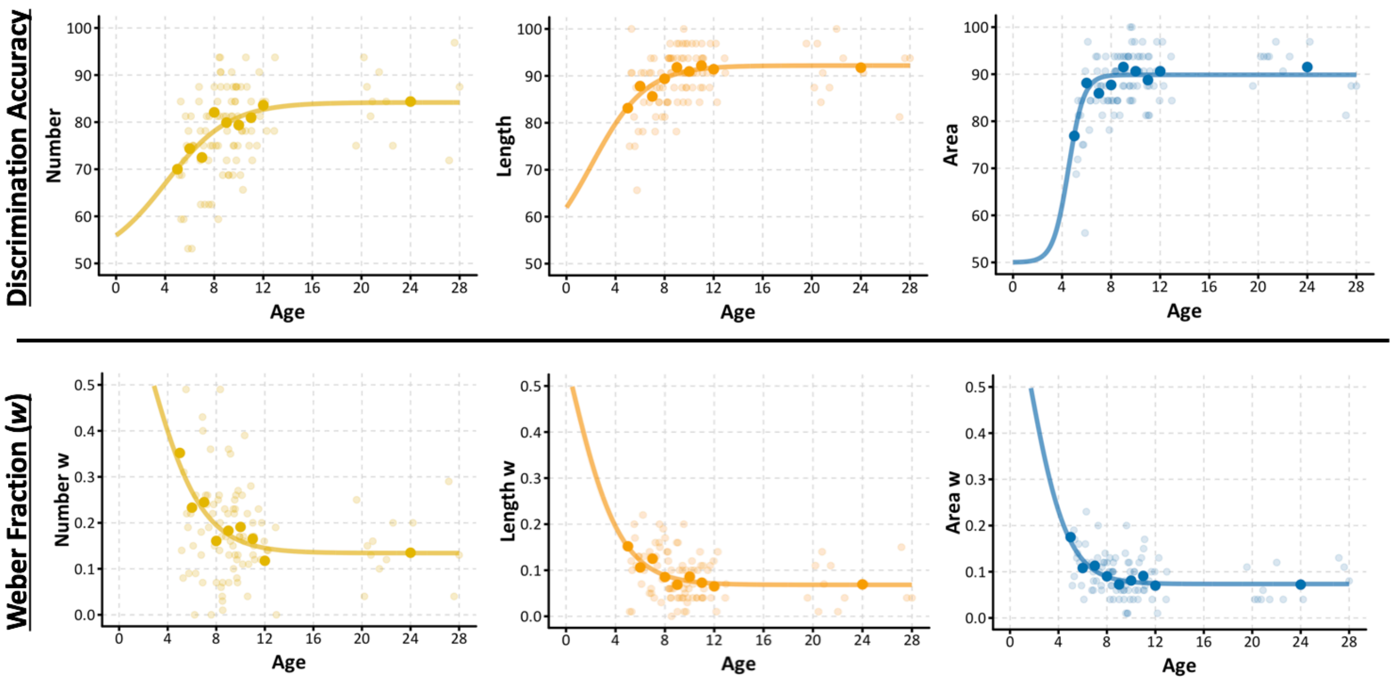

Figure 3. Discrimination performance modeled across development and dimension. The bold dots indicate group means, while the faded dots are individual participants. Lines indicate the best fit logistic growth models. For aesthetics, the figures only show data up to age 28.

The estimated approximate age of maturity for each of the three dimensions for both tasks is shown in Table 2 and 3. Replicating previous results (e.g., Odic, 2018), we find that—in the case of Discrimination accuracy—Number reaches maturity last in middle adolescence (M = 16;1, SE = 2.31) compared to Length (M = 12;0, SE = 0.99) and Area (M = 7;1, SE = 1.48), which peak around the elementary school years (see Figure 3). This difference can be tested statistically through Fisher's Z test. We find that the age of maturity for Number is significantly different from Length (Z = 3.29, p < .001) and Area (Z = 2.77, p < .01) but that Area and Length peak at the same age (Z = 1.63, p = .10).

In contrast to Discrimination, we find the reverse pattern in Estimation accuracy and variability. As can be seen in Table 3 and Figure 4, estimation accuracy for Number peaks very early (M = 6;6, SE = 0.63), followed by Length (M = 14;0, SE = 0.97; Z = 6.53, p < .001) and

Area (M = 17;0, SE = 1.58; Z = 6.20, p < .001), which are, in turn, not significantly different from each other (Z = 1.60, p = .11). We find the same pattern of results for estimation variability, with Number peaking much sooner than Length and Area.

