THE DEVELOPMENT OF COMPLEX MEANING: LEXICAL REPRESENTATION AND ACCESS IN ADULTS, TYPICAL CHILDREN, AND CHILDREN WITH AUTISM

by

NORA ISACOFF

A Dissertation submitted to the

Graduate School-New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Psychology

written under the direction of

Karin Stromswold

and approved by

____

____

____

____

New Brunswick, New Jersey

OCTOBER, 2014

ABSTRACT OF THE DISSERTATION

The Development of Complex Meaning: Lexical Representation and Access in Adults, Typical Children, and Children with Autism

By NORA M. ISACOFF

Dissertation Director:

Karin Stromswold

This dissertation explores the development of words and categories with the intention of better understanding the organization of the lexicon. Study I investigates lexical organization, lexical access, and categorization in adults. In particular, I contrast the way adults access lexical items with the way they explicitly categorize these items. By testing speakers of two different languages—English and Mandarin—I also explore which aspects of lexical organization are universal and which are linguistically relative. Study II investigates the nature of lexical representation, access, and categorization in two- through nine-year-old typically-developing children. Additionally, I investigate the way developments in categorization do and do not coincide with development of the ability to articulate these categories, with the acquisition of relevant declarative knowledge, and with developments in other cognitive abilities. The results of the first two studies suggest that lexical organization and access are qualitatively different in young children than in adults. Study III investigates the nature of the lexicon and categories in children with autism. I find that although autistic children have many cognitive deficits, they seem to follow typical patterns of category development. Together, these studies improve our

understanding of the nature and development of lexical organization, lexical access, and categorization.

Acknowledgements

I would like to thank my advisor, Karin Stromswold, for giving me the freedom to spend six years working on the “intractable” topic of meaning. Her faith in me has helped me develop an intellectual independence that I am sure will serve me well throughout my research career. Together, I think we may have made some progress on our intractable topic. I would also like to thank my committee members, Kimberly Brenneman, Rochel Gelman, and Lila Gleitman for their invaluable support and guidance. I am also forever indebted to Janet Fodor, who gave me a research assistant position in her lab when all I knew was that I wanted to get my hands dirty and learn about language.

To the members of the Language Acquisition and Language Processing lab, including my fellow graduate students (Choonkyu Lee, Gaurav Kharkwall, and Gwendolyn Rehrig) and all of the undergraduates, I am very grateful for your help and suggestions over the past six years. I especially want to thank my research assistant Mackenzie Stabile. Mackenzie, you have undoubtedly been my partner in crime on this project. For all of your dedication and thoughtfulness, and for the many late nights (early mornings) working with me in the lab, I will forever be grateful.

My fellow graduate students (inside and outside of RUCCS) have been my community for the past six years. It has been a blessing to have so many wonderful colleagues and friends experiencing the highs and lows of graduate school with me. In large part, I call myself an interdisciplinary researcher because of the challenging and enriching conversations I have had with my fellow graduate students

To all of my friends and family—especially my parents Adrienne & Stuart and my sister Rachel—I have only accomplished this dissertation with your unwavering love and support. For that, and for filling my life with joy and laughter, this is your achievement as much as it is mine.

My research was supported in part by grants from the the Bamford-Lahey Children's Foundation, the Charles and Johanna Busch Biomedical Research Fund, the National Science Foundation (BCS-0446850, DGE-0549115), and a Rutgers University Special Study Award. Additionally, my research would not have been possible without the many children, families, and schools who chose to participate.

Table of Contents

Abstract ……………………………………………………………………………............... ii

Acknowledgments ………………………………………………………………………….. iv

List of Tables ……………………………………………………………………………….. vi

List of Figures ……………………………………………………………………………….. x

Forward ……………………………………………………………………………………… 1

I. Lexical representation, access, and categorization in adults ……..................................... 4

II. Lexical representation, access, and categorization in typical children …………………. 32

III. Lexical representation, access, and categorization in children with autism …………… 89

IV. General discussion ……………………………………………………………………. 147

Appendix A. Animal declarative knowledge questions …………………………………... 154

Appendix B. Metamemory task questions ……………………………………….............. 155

References ………………………………………………………………………………… 156

List of Tables

Table 1.1 Number of explicit grouping participants who grouped by each feature ………….... 14

Table 1.2 Percent of verbal fluency animals tagged with each explicit grouping feature ……... 16

Table 1.3A Comparison of clustering in toy example ………………………………………. 18

Table 1.3B Randomized version of animal list ……………………………………………… 18

Table 1.4 Explicit grouping features used in clustering …………………………………….. 19

Table 2.1 Logistic regression analysis of intermediate level taxonomic justification with superordinate level taxonomic justification and age as predictors …………….……. 76

Table 2.2 Logistic regression analysis of intermediate level taxonomic justification with superordinate level thematic justification and age as predictors …………………..... 77

Table 3.1 Logistic regression analysis of perceptual justification with MA and group as predictors ………………………………………………………………………. 108

Table 3.2 Logistic regression analysis of perceptual justification with CA and group as predictors ………………………………………………………………………. 109

Table 3.3 Logistic regression analysis of shape justification on the perceptual grouping task with MA and group as predictors …………..…………………………………… 110

Table 3.4 Logistic regression analysis of shape justification on the perceptual grouping task with CA and group as predictors …………………………………………………… 110

Table 3.5 Logistic regression analysis of switching perceptual grouping dimensions with MA and group as predictors ……………………………………………………… 112

Table 3.6 Logistic regression analysis of switching perceptual grouping dimensions with CA and group as predictors ……………………………………………………...... 113

Table 3.7 Logistic regression analysis of switching perceptual grouping justification dimensions with MA and group as predictors ………………………………….... 114

Table 3.8 Logistic regression analysis of switching perceptual grouping justification dimensions with CA and group as predictors ………………………………..... 115

Table 3.9 Logistic regression analysis of taxonomic grouping on the superordinate level task with MA and group as predictors …………………………………………… 119

Table 3.10 Logistic regression analysis of taxonomic grouping on the superordinate level task with CA and group as predictors ……………………….…………………… 119

Table 3.11 Logistic regression analysis of superordinate level taxonomic justification with MA and group as predictors ……………………………………………………… 121

Table 3.12 Logistic regression analysis of superordinate level taxonomic justification with CA and group as predictors ……………………………………………………..… 121

Table 3.13 Logistic regression analysis of superordinate level thematic justification with MA and group as predictors …………………………………………………………... 123

Table 3.14 Logistic regression analysis of superordinate level thematic justification with CA and group as predictors …………………………………………………………... 124

Table 3.15 Logistic regression analysis of intermediate level taxonomic justification with MA and group as predictors ……………………………………………..……….. 125

Table 3.16 Logistic regression analysis of intermediate level taxonomic justification with CA and group as predictors ……………………………………………...…….. 126

Table 3.17 Logistic regression analysis of intermediate level perceptual justification with MA and group as predictors ……………………………………………..……….. 127

Table 3.18 Logistic regression analysis of intermediate level perceptual justification with CA and group as predictors ……………………………………………..……….. 128

Table 3.19 Logistic regression analysis of basic level grouping with MA and group as predictors ………………………………..…………………………………… 130

Table 3.20 Logistic regression analysis of basic level grouping with CA and group as predictors ………………………………..…………………………………… 131

Table 3.21 Logistic regression analysis of basic level justification with MA and group as predictors …………………………………………………………………..… 132

Table 3.22 Logistic regression analysis of basic level justification with CA and group as predictors …………………………………………………………………..… 133

List of Figures

Figure 1.1 Mean cluster size …………………………………………………………………… 19

Figure 1.2 Overlap in importance of features in Explicit Grouping and Verbal Fluency tasks .. 21

Figure 2.1 Familiarization task materials (Disney) …………………………………………….. 49

Figure 2.2 Familiarization task materials (Sesame Street) ………………………...……… 49

Figure 2.3 Familiarization task materials (Mickey Mouse) ………………………………… 50

Figure 2.4 Superordinate level cards ………………………………………………………... 51

Figure 2.5 Perceptual grouping cards ……………………………………………………….. 54

Figure 2.6 Intermediate level cards …………………………………………………………. 57

Figure 2.7 Basic level cards ………………………………………………………………… 58

Figure 2.8 Productive syntax stimulus ………………………………………………………… 61

Figure 2.9 Age and vocabulary scores ………………………………………………………… 62

Figure 2.10 Age and MLU on the productive syntax task …………………………………….. 63

Figure 2.11 Age on grouping and justification on the perceptual task ………..……………. 64

Figure 2.12 Age on the animal declarative knowledge task ……………………...………. 65

Figure 2.13 Age on the metamemory task …………………………………………………….. 66

Figure 2.14 Age on grouping and justification type on the superordinate level task .………… 68

Figure 2.15 Age and justification type on the intermediate level task ………………………… 69

Figure 2.16 Age on grouping and justification on the basic level task ………………………… 70

Figure 2.17 Taxonomic grouping vs. justification on the superordinate level task …..……….. 73

Figure 2.18 Taxonomic grouping vs. justification on the basic level task …………………….. 74

Figure 2.19 Grouping vs. justification on the perceptual grouping task ………………………. 74

Figure 2.20 Percent of children with 10+ on the animal declarative knowledge task and percent justifying taxonomically on the intermediate level task ……….……………… 75

Figure 3.1 Age and superordinate vocabulary in TD and ASD ……………………....…… 104

Figure 3.2 Age and intermediate vocabulary in TD and ASD ………………………………... 105

Figure 3.3 Age and MLU in TD and ASD …………………………………………………. 107

Figure 3.4 Age and dimension-switching in perceptual task groups in TD and ASD ………... 112

Figure 3.5 Age and dimension-switching in perceptual task justifications in TD and ASD 114

Figure 3.6 Age and animal declarative knowledge in TD and ASD 116

Figure 3.7 Age and metamemory in TD and ASD 117

Figure 3.8 Age and taxonomic grouping on the superordinate level task in TD and ASD 119

Figure 3.9 Age and taxonomic justification on the superordinate level task in TD and ASD ... 120

Figure 3.10 Age and thematic justification on the superordinate level task in TD 123

Figure 3.11 Age and taxonomic justification on the intermediate level task in TD and ASD .. 125

Figure 3.12 Age and perceptual justification on the intermediate level task in TD and ASD .. 127

Figure 3.13 Age and justification type on the intermediate level task in TD 129

Figure 3.14 Age and justification type on the intermediate level task in ASD 129

Figure 3.15 Age and taxonomic grouping on basic level task 130

Figure 3.16 Age and taxonomic justification on the basic level task in TD and ASD 132

Figure 3.17 Age and taxonomic grouping and justification on the superordinate level task in

TD 134

Figure 3.18 Age and taxonomic grouping and justification on the superordinate level task in

ASD 135

Figure 3.19 Age and taxonomic grouping and justification on the basic level task in TD 136

Figure 3.20 Age and taxonomic grouping and justification on the basic level task in ASD 136

Figure 3.21 Age and grouping and justification on the perceptual grouping task in TD 137

Figure 3.22 Age and grouping and justification on the perceptual grouping task in ASD 137

Figure 3.23 KBIT-2 Matrices sample A: Taxonomic association 142

Figure 3.24 KBIT-2 Matrices sample B: Thematic association 142

Figure 3.25 KBIT-2 Matrices item 24 143

Figure 3.26 KBIT-2 Matrices item 25 143

Figure 3.27 KBIT-2 Matrices item 26 144

Forward

This dissertation explores the development of words and categories with the intention of better understanding why only some types of knowledge seem integral to meaning. For example, both children and adults have the strong intuition that what makes something a dog is not its prototypical features but something intangible, a “dogginess,” what a biologist would call its DNA. However people’s intuitions about meaning often conflict with the way they use these meanings. For example, adults agree that *odd number* has a set of necessary and sufficient conditions—a precise definition—yet prototypicality effects hold for odd numbers, such that adults feel that some odd numbers are more prototypical than others and are faster to judge some numbers as odder than others (Armstrong, Gleitman, & Gleitman, 1983). Additionally, people’s intuitions about meaning often conflict with their declarative knowledge. For example, both children and adults believe that “ducks lay eggs” is true and that “ducks are females” is false, even though they know that the number of ducks who lay eggs is not greater than the number of ducks who are female (Leslie, Khemlani, & Glucksberg, 2010). The philosophical question then is what counts as the meaning of a word. The psychological question is how all of the information about a thing (e.g., an odd number or a duck) is mentally organized and how it is retrieved such that different information is more salient or appears more important in different situations. This dissertation is primarily concerned with psychological rather than philosophical meaning. However, by using empirical evidence to elucidate the complexity of psychological meaning, I hope to contribute a clearer understanding of the complexity of philosophical meaning.

Throughout this dissertation, I argue that although word meanings themselves are atomic—that is "duck" means DUCK—words are linked in a network such that words sharing a relevant feature prime each other. Please note that I am using the word "feature," for lack of a better word, to mean a property heading by which the item described by the word could typically be described, rather than to mean a necessary element of meaning. For example, a DOG is often a PET. If it is the case that "dog" primes other typical pet words (e.g., cat), then, by definition, PET is a feature of "dog" and of "cat." This does not mean that "pet" is part of the meaning of "dog" or that something must be a pet to be a dog. By evaluating which words activate each other in different contexts we can better understand how words are organized, how this organization is linked to non-lexical aspects of cognition, and how this organization develops.

In the following chapters, I explore the nature of lexical organization in several ways. Chapter I investigates lexical organization, lexical access, and categorization in adults. In particular, I contrast the way adults access lexical items with the way they explicitly categorize these items. By testing speakers of two different languages—English and Mandarin—I also explore which aspects of word meaning are universal and which are linguistically relative.

Chapter II investigates similar questions in typically developing children ages two through nine years. Here, I explore the relationship between developments in lexical representation and access with developments in three types of categories: taxonomic, thematic, and perceptual. *Taxonomic categories* share abstract properties (e.g., "cat" and "mouse" are both members of the category ANIMALS, whose members share the property of being self-propelled). *Thematic categories* share relations (e.g., “cat” and “mouse” are both members of the category CATS CHASE MICE. *Perceptual categories* share physical attributes (e.g., “a grey cat” and “a grey mouse” are both members of the category GREY). In all of these cases, if two words (e.g., “dog” and “cat”) are linked by a node a in the lexicon (e.g., by DOGS CHASE CATS), and if it is the case that “dog” primes “cat” through that node, then by definition, that node is a feature of both “dog” and “cat.”

This chapter investigates the way developments in categorization do and do not coincide with development of the ability to articulate these categories, with the acquisition of relevant declarative knowledge, and with developments in other cognitive abilities. I conclude this chapter by arguing that lexical organization is qualitatively different in young children than in adults, and I connect this finding to my categorization results.

Chapter III investigates the nature of words and categories in children with autism spectrum disorders. In this chapter, I find that although autistic children have many cognitive deficits, they follow typical patterns of category development. Finally, the Conclusion summarizes how this dissertation improves our understanding of the nature and development of lexical organization, lexical access, and categorization, and I propose a new experiment to build on current findings.

