Discovering Quantification and Number in a Role-Filler Model

Aaron J. Hamer (ahamer@hawaii.edu)
University of Hawaii at Manoa, Linguistics
1890 East-West Road, Honolulu, HI 96822 USA

Leonidas A. A. Doumas (leonidas@hawaii.edu)
University of Hawaii at Manoa, Psychology
2530 Dole Street, Honolulu, HI 96822 USA

Abstract

Quantification plays a central role in human reasoning and models thereof, but the discovery and development of quantification remains an open question. We present a theory of how such concepts are learned from experience in the DORA model, a neurally-plausible computational model of relational learning and reasoning (Doumas et al., 2008). The same theory accounts for how concepts of number are acquired in this class of model. We are unaware of any prior model that accounts for the development of both quantification and number from unstructured (e.g., perceptual) input.

Keywords: number; quantification; relational discovery; computational modeling.

Introduction

Quantification and number are key representational constructs in human cognition. These concepts are foundational in science, mathematics, music, and many other domains of human achievement. Many models of cognition rely on these representational primitives (e.g., any symbolic model that relies on first-order predicate calculus, many Bayesian models such as Piantadosi, Tenenbaum, & Goodman's (2013) model of quantifier discovery, etc.), but as Carey (2009, p. 456) notes, "There is no proposal I know for a learning mechanism available to non-linguistic creatures that can create representations of objects, number, agency, or causality from perceptual primitives."

These concepts share significant semantic overlap ranging from their function as predicates over sets of objects (Barwise & Cooper, 1981) to innate, scalar ordering (e.g., one, two, three & some, many, most; Horn, 1972). Both sets of concepts can be derived from a small set of axioms via set theory (i.e., set membership, identity; Van Heijenoort, 1977). It does not seem unreasonable to consider the problem of their acquisition jointly. While there have been attempts to explain their acquisition in terms of a developmental trajectory from number to quantifiers or vice versa (e.g., Gelman & Gallistel, 1978; Carey, 2004), we are unaware of any existing model that accounts for the development of representations of both quantification and number from unstructured (i.e., perceptual) input.

Behavioral Data

Quantification Facts

Behavioral evidence suggests that there are three broad areas of difficulty with the acquisition of quantification: quantifier spreading, mapping issues, and superlative quantifiers.

Quantifier Spreading Philip and his colleagues (1991a, 1991b) popularized the term *quantifier spreading* to describe a phenomenon first reported by Inhelder & Piaget (1964). Children aged six to seven were unable to restrict universal quantifiers to a subset of items present in an array based on a shared feature. When presented with three purple triangles and a purple circle and asked "Are all the triangles purple?" the children would respond in the negative. When asked for an explanation, a typical response was "The circle is purple, too."

Mapping Issues Brooks & Braine (1996) demonstrated that children have more rigid mappings for the quantifiers *all* and *each* than adults. Children preferred a grouped interpretation of *all* in scenarios such as "All of the roses are in a vase" and a distributed, one-to-one interpretation of *each* in scenarios such as "Each of the roses is in a vase". They interpreted scenes where roses were distributed over more than one vase as false for the *all* quantifier and scenes where there was not a one-to-one mapping of roses-to-vases (e.g., more roses than vases, more vases than roses) as false for the *each* quantifier. Children achieve adult-like performance reasoning about *all* at around age five but do not reach adult-like performance reasoning about *each* until age nine.

Superlative Quantifiers Scalar quantifiers can be divided into two types: superlative quantifiers that include their endpoints (e.g., at most three, three or more) and comparative quantifiers that exclude their endpoints (e.g., less than four, more than two). Musolino (2004) showed that five-year-old children performed worse on tasks relying on superlative quantifiers versus comparative quantifiers. Geurts et al. (2010) investigated this phenomenon further and showed that the difficulty of acquiring superlative quantifiers extended to 11-year-old children. Geurts et al. also showed that superlative quantifiers were more difficult for adults to process (as shown by higher RTs). Hurewitz et al. (2006) found that three-year-olds interpret *some* as inclusive of *all*. This result suggests that *some* undergoes a transition from a superlative quantifier to a comparative quantifier at some point in development.

Number Facts

Our discussion of the behavioral data on the acquisition of number will focus on three areas: numerosity and counting, the linear shift, and operational momentum.