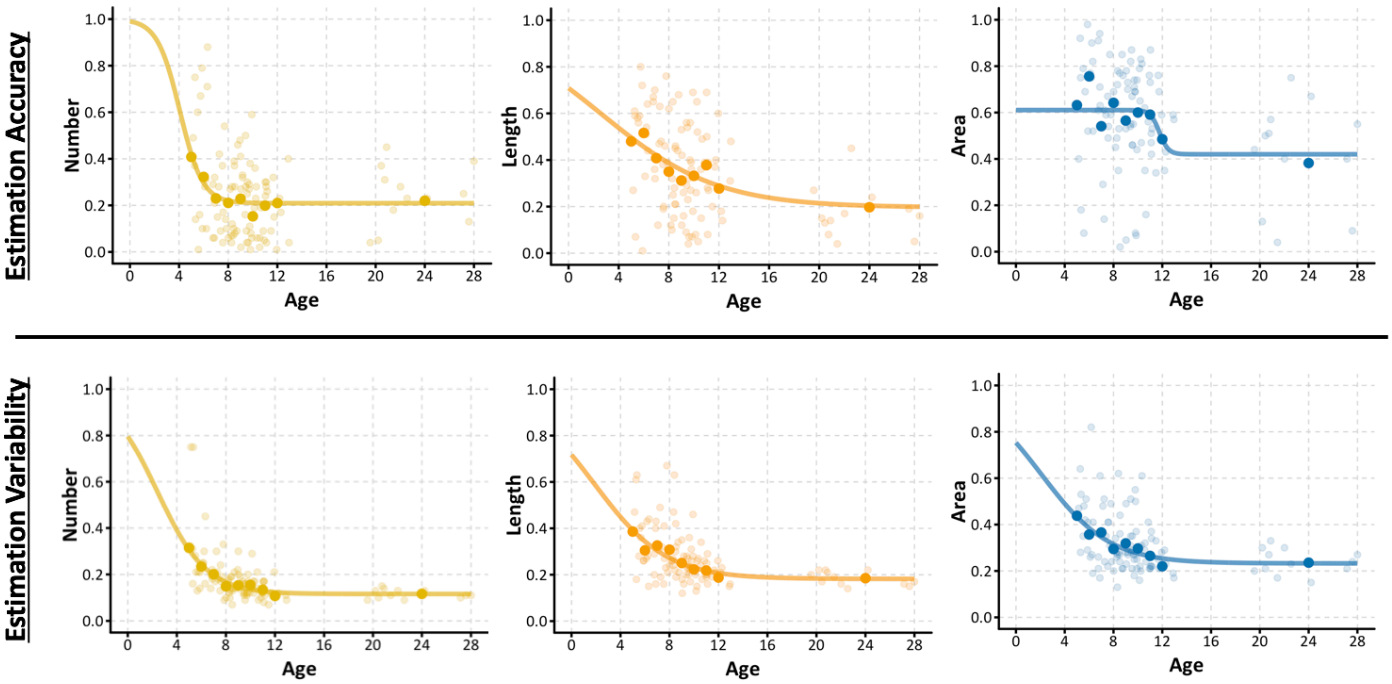

Figure 4. Estimation performance modeled across development and dimension. The bold dots indicate group means, while the faded dots are individual participants. Lines indicate the best fit logistic growth models. For aesthetics, the figures only show data up to age 28.

These results suggest that the interface between number words and perceptual quantity is not built in piecemeal, as the age of maturity for Area and Length are reached at roughly the same time (Table 3). However, these findings also suggest that Number holds a privileged status within this interface, as it peaks significantly before the other dimensions. In the final section, we further examine whether the interface between these dimensions and language is domain-specific or domain-general by turning to correlations.

Correlations Within and Across Tasks.

To examine whether our perceptual sense of number, length, and area are domain-general, we examined correlations of *w* values while controlling for age (because the heterogeneous age of the adults may disproportionally affect the correlation coefficients, we report the data both with and without adults included). As positive correlations can be induced by task-general aspects, such as guessing behaviour due to fatigue, we focused primarily on *w,* which statistically removes the effects of guessing, rather than Discrimination accuracy.