I. Lexical representation, access, and categorization in adults

If it walks like a duck and quacks like a duck, is it a duck? Adults reason that an animal can lack all of the characteristic features of a duck and still be a duck (Carey, 1985). Previous research has demonstrated that concept-specific features are not necessary for people's judgments of category inclusion, but rather, that domain-general features such as parenthood and DNA (for natural kinds) and functionality (for artifacts) govern these judgments (e.g., S.A. Gelman & Markman, 1987; Keil, 1986), and that for the category ANIMAL, these judgments hold in infancy (Setoh, Baillergeon, & R. Gelman, 2013). In particular, researchers (e.g., Carey, 1985; S.A. Gelman, 2004; Strevens, 2000) have argued that people believe natural kinds (including, for adults, specific animals) have a "causal essence" (Hirschfeld, 1996), empirically discoverable but beyond intuitive grasp. This claim meshes nicely with the intuition that most if not all of the facts people believe about a kind—e.g., that ducks live on the Earth, or that they are not made of jello—do not seem integral to conceptual constitution. Additionally, Malt, Slobin, & Gennari (2003) have demonstrated that linguistic categories are not necessarily isomorphic to conceptual categories, and Genome & Lombrozo (2012) have argued that neither description nor causal information can fully account for judgments about concept reference.

Given the vagueness of "causal essence" and the fact that neither world knowledge nor conceptual content equate with word meaning, how are words mentally represented, how are they organized, and how are they accessed during lexical retrieval?

In this chapter, I investigate the organization of the adult lexicon in two tasks: one more automatic and one more explicit. In the more automatic task, I find that words sharing more personally-relevant features (e.g., SCARY) tend to activate each other, whereas in the more explicit task, I find that words sharing explicitly taught features (e.g., MAMMAL) tend to activate each other.

Structure of the lexicon

The structural relationship between a word's linguistic content and the concept it denotes has been a central point of debate in linguistics, philosophy, and psychology for decades. Jackendoff (e.g., 1985, 2010) argues for direct links between phonological, syntactic, and conceptual information, with no additional word-specific semantic meaning. Conversely, Katz and Fodor (1963) initially contended that grammatical and conceptual meaning are straddled by an additional semantic layer. However, Fodor has since argued that lexical concepts are innate, atomic wholes (e.g., Fodor, 1970). Taking the opposite perspective, some psychologists (e.g., Collins & Quillian, 1969, Landauer & Dumais, 1997) have operationalized a word's meaning as its relationship to semantically related words linked with them in a network.[1]

Although the explicitness and cognitive economy of networks makes them appealing, one problem with the claim that meaning is *reducible* to links in a network is

[1] Although it is beyond the scope of this dissertation, one problem with semantic networks is the grounding problem. That is, if in a semantic network, the meaning of A is its relationship to B, and the meaning of B is its relationship to A, it is unclear how meaning could exist within this network. A possible solution is to ground one or more nodes to something meaningful outside of the symbolic system (e.g., to perception, see Harnad, for discussion). The grounding problem is a primary motivation for embodied cognition, and several researchers are working on embodied accounts of word meaning (e.g., Gallese & Lakoff, 2005). See Hauk and Tschentscher (2013) for a discussion of the current evidence for and against embodied semantics.

that some relationships are more integral to our intuitive sense of word meaning.[2] "Barking" is a prototypical but non-necessary feature of DOG; "having a liver" is necessary but not prototypical; and neither seems to approach the meaning of the word. At the very least, it seems necessary to allow for different distances in the links between words (as suggested by Collins & Loftus, 1975), but some researchers have argued that such a model is not falsifiable (Johnson-Laird, Herrmann & Chaffin, 1984). The distance between words could depend on their co-occurrence frequency (Griffiths, Steyvers, & Tenenbaum, 2007), but, although co-occurrence could be one important element of word meaning, it does not seem to capture meaning fully. For example, although DOG and FISH more frequently co-occur, DOG and WOLF are intuitively closer in meaning, and both FISH and WOLF prime DOG (Ferrand & New, 2003).

My proposal attempts to reconcile the "atomic whole" perspective with the "items in a network" perspective by linking atomic items (i.e., words) in a network in which the links are describable by features (e.g., PET). Further, the parts of the network that are activated are malleable both through development and by context (i.e., task), which seeks to address the need to differentiate between links (i.e., BARKING and HAVING A LIVER are likely relevant in different contexts).

Previous methods

A natural way to divide the literature on lexical organization is to distinguish between the features that participants give when explicitly asked to group items or to give their intuitions about essential characteristics and those features that emerge during more

[2] See Discussion section for a consideration of the role of intuition in investigating mental content.

automatic tasks. To investigate how people determine membership to a lexical category when they are explicitly asked to do so, some researchers (e.g., Guastavino, 2007) have used grouping tasks in which participants are asked to sort a list of words, or cards containing words, into groups by features of their choosing (see Explicit Grouping task below). Other researchers (e.g., Rosch, 1973, Schmitz & Wentura, 2012) have used semantic verification tasks in which participants are asked whether an item is a member of a category given by the researcher. Finally, to investigate the neurological correlates of explicit lexical meaning, researchers have employed several cognitive neuroscience techniques during categorization tasks including PET (Sergent, Zuck, Levesque, & MacDonald, 1992), ERP (Mari-Beffa, et al., 2005), MEG (Low, et al., 2003), and fMRI (Mahon & Caramazza, 2010).

Similarly, psychologists have investigated more automatic lexical access (and by extension, the organization of the lexicon) via a variety of techniques including priming (e.g., Forster, Mohan, & Hector, 2003); lexical decision tasks (e.g., Meyer & Schvanevelt, 1971); analyses of retrieval failures (e.g., Brown & McNeil's (1966) work on "tip of the tongue" and Fromkin's (1980) work on slips of the tongue); and neuroimaging techniques including PET (e.g., Frith, Friston, Liddle & Frackowiak, 1991), ERP (e.g., Federmeier, McLennan, de Ochoa & Kutas, 2002), MEG (e.g., Amunts, Weiss, Mohlberg, et al., 2004), fMRI (e.g., Gauthier, Duyme, Zanca, & Capron, 2009), and NIRS (e.g. Takahashi et al., 2011).

Together, these studies suggest that there are multiple types of features including semantic (e.g., Frenck-Mestre & Bueno, 1999; Troyer, 2000), but also associative (e.g., Buchanan, Brown, Cabeza, & Maitson, 1999), phonological (e.g., Slowiaczek &

Hamburger, 1992, Yee & Sedivy, 2006), orthographic (e.g., Chereau, Gaskell, & Dumay, 2007) and frequency (e.g., Segui, Frauenfelder, & Morton, 1982).

Taxonomic hierarchies

Many researchers have hypothesized that words are organized into taxonomic hierarchies. For example, consider that if someone pointed to a thing that has four legs, a tail, and barks and asked what that thing is called, correct responses would include “dog,” mammal,” and “animal,” because that thing would be a member of all three categories. The lowest of these categories is usually called the *basic level*, whose word is most frequently used to describe an object (i.e., for most adults, that thing to which one could point is most naturally described as a “dog” rather than as a “mammal” or an “animal”). The highest category (e.g., animal) is called the *superordinate level*. In this paper, I will call the middle level (e.g., mammal) the *intermediate level*. Note that there is nothing special distinguishing the superordinate from intermediate levels but that I am introducing this term for ease of discussion.

I propose that higher level categories (e.g., the superordinate level category ANIMALS can serve as features connecting lower level words (e.g, the intermediate level word “mammal” and the basic level word “dog). (Similarly, the intermediate level word MAMMAL could serve as a feature connecting the basic level words “dog” and cat.” Thus, *mammal* is both a word (“mammal) and a feature (MAMMAL).

This system is efficient for learning because once one has learned a fact about a higher level category (e.g., animals are self-propelling), one does not need to relearn that fact about new subcategories (e.g., a new animal). A hierarchical system is also efficient for retrieval. That is, if one wants to buy a new pet (an Intermediate level category) and wants to consider the options, one can traverse this category to come across DOG and CAT.

Verbal fluency tasks

A technique that capitalizes on the efficiency of lexical retrieval from taxonomic hierarchies is the verbal fluency task, in which participants are given a short period to name basic level members of a superordinate level category (e.g., they are asked to name animals or foods), and the order in which items are named is taken to reflect lexical organization. For example, if three quarters of the jungle animals that participants name are consecutive, but the brown animals that participants name are distributed throughout their lists, this would suggest that animals in the lexicon are organized by habitat (i.e., there is an intermediate level JUNGLE category) but not by color features.

Verbal fluency data are generally analyzed in one of two ways. In the first, researchers examine participants' lists (e.g., lists of animals or foods) and search for consecutive responses that intuitively share some feature. For example, Troyer, Moscovitch, & Winocur (1997) evaluated individual participants' lists of animals and, *post hoc*, identified clusters of what they argued were "obvious" animal subcategories based on biological type, habitat, domesticity, and other semantic subcategories. In identifying clusters, Troyer et al. gave "participants the benefit of the doubt regarding their use of clusters" Troyer et al., 1997, pp. 140). Subsequently, da Silva, Petersson, Faisca, Ingvar, & Reis (2004) employed Troyer et al.'s method to evaluate verbal fluency of both animals and supermarket items and found that literate and non-literate populations produced quantitatively and qualitatively similar clusters. In another study using a similar method, participants named supermarket items, and Sauzeon, Lestage, Raboutet, N'Kaoua, & Claverie (2004) identified semantic clusters by sorting the items into one of ten predetermined categories including fruits, meats, and desserts. Troyer et al.'s criteria have also been used in neuropsychological assessments that analyze clustering and switching, that is, clustering by one category (e.g., fruits) before switching to another category (e.g., meats, see Strauss, Sherman, & Spreen, 2006).

Although studies that rely on researchers' intuitions have been invaluable in providing the groundwork for using verbal fluency tasks to investigate lexical structure, one concern with this method is that researchers may over-identify features by observing a feature that is not used in lexical retrieval and that does not reflect lexical structure. For example, if three quarters of the animals that participants name are mammals, researchers are likely to observe several consecutive mammals and declare the existence of a mammal cluster, even though a participant is likely to name several consecutive mammals by chance alone. Conversely, researchers could under-identify features: ignoring an unintuitive but psychologically important feature. For example, if researchers do not entertain the possibility that SCARY is a feature, they would fail to notice if people consecutively name scary animals.

An additional concern with the *implementation* of this technique is that researchers often do not tag items with multiple features. For example, in their study of supermarket items, Sauzeon et al. could have tagged "milk" as both DAIRY and DRINK—two of their features—but, given that they only allowed each item to fit in one category, they only tagged "milk" as DAIRY. Limiting an item to a single category does not *a priori* yield the most psychologically plausible model. The gravest concern with researchers identifying features based on their intuitions is that if intuition were *sufficient* for uncovering lexical meaning, there would be no reason to conduct an experiment.

In the second general approach, researchers use automated clustering algorithms to analyze verbal fluency data. Employing the information theoretic paradigm initially adopted in psychology to investigate memory (e.g., Tulving, 1962), researchers have analyzed verbal fluency data via a variety of techniques including a next-to similarity matrix (Rubin & Olson, 1980), latent semantic analysis (e.g., Landauer, Foltz, & Laham, 1998), correspondence analysis and hierarchical clustering (Schwartz, Baldo, Graves, & Brugger, 2003), dynamical models such as the random inheritance model (Borge-Holthoefer & Arenas, 2009), and network theory (Goni et al., 2011). Although each technique is computationally distinct, they are similar in that they compute co-occurrence frequencies for items in verbal fluency lists, generating a multidimensional map of clusters.

A major advantage of clustering algorithms is that they detect patterns without projecting preexisting notions of what features—if any—people use to retrieve lexical items. The disadvantage, given my research interest of understanding *why* some words prime others, is that the output is merely *descriptive.* Once the model outputs clusters of items, the researcher must label the clusters (or at least conjecture why people tend to name some items together). For researchers concerned exclusively with modeling, lack of explicit features may not be a disadvantage at all. However, for researchers such as myself who seek an explanation for the underlying structure of the lexicon, this methodology is not ideal. Researchers who do label the clusters created by these computational models often provide labels that are not intuitively compelling. For example, Goni and colleagues tagged both “brown bear” and “starfish” with their BEAR AND POLAR feature (which includes any bear *or* polar animal) and assign UNCLASSIFIABLE as the feature linking items for which they could not decipher a common thread. In sum, although verbal fluency tasks present a promising avenue for exploring lexical structure, there are limitations to current methods for uncovering clusters from verbal fluency data.

Explicit vs. automatic access

Previous research suggests *some* overlap between the order in which participants name items and measures of more explicit lexical meaning. For example, Henley (1969) demonstrated that the proximity of animals named in verbal fluency tasks was highly correlated with both the similarity ranking that participants gave pairs of animals and also with which triads of animals participants chose as most similar when given a larger set of animals. Similarly, Rosch, Simpson, & Miller (1976) found a correlation between the order in which participants named items and other participants’ prototypicality ratings of those items. However, verbal fluency data have generally not been analyzed in conjunction with data from explicit tasks.

There are two primary goals of this chapter. The first is to present a new technique for extracting semantic clusters from verbal fluency data that reduces some of the problems with currently existing techniques. The second is to elucidate the differences between the semantic features people use when they are explicitly asked to group and those they use during automatic lexical retrieval.

Explicit vs. implicit lexical access.

For a previous study (see Isacoff, Liu, Hou-Imerman, & Stromswold, 2013), I tested over 140 monolingual English-speaking adults. Half of the participants (the “Verbal Fluency” participants) performed a 60 second verbal fluency (i.e., they had 60 seconds to list animals). The other half (the “Explicit Grouping” participants) was given a list of the 20 most frequent Verbal Fluency animals. They were told to write down categories in which they could group the animals and to write down the appropriate animals next to each category. An experimenter instructed the participants that there were no right or wrong answers, that an animal could go in multiple groups or no groups, and that a group could have any number of members. Participants were not guided on what types of features or how many features to use.

The Explicit Grouping participants cumulatively produced forty-five distinct features. There was considerable overlap in participants’ groupings. Although I did not specify which types of features to use, all but three of the forty-five generated features were (broadly-speaking) semantic. [3] Although studies of lexical retrieval failures such as speech error studies (e.g. Fromkin, 1980; Moller, Jansma, Rodriguez-Fornells, & Munte, 2007) and tip-of-the-tongue studies (e.g. Brown & McNeil, 1966; Schwartz & Metcalfe, 2011) suggest the existence of a phonological route to *lexical retrieval*, not one Explicit Grouping participant used a truly phonological feature (e.g., onsets, number of syllables, stress). Furthermore, no participant used the perceptual features color, shape, smell, or touch (e.g., soft) to group animals.

[3] Only one participant used an orthographic feature (NUMBER OF VOWELS). One participant used a grammatical/phonological feature (animals that SOUND THE SAME SINGULAR OR PLURAL, i.e., animals like *fish* and *deer* that have the same singular and plural form); and one participant used sound (DISTINCT SOUND, i.e., animals that make distinctive sounds).

Further, the feature PET was produced by over half of the participants. Twelve features were produced by ten or more participants, and fifteen features were produced by five or more participants. Twenty-two features were produced by more than one participant. Henceforth, I will refer to these 22 features as the Explicit Grouping Features. Table 1.1 shows the number of Explicit Grouping participants who grouped by each of these 22 feature (e.g., 19 participants grouped by MAMMAL).

Although all Explicit Grouping features were semantic, the features otherwise varied widely. Roughly speaking, there were biological, features (MAMMAL, REPTILE, BIRD, RODENT, FELINE, APE, HERBIVORE, CARNIVORE, QUADRUPED); habitat features (FARM, HOUSEHOLD PET, BACKYARD, CIRCUS, AFRICA, WATER); a behavior feature (FLIES); human use features (EATEN, RIDDEN); and descriptive features (WILD, SCARY, DISGUSTING, LARGE).