Numerosity and Counting Children as young as two-years-old can subitize, or determine the numerosity of small sets without counting (Gelman & Gallistel, 1978). However, three-year-olds struggle with the foundations of counting (Grinstead et al., 1997), and have difficulty with cardinality (Wynn 1990, 1992). By three-and-a-half, most children demonstrate exact judgments of numbers up to four and the ability to count to similar magnitudes (Gelman & Gallistel, 1978; Hurewitz et al., 2006).

The Linear Shift Children initially estimate numerical quantities based on a logarithmic scale before undergoing a shift to using a linear scale at approximately 12 years of age (Siegler & Opfer, 2003). Logarithmic estimations of quantity are consistent with a perceptual system that obeys the Weber-Fechner law (Fechner, 1860).

Operational Momentum McCrink et al. (2007) showed that adults overestimate sums and underestimate differences, a phenomenon referred to as *operational momentum*. The pattern of errors fits a Gaussian distribution if magnitudes are represented logarithmically rather than linearly.

Summary of Behavioral Data

Children struggle with the acquisition of concepts of quantification and number. Some abilities are present early (e.g., subitization at two years) and others develop quickly (e.g., developing counting between ages three and three-and-a-half). Other abilities develop more gradually (e.g., restriction of quantifiers) and some developmental trajectories extend into adolescence (e.g., the linear shift). In some cases earlier points on the developmental trajectory are more compatible with formal logic than the adult norm (e.g., *some* as a superlative quantifier).

Developmental Accounts

Theories of Quantification

Existing accounts of the development of quantification can be grouped into three broad categories: connectionist models, symbolic models, and Bayesian models (e.g., Clark, 1996; Carey, 2004; and Piantadosi, Tenenbaum, & Goodman, 2013, respectively). Existing connectionist models model the association of externally supplied symbols such as words with first-order quantifiers. We have not found an account that does not assume pre-existing symbolic representations such as number[1] (Carey, 2004) or the set theoretic equivalents of number, the existential quantifier, the universal quantifier, or formally equivalent items (i.e., cardinality, non-exhaustion, exhaustion, and membership & identity, respectively; Piantadosi et al., 2013; Van Heijenoort, 1977).

Theories of Number

We will examine four classes of models of the acquisition of number: connectionist models, spiking-neuron models, symbolic models, and Bayesian models.

Connectionist Models of Number Existing connectionist models provide an excellent account for the development of subitization via associative learning or summation encoding (e.g., Ahmad, Casey, & Bale, 2002; Dehaene & Changeux, 1993; and Verguts & Fias, 2004). Various models have provided an account for innate ordering via unsupervised competitive recurrent back-propagation networks (e.g., Ahmad et al., 2002) and the association of external symbols with existing representations of number via co-occurrence (Verguts & Fias, 2004). These models do not address phenomena that occur later in development, nor do they provide an account for the emergence of symbolic representations.

Spiking-Neuron Models of Number These models focus on tying specific abilities or developmental processes to what is known about neuronal behavior. Examples include modeling number as a consequence of gamma oscillations[2] that predicts subitization behavior that obeys the Weber-Fechner law (Miller & Kenyon, 2007) and a tuning function based on neuronal spike trains that accounts for both operational momentum and the linear shift (Prather, 2012).

Symbolic Models of Number Existing symbolic accounts either require "explicit external symbols" (e.g., Carey, 2009) or assume an existing set of quantifier representations (e.g., Gelman & Gallistel, 1978). While these models account for many developmental phenomena, they openly assume a pre-existing cache of symbolic currency to build upon.

Bayesian Models of Number Extant Bayesian models of the acquisition of number share the flaws of Bayesian models of quantification – they assume set theoretic equivalents of number, the existential quantifier, the universal quantifier, or formally equivalent items (i.e., cardinality, non-exhaustion, exhaustion, and membership & identity, respectively; Piantadosi, Tenenbaum, & Goodman, 2012; Van Heijenoort, 1977).

[1] Set theory has demonstrated that quantifiers can be derived from number, and vice versa (Van Heijenoort, 1977).

[2] Gamma-band oscillations have been advanced as a candidate for carrying binding information in object representations (Knowlton et al., 2012).

Summary of Developmental Accounts

Existing accounts of the development of quantification and number can be grouped into connectionist, symbolic, and Bayesian models. While each class of model has strengths, all existing models fail to account for the development of the symbolic currency such as predicates or set operations that they either map to or build upon. Furthermore, no existing model has accounted for both domains of concepts or all of the key developmental trajectories within a single domain.