We found that Number and Length *w* values were not significantly correlated (rho_{adults} = -.01, n = 104, p = .90), nor was Number *w* correlated with Area *w* (rho_{adults} = .05, n = 104, p = .64). However, we found that there was a significant correlation between the Length and Area *w* values (rho_{adults} = .25, n = 104, p = .008). We find the same effects, when adults are excluded: Number *w* does not correlate with Length *w* ($rho_{noadults}$ = -.02, n = 90, p = .85) nor Area *w* ($rho_{noadults}$ = .03, n = 90, p = .75), but Length *w* and Area *w* are correlated ($rho_{noadults}$ = .33, n = 90, p = .001). These results replicate previous work suggesting a domain-specific number sense (e.g., Odic, 2018; Odic et al., 2013), and further new insights about the correlation between Length and Area perception, which has not been previously tested.

In contrast, we found that estimation accuracy was significantly correlated across all dimensions when adults were excluded: Number and Length ($rho_{noadults}$ = .26, n = 90, p = .01), Number and Area ($rho_{noadults}$ = .33, n = 90, p = .001), and Length and Area ($rho_{noadults}$ = .46, n = 90, p < .001). With adults, included in the analyses, we find the same effects with Number and Area (rho_{adults} = .32, n = 104, p < .001) and Length and Area (rho_{adults} = .47, n = 104, p = < .001), and a marginal significance between Number and Length (rho_{adults} = .17, n = 104, p = .08). When examining estimation variability, we find that performance was significantly correlated across all dimensions, with and without adults: Number and Length (rho_{adults} = .31, n = 104, p < .001; $rho_{noadults}$ = .32, n = 90, p = .002), Number and Area (rho_{adults} = .35, n = 104, p < .001; $rho_{noadults}$ = .35, n = 90, p < .001), and Length and Area (rho_{adults} = .46, n = 104, p < .001; $rho_{noadults}$ = .49, n = 90, p = < .001).

Lastly, we explored whether there were any relationships across the two tasks. We found that when controlling for age-related improvements, discrimination accuracy did not correlate with estimation accuracy in Number (rho_{adults} = -.07, n = 104, p = .49; $rho_{noadults}$ = -.08, n = 90, p = .43), Length (rho_{adults} = -.05, n = 104, p = .65; $rho_{noadults}$ = -.13, n = 90, p = .24), nor Area (rho_{adults} = -.07, n = 104, p = .51; $rho_{noadults}$ = -.12, n = 90, p = .28). These results replicated when examining w and estimation accuracy: Number (rho_{adults} = .03, n = 104, p = .76; $rho_{noadults}$ = .01, n = 90, p = .89), Length (rho_{adults} = -.01, n = 104, p = .93; $rho_{noadults}$ = .05, n = 90, p = .66), Area (rho_{adults} = .03, n = 104, p = .79; $rho_{noadults}$ = .09, n = 90, p = .42).

Likewise, we find no significant correlation between discrimination accuracy and estimation variability in Length (rho_{adults} = -.03, n = 104, p = .74; $rho_{noadults}$ = -.02, n = 90, p = .84) nor Area (rho_{adults} = -.06, n = 104, p = .54; $rho_{noadults}$ = -.06, n = 90, p = .57), but we do find a correlation between these factors in Number (rho_{adults} = -.20, n = 104, p = .04; $rho_{noadults}$ = -.27, n = 90, p = .01). We find a similar pattern of results when examining the relationship between w values and estimation variability: no significant correlation in Length (rho_{adults} = -.001, n = 104, p = .99; $rho_{noadults}$ = -.02, n = 90, p = .84) nor Area (rho_{adults} = .01, n = 104, p = .90; $rho_{noadults}$ = .009, n = 90, p = .94), and a significant correlation between these factors in Number, but only when adults were excluded (rho_{adults} = .18, n = 104, p = .07; $rho_{noadults}$ = .24, n = 90, p = .02). These findings largely suggest that the perceptual acuity of these intuitive representations did not predict participants' abilities to interface language (i.e., number words) with them, with the apparent exception of Number estimation variability.

General Discussion

By their fifth birthday, children have formed an interface between their intuitive, approximate sense of number and number words. This thesis examined when and how this interface is extended to other perceptual magnitudes: is the interface formed at once for all dimensions (e.g., number, length, and area) due to a domain-general structure mapping process, or is the interface formed in a slow, piecemeal fashion due to associative mapping between number words and each independent dimension?