Table 1.1 Number of Explicit Grouping participants who grouped by each feature

Feature	**% participants grouping by each feature**
PET	51
WILD	30
LARGE	27
MAMMAL	27
FARM	25
AFRICA	21
REPTILE	20

SCARY	20
CARNIVORE	18
WATER	18
HERB	17
QUADRUPED	14
FLIES	10
EATEN	8
RODENT	7
FELINE	6
BIRD	4
APE	3
BACKYARD	3
CIRCUS	3
DISGUSTING	3
RIDDEN	3

Collectively, the Verbal Fluency group named 174 distinct animals, with participants naming an average of 18 animals. Given that the Explicit Grouping participants only grouped the twenty animals, I manually tagged the remaining 154 animals with the Explicit Grouping features. In tagging the animals, I used encyclopedic information when possible (e.g., to tag mammals) and my own intuitions when encyclopedic information was not available (e.g., to tag disgusting animals). To test the generalizability of the way I tagged animals, 17 additional native-English-speaking adults were given a semantic verification task in which they were asked whether each of the 22 Explicit Grouping features applied to each of the most frequent 64 animals (e.g., whether a dog was a pet). (The most frequent 64 animals accounted for 85% of all instances of animals named by Verbal Fluency participants.) The concordance rate between these judgments and mine was 94%, suggesting that features are consistent between individuals.

Table 1.2 shows the number of distinct animals tagged with each of the Explicit Grouping features. For example, two-thirds of animals were mammals. Statistically speaking, then, it is likely that participants would have named multiple mammals consecutively even if they named animals randomly. In contrast, only 5% of animals were from the ape family. Therefore, it is unlikely that participants would have named multiple apes consecutively if they were naming animals randomly.

Table 1.2. Percent of Verbal Fluency animals tagged with each Explicit Grouping feature

Feature	**% animals with each feature**
CARNIVORE	68
MAMMAL	65
HERBIVORE	64
QUADRUPED	63
WILD	58
LARGE	45

SCARY	33
PET	21
BACKYARD	20
WATER	20
DISGUSTING	16
AFRICA	15
FLIES	15
BIRD	14
EATEN	14
FELINE	10
RODENT	9
FARM	8
REPTILE	7
RIDDEN	6
APE	5
CIRCUS	5

Clustering analysis

I transformed each of the 174 verbal fluency animals into a set of twenty-two binary values, corresponding to the twenty-two Explicit Grouping features (e.g. WHALE = +MAMMAL, -PET, -FELINE, +WATER, etc.). I operationalized clustering as two or more consecutive positive instances of a single feature and, as demonstrated in Figure 1.3, calculated mean cluster size.

To test whether Verbal Fluency participants semantically clustered at above chance level, I compared the mean cluster sizes of actual lists with randomized lists.

Table 1.3A represents a toy example of a single participant's list of animals. In this example, the participant generated five clusters: two consecutive mammals, another six consecutive mammals, three consecutive pets, four consecutive felines, and four consecutive water animals. The participant's mean cluster size is 3.8 ((2 mammals + 6 mammals + 3 pets + 4 felines + 4 water)/5 clusters). Table 1.3B represents the randomized version of the participant's list of animals. In the randomized toy example, the mean cluster size is 3.33 ((5 mammals +3 mammals +2 water)/3 clusters).

Table 1.3A. Comparison of clustering in toy example

	Mammal	Pet	Feline	Water
Whale	✓			✓
Dolphin	✓			✓
Shark				✓
Fish		✓		✓
Dog	✓	✓		
Cat	✓	✓	✓	
Lion	✓		✓	
Tiger	✓		✓	
Cheetah	✓		✓	
Elephant	✓			

Table 1.3B. Randomized version of animal list

	Mammal	Pet	Feline	Water
Lion	✓			
Whale	✓			
Cat	✓			
Shark				✓
Fish				✓
Tiger	✓			
Elephant	✓			
Dolphin	✓			
Dog	✓			
Cheetah	✓			

Returning to the actual data, I randomized each participant's animal list and calculated mean cluster size for each of these randomized lists. I conducted paired *t*-tests comparing these randomized cluster indices with actual cluster indices. As shown in Figure 1.1, these analyses revealed that collapsing across features, mean cluster size was significantly greater in the actual lists than in the randomized lists (3.7 & 3.3, respectively, *t* (71) = 5.7, *p* = .001).).

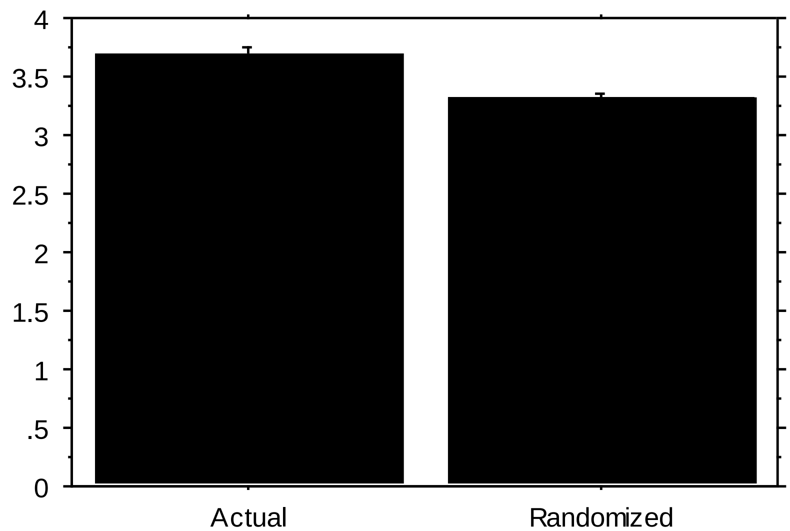


Fig 1.**1** Mean cluster size (error bars = SEM, *p* = .001)

Clustering of Individual Semantic Features. As shown in Table 1.4, similar analyses conducted on each individual feature revealed significant clustering for 12 features, with two features (WILD, LARGE) playing a large role in lexical access, seven features (PET, FELINE, AFRICA, RODENT, SCARY, FARM, EATEN) playing a moderate role, and three features (DISGUSTING, REPTILE, WATER) playing a modest role.

Table 1.4. Explicit Grouping features used in clustering. ($p < .001$)

	Mean Cluster Size	
FEATURE	**T statistic**	**Cohen's d**
WILD	5.53	0.90
LARGE	5.98	0.86

PET	7.18	0.79
FELINE	5.54	0.78
AFRICA	6.22	0.74
RODENT	4.46	0.61
SCARY	4.52	0.58
FARM	4.54	0.57
EATEN	4.09	0.50
DISGUSTING	4.28	0.49
REPTILE	3.40	0.49
WATER	3.84	0.47
CARNIVORE	2.80	0.37

Relationship between features explicitly named and features implicitly used. Finally, I investigated the degree of overlap between the explicit features used in the Explicit Grouping task and the implicit features used in the Verbal Fluency task. As shown in Figure 1.2, a regression analysis revealed a significant but not perfect overlap between the number of Explicit Grouping participants who used a feature to group animals and that feature's effect size in the verbal fluency task, with about a third of the variance of the Verbal Fluency effect sizes accounted for by the frequency of the Explicit Grouping features. Even if the outlier corresponding to PET is removed, more than half of the variance of the Verbal Fluency data is still unaccounted for by the Explicit Grouping data, suggesting that to some extent, different features are used for explicit and implicit lexical access.

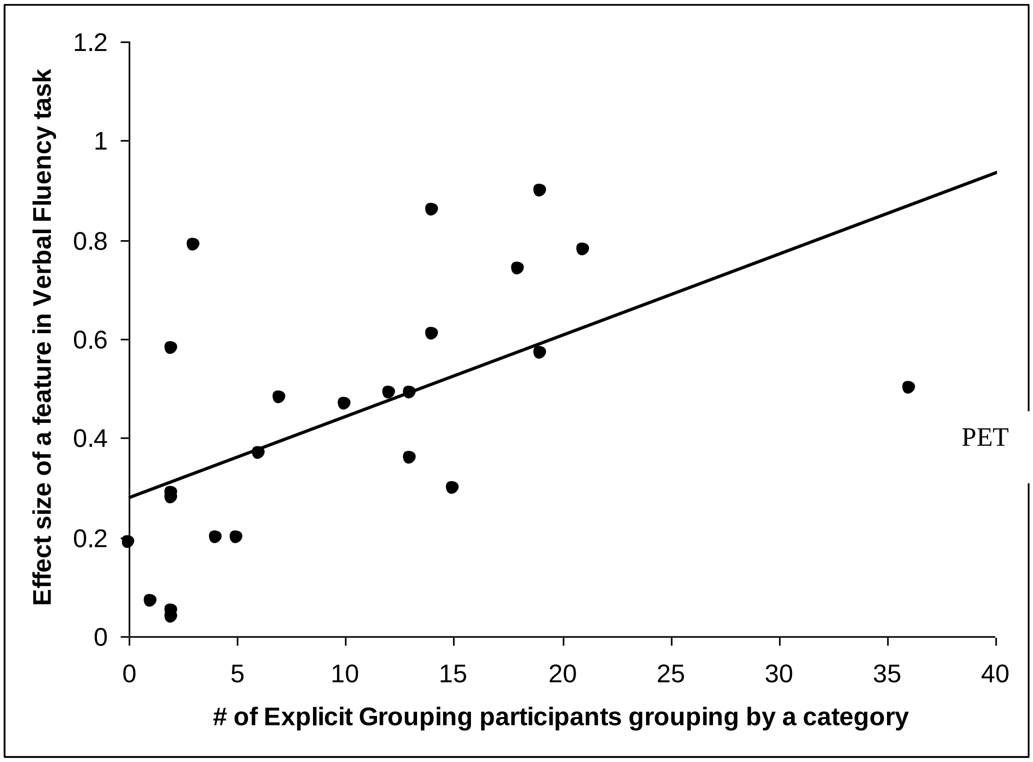


Fig 1.2. Overlap in importance of features in Explicit Grouping and Verbal Fluency tasks

Discussion

Intuition is a widely-used tool for investigating mental content in linguistics, philosophy, and psychology (see Goldman, 2007, for a review), and I struggled with how much weight should be given to features being intuitively compelling. Since the fifteenth century when Descartes declared the mind fundamentally knowable by self-reflection, some philosophers have argued that intuition is a valid window into cognition. Indeed, Kripke (1980) argued that intuition is ultimately the most conclusive evidence for investigating mental content. Similarly, many linguists (e.g., Chomsky, 1965) have relied heavily upon native speaker intuitions, and cognitive psychologists have frequently employed participant judgments in investigating the structure of concepts, categories, and the lexicon. After all, Rosch's prototypes would hold little weight if people did not concur that APPLE is a more typical fruit than OLIVE. Gelman's Essentialism would fall flat if people did not agree that natural kinds carry a causal essence. And so on.

On the other hand, psychological science prides itself on poking holes in folk psychology. If Descartes were correct that intuition is "indubitable," psychology would be superfluous. Consequently, epistemologists past and present have cautioned against an over-reliance on intuition (see, for example, papers in Depaul & Ramsey, 1998). Crucially, these admonitions are consistent in their criticism against intuition uncorroborated by empirical evidence, rather than against any appeal to intuition. In my approach, I sought to use empirical measures to account for intuitions. Given the enormity of the intuition problem, I do not pretend that my approach is perfect. Rather, I argue that I have addressed some of the limitations of previous work.

In particular, my method of analyzing verbal fluency data makes a novel contribution to the study of semantic clustering. One problem I identified with some previous work is that researchers intuited the existence of semantic clusters and/or imposed intuitive feature labels on clusters without empirical support. As I argued earlier in the chapter, the problem with relying solely on intuitions in investigating semantic clusters or features is that doing so can lead to over-identifying or under-identifying clusters. My results suggest that my concern is valid. In particular, studies using both of the methods of analyses described in the introduction (e.g., Borge-Halthoefer, et al., 2009; Troyer et al., 1997) suggested that people cluster using the features MAMMAL and BIRD. My results suggest that this is not the case. Furthermore, I demonstrated that people do cluster using features previously overlooked by other researchers (e.g., WILD, LARGE).

Also worrisome is research that does not consider intuition at all. For example, as discussed in the introduction, Goni et. al's BEAR AND POLAR feature does not mesh with my intuitions about category structure, and the feature UNCLASSIFIABLE does not seem to mesh with Goni et. al's intuitions, either.

My finding that adults semantically cluster by features suggests that neither a purely atomic model of lexical meaning (e.g., Fodor, 1970) nor a network model that does not include node labels (e.g., Collins & Quillian, 1969) fully captures the data. Rather, the data is consistent with a model in which atomic words are linked in a network through features.

In my Explicit Grouping task, people overwhelmingly used semantic features to group animals. Although the features most frequently used to group animals tended to be the same features used in automatic lexical retrieval, there were some notable exceptions. Whereas MAMMAL was the third most frequent feature used in the Explicit Grouping task (produced by over a quarter of participants), it was not important in the Verbal Fluency task. Conversely, only two Explicit Grouping participants grouped animals by the feature DISGUSTING, but this feature was important in the Verbal Fluency task. This contrast is consistent with my general finding that participants in the Verbal Fluency task appeared not to have relied on biological features (e.g., MAMMAL, BIRD), nor on purely functional features (e.g., RIDDEN, CIRCUS), but rather on the most personally-relevant features (e.g., WILD, LARGE, PET).

My feature set contains not only IS-A features (e.g., a DOG IS-A MAMMAL), but also descriptions (e.g., SCARY), habitats (e.g., BACKYARD), things animals do (e.g., FLIES) and things done to animals (e.g.. EATEN, RIDDEN). The evidence that word meaning includes a network of IS-A links is mixed. Collins and Quillian (1969) found that people are generally faster at verifying statements that require traversal of only one IS-A link (e.g., “A robin is a bird”) than two IS-A LINKS (e.g., “A robin is an animal”). However, in a similar study, Rips, Shoben, and Smith (1973) found an exception: participants were faster to verify “a dog is an animal” than “a dog is a mammal.” They argued that items whose prototypes are more similar are easier to verify as being of the same type, and that items with fewer intervening IS-A links typically have closer prototypes, but that their results follow from the fact that DOG is a more typical animal than mammal.

In keeping with my claim that higher-level hierarchical categories can serve as features for lower level category words, it seems plausible that these IS-A links exist for explicit reasoning, but that when asked to verify sentences swiftly, people use probabilistic information as suggested by Rips and colleagues. Another way of viewing the IS-A/non IS-A feature distinction is that *ontological* features can only be instantiated as IS-A links. What an animal does and where it does it are not relevant to its ontological status. It is notable that the Explicit Grouping features (including those that were significant in the Verbal Fluency task) are composed of both more ontological/less salient and less ontological/more salient features. This finding is consistent with previous research showing that adults’ similarity judgments are not constrained by their notions of ontological similarity. For example, when adults are presented with triplets of items containing a target item, a taxonomically-related match, and a thematically-related match and are asked to use knowledge about the target item to make inferences about the matches, people can make inferences along both lines (e.g., Ross & Murphy, 1999).

In contrast with some previous methods for analyzing verbal fluency data that only allowed an item to be tagged with a single feature, the flexibility of my system permitted animals to be tagged with as many features as my Explicit Grouping participants saw fit. This is an advantage because there is no reason to believe, for example, that "whale could only be organized as a MAMMAL or as a WATER ANIMAL, but not both. Additionally, previous research has suggested that the ability to switch between features in a verbal fluency task is a sign of normal cognitive functioning (e.g., Unsworth, Spillers, & Brewer, 2011). One aspect of the switching mechanism is particularly compatible with the multiple-feature assumption that underlies my method. It seems likely that participants sometimes transition between clusters via an item that shares one feature with a previous cluster and one feature with a subsequent cluster. For example, a participant could begin with a PET cluster (DOG, CAT, FISH) and then transition into a WATER animal cluster (FISH, WHALE, CRAB), with FISH fitting into both of these clusters. Use of multiple features captures this phenomenon, whereas more traditional methods (e.g., Troyer et al., 1997) cannot capture gradual switching.