The DORA Model

Overview

The DORA model is a symbolic connectionist architecture: a computational model using a neural network to store and manipulate structured representations. DORA represents objects and roles in a distributed fashion - that is, as patterns of simultaneous activation over units (analogous to groups of neurons) that represent the semantic features of the item being encoded.

DORA learns structured representations of properties shared between objects by comparing them. Features shared between objects receive input from multiple sources and are isolated via simple Hebbian learning. The resulting representations are comprised of these shared features, are independent of any specific objects, and can be bound to novel objects encountered in the future.[3] When DORA compares instances of objects searching for another (e.g., a cat searching for a mouse and a sister searching for her brother) it learns representations of *searcher* (comparing the cat and sister) and *sought* (comparing the mouse and the brother). When observing a new instance of searching the existing representation of *sought* can apply (i.e., be bound) to the sought object.

The representations DORA learns are functionally equivalent to single-place predicates that take novel arguments. Although the initial representations that DORA learns contain extraneous features (e.g., the shared features of the cat and sister irrelevant to *searcher*), comparisons between different instances produce representations that are progressively more refined (i.e., comparing representations *searcher* learned from different instances causes context-specific features to wash out).

The DORA model represents multi-place relations by combining sets of these single-place predicates - e.g., after learning representations of *searcher* and *sought* they can be combined to form a representation of the multi-place relation *searching*. If there is anything invariant about a concept (and there must be for us to recognize it), DORA can learn a structured representation of it.

Discovery of Quantification and Number

The DORA model learns new representations through a process of iterated comparison of items in the object and role layer, where featural overlaps[4] are learned as new representations. This process allows for refinement of existing representations by comparing them to other existing representations or new input.

All quantifiers are learned by comparing instances of countable items and extracting numerosity features. There are many accounts of how a connectionist model can acquire basic numerosity features (e.g., Ahmad et al., 2002; Dehaene & Changeux, 1993; and Verguts & Fias, 2004); DORA implements a version of the Metric Array Module (Hummel & Holyoak, 2001) which extracts magnitude features for any metric dimension, such as numerosity or length.

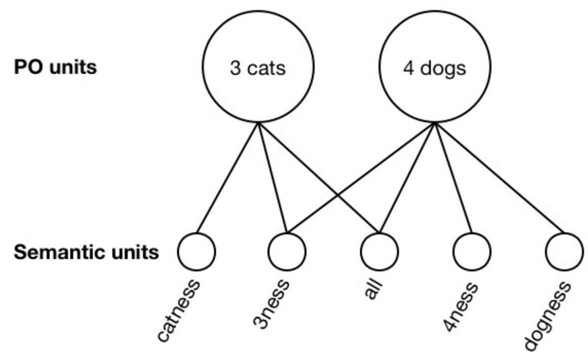

Figure 1: An example of comparing instances of countable items. Note that the featural representation of *3ness* is active for higher cardinality sets, at least in quantities where subitization is an effective strategy to extract numerosity features.

Initial comparisons, especially when the arity of compared sets differ, will result in representations of quantifiers such as the *all* node in Figure 1. Note that the initial representation in this example contains the *3ness* node as well. This process allows for the extraction of quantifiers such as *all*, and through additional experience, quantifiers such as *some*. The nodes *3ness* and *all* referenced here are purely expository and stand in for the perceptual features that map to these concepts just as the

[3] DORA binds representations of roles to fillers (e.g., objects) dynamically (i.e., the binding of a role to a filler is temporary so that the same role can be bound to different fillers in different contexts) via systematic asynchrony of firing (Doumas et al., 2008). In asynchrony-based binding, roles are bound to their fillers by proximity of firing, with bound roles and fillers firing in direct sequence. For example, to bind *searcher* to cat, and *sought* to mouse, the units coding the *searcher* role fire, followed by the units coding cat. Next, the units coding the *sought* role fire followed by the units coding for mouse.

[4] As well as non-overlaps, though not as quickly.

nodes for *catness* and *dogness* are collapsed representations of the featural invariants present in cats and dogs.