Under the associative mapping account, children should learn to interface precise number words with each dimension separately, through independent experiences of how a number word, like "five," maps to a particular representation of dots (number), lines (length), or blobs (area). Hence, we should expect that: (1) children form the interface for each dimension at different times (in proportion to the relative experience they gain with number, length, and area in their daily lives), and (2) that the quality of the interface for one dimension should not predict the quality of another. Contrary to this—and consistent with the structure mapping account—our results show that, unlike in discrimination, number, length, and area estimation are strongly correlated and show a similar age of maturation in both estimation accuracy and variability. At the same time, however, we find that number estimation abilities peak significantly sooner compared to length and area, suggesting a privileged status for children's interface between number words and the Approximate Number System (ANS). Moreover, the earlier emergence of the interface with number is not merely due to inherent differences at the perceptual level, as in discrimination, number acuity peaks last. These results hold a number of important implications for our understanding of how our continuous and universal perceptual sense of magnitude interacts with human-specific and culturally acquired number words.

First, the asymmetry in the age of maturation for number discrimination vs. estimation suggests an important and unique status for the interface between number words and the ANS throughout early development. Why do number estimation abilities peak so much sooner than the ANS? One possibility explaining this asymmetry is that, despite the higher perceptual noise and later maturation, ANS representations are highly salient and lend themselves naturally to children forming categorical number word labels over them. Consistent with this idea, work by Ferrigno, Jara-Ettinger, Piantadosi, and Cantlon (2017) has shown that adults, children, and even monkeys will spontaneously choose to categorize stimuli by number over other dimensions (e.g., size, area; see also Cantlon, Safford, & Brannon, 2010), and that these patterns are observed in human cultures with both low- and high-numeracy knowledge (i.e., Tsimane' adults in Bolivia and U.S. adults, respectively). Thus, number could be a naturally salient, prominent dimension, independent of species or culture (see also Cantlon & Brannon, 2007a, 2007b; Dehaene, Dehaene-Lambertz, & Cohen, 1998; Boysen, Berntson, Hannan, & Cacioppo, 1996; Xu & Spelke, 2000), contributing to the earlier emergence its interface with precise and discrete number words.

Alternatively, the interface with number may peak first as a natural by-product of experience and practice. After all, children are first introduced to number words through counting (Sarnecka & Carey, 2008; Wynn, 1992; Benoit et al., 2004; Wiese, 2007; Le Corre & Carey, 2007), and it is not until much later in school that they begin to learn about units of measurement in length and area (e.g., cm, mm^2). Thus, by the time children reach age five, they may have already accumulated more experience with reasoning in number relative to other dimensions, allowing them to map number words onto their ANS significantly sooner. One potential way of disentangling these two hypotheses would be to provide children with number estimation units that they are not well practiced with, such as with a set of three dots labelled as "a toma". If children's number estimation abilities are reduced in this task, we would have clear evidence that the earlier peak of this interface between the ANS and number words is due to experience and practice in the number domain with single object units. Alternatively, observing a continued benefit for number estimation in this task would point towards a generally better interface between number words and the ANS, independent of practice.

Second, despite the fact that the interface between the ANS and number words may be privileged, the high correlations amongst all three dimensions in estimation accuracy and variability demonstrate that some common mechanism underlies children and adults' abilities to map number words to their perceptual magnitude representations. As a result, our data have strong implications for theories of how our perceptual representations of quantity interface with number words.

To date, two major theories have been proposed to explain how a structure mapped interface may itself function. Under the first, children learn what portions of their internal perceptual representations correspond to particular units (e.g., mapping *this* much of their internal perception of length to 1 unit of length), then subsequently perform ratio comparisons between their mapping to that unit and their observations (Krantz, 1972; Stevens, 1946). Thus, a child should respond with "five" when shown an object that appears to be 5 times longer than the memorized unit mapped to her representation of length. This straightforward ratio computation could explain why we observe a common age of maturation for area and length: once a child has learned now to estimate through ratio computation for number, they could straightforwardly extend that knowledge to other domains. At the same time, however, ratio computations should be localized within each perceptual system itself, making it unclear why we observed strong correlations in estimation while finding few in discrimination. While some previous work has suggested that ratio computations may themselves be domain-general and bridge across perceptual domains (e.g., Bonn & Cantlon, 2017), future work is required to empirically validate this assumption.