Although feature lists are beneficial for highlighting semantic clusters, they are limited in that they do not capture the causal relations between features (Barsalou & Hale, 1993). Murphy and colleagues found that concepts composed of causally-related features (e.g., DRIVES IN JUNGLES fits with MADE IN AFRICA better than with MADE IN THE ARCTIC) are easiest to learn (Murphy & Wisniewski, 1989, Murphy & Allopenna, 1994). However, Medin and colleagues (Medin, Wattenmaker, & Hampson, 1987) demonstrated that in tasks in which participants grouped novel stimuli, participants tended only to group along one dimension at a time, even when the experiment is rigged to encourage people to group along more than one dimension simultaneously (e.g., by including some fuzzy categories or by having experimenters highlight the causal relationships between features). Medin et al's findings are consistent with the developmental finding that the ability to categorize along multiple dimensions simultaneously (e.g., grouping things of the same shape and color) develops much later than the ability to categorize along a single dimension (e.g., grouping by just shape or just color, see Cartwright, 2002). These findings suggest that my Verbal Fluency participants could truly have been clustering along unidimensional lines, which my methodology would capture. Nonetheless, there are clear relationships between some of my features (e.g., BIRD and FLIES), and it is possible that feature relationships could be incorporated into a future model (e.g., through the use of principal component analysis). One ramification of such a model could be a reduction in the number of significant features, in that it is possible that more than one of my current features could be subsumed under a single feature heading.

An additional limitation of my clustering method is that I used binary features, which do not capture the graded nature of category membership (see Rosch et al., 1976). For example, it could be that "dog" is a better example of a pet than "turtle" which is a better example of pet than "lion." Future work could have participants rate animals on a non-binary scale for each feature and incorporate these rankings into the search for clusters.

A comprehensive theory of meaning must reconcile stable lexical representation with flexible word use. Lakoff and Johnson (1980) observed that a waiter can use "the ham sandwich" to refer to the person who ordered it, and Barsalou (1983) further demonstrated that people can use words as *ad hoc* metaphors. Perhaps future research could investigate whether people more frequently employ explicit or implicit features in metaphor use. Additionally, the intractability of lexical meaning has been a central theme in the philosophy of language and related scholarship. Frege (1884) asserted that words only have meaning within the context of a proposition. Wittgenstein (1921) assented with his thought experiment demonstrating the indefinability of "game," and subsequently, linguists (e.g., Labov, 1973) and psychologists (e.g., Malt, 1994) empirically demonstrated that people use words like "cup" and "water" in intractable ways. I suggest that the problem of meaning becomes more tractable when the links between words are malleable, changing with context. Exploring which types of features are important in different pragmatic contexts would provide further insight into the nature of lexical representation and access.

Finally, one notable finding of my study is that there was a great deal of overlap in the features by which participants grouped animals, both in the Explicit Grouping task and in the Concordance task. These results suggested that findings from these tasks could be generalized to a broader population, at least within the same language and culture. However, it was less clear whether these findings would generalize to a different linguistic and cultural population. Therefore, to determine the effects of language and culture on the Explicit Grouping and Verbal Fluency features, I conducted a follow up experiment.

Lexical access in Mandarin-speaking adults

Linguistic relativity continues to be a hot-button issue. Pro-Whorfian scholars argue that lexical differences between languages cause speakers of these languages to think differently about objects (e.g., Lucy & Gaskins, 2001), time (e.g., Boroditsky, 2001), space (e.g., Casasanto, 2008), and color (e.g., Mo, Xu, Kay, & Tan, 2011). Anti-Whorfian scholars counter that in their own work, they have not found conceptual differences between speakers of different languages (e.g., Iwaski, Vinson, & Vigliocco, 2010); that lexical differences found in Pro-Whorfian studies do not extend to conceptual differences (e.g., Slobin, 1987; Vigliocco, Vinson, Paganelli, & Dworzynski, 2005); that lexical effects on spatial representations are dynamic rather than permanent (e.g., Landau, Desselegn, & Goldberg, 2010) or are task-specific (e.g., Gennari, Sloman, Malt, & Fitch, 2002); that specific Pro-Whorfian studies are methodologically-flawed (e.g., January & Kako, 2007) or do not replicate (e.g., Chen, 2001); or that there are other explanations for cross-linguistic differences (e.g., Li, Abarbanell, Gleitman, & Papagragou, 2011).

To investigate the extent to which the features used for explicit grouping and lexical access are similar across languages and cultures, I had over 100 Mandarin-speaking adults complete either the Explicit Grouping or Verbal Fluency task (See Isacoff et al., 2013). Participants were Mandarin-English bilinguals who were dominant in and completed the experiment in Mandarin. Using both English and Mandarin Explicit Grouping features to analyze the Mandarin Verbal Fluency data, mean cluster size was greater in Mandarin actual lists than in randomized lists. Ten features played a significant role in lexical access. Of these 10 features, one (POULTRY) had a large effect on lexical access; eight had a moderate effect (LIVESTOCK, FARM, ZODIAC, SCARY, TRANSPORTATION, BIRD, LIVESTOCK, BEAST); and one (BIPED) had a modest effect. As shown in Figure 1.3, in contrast with the English data, the number of Mandarin Explicit Grouping participants who used a Mandarin Explicit Grouping feature to group animals was not significantly correlated with that feature's Verbal Fluency effect size (r = .17, p = .50). That is, the distinction between explicit and implicit features holds in the Mandarin data, although this distinction is more pronounced in the Mandarin data than in the English data.

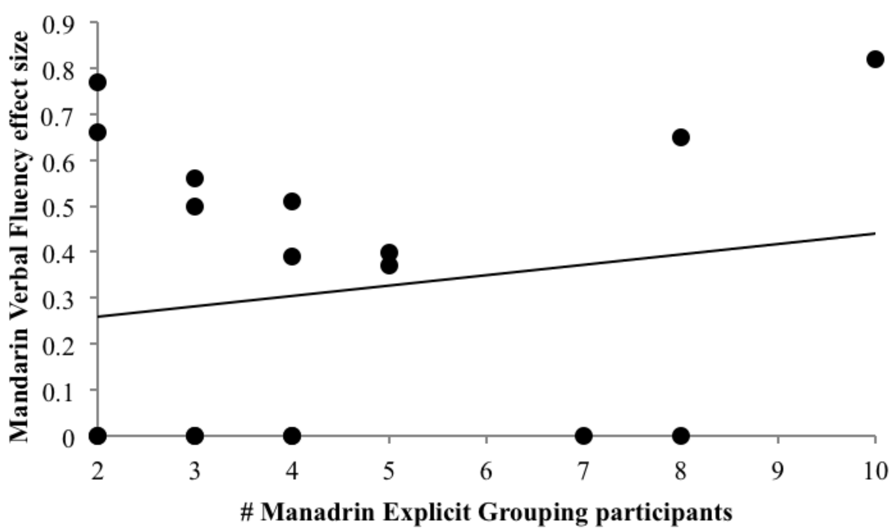


Figure 1.3. Overlap in importance of features in Explicit Grouping and Verbal Fluency tasks in Mandarin

Additionally, we reanalyzed the English Verbal Fluency data with the Explicit Grouping features provided only by the Mandarin participants. Two of these features—BEAST and LIVESTOCK—had a moderate effect on lexical access.

Comparing the English and Mandarin data, there was some overlap and some discrepancies. In the Explicit Grouping task, speakers of both languages grouped by traditional taxonomic features (MAMMAL, REPTILE, BIRD, FELINE, CARNIVORE, HERBIVORE). However, whereas English speakers also grouped by emotional features

(WILD, LARGE, SCARY, DISGUSTING), Mandarin speakers tended to group by cultural features (ZODIAC, MYTHICAL). In the Verbal Fluency task, speakers of both languages clustered by some emotional features (SCARY, BEAST) and by some utility features (LIVESTOCK, FARM). However, consistent with the Explicit Grouping results, English speakers tended to cluster by emotional features and Mandarin speakers by cultural features.

Results also suggest a cross-linguistic incongruity between the features used for explicit grouping and those used for automatically accessing words during a verbal fluency task. In both English and Mandarin, the most important features differed across tasks. For example, MAMMAL was named by over half of English Explicit Grouping participants and over half of Mandarin Explicit Grouping participants; however, it was not used for lexical retrieval in either language. Conversely, DISGUSTING was rarely named by English Explicit Grouping participants but was a significant English Verbal Fluency feature.

Surprisingly, in Mandarin, BIRD, WATER, and BIPED were rarely named by Explicit Grouping participants but were significant Verbal Fluency features. Similarly, there were two significant English Verbal Fluency features (LIVESTOCK, BEAST) that were not named by any English Explicit Grouping participant, and there were three significant Mandarin Verbal Fluency features (FARM, SCARY, TRANSPORTATION) that were not named by any Mandarin Explicit Grouping participant. That is, there were features that speakers of a given language did not generate in the Explicit Grouping experiment but by which speakers of that language clustered in the Verbal Fluency experiment. These data demonstrate that across languages, explicit lexical access features are not isomorphic to implicit lexical access features and that lexical access features are context-dependent.

II. Lexical representation, access, and categorization in typical children

In the previous chapter, I reported that in a verbal fluency task, adults clustered by semantic features. These results suggested that for adults, words (or at least, animal words) are organized around intermediate-level features and that adults can access lexical items by traversing these features.

To investigate whether similar results obtain in children, in previous work (Isacoff & Stromswold, 2011), I used the adult English Explicit Grouping features to analyze verbal fluency data from 375 typically-developing, monolingual English-speaking three- to five-year-olds recruited from the Perinatal Environment and Genetics Interaction (PEGI) study (Stromswold, 2006). Children named an average of six animals (range = 3 – 15). Collapsing across features, none of the individual age groups—three-, four-, or five-year-olds—nor all ages grouped together, clustered more in the actual lists than in the randomized versions of these lists (all paired *t*-test *p*s > .8). Furthermore, none of the age groups, nor all ages grouped together, clustered by any individual feature in the actual lists more than in the randomized versions of these lists at the .01 level, although one feature, BIRD, approached significance for five-year-olds ($p = .03$).

There are at least four possible accounts for why I did not find clustering in the three- to-five-year-olds' data.

Account I: Within the animal domain, preschool-age children lack the declarative knowledge to access animal names at the intermediate level. In other words, children do not access multiple consecutive pets because they do not know what pets are. This account is consistent with two possibilities. First, young children may lack the structure necessary to organize animal words in an adult-like fashion, and this structure may develop with the acquisition of intermediate level knowledge. This possibility is consistent with evidence that knowledge acquisition affects category structure (e.g., Carey, 1999; Sheng, McGregor, & Marion, 2006). This account is also consistent with the possibility that young children have an adult-like structure but that they do not have the knowledge to fill this structure. This possibility is consistent with evidence that knowledge acquisition does *not* affect underlying structure (e.g., Spelke, 1991).

Account II: Preschool-aged children do have intermediate level declarative knowledge about animals but animal words are not organized around intermediate level categories and/or they do not access words via intermediate level categories.

Account III: Preschool-aged children do have intermediate level categories in their lexicons and/or they access animal names by traversing intermediate level categories (i.e., they cluster). However, they do so by different features than adults. Recall that I looked for clusters of animals sharing features generated by *adults* in the Explicit Categorization task. It is possible that children would have generated different features in an Explicit Categorization task and that had I input different features into my clustering algorithm, I would have found that children do indeed cluster by at least some of these features. In particular, one possibility is that children cluster by thematic rather than taxonomic features. In this chapter, I empirically investigate these accounts. Doing so will provide insight into developments in declarative knowledge, the lexicon/lexical access, and the nature of children's categories.

Thematic-taxonomic shift

Two types of categories have been widely investigated in the cognitive development literature. As introduced in the Forward, thematic categories are organized around relations (e.g., DOGS eat BONES), whereas taxonomic categories are organized around shared properties (DOGS and CATS both share the properties of animals). Although both taxonomic and thematic categories are stored in semantic memory (rather than in episodic memory), there is evidence that these types of categories are stored differently. For example, in a study in which aphasic patients named pictures, Schwartz et al. (2011) found that patients with damage to the left anterior temporal lobe (associated with naming) were more likely to make taxonomic errors (e.g., calling an apple "pear"), whereas patients with damage to the left temporoparietal junction (associated with mental states) were more likely to make thematic errors (e.g., calling an apple "worm"). Similarly, in a picture-matching task with unimpaired adults, Kalenine et al. (2009) found that the temporoparietal region was only activated when participants were making thematic matches. These neurocognitive results support behavioral evidence (described below) for a thematic-taxonomic distinction.

A longstanding question in cognitive development is whether children around age seven undergo a shift from thematic-to-taxonomic categorical structure (e.g., Piaget, 1962, Vygotsky, 1962). Markman (1989) characterizes this development as a shift from placing objects together using criteria that are relational (e.g. spatial, temporal, or causal) to placing objects of the same kind together.

Although even infants have some taxonomic categorization (Quinn, Eimas, & Rozenkrantz, 1993) and even adults have some thematic organization (Murphy, 2001), there is strong converging evidence for a shift in category *preference*, which is apparent in children's performance on word association tasks (e.g., Cronin et al., 1986), picture-pairing tasks (e.g., Denney & Moulton, 1976), match-to-sample-tasks (e.g., Dunham & Dunham, 1995), and lexical decision tasks (e.g., Nation & Snowling, 1999).

In the Explicit Categorization task outlined in the previous chapter, adults did not generate any thematic features. Therefore, if preschooler verbal fluency participants exclusively accessed animals via thematic categories (e.g., accessing "cat" followed by "mouse" because both are part of a CATS CHASE MICE category), this would explain why I did not find clustering in their data when I used the adult features.

Why thematic?

Inhelder & Piaget (1959, 1964), Nelson (1979), and Fenson, Vella, and Kennedy (1989) argue for the primacy of thematic associations due to their significance in everyday life (i.e., their grounding in schemata) and because children practice thematic relations in their spontaneous play (Nelson & Seidman, 1984). A thematic preference in young children is consistent with their mental representations being organized around events (i.e., episodic memory) rather than around abstract knowledge (i.e., semantic memory, see Mandler, 1979).

Why taxonomic?

For several reasons, one might predict a *taxonomic*, rather than thematic, preference in children. First, taxonomic relationships are important for efficient mental processing (Fenson, Vella, & Kennedy, 1989; Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976). For example, Melkman, Tversky, & Baratz (1981) suggest that as children develop, they increasingly encode events in anticipation of retrieval, for which abstract information is necessary, and thus they eventually prefer a conceptually-based taxonomic organization. Second, taxonomic categories are more consistent with human language than thematic categories (see Anglin, 1977; Markman, 1989; Waxman, 1991). Finally, unlike most thematically related items, taxonomically related items tend to share perceptual features, which, many researchers have argued, are particularly salient to young children (see Clark, 1983). This perceptual similarity is especially strong for items sharing the basic level or *subordinate* level[4] (see Smith & Heise, 1992; Tversky, 1985).