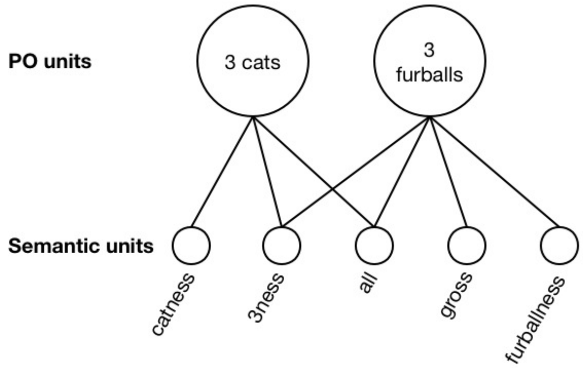

Figure 2: Extraction of the quantity 3.

The same process accounts for the extraction of number representations. As a consequence of this process, concepts that are encountered more frequently (*all*, *one*, *some*) will be learned before concepts that are encountered less frequently (*fifteen*, *at least*), and previously learned concepts can be used to bootstrap the learning of future concepts. Eventually, pure conceptual representations of frequently encountered quantifiers and numbers are extracted through repeated comparison.

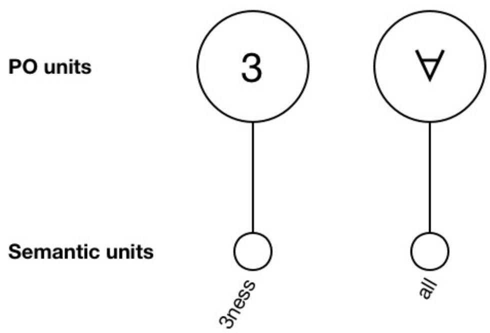

Figure 3: The resulting representations for frequently encountered quantities.

Representational Consequences

The representations shown in Figure 3 are pure set representations, suitable for set operations. They can be bound to other relations to create bound sets (solving the quantifier spreading problem, assuming that the cognitive system has developed both these representations and scope-limiting representations and has enough WM to bind them together). There are some other significant consequences of this manner of representation.

Cardinality of the Universal Quantifier All quantifiers are learned through experience; there is never a time when a quantifier is perceived without being predicated over some set. Consequently, the universal quantifier is cardinal. While the cardinality of the universal (and other quantifiers) will change based on the specific context it is experienced or represented in, it will always possess cardinality. This underscores the results from set theory that suggest that numbers and quantifiers are formally equivalent (Van Heijenoort, 1977).

Place-Value Notation Numeral Systems Commonly encountered quantities will be explicitly represented in such a system. It is likely that specific quantifiers for the numbers one through ten exist in such a system. However, it is extremely unlikely that such a system learns a specific representation for quantities such as 347. However, such representations can be built form the representational currency of lower-order numbers such as three, four, and seven, and a representation for place that takes on features of the base of the numeral system (e.g., 10 for Arabic numerals) and magnitude of the base (e.g., two for the hundreds place), and so on.

The Way Forward - Count on DORA to Quantify Development

Our theory of quantification and number development handles three major issues not addressed in current models. First, we account for both domains within a single model using a small set of principles (e.g., comparison-based learning, building complex representations from single-place predicates) and processes. Furthermore, we provide an account for how these symbolic representations are developed and structured as a consequence without drawing from an existing cache of symbolic currency. Finally, our model accounts for a wide variety of developmental trajectories within each domain using the same set of basic parameters and processes, as well as a wide variety of other developmental trajectories.

Unifying Quantification and Number

One of the core goals of framing the acquisition of quantification and number within the DORA framework is to provide a unified account of their development. Unifying both domains as opposite endpoints of a developmental trajectory has been attempted (e.g., Gelman & Gallistel, 1978; Carey, 2004) but such attempts fail to account for the intertwined developmental trajectories as they are built on assumptions of mastery within a domain as a foundation on which to build mastery of the other. The most successful Bayesian modeling attempts to account for the development of quantification and number are currently instantiated as separate Bayesian models built on the same set of priors (Piantadosi et al., 2012, 2013). While unifying Bayesian models built on the same set of priors is relatively simple , it remains to be done.

Our account of the development of quantification and number captures key developmental trajectories in both domains as a consequence of comparison-based learning iterating over previously learned concepts and new experience. The interactions between the developmental trajectories of quantification and number are captured because they arise as a natural consequence of learning both domains at the same time. These interaction effects forced us to deal with both domains simultaneously as modeling either quantification or number learning in isolation failed to account the developmental facts for either domain. DORA cannot model either quantification or number in isolation as successfully as it can account for both together.