The alternative view for how structure mapping allows for the interface between number words and the ANS is that observers form a categorical “response grid”, learning which boundaries of the ANS representations correspond to each number word, much like children learn the boundaries of the continuous colour space that correspond to “yellow”, “orange”, “blue”, etc. (Izard & Dehaene, 2008; Sullivan & Barner, 2014). Crucially, the child does not need to learn the boundary for each and every number word: once they realize that a particular range of ANS values all correspond to the same number word, the child can assume that each number word captures an equal amount of the internal ANS scale, allowing them to extend their response grid to portions of the ANS that they have rarely or never experienced. Under this view, once children have formed a response grid for the ANS and number words, they could extend this logic to other perceptual dimensions of quantity (e.g., length, area) by virtue of the shared representational format that underlies them (e.g., linear dimensions, where higher values represent greater quantities). This may account for why we then find the near-simultaneous emergence of the interface with length and area following the interface with number; and because the response grid is itself fluid and susceptible to internal noise (Izard & Dehaene, 2008), this could explain the strong correlations we see between the three dimensions in the estimation task.

Although future work is required to formally disentangle how the interface between number words and perceptual magnitudes functions (i.e., via ratio comparison or response grids), the current data suggests that any model of this process that functions for number representations must be able to easily extend to other perceptual dimensions, as well.

Thus far, we have only discussed the implications of our data to the literature on estimation, but importantly our data also replicate and extend the existing work on the nature of perceptual magnitude representations in discrimination tasks. First, we replicate previous work showing that number peaks significantly later and is uncorrelated with our perception of area and length (Odic et al., 2013; Odic, 2018), suggesting that the ANS is a domain-specific system for representing number. At the same time, this work is the first to show a correlation between length and area perception, which suggests some shared component within these dimensions. This is worthy of further research, as it is unclear whether the observed correlations stem from actually shared domain-general perceptual representations of quantity (e.g., Leibovich et al., 2017; Lourenco, 2015), or whether they emerge due to some shared low-level features used to extract length and area information from a visual display. For example, some previous work has shown that observers sometimes compute area in a 2D space by taking the diameter or radius of the object as a proxy (Nachmias, 2008). Thus, one possibility is that a correlation between dimensions is observed when participants elect to use length as a partial proxy for area. Nevertheless, one of the most important findings of this thesis are that number representations are not correlated with neither length nor area in discrimination, but are strongly correlated in estimation, suggesting that a shared mechanism—outside of the domain of perception—underlies the interface of these representations with number words.

Taken together, these findings provide several novel insights about language and perception. While we share with other animals several intuitive perceptual capacities—including an intuitive ability to reason about quantity—we can think and talk about our representations in ways that far exceed that of our animal counterparts. Understanding when and how children map number words to a broader spectrum of quantity representations—beyond just number—is critical not only to our understanding of development, but because we, as adults, engage in these complex reasoning abilities every day. Furthermore, examining the similarities and differences in the acquisition of this interface across a broad range of dimensions—which have previously not been tested—contributes to our theories of how children learn to navigate and reason about their rich perceptual worlds.

References

Barth, H., Starr, A., & Sullivan, J. (2009). Children's mappings of large number words to numerosities. *Cognitive Development*, *24*(3), 248–264.

Benoit, L., Lehalle, H., & Jouen, F. (2004). Do young children acquire number words through subitizing or counting? *Cognitive Development*, *19*(3), 291–307.

Beran, M. J., Rumbaugh, D. M., & Savage-Rumbaugh, E. S. (1998). Chimpanzee (Pan troglodytes) counting in a computerized testing paradigm. *The Psychological Record*, *48*(1), 3–19.

Bonn, C. D., & Cantlon, J. F. (2017). Spontaneous, modality-general abstraction of a ratio scale. *Cognition*, *169*, 36–45.

Bonny, J. W., & Lourenco, S. F. (2013). The approximate number system and its relation to early math achievement: Evidence from the preschool years. *Journal of Experimental Child Psychology*, *114*(3), 375–388.

Boysen, S. T., Berntson, G. G., Hannan, M. B., & Cacioppo, J. T. (1996). Quantity-based interference and symbolic representations in chimpanzees (Pan troglodytes). *Journal of Experimental Psychology: Animal Behavior Processes*, *22*(1), 76.