There is some evidence for two shifts: a taxonomic/perceptual-thematic shift around age three, and a thematic-taxonomic shift around age seven. In one study, Daehler, Lonardo, & Bukatko (1979, Experiment 3) held up a real or toy object (e.g. a spoon) for twenty-two-, twenty-seven-, and thirty-three-month-olds and asked them to "find the one that goes with this one" from among four other objects. The four objects consisted of three unrelated items and one target item, which was either identical to the standard; shared the basic level with the standard; shared the superordinate level with the standard; or was thematically related to the standard. Performance in all conditions improved with age, but in each age group, children chose the target item most often in the identity condition, followed by the basic level condition, then the superordinate level condition, and finally the thematic condition. Daehler et al.'s results suggest that taxonomic relations are more salient than thematic relations to toddlers. One limitation of the study was that children were initially trained in the identity condition. Scott, Serchuk, & Mundy (1982) suggest that having children respond in different ways within the same condition in the same experimental session could bias results and that training children in

[4] The subordinate level is beneath the basic level (e.g., COLLIE is more specific than DOG). This dissertation does not investigate subordinate categories.

the identity condition could bias children toward perceptual matching. This bias would lead to taxonomic responses, given that taxonomically related items tend to share perceptual characteristic more than thematically related items do.

Results from a study by Waxman and Namy (1977) suggest a taxonomic preference in two-year-olds but a thematic preference in four-year-olds. In their match-to-sample task, two-year-olds selected the taxonomic choice more than would be expected by chance. However, four-year-olds selected the thematic choice more often than three-year-olds.

There is even some evidence that a taxonomic bias persists in three-year-olds. For example, in Dunham & Dunham's (1995) match-to-sample task, three-year-olds chose the taxonomic item 63% of the time. In a modified version of this task, Dunham & Dunham added an additional choice that was unrelated to the target item. In this condition, 72% of three-year-olds chose the taxonomic item most frequently, significantly more than in the previous task. One difficulty in interpreting these results is that the children's choices could have been based on either *perceptual* similarity or on an abstract, *conceptual* understanding of taxonomic relationships.

Is perceptual part of taxonomic? Many researchers define taxonomic categories as those grounded in either perceptual or abstract features.[5] The argument for subsuming perceptual under taxonomic is that the relationship between perceptual and conceptual features is often non-arbitrary, and these features are sometimes inextricable. (Prototypical) birds appear to have wings (a more perceptual-y feature) because they use these wings in flight (a more conceptual-y feature). (Prototypical) tables have flat

[5] I do *not* subsume perceptual under taxonomic in the studies in this dissertation.

surfaces (a more perceptual-y feature) because flat surfaces are more conducive to having things placed on them (a more conceptual-y feature).

Furthermore, as Mervis & Rosch (1981, pp. 92) note, “the basic [level, a conceptual level,] is the most general level at which (a) a person uses similar motor actions for interacting with category members, (b) category members have similar overall shapes, and (c) a mental image can reflect the entire category.” In other words, perceptual and conceptual features are often inseparable.

This conflation of perceptual and taxonomic categorization in some but not other studies could account for disagreements in the literature about whether children undergo a taxonomic-thematic shift around age three. For example, Daehler et al. (1979) found that two-year-olds prefer taxonomic relations, whereas Scott et al. (1982) found that two-year-olds prefer thematic relations. Fenson et al. (1989) point out that Daehler et al. used perceptually similar taxonomically related items (e.g., FORK and SPOON), whereas Scott et al. used perceptually-dissimilar taxonomically related items (e.g., BLOCKS and DOLLS). Given children’s sensitivity to perceptual similarity (see Clark, 1983), it is possible that when items are perceptually different, taxonomic relations are not salient and children group thematically. However, when items are perceptually similar, children use these perceptual cues to group taxonomically.

Despite the close relationship between perceptual and conceptual features, several studies have attempted to tease apart the roles of these two types of features. Fenson, et al. (1989) asked two- and three-year-olds to match a standard picture with one of five target pictures. The targets were 1. Perceptually-similar on the same basic level, 2. Perceptually-dissimilar on the same basic level, 3. Perceptually-similar on the same superordinate level, 4. Perceptually-dissimilar on the same superordinate level, or 5. Perceptually-dissimilar but thematically related. At both ages, children chose the basic level matches most frequently, the perceptually-similar superordinate level match and the (perceptually-dissimilar) thematic match occasionally, and the perceptually-dissimilar superordinate level match most rarely. These results suggest that although two-year-olds can recognize both basic level and superordinate level relationships, this ability is largely perceptually bound.

Furthermore, there is evidence that children's perceptual boundedness persists when they hear a novel noun attached to an item, an experimental manipulation generally thought to make the abstract, core properties of an object more salient (e.g., Deak & Bauer, 1996; S.A. Gelman & Coley, 1990; Gentner & Namy, 1999). For example, in their study of three- and five-year-olds, Imai, Gentner & Uchida (1994) either assigned a standard object (e.g. an APPLE) a novel name (e.g. "dax") or just presented the object to children and asked them, "Which one does it go with?" The children chose between an object from the same superordinate category (e.g. a BANANA), an object of the same shape (e.g. a TENNIS BALL), or a thematically related item (e.g. an APPLE TREE). As expected, five-year-olds in the no name condition chose the thematic option more often than the children in the novel name condition, and five-year-olds in the novel name condition were more likely than three-year-olds in this condition to choose the object with the same shape. However, an unexpected finding was that in both conditions, children of both ages chose the perceptual choice most often, and furthermore, three-year-olds chose the category alternative (e.g., BANANA) more often in the no word condition than in the novel word condition. Although the authors do not consider these possibilities, there are at least two possible explanations for these surprising results. One is that a mutual exclusivity constraint could have caused children to be uncomfortable using novel labels to describe objects for which they already have words (see Markman & Wachtel, 1988). A second possibility is that hearing a word could have led children to search for a basic level match (e.g. another apple), which was not an option is this study. Either of these possibilities would have pushed children away from the taxonomic choice in the novel word condition.

Why a shift? Regardless of whether toddlers have a taxonomic preference, it is well established that even infants have some taxonomic categories such as ANIMATE and INANIMATE (Rostad, Yott, & Poulin-Dubois, 2012) and that word learning relies on taxonomic categorization (see Nelson, 1977). Therefore, it is highly unlikely that preschoolers have no ability to categorize taxonomically. A more likely possibility is that around age seven, children undergo a shift in preference towards taxonomic categorization. This shift could be driven by developments in structure and/or access.

One possibility is that the thematic-taxonomic shift is driven by development in one or more specific areas, such as declarative knowledge, language, cognitive flexibility, or metamemorial awareness. Another possibility is that the shift occurs as part of more general cognitive development. In this section, I describe some factors that could play important roles in the thematic-taxonomic shift.

Declarative knowledge. Sheng et al. (2006) propose that the thematic-taxonomic shift is caused by an increase in "world knowledge," although they do not test this hypothesis. World knowledge, or declarative knowledge, is certainly necessary for taxonomic categorization. For example, to group mammals together, a child must either *know* what a mammal is and which animals are/are not mammals, or the child must somehow be able to infer this category membership (e.g., by *knowing* properties shared by members of the category). However, it is unclear whether development in declarative knowledge is instrumental in the thematic-taxonomic shift. If young children who have a wealth of domain-relevant declarative knowledge still have a thematic preference, this would suggest that declarative knowledge is not the critical factor driving the thematic-taxonomic shift.

Language. Another possibility is that the thematic-taxonomic shift is driven by linguistic development. In particular, advances in syntactic sophistication and vocabulary may play roles. Evidence that language drives the thematic-taxonomic shift derives from work on a related shift—the syntagmatic-paradigmatic shift. Woodworth and Schlosberg (1954) were the first to note that when given a word association task, children "tell something. . . about the stimulus word," whereas "adults jump to a related, parallel idea." For example, Woodworth and Schlosberg noted that in response to "table," children said "eat;" adults said "chair." In response to "man;" children said "work;" adults said "woman." Brown and Berko (1960) coined the phrase "syntagmatic-paradigmatic shift," noting that in word association tasks, children tend to respond with words found in syntactic contiguity, whereas adults tend to respond with words from the same grammatical class. The syntagmatic-paradigmatic shift is clearly related to the thematic-taxonomic shift, in that syntagmatic responses reflect events whereas paradigmatic responses reflect like kinds.

Brown and Berko argued that syntactic sophistication predicts extent of paradigmatic responding on a word association task. They gave first- through third-graders and adults a word association task and a "usage test" in which participants were given a nonsense word in a sentence and asked to insert the word in a new sentence. Scores on the usage test were highly correlated with percent of paradigmatic responses in the word association task, suggesting that scores on both tasks reflect development in the organization of parts of speech. It should be noted that both extent of paradigmatic responding and performance on the usage test were highly correlated with age, as are increases in many cognitive abilities in childhood, and therefore, it is possible that paradigmatic responding and syntax are not causally related. It is also possible that syntax development causes a language-specific syntagmatic-paradigmatic shift but does not cause the thematic-taxonomic shift.

However, there are several reasons to believe that syntax could play an important role in the thematic-taxonomic shift. The first is that an increased understanding that *words* fall into categories sharing syntactic properties could lead to a greater awareness that *the things those words represent* fall into categories sharing other conceptual properties (i.e., taxonomic categories). The second is that complex syntax is necessarily hierarchical, and therefore, proficiency in generating syntactically complex sentences could facilitate a preference for categories fitting into a complex taxonomic hierarchy. The third reason is specific to justification tasks, a common measure of categorization development in which participants are asked to justify their categories (e.g., "Why did you put the lion with the tiger?). Increased syntactic facility could make it easier to *justify* why items are members of the same taxonomic category.

Vocabulary is another potential language-based mechanism of change in the syntagmatic-paradigmatic shift and therefore could also play a role in the thematic-taxonomic shift. Cronin (2002) found that rate of paradigmatic responding was correlated with reading level but not with mental or chronological age. In a year-long longitudinal study of first graders, Cronin compared two measures of literacy. The first was a word comprehension task. The second was the Woodcock Reading Mastery Test (Woodcock, 1973), which assesses a child's ability to name letters written in unusual fonts and to read aloud words and pseudowords. Cronin found that when word comprehension was used as a measure of reading ability, but not when the Woodcock score was used, reading level correlated with paradigmatic responding. Cronin suggests that deep knowledge of word meaning leads both to better reading comprehension and to paradigmatic responding, which requires a person to know about the underlying properties of a word's meaning.

Cognitive flexibility and metamemory. Cognitive flexibility is the aspect of executive function that enables a person to think about multiple things at a time or to switch between modes of thinking (Scott, 1962). It encompasses both representation (e.g., knowing that a dog is both a pet and a mammal) and access (e.g., being able to name pets and then name mammals (see Spiro & Jehng, 1990). Metamemory is one's knowledge and awareness of his/her own memory, including information storage and retrieval (Flavell & Wellman, 1977). As discussed in the previous chapter, many researchers have argued that success on verbal fluency tasks is dependent in part on the ability to cluster and switch (e.g., Troyer et al., 1997). Therefore, increased cognitive flexibility could aid in the *ability* to switch between clusters. Additionally, increased metamemorial awareness could result in better clustering and switching *strategies*. For example, a child with metamemorial awareness might consciously think of a category that he knows contains many members (e.g., pets) and consciously name members of that category (e.g., dog, cat, fish). When the child cannot think of more pets, he might consciously think of a different category (e.g., farm animals) and consciously name members of that category (e.g., cow, horse, pig), and so on. One possible explanation for why children did not cluster in my verbal fluency task is that they lacked the cognitive flexibility and/or metamemorial awareness to cluster and switch. This possibility is consistent with my "lexical access" account; that is, preschoolers do not access words from their lexicons the same way adults do.

In terms of the types of tasks generally used to investigate the thematic-taxonomic shift (including the tasks used in the present study), one possibility is that cognitive flexibility is required to create *ad hoc* categories and/or to generate *post hoc* justifications. Under this hypothesis, the ability to generate categories and to describe a relationship between two items speaks more to a child's degree of cognitive flexibility and metamemorial awareness than to the nature of the child's conceptual organization. (See the Discussion section of this chapter for further discussion of *ad hoc* categories and *post hoc* justifications.)

Degree of cognitive flexibility may play a particularly important role in how children categorize when they are given more constraints than in classic match-to-sample or grouping tasks. In one study, Blaye & Bonthoux (2001) first gave children a target (e.g., a mouse) and had children choose between a taxonomic match (e.g., a bird) or a thematic match (e.g., cheese). In the control condition, Blaye & Bonthoux next had children match the target again. In the experimental condition, if children chose the thematic match, Blaye & Bonthoux gave children a scene to prompt taxonomic categorization (e.g., an animal book). If children chose the taxonomic match, Blaye &

Bonthoux gave children a scene to prompt *thematic* categorization (e.g., a mousetrap). Children then had the opportunity to re-match the target based on the new information. Five-year-olds tended to group flexibly (i.e., they grouped the same item taxonomically in one case and thematically in the other) only in the experimental condition, whereas three-year-olds tended to group flexibly in both conditions, although not to the same extent as the five-year-olds in the experimental condition. Four-year-olds tended to group thematically in both trials of both conditions. These results suggest a development from "spontaneous variability to adaptive switching" (Blaye & Bonthoux, 2001, pp. 403).

I am not aware of previous research on the relationship between metamemory and the thematic-taxonomic shift. However, if children's knowledge is organized similarly to adults' but children lack the ability to access this knowledge, an increase in metamemorial awareness could lead to an increased ability to access this knowledge (i.e., to generate taxonomic categories and to justify these categories taxonomically). For example, in deciding how to group items, a child could consciously think about how his lexicon is organized (e.g., "I know that I know that some things are animals"), which could prompt him to group animals together.

Effect of levels

In terms of taxonomic categorization, there is mixed evidence for which develops first, the superordinate level or the basic level (See Hajibayova, 2013, for review). Several studies using both known stimuli (e.g., Daehler, Lonardo, & Bukakto, 1979; Rosch et al., 1976) and novel stimuli (e.g., Mervis & Crisafi, 1982) have found that two-year-olds are more accurate at grouping items sharing the same basic level than those sharing only the same superordinate level. However, Mandler & Bauer (1988) found that 16- to 20-month olds were only able to differentiate items in different superordinate level categories (e.g., dogs vs. cars) not items in different basic level categories (e.g., dogs vs. horses), suggesting that superordinate level categories emerge first. Several factors could account for the disagreement in the literature. First, some of the discrepancy between findings is attributable to coding differences. For example, Rosch et al (1976) did not give a child credit for a grouping if the child only included some members of a category (e.g., if the child only included SHIRT and PANTS but not SOCK or SHOES in a clothing grouping). Mandler, Bauer, & McDonough (1991) argued that Rosch et al's coding scheme was overly strict. Second, some studies (e.g., Mandler et al., 1991) control for perceptual similarity, whereas others (e.g., Rosch et al., 1976) do not. As noted above, basic level category members are often more perceptually similar to one another than superordinate level category members. Therefore, if children categorize based on perceptual similarity, one would expect there to be more of a difference in children's performance on the two levels in experiments *not* controlling for perceptual similarity than in experiments controlling for perceptual similarity. Indeed, children were comparatively better at categorizing at the basic level in Rosch et al's study, which did not control for perceptual similarity than in Mandler et al's study, which did. Finally, differences in task demands could account for the different findings. For example, if young children have superordinate level categories (i.e., they have this competence), they could be able to succeed in Mandler et al's (1988 & 1991) object manipulation tasks (which do not require conscious categorization) but still unable to group in Rosch et al's grouping task (which do require conscious categorization). That is, the discrepancy between these findings could be due to performance demands.

In the current study, I investigate why children in the verbal fluency study did not semantically cluster. I also investigate the development of children's categories and their justifications for these categories. Finally, I explore how categorization development relates to developments in other cognitive abilities.