Symbolic Structure Developed, Not Borrowed

Most accounts of cognition fail to explain where the structured symbolic representations they use to solve problems come from. Such structures range from predicates, set operators, and even quantifiers and cardinality. The core function of the DORA model is to extract invariance from unstructured (e.g., perceptual) input via comparison. Using a comparison-based learning mechanism not only explains how such structure arises, but also what this structure looks like. This mechanism creates the representations that many models rely upon.

Bayesian models of development rely on an external source of structured symbols to build a foundation upon. While Bayesian models provide an excellent way to model competency, when modeling development they run into more fundamental issues than failing to account for where the structures they rely on come from. The most successful Bayesian models of the development of quantification and number competency in people (i.e., Piantadosi et al., 2012, 2013) rely on priors that are a superset of the concepts they claim to develop. Put simply, they start with the assumption that people can already count to three and use the quantifiers for existence, some, all, and none. We find it difficult to characterize a model as developmental when it assumes its outputs as priors.

Modeling Developmental Trajectories

We have provided a brief overview of how DORA learns cardinality and number from experience, but we have not yet laid out how our model handles the developmental trajectories at play.

DORA begins subitizing using the Metric Array Module, a simple, neurally plausible mechanism that could easily be available to two-year-old children. This mechanism outputs magnitude judgments that obey the Weber-Fechner law. Logarithmic judgments of magnitude explain why children treat numbers and analogous quantifiers such as *some* as superlative quantifiers initially because a point on a logarithmic scale corresponds to a range on a linear scale. As DORA is exposed to many instances of small sets (as children are) it quickly learns to represent small cardinal numbers explicitly. These explicit representations do not rely on logarithmic magnitude features; consequently, children no longer treat these numbers as superlative quantifiers.

Children gain working memory as the prefrontal cortex matures. Quantifier spreading disappears as children are able to marshall the working memory needed to build the complex representations required to simultaneously bind a quantifier to a scope-limiting representation and match that representation to a particular situation. The representations for cardinal numbers continue to develop throughout childhood as larger and larger numbers become explicitly represented, accounting for the linear shift in early puberty.

We account for all of these developmental facts with a single set of parameters and simple processes. DORA also accounts for over 35 findings surrounding the development of relational thinking (Doumas & Hummel, 2010; Doumas et al., 2006; Doumas et al., 2008; Sandhofer & Doumas, 2008), including the relational shift (Rattermann & Gentner, 1991), the development of relational representations (Smith, 1984), and the development of shape bias (Abecassis et al., 2001).

Conclusion

Our proposal is a promising account of how concepts of quantifiers and number can be learned from perceptual input. The DORA model's working memory constraints allow a developmental trajectory to be modeled, and make specific predictions about how specific types of quantified reasoning will fail based on working memory demands, such as differing magnitudes of n-back tasks. We are exploring these predictions with human participants. Crucially, our model accomplishes these goals using the same parameters and processes that have allowed us to successfully account for more than 35 developmental phenomena in other domains.

Acknowledgments

The first author would like to thank Daft Punk for the Alive 2007 set.

References

Abecassis, M., Sera, M. D., Yonas, A., & Schwade, J. (2001). What's in a shape? Children represent shape variability differently than adults when naming objects. *Journal of Experimental Child Psychology*, 78, 303-326.

Ahmad, K., Casey, M., & Bale, T. (2002). Connectionist simulation of quantification skills. *Connection Science*, 14 (3), 165-201.

Barwise, J., & Cooper, R. (1981). Generalized quantifiers and natural language. *Linguistics and Philosophy*, 4, 159–219.

Brooks, P. J. & Braine, M. D. S. (1996). What do children known about the universal quantifiers all and each? *Cognition*, 60, 235-268.

Carey, S. (2004). Bootstrapping and the origins of concepts. *Daedalus*, 133(1), 59–68.

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

Clark, R. (1996). Learning first order quantifier denotations: An essay in semantic learnability. *IRCS Technical Report 96-19*.

Dehaene, S. & Changeux, J. (1993). Development of elementary numerical abilities: A neuronal model. *Journal of Cognitive Neuroscience*, 5(4), 390-407.

Doumas, L. A. A. & Hummel, J. E. (2010). A computational account of the development of representations underlying object recognition. *Cognitive Science*, 34, 698-712.

Doumas, L. A. A., Hummel, J. E., & Sandhofer, C. M. (2008). A theory of the discovery and predication of relational concepts. *Psychological Review*, 115, 1-43.

Fechner, G. T. (1860). *Elemente der Psychophysik*. Leipzig: Breitkopf and Hartel.