Boysen, S. T., & Capaldi, E. J. (2014). *The development of numerical competence: Animal and human models*. Psychology Press.

Burr, D., & Ross, J. (2008). A visual sense of number. *Current Biology*, *18*(6), 425–428.

Cantlon, J. F. (2018). How Evolution Constrains Human Numerical Concepts. *Child Development Perspectives*, *12*(1), 65–71.

Cantlon, J. F., & Brannon, E. M. (2007a). Basic math in monkeys and college students. *PLoS Biology*, *5*(12), e328.

Cantlon, J. F., & Brannon, E. M. (2007b). How much does number matter to a monkey (Macaca mulatta)? *Journal of Experimental Psychology: Animal Behavior Processes*, *33*(1), 32.

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.

Cordes, S., & Brannon, E. M. (2008). Quantitative competencies in infancy. *Developmental Science*, *11*(6), 803–808.

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*, 65.

Cordes, S., Gallistel, C. R., Gelman, R., & Latham, P. (2007). Nonverbal arithmetic in humans: Light from noise. *Perception & Psychophysics*, *69*(7), 1185–1203.

Dakin, S. C., Tibber, M. S., Greenwood, J. A., & Morgan, M. J. (2011). A common visual metric for approximate number and density. *Proceedings of the National Academy of Sciences*, *108*(49), 19552–19557.

Dehaene, S. (2001). Précis of the number sense. *Mind & Language*, *16*(1), 16–36.

Dehaene, S. (2009). Origins of mathematical intuitions. *Annals of the New York Academy of Sciences*, *1156*(1), 232–259.

Dehaene, S. (2011). *The number sense: How the mind creates mathematics*. OUP USA.

Dehaene, S., Dehaene-Lambertz, G., & Cohen, L. (1998). Abstract representations of numbers in the animal and human brain. *Trends in Neurosciences*, *21*(8), 355–361.

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

Ferrigno, S., Jara-Ettinger, J., Piantadosi, S. T., & Cantlon, J. F. (2017). Universal and uniquely human factors in spontaneous number perception. *Nature Communications*, *8*, 13968.

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.

Gebuis, T., & Reynvoet, B. (2012). The interplay between nonsymbolic number and its continuous visual properties. *Journal of Experimental Psychology: General*, *141*(4), 642.

Gebuis, T., & Van Der Smagt, M. J. (2011). False approximations of the approximate number system? *PLoS One*, *6*(10), e25405.

Halberda, J., & Feigenson, L. (2008). 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.

Halberda, J., Ly, R., Wilmer, J. B., Naiman, D. Q., & Germine, L. (2012). Number sense across the lifespan as revealed by a massive Internet-based sample. *Proceedings of the National Academy of Sciences*, *109*(28), 11116–11120.

Holden, C. (2004). *Life without numbers in the Amazon*. American Association for the Advancement of Science.

Howard, S. R., Avarguès-Weber, A., Garcia, J. E., Greentree, A. D., & Dyer, A. G. (2018). Numerical ordering of zero in honey bees. *Science*, *360*(6393), 1124–1126.

Izard, V., & Dehaene, S. (2008). Calibrating the mental number line. *Cognition*, *106*(3), 1221–1247.

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.

Jordan, K. E., & Brannon, E. M. (2006). Weber's Law influences numerical representations in rhesus macaques (Macaca mulatta). *Animal Cognition*, *9*(3), 159–172.

Kersey, A. J., Braham, E. J., Csumitta, K. D., Libertus, M. E., & Cantlon, J. F. (2018). No intrinsic gender differences in children's earliest numerical abilities. *Npj Science of Learning*, *3*(1), 12. https://doi.org/10.1038/s41539-018-0028-7

Krantz, D. H. (1972). A theory of magnitude estimation and cross-modality matching. *Journal of Mathematical Psychology*, *9*(2), 168–199. https://doi.org/10.1016/0022-2496(72)90025-9

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.

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 Sciences*, *40*.

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.

Libertus, M. E., Feigenson, L., & Halberda, J. (2013). Is approximate number precision a stable predictor of math ability? *Learning and Individual Differences*, *25*, 126–133.