Methods

Participants

Participants were 91 typically-developing, monolingual English-speaking children ranging in age from 2.19 years to 8.99 years (mean = 5.77 years, *SE* = .17). Forty-nine children were females, and 42 children were males. One additional child was excluded after giving no responses on the first two tasks. Children were recruited from New Jersey schools. The protocol was approved by the Rutgers University Institutional Review Board for Human Subjects.

Procedure & coding

Children were tested in a quiet room with no other children present. Each child completed the entire battery of tasks, in the order presented below, in one twenty-five minute session. All sessions were video recorded for later coding.

General procedure. In my study, I presented children with picture cards and trading card holders and asked them to "put the ones that are kind of the same together." We used this methodology and instructions after a careful review of the procedures used in previous studies. Many researchers have investigated the thematic-taxonomic shift by using a match-to-sample task. In the match-to-sample paradigm, children are presented with a target item, a taxonomic match, and a thematic match and asked which one is "the same" as or which one "goes with" the target. However, as several researchers have noted, the match-to-sample task can bias children towards the thematic choice, because thematic but not taxonomic pairs tend to involve a functional relationship between exactly two items (see Blaye et al., 2010, for discussion). To avoid this bias, and also to more closely approximate the Explicit Grouping task that adults completed (see previous chapter), I instead used a free grouping task.

Researchers who employ either a match-to-sample task or a free sorting task tend to give instructions in one of two ways; they ask children to group items that are either "the same" or items that "go together." Asking children to group items that "go together" is problematic because it can bias them to behave thematically (see Denney & Moulton, 1976; Waxman & Namy, 1997). However, during pilot testing for my study, children seemed to think that "the same" meant "exactly the same." Therefore, I modified my instructions and asked children to group items that were "kind of the same." I used trading card holders because Markman, Cox, and Machida (1981) found that using a "spatially-extended surface" such as a table with no compartments biased preschoolers towards thinking thematically.

Familiarization task. In this task, the child was shown two trading card holders, one filled with Disney cards (see Figure 2.1) and the other with Sesame Street cards (see Figure 2.2).

Fig 2.1. Familiarization task materials (Disney)

Fig 2.2. Familiarization task materials (Sesame Street)

The child was directed to the Disney holder and asked, “Do you know who any of these characters are?” If the child could name at least one of the characters, the experimenter responded, “Good!” If not, the experimenter responded, “That’s okay!” The experimenter then asked the child where the characters were from. If the child did not know the answer, the experimenter explained that they were all Disney characters. Next, the experimenter directed the child’s attention to the Sesame Street card holder and asked the same set of questions. Finally, the experimenter revealed a Mickey Mouse card (see Figure 2.3) and asked the child to name the character on the card. If the child did not say Mickey Mouse, the experimenter said, “This is Mickey Mouse!”

Fig 2.3. Familiarization task materials (Mickey Mouse)

Then the experimenter said, “In this game, we want to put the ones that are kind of the same together. Is Mickey kind of the same as the Disney cards (pointing) or the Sesame Street cards (pointing)? If the child responded “Disney,” the experimenter said, “That’s great! Now you know how to play the game. Let’s try another one.” If the child did not give the correct answer, the experimenter said, “Mickey is kind of the same as the Disney cards. Now you know how to play the game. Let’s try another one.”

Superordinate level task. In this task, as shown in Figure 2.4, children were asked to group eight cards depicting black and white line drawings of items that fall into two traditional superordinate categories: food (banana, carrot, sandwich, spaghetti) and clothing (dress, pants, shirt, skirt).

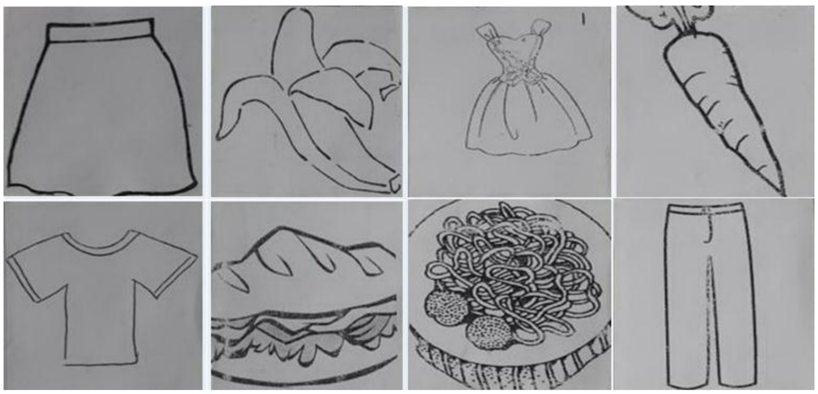


Fig 2.4. Superordinate level cards

Two randomized presentation orders were created, and children were randomly assigned to an order. This task contained three subparts: vocabulary[6], grouping, and justification.

I. Procedure

a. Vocabulary. At the beginning of the task, a stack of empty card holders was placed in front of the child. The experimenter then put down one card, asking the child, “Do you know what this is?” If the child correctly identified the item, the experimenter said, “Good” and put down the next card. If the child answered incorrectly or did not answer, the experimenter said, “It’s a (banana)” and put down the next card. After all eight cards were laid out in front of the child, the experimenter again asked the child to

[6] One reason we had children label the items was because Deak & Bauer (1996) found that preschool children often label line drawings at the Superordinate level (e.g., labeling a banana as “food”), and we wanted to be sure that children understood what we intended our pictures to represent and to clarify if necessary. However, none of our participants labeled any picture at the superordinate level on any task.

name any of the items that the child did not initially name correctly. Again, the experimenter responded with “Good” or “It’s a (banana”).

b. Grouping and justification. Next, the experimenter said, “Remember, in this game, we want to put the ones that are kind of the same together.” The experimenter then took an empty card holder off the stack and asked the child, “Which ones should we put in here?” and allowed the child to place cards into the holder. When the child stopped putting cards into the holder for three seconds, the experimenter said, “Any more or just those?” and gave the child the opportunity to add more cards. This question was repeated until the child did not want to add more cards.

The experimenter then pointed to the filled card holder and said, “Great, how are these kind of the same?” to elicit the child’s justification. The experimenter then picked another empty card holder off the stack and said, “Remember, we want to put the ones that are kind of the same together. Are any more kind of the same?” while directing the child’s attention to the remaining cards. The above procedure was repeated until all cards were used or until the child said that there were *not* additional cards that were kind of the same.

II. Coding.

a. **Vocabulary.** Superordinate vocabulary score was coded as the number of items out of eight that children named correctly on the first attempt.

b. Grouping. FOOD and CLOTHING were determined *a priori* to be the two taxonomic features by which children could group in this task. A child was coded as grouping taxonomically if at least one of his/her groupings of two or more cards contained only food items or only clothing items.[7] No thematic or perceptual features were identified *a priori* and thus it was not possible to code children's groupings by these types of features. In this task and all subsequent grouping tasks, a grouping was not counted if the child appeared to be arbitrarily grouping cards. Specifically, groupings were excluded if a child chose consecutive cards in a row or column without justifying this grouping; if a child grabbed several cards at once without justifying this grouping; or if a child included all items in one grouping.

c. Justification. Each justification was coded in one of four ways: taxonomic (e.g., "they're foods", "they're fruits and vegetables"); thematic (e.g., "I use them in the morning"); perceptual (e.g., "They're long and skinny"); or no justification given (if a child gave no response or gave a non-justification such as "They're just the same.") All justifications from twenty randomly selected participants were independently coded by two coders. Coders agreed on all but one justification (Cohen's kappa = .95).

Perceptual Grouping task. In this task, as shown in Figure 2.5, children grouped twenty "Set" cards, which varied by four dimensions. These dimensions were shape (diamond, oval, squiggly); color (red, green, purple); texture (solid, striped, blank); and number (one, two, three).

[7] In all Grouping tasks, we followed Mandler et al.'s (1991) recommendation of using a minimum of two cards to constitute a group.

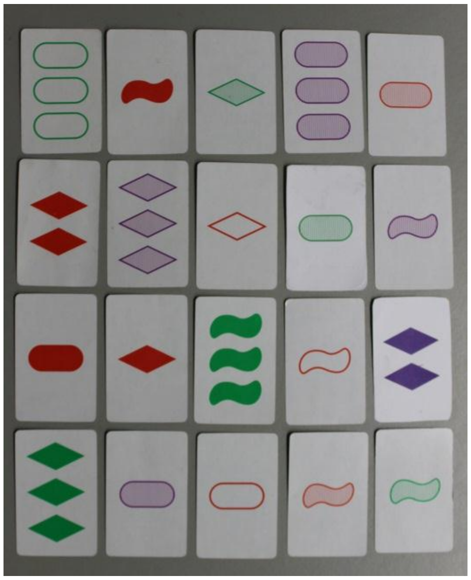


Fig 2.5. Perceptual Grouping cards

Two randomized orders were created and children were randomly assigned to an order.

I. Procedure. The grouping procedure was the same as in the Superordinate level task but without a vocabulary component. The experimenter began by laying out the cards next to the stack of empty card holders. Then the experimenter said, “Remember, in this game, we want to put the ones that are kind of the same together.” The experimenter then took a card holder off the stack and said, “Which ones should we put in here?” The experimenter followed the same procedure as above, giving the child the opportunity to add more cards until the child was finished and then asking for the child’s justification. However, after the child’s justification, the experimenter praised the child and then removed the cards from the holder and added them back to the table saying, “Is there another way that some are kind of the same?” This procedure was repeated until the child said that there were no more groupings.

II. Coding. Responses were coded in several ways.

a. **Grouping.** I coded whether a child grouped at all (i.e., all cards in a least one grouping contained only cards sharing the same member of at least one of the four dimensions (e.g., only RED cards).

b. Switching grouping dimensions. I coded whether a child grouped by different dimensions in different groupings. For example, a child received credit for grouping cards sharing a color (e.g., two red cards) in one grouping and two or more cards sharing a shape (e.g., two diamonds) in another grouping. A child also received credit for grouping by color and shape (e.g., red diamonds) in one grouping and by texture and shape (e.g., striped diamonds) in another grouping. Note that grouping first by “red” and then by “green” was not considered switching dimensions because both red and green are members of the dimension “color”. Additionally, grouping by color and shape (e.g., red diamonds) in one grouping and again by color and shape (e.g., green ovals) in another grouping was not considered switching, because these groupings contain the same dimensions.

3. Justification. I coded whether a child justified by any of the four dimensions (e.g., “They’re all the same color”) or by any member of the four dimensions (e.g., “They’re all red”).

4. Justifying by multiple dimensions. I coded whether a child justified by more than one dimension either within or across groupings. For example, if a child said, “They’re red diamonds,” this was considered justifying by more than one dimension. Additionally, if a child said, “They’re red” for the first grouping and “They’re diamonds” for the second grouping, this was also considered justifying by more than one dimension. However, if a child said, “They’re red” for the first grouping and “They’re green” for the second, this was not considered justifying by multiple dimensions because red and green are both members of the same dimension—color.

Switching grouping dimensions and *justifying by multiple dimensions* were modified versions of the Dimensional Change Card Sort Task (DCCS, see Zelazo, 2006), which is a classic measure of flexible thinking. In the DCCS, children are shown cards (e.g., red circles, green circles, red squares, green squares) that can be grouped in two ways (e.g., by color or shape). In the task, children are given grouping criteria (e.g., put the red ones *here* and the green ones *here*) and then given different grouping criteria (e.g., now put the circles *here* and the squares *here*). The task is taken as a measure of flexible thinking because, in order to succeed with the second grouping criteria, children must inhibit the initial grouping criteria and adapt to the new criteria. My task differs from the DCCS in that children are not given grouping criteria. However, the ability to switch grouping dimensions or to justify by multiple dimensions on my task similarly requires inhibition of previous dimensions and adaptation to new dimensions. Thus, I take these behaviors as indications of flexible thinking. See the Discussion section for a further consideration of how these measures differ from the DCCS.

Intermediate level task. In this task, as shown in Figure 2.6, children grouped twenty cards depicting black and white line drawings of distinct animals: dog, bird, tiger, fish, bear, frog, monkey, hippo, horse, sheep, cow, mouse, giraffe, chicken, duck, pig, cat, worm, elephant, lion. The depicted animals were the twenty most frequent animals named by the 375 three- to five-year-olds who completed the verbal fluency task described above (see Isacoff & Stromswold, 2014).

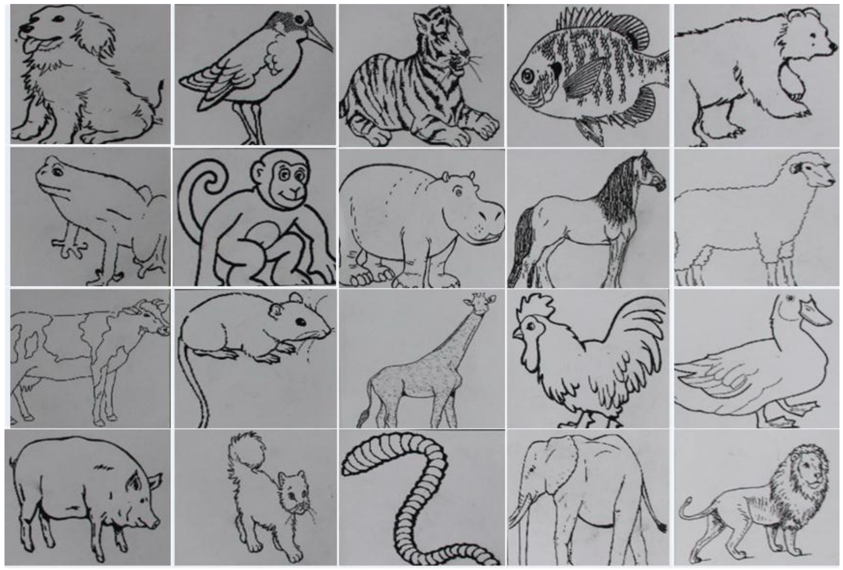


Fig 2.6. Intermediate level cards

Two randomized orders were created, and children were randomly assigned to an order. Like the Superordinate level task, the Intermediate level task contained three components: vocabulary, grouping, and justification.

I. Procedure. To begin, the experimenter elicited the names of the animals from the child using the same procedure as in the Superordinate level task. Next, the experimenter reminded the child, “Remember, we want to put the ones that are kind of the same together” and then used the same grouping procedure as in the Perceptual-level task to elicit groupings and justifications.

II. Coding. Vocabulary score was coded as the number of animals correctly identified on the first try. Given that I did not posit grouping features *a priori*, I did not code children’s groupings. However, just as in the Superordinate level task, justifications were coded as taxonomic (e.g., “They’re mammals,” “They’re pets,” “They’re farm animals”); thematic (e.g., “Cats chase mice,” “Birds eat worms”); perceptual (e.g., "They're round," "They have big ears"); or none (no meaningful justification).[8] All justifications from twenty randomly selected participants were independently coded by two coders. Coders agreed on all justifications.

Basic level task.

I. Procedure. In this task, as shown in Figure 2.7, children grouped black and white line drawings depicting multiple distinct exemplars of the same traditional basic level categories (two cats, two dogs, two giraffes, three elephants, four fish).