Gelman, R., & Gallistel, C. R. (1978). *The Child's Understanding of Number*. Cambridge, MA: Harvard University Press.

Geurts, B., Katsos, N., Cummins, C., Moons, J., & Noordman, L. (2010). Scalar quantifiers: Logic, acquisition, and processing. *Language and Cognitive Processes*, 25(1), 130-148.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics*: Vol. 3. New York, NY: Academic.

Grice, H. P. (1989). *Studies in the way of words*. Cambridge, MA: Harvard University Press.

Grinstead, J., MacSwan, J., Curtiss, S., & Gelman, R. (1997). The independence of language and number. In Greenhill, A., Hughes, M., Littlefield, H., & Walsh, H. (Eds.), *Proceedings of the 22nd Boston University Conference on Language Development*, 1. Somerville, MA: Cascadilla.

Horn, L. (1972). *On the semantic properties of the logical operators in English*. (Unpublished doctoral dissertation). University of California, Los Angeles, CA.

Hummel, J. E., and Holyoak, K. J. (2001). Toward an understanding of analogy within a biological symbol system. In D. Gentner, K. J. Holyoak, & B. N. Kokinov (Eds.), *The Analogical Mind: Perspectives from Cognitive Science*. Cambridge, MA: MIT Press.

Hurewitz, F., Papafragou, A., Gleitman, L., & Gelman, R. (2006). Asymmetries in the acquisition of numbers and quantifiers. *Language Learning and Development*, 2(2), 77-96.

Inhelder, B., & Piaget, J. (1964). *The Early Growth of Logic in the Child*. London: Routledge and Kegan Paul.

Knowlton, B. J., Morrison, R. G., Hummel, J. E., & Holyoak, K. J. (2012). A neurocomputational system for relational reasoning. *Trends in Cognitive Sciences*, 16(7), 373-381.

McCrink, K., Dehaene, S., & Dehaene-Lambertz, G. (2007). Moving along the number line: operational momentum in nonsymbolic arithmetic. *Perception & Psychophysics*, 69 (8), 1324–1333.

Miller, J. A., & Kenyon, G. T. (2007). Extracting number-selective responses from coherent oscillations in a computer model. *Neural Computation*, 19(7), 1766-1797.

Musolino, J. (2004). The semantics and acquisition of number words: Integrating linguistic and developmental perspectives. *Cognition*, 93, 1-41.

Philip, W. & Aurelio, S. (1991). Quantifier spreading: Pilot study of preschooler's "every". In Maxfield, T. & Plunkett, B. (Eds.), *Papers in the Acquisition of WH: Proceedings of the UMass Roundtable, May 1996*. Amherst, MA: Graduate Linguistics Student Association.

Philip, W. & Takahashi, M. (1991). Quantifier spreading in the acquisition of "every". In Maxfield, T. & Plunkett, B. (Eds.), *Papers in the Acquisition of WH: Proceedings of the UMass Roundtable, May 1996*. Amherst, MA: Graduate Linguistics Student Association.

Piantadosi, S. T., Tenenbaum, J. B., & Goodman, N. D. (2012). Bootstrapping in a language of thought: A formal model of numerical concept learning. *Cognition*, 123, 199-217.

Piantadosi, S. T., Tenenbaum, J. B., & Goodman, N. D. (2013). *Modeling the acquisition of quantifier semantics: a case study in function word learnability*. Manuscript submitted for publication.

Prather, R. (2012). Connecting neural coding to number cognition: a computational account. *Developmental Science*, 15(4), 589-600.

Rattermann, M. J. & Gentner, D. (1998). More evidence for a relational shift in the development of analogy: Children's performance on a casual-mapping task. *Cognitive Development*, 13, 453-478.

Sandhofer, C. M. & Doumas, L. A. A. (2008). Order and presentation effects in learning categories. *Journal of Cognition and Development*, 9, 194-221.

Siegler, R. S. & Opfer, J.E. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological Science*, 14, 237-243.

Smith, L. B. (1984). Young children's understanding of attributes and dimensions. *Child Development*, 55, 363-380.

Van Heijenoort, J. (1977). *From Frege to Gödel: a source book in mathematical logic, 1879-1931*. Cambridge, MA: Harvard University Press.

Verguts, T., & Fias, W. (2004). Representation of number in animals and humans: a neural model. *Journal of Cognitive Neuroscience*, 16, 1493–1504.

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

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