Libertus, M. E., Odic, D., & Halberda, J. (2012). Intuitive sense of number correlates with math scores on college-entrance examination. *Acta Psychologica*, *141*(3), 373–379. https://doi.org/10.1016/j.actpsy.2012.09.009

Lidz, J., Pietroski, P., Halberda, J., & Hunter, T. (2011). Interface transparency and the psychosemantics of most. *Natural Language Semantics*, *19*(3), 227–256.

Lourenco, S. F. (2015). On the relation between numerical and non-numerical magnitudes: Evidence for a general magnitude system. In *Mathematical Cognition and Learning* (Vol. 1, pp. 145–174). Elsevier.

Lucon-Xiccato, T., Gatto, E., & Bisazza, A. (2018). Quantity discrimination by treefrogs. *Animal Behaviour*, *139*, 61–69.

Marceau, K., Ram, N., Houts, R. M., Grimm, K. J., & Susman, E. J. (2011). Individual differences in boys' and girls' timing and tempo of puberty: Modeling development with nonlinear growth models. *Developmental Psychology*, *47*(5), 1389.

Marques, J. F., & Dehaene, S. (2004). Developing intuition for prices in euros: rescaling or relearning prices? *Journal of Experimental Psychology: Applied*, *10*(3), 148.

Nachmias, J. (2008). Judging spatial properties of simple figures. *Vision Research*, *48*(11), 1290–1296.

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.

Odic, D., Le Corre, M., & Halberda, J. (2015). Children's mappings between number words and the approximate number system. *Cognition*, *138*, 102–121.

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.

Odic, D., & Starr, A. (2018). An Introduction to the Approximate Number System. *Child Development Perspectives*.

Piantadosi, S. T. (2016). A rational analysis of the approximate number system. *Psychonomic Bulletin & Review*, *23*(3), 877–886. https://doi.org/10.3758/s13423-015-0963-8

Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science*, *306*(5695), 499–503.

Piffer, L., Petrazzini, M. E. M., & Agrillo, C. (2013). Large Number Discrimination in Newborn Fish. *PLOS ONE*, *8*(4), e62466. https://doi.org/10.1371/journal.pone.0062466

Platt, J. R., & Johnson, D. M. (1971). Localization of position within a homogeneous behavior chain: Effects of error contingencies. *Learning and Motivation*, *2*(4), 386–414.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, *1*(4), 515–526.

Ram, N., & Grimm, K. J. (2015). Growth curve modeling and longitudinal factor analysis. *Handbook of Child Psychology and Developmental Science*, 1–31.

Ross, J., & Burr, D. C. (2010). Vision senses number directly. *Journal of Vision*, *10*(2), 10–10.

Sarnecka, B. W., & Carey, S. (2008). How counting represents number: What children must learn and when they learn it. *Cognition*, *108*(3), 662–674.

Schneider, M., & Siegler, R. S. (2010). Representations of the magnitudes of fractions. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(5), 1227.

Siegler, R. S., Fazio, L. K., Bailey, D. H., & Zhou, X. (2013). Fractions: The new frontier for theories of numerical development. *Trends in Cognitive Sciences*, *17*(1), 13–19.

Siegler, R. S., Thompson, C. A., & Schneider, M. (2011). An integrated theory of whole number and fractions development. *Cognitive Psychology*, *62*(4), 273–296.

Stevens, S. S. (1946). On the theory of scales of measurement.

Sullivan, J., & Barner, D. (2010). Mapping number words to approximate magnitudes: associative learning or structure mapping? In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 32).

Sullivan, J., & Barner, D. (2011). Children's use of Structure Mapping in numerical estimation. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 33).

Sullivan, J., & Barner, D. (2013). How are number words mapped to approximate magnitudes? *The Quarterly Journal of Experimental Psychology*, *66*(2), 389–402.

Sullivan, J., & Barner, D. (2014). Inference and association in children's early numerical estimation. *Child Development*, *85*(4), 1740–1755.

Tomonaga, M. (2008). Relative numerosity discrimination by chimpanzees (Pan troglodytes): evidence for approximate numerical representations. *Animal Cognition*, *11*(1), 43–57.

Wiese, H. (2007). The co-evolution of number concepts and counting words. *Lingua*, *117*(5), 758–772.

Wynn, K. (1992). Children's acquisition of the number words and the counting system. *Cognitive Psychology*, *24*(2), 220–251.

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition*, *74*(1), B1–B11.