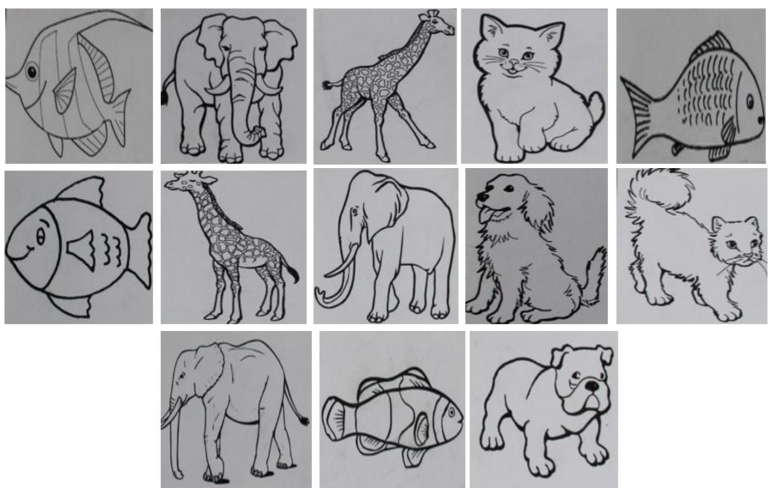


Fig 2.7. Basic level cards

Two randomized orders were created, and children were randomly assigned to an order. Children were not asked to name the animals in this task in order to avoid increasing the salience of the basic level groupings (e.g., saying "dog" could increase the salience of the category DOG). In all other respects, the procedure was identical to that described under *Superordinate level task*.

[8] Children did not realize credit for justifying with, "They're animals," because this justification did not distinguish a grouping from the other cards.

II. Coding.

a. Grouping. Five groupings—CAT, DOG, ELEPHANT, FISH, and GIRAFFE were determined *a priori* to be the groupings by which children could taxonomically group at the basic-object level. Just as in the Superordinate level task, participants were given credit for taxonomically grouping if at least one of their groupings contained only members from one of these categories.

b. Justification. Children's justifications were coded as Basic level taxonomic (e.g., "They're dogs"), other taxonomic (e.g., "They're pets), thematic (e.g., "Cats hate dogs), or Perceptual (e.g., They're furry).

Animal Declarative Knowledge task. In this task, the intermediate-grouping task was reversed in that children were given features and asked to identify appropriate animals. For example, children were asked, "Which ones are pets? Which ones are scary?" and so on. Children were asked a total of twenty questions based on the features generated by adults in the Explicit Grouping task described in the previous chapter

I. Procedure. The experimenter laid out the cards in front of the child in the same order as in the Intermediate level task (see Figure 2.6, above). The experimenter then said, "This game is a little different. I am going to tell you things about animals, and you tell me which animals I'm talking about." The same exclusion criteria applied as in the grouping tasks (i.e., if children only pointed to or named three or more animals in a row, grabbed at piles of animals haphazardly, or included all of the animals, the answer was excluded). Two randomized orders of questions were generated, and children were randomly assigned to an order.

II. Coding. Animal Declarative Knowledge score was coded as the number of questions (out of twenty) for which the child correctly identified at least one animal and did not misidentify any animals. For example, in response to "Which ones are mammals?" a child would receive credit for selecting the lion but not for selecting both the lion and the bird. Children could identify an animal by naming it or by pointing at it. In order to verify which animals counted for each question, I used adult subjects' responses from the verification task described in the previous chapter.

Metamemory task. This task was adapted from Flavell's (1976) test of metamemory. In this task, the child was introduced to two puppets. The experimenter said, "This is Megan, and this is Henry. They are trying to learn some new words. I'm going to tell you some things about them, and you tell me who has the harder job." The child then heard five sets of facts about Megan and Henry. For example, the child was told, "Megan (pointing) is trying to learn 18 new words. Henry (pointing) is trying to learn three new words," for which the correct answer is that Megan has the harder task. (See appendix for list of Metamemory items.) The facts were randomly assigned to each puppet.

Productive Syntax. In this task, the child saw a colored drawing of a playground, shown in Figure 2.8. The child was told, "Here is a picture of a playground. Can you tell me what you see?" After the child described the picture, or if the child did not provide a description, the child was asked, "Can you tell me a story about the picture?" Mean number of morphemes per utterance (i.e., mean length of utterance, or MLU) was calculated using the procedure outlined in Brown (1973).

Fig 2.8. Productive syntax stimulus

Results

Development of cognitive predictors

One goal of this study was to determine the relationship between children's categories and children's other cognitive abilities (language, ability to group and justify perceptual groups, flexible thinking, declarative knowledge, metamemorial awareness). We first investigated the relationships between age and each of these cognitive abilities so that we could later determine whether any of these abilities predict categorization behavior beyond what is predicted by age. Given the number of analyses in this study, we set α = .01

Vocabulary. On the Superordinate level task, children correctly identified an average of 7.37/8 items (92%) on the first attempt (range = 4 – 8, *SE* = .11). As shown in Figure 2.9A, age and Superordinate level vocabulary were highly correlated ($r = .54$, $p < .0005$). On the Intermediate level task, children correctly identified an average of

19.02/20 items (95%) on the first attempt (range = 10 – 20, *SE* = .17). As shown in Figure 2.9B, Intermediate level vocabulary and age were highly correlated ($r = .51$, $p < .0005$). Exclusion of the outlier (vocabulary score = 10) did not change the correlation coefficient or the significance level.

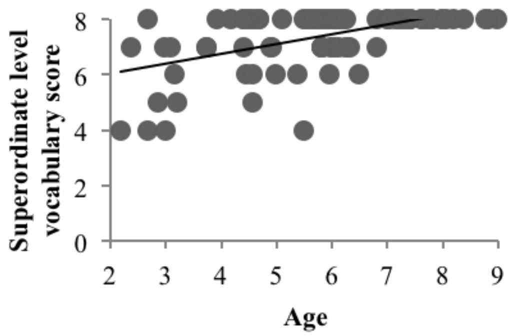


Fig 2.9A. Superordinate vocabulary

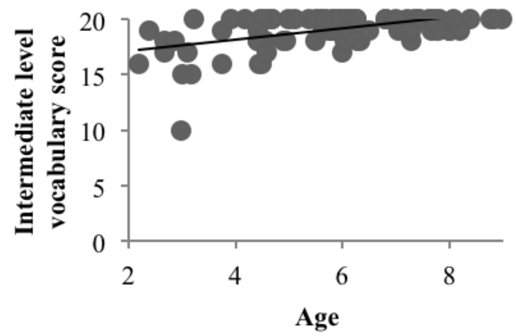


Fig 2.9B. Intermediate vocabulary

Fig 2.9. Age and vocabulary scores

Productive syntax. On the Productive Syntax task, children's average mean length of utterance (MLU) was 4.41 morphemes per utterance (*SE* = .30, range = 0 – 16.25). As shown in Figure 2.10, age and MLU were moderately correlated ($r = .43$, $p < .0005$). Exclusion of the two outliers (MLU = 14.50 & 16.25) had no effect on the correlation coefficient or significance level.

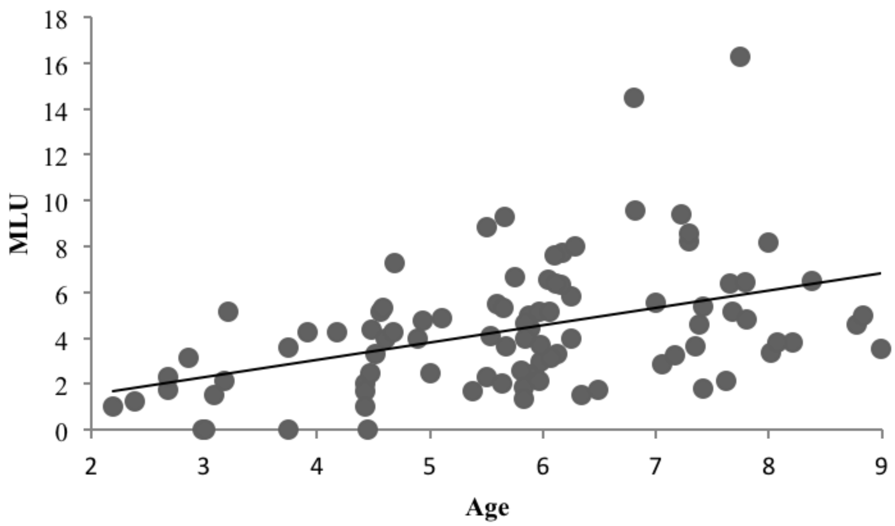


Fig 2.10. Age and MLU on the Productive Syntax task

Perceptual Grouping measures. For this task, we used logistic regression analyses to accommodate our use of binary dependent variables.

1. Grouping. On the Perceptual grouping task, 85 children grouped by at least one of the four dimensions, with seventy-three children grouping by shape, 58 by color, 54 by texture, and 50 by number. As shown in Figure 2.11A, a logistic regression analysis revealed that age did not significantly predict grouping *(p = .10)*[9]. Furthermore, a set of logistic regression analyses with age as the independent variable and each dimension (shape, color, texture, and number) as the dependent variable revealed that there was no effect of age on the dimensions by which children grouped.

2. Switching grouping dimensions. Sixty-seven children switched the dimension by which they grouped. As shown in Figure 2.11B, a logistic regression analysis revealed

[9] Note that although in the graphs we divided the participants into half-year age brackets, in the logistic regression analyses we entered the participants' exact age. Additionally, the depiction of percentages in the figures is merely for clarity. The analyses were conducted as described in the text.

that age had a positive linear effect on switching grouping dimensions (Log likelihood = -42.81, Nagelkerke R^2 = .28, Wald $\chi2$ (1, N = 91) = 14.60, p < .0005, Odds ratio = 2.12).

3. Justification. Eighty-two children justified by at least one dimension, with sixty-five children justifying by shape, 45 by color, 35 by texture, and 24 by number. As shown in Figure 2.11C, a logistic regression analysis revealed that age had a positive linear effect on justification (Log likelihood = -30.05, Nagelkerke R^2 = .26, Wald $\chi2$ (1, N = 91) = 11.33, p < .001, Odds ratio = 2.21). Furthermore, a set of logistic regression analyses with age as the independent variable and each justification dimension as the dependent variable revealed that age had a positive linear effect on children's likelihood to justify by two of the dimensions: Shape (Log likelihood = -49.82, Nagelkerke R^2 = .16, Wald $\chi2$ (1, N = 91) = 9.50, p = .002, Odds ratio = 11.67) and Texture (Log likelihood = -49.98, Nagelkerke R^2 = .25, Wald $\chi2$ (1, N = 91) = 13.52, p < .0005, Odds ratio = 2.00).

4. Multiple justifications. Fifty-five children justified by more than one dimension. As shown in Figure 2.11D, age had a positive linear effect on switching justification dimensions (Log likelihood = -44.16, Nagelkerke R^2 = .42, Wald $\chi2$ (1, N = 91) = 19.98, p < .0005, Odds ratio = 2.79).

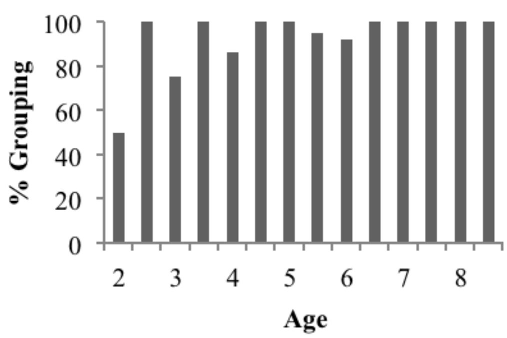


Fig 2.11A. Grouping

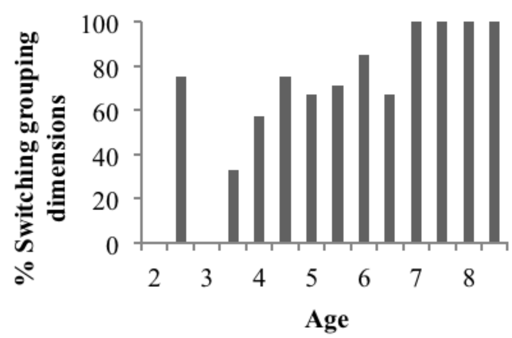


Fig 2.11B. Switching group dimensions

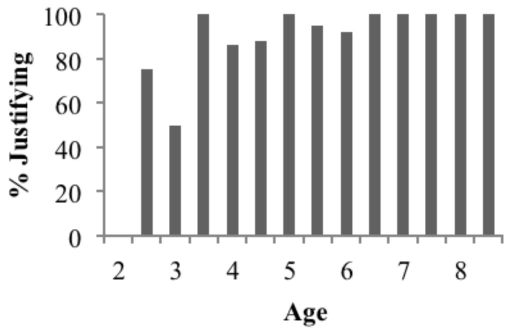


Fig 2.11C. Justification

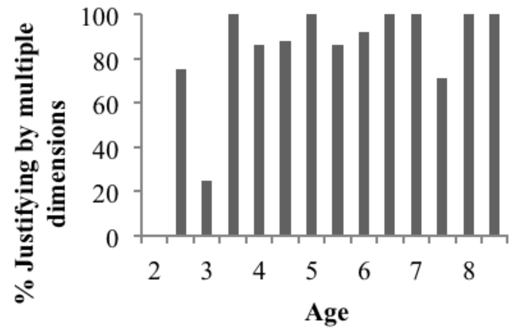


Fig 2.11D. Switching justification

Fig 2.11. Age on grouping & justification on the Perceptual task

Animal Declarative Knowledge. On the Animal Declarative Knowledge task, children correctly answered an average of 15.44/ 20 questions (*SE* = .43, range = 0 – 20). All but one child correctly identified animals on at least four questions, and 81 children correctly identified animals on at least ten questions. As shown in Figure 2.12, age and Animal Declaritive Knowledge score were highly correlated ($r = .66$, $p < .0005$).

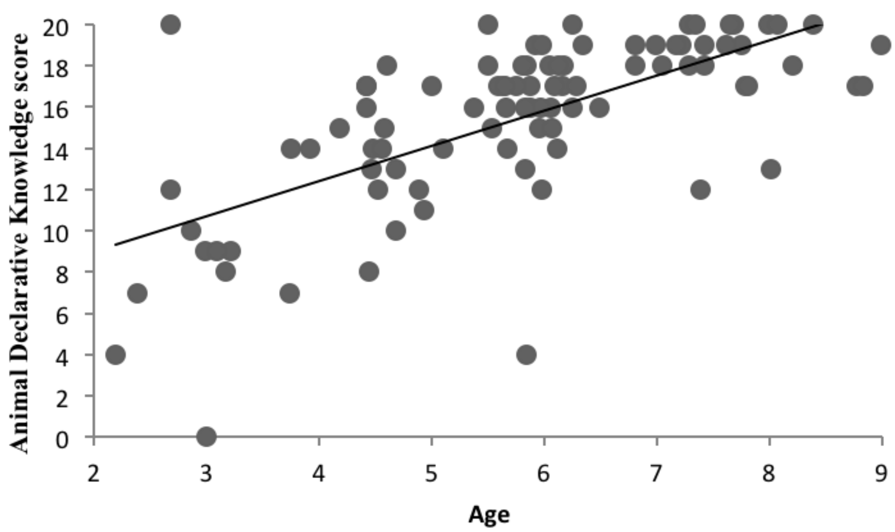


Fig 2.12. Age on the Animal Declarative Knowledge task.

Metamemory. On the Metamemory task, children correctly answered an average of 3.11/5 questions (*SE* = .12, range = 0 – 5). As shown in Figure 2.13, age and Metamemory score were highly correlated ($r = .54$, $p < .0005$).

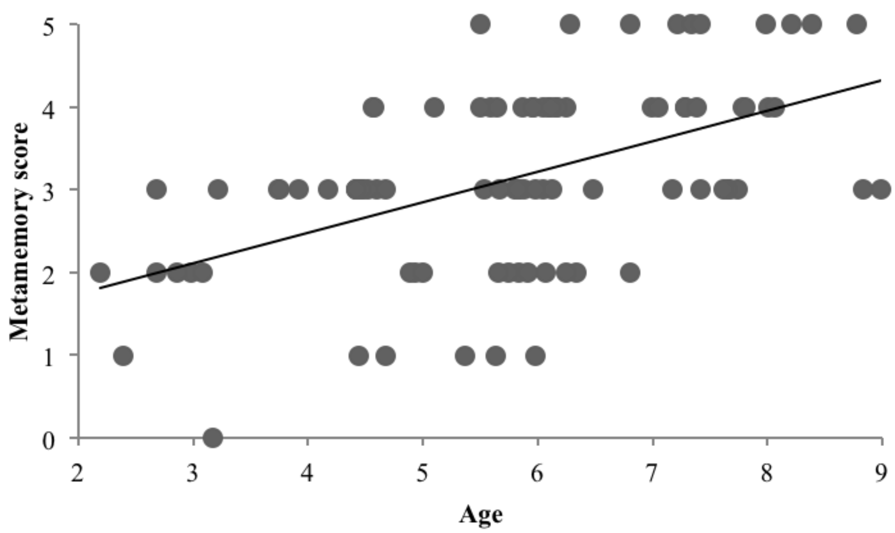


Fig 2.13. Age on the Metamemory task.

Categorization development

To investigate how well age predicted categorization on the Superordinate level, Intermediate level, and Basic level tasks, we conducted a series of simple logistic regression analyses with age as the independent variable and performance on each task as the dependent variable.

Superordinate level task. Seventy-four children grouped taxonomically on the Superordinate level task. As shown in Figure 2.14A, there was a positive linear effect of age on taxonomic grouping (Log likelihood = -34.48, Nagelkerke $R^2 = .30$, Wald $\chi 2$ (1, $N = 91$) = 13.89, $p < .0005$, Odds ratio = 2.30). Of the seventy-four children who grouped taxonomically, fifty-four children justified taxonomically. As shown in Figure 2.14B, there was also a positive linear effect of age on taxonomic justification (Log likelihood = -53.58, Nagelkerke R^2 = .22, Wald $\chi 2$ (1, N = 91) = 12.65, p < .0005, Odds ratio = 1.81).

Recall that we did not specify *a priori* thematic or perceptual groupings at the superordinate level. Therefore, we could only examine thematic and perceptual *justifications.* Twenty-five children justified thematically on the Superordinate level task. logistic regression analysis revealed that age did not play a significant role in whether children thematically justified (p = .89). Visual inspection of Figure 2.14C revealed an apparent inverted U-shaped effect of age, with children between the ages of four and six providing thematic justifications more than younger or older children. To determine *post hoc* whether the apparent inverted U- shaped effect of age on thematic justification was significant, we conducted a logistic regression with the independent variable being the negative absolute value of the z score of age and the dependent variable being thematic justification. This model was a good fit for the data (Log likelihood = -50.94, Nagelkerke R^2 = .13, Wald $\chi 2$ (1, N = 91) = 7.31, p = .007, Odds ratio = 3.51).

Six children justified perceptually on the Superordinate level task. Given that so few children justified perceptually, it is not surprising that, as shown in Figure 2.14D, logistic regression analyses did not reveal a linear (p = .29) or U-shaped effect (p = .19) on perceptual justification.

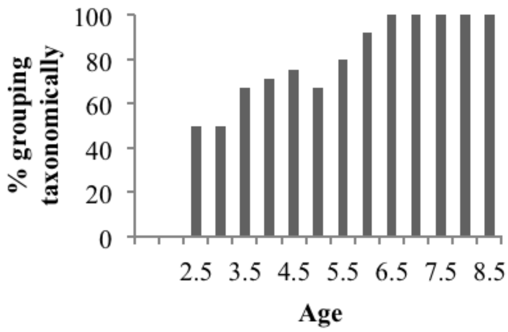


Fig 2.14A. Taxonomic grouping

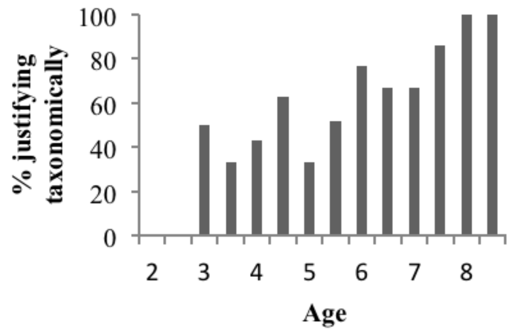


Fig 2.14B. Taxonomic justification

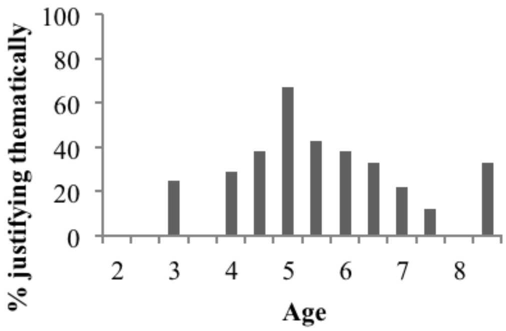


Fig 2.14C. Thematic justification

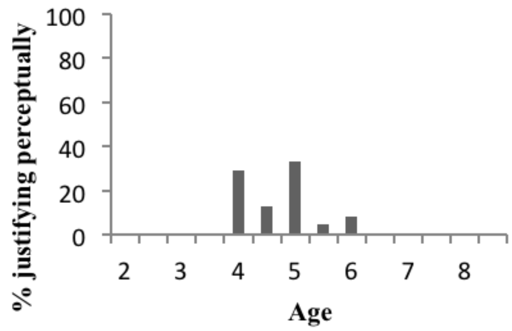


Fig 2.14D. Perceptual justification

Fig 2.14. Age on grouping & justification type on the Superordinate level task.

Intermediate level task. Fifty-one children justified taxonomically on the Intermediate level task. As shown in Figure 2.15A, age had a positive linear effect on justifying taxonomically (Log likelihood = -52.16, Nagelkerke R^2 = .27, Wald $\chi 2$ (1, N = 91) = 15.20, p < .0005, Odds ratio = 2.01),

Eighten children justified thematically on the Intermediate level task. Using the same logistic regression analyses we used to analyze the Superordinate level thematic data, we found that age did not have a linear effect on thematic justification (p = .59), but did have an inverted U-shape effect (Log likelihood = -34.89, Nagelkerke R^2 = .32, Wald

$\chi 2$ (1, N = 91) = 11.31, p = .001, Odds ratio = 18.39). As shown in Figure 2.15B, only children between the ages of four and seven provided thematic justifications.

Fifty-five children justified perceptually on the Intermediate level task. As shown in Figure 2.15C, logistic regression analyses revealed both a significant linear effect of age on perceptual justification (Log likelihood = -58.35, Nagelkerke R^2 = .12, Wald $\chi 2$ (1, N = 91) = 7.61, p = .006, Odds ratio = 1.51) and a significant inverted U-shape effect of age on perceptual justification (Log likelihood = -57.64, Nagelkerke R^2 = .14, Wald $\chi 2$ (1, N = 91) = 8.92, p = .003, Odds ratio = 3.10). To compare the linear and U-shaped models, we compared the Akaike Information Criterion (AIC) for each of these models.[10] Comparison of the Akaike Information Criterion (AIC) for the two models revealed that the model with an inverted U-shaped effect of age was a better fit than the model with the linear effect of age (ΔAIC = 4.00).

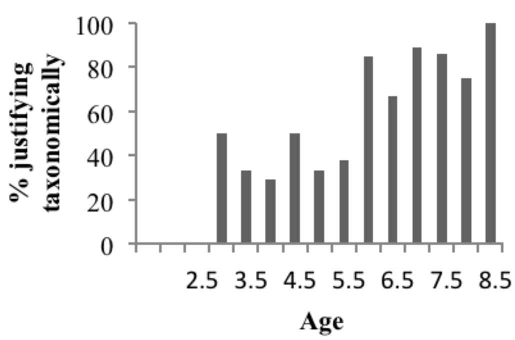
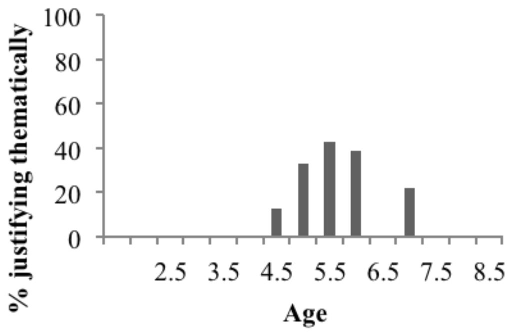


Fig 2.15A. Taxonomic justification

[10] The AIC allows for comparison of nested or non-nested models by measuring the goodness of fit (i.e., log likelihood) and adding a penalty for each added parameter (see Burnham & Anderson, 2002).[10] One model whose AIC is two less than that of another model (i.e., ΔAIC $\geq$ 2) is considered to be significantly better. Although the AIC was originally devised to compare nested models, the AIC has been shown to be valid for, and is now commonly used for comparison of non-nested models (see Glatting, Kletting, Reske, Hohl, & Ring, 2007, for discussion).



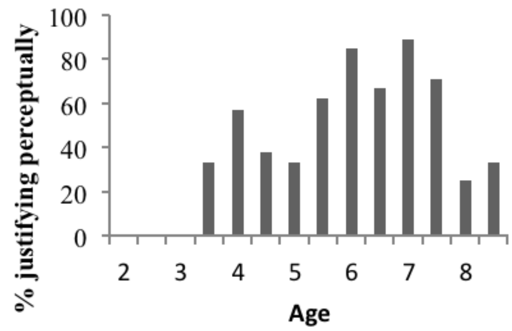
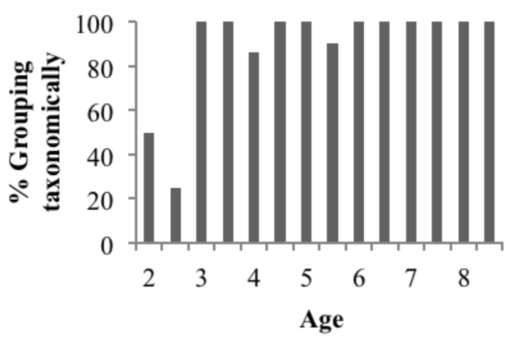


Fig 2.15B. Thematic justification **Fig 2.15C.** Perceptual justification

Fig 2.15. Age and justification type on the Intermediate level task

Basic level task. Eighty-four children grouped by basic level taxonomic categories on the Basic level task. As shown in Figure 2.16A, age did not have a linear effect on taxonomic grouping (Log likelihood = -24.67, Nagelkerke $R^2 < .0005$, Wald $\chi 2$ (1, N = 91) = .01, p = .91, Odds ratio = 1.03).

Seventy-nine children justified by basic level taxonomic categories on the Basic level task. As shown in Figure 2.16B, there was not a linear effect of age on taxonomic justification (Log likelihood = -35.44, Nagelkerke $R^2 = .002$, Wald $\chi 2$ (1, N = 91) = .10, p = .76, Odds ratio = 1.06).



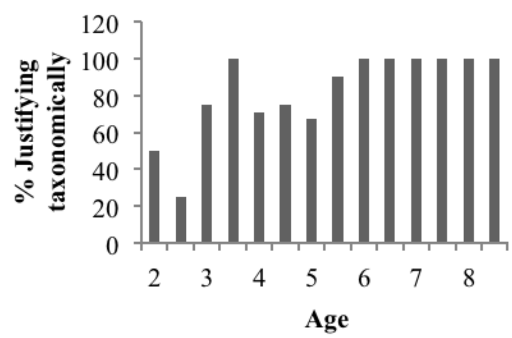


Fig 2.16A. Taxonomic grouping **Fig 2.16B.** Taxonomic justification

Figure 2.16. Age on grouping and justification on the Basic level task

Four children provided other types of justifications. Two six-year-olds provided only non-basic level taxonomic justifications (e.g., putting the fish together because, “They are water animals”; or putting the dogs and cats together because, “They’re pets”). One six-year-old provided both non-basic level taxonomic and perceptual justifications (putting the dogs and cats together because they “have tails, run fast, and have fur”). And one eight-year-old provided a perceptual justification (putting the giraffes and elephants together because “They’re tall”).

To investigate whether any of our measures predicted Superordinate level or Intermediate level taxonomic justification beyond what is predicted by age, we compared full models with age and one other predictor (Superordinate level vocabulary score, switching dimensions on the Perceptual task, Intermediate level vocabulary score, Animal Declarative Knowledge score, Metamemory score, and MLU) to a reduced version of each of these models that only had age as a predictor. We compared each pair of nested models by computing the Akaike Information Criterion (AIC) for the one parameter model (age only) and the two parameter model (age and one other independent variable). We found no cases in which the ΔAIC for the full and reduced model was greater than two. In other words, none of our measures predicted Superordinate or Intermediate taxonomic justification better than age.. However, the data do provide several important insights into the development of taxonomic categories.

Effect of MLU. If it were the case that none of the cognitive abilities we measured play independent roles in the thematic-taxonomic shift but rather, that this shift is driven by more general cognitive development (for which age is an excellent proxy), we would expect that the measures most correlated with age would also be those that best predict taxonomic grouping and justification. However, this was not the case. Recall that metamemory score was more highly correlated with age (r = .54, $p < .0005$) than MLU was with age (r = .43, $p < .0005$). However MLU (Log likelihood = -58.80, Nagelkerke $R^2 = .08$, Wald $\chi 2$ (1, N = 91) = 4.52, $p = .03$, Odds ratio = 1.22) is a better predictor than Metamemory score (Log likelihood = -59.36, Nagelkerke $R^2 = .06$, Wald $\chi 2$ (1, N = 91) = 4.00, $p = .05$, Odds ratio = 1.48) of justifying taxonomically on the Superordinate level task, and the AIC is significantly smaller for MLU than for Metamemory (ΔAIC = 33.24).

Similarly, MLU (Log likelihood = -57.18, Nagelkerke $R^2 = .15$, Wald $\chi 2$ (1, N = 91) = 8.09, $p = .004$, Odds ratio = 1.34) is a better predictor than Metamemory score (Log likelihood = -60.19, Nagelkerke $R^2 = .06$, Wald $\chi 2$ (1, N = 91) = 4.15, $p = .04$, Odds ratio = 1.49) of justifying taxonomically on the Intermediate level task, and the AIC is significantly smaller for MLU than for Metamemory (ΔAIC = 52.21). These results suggest that language may play a role in the thematic-taxonomic shift.

At this time, we are not able to reconcile the facts that 1. MLU and age together do not predict taxonomic justification better than just age. 2. MLU is a better predictor of taxonomic justification than metamemory even though it is less highly correlated with age.

Grouping vs. justification. Grouping and justifying are both commonly used to investigate categories and categorization, but they require different cognitive abilities. For example grouping is a nonverbal measure, whereas justification is a verbal measure. Additionally, grouping can be done without a conscious understanding of how category members are similar, whereas justification cannot be done without this conscious understanding. To determine the relationship between grouping and justification, and to compare the developmental trajectories of each of these abilities, we compared grouping to justification on three of our tasks. For two of these tasks—the Superordinate level task and the Basic level task, we compared taxonomic grouping to taxonomic justification. On the third task—the Perceptual Grouping task—we compared perceptual grouping to perceptual justification.

As shown in Figure 2.17, on the Superordinate level task, 21 children (23%) grouped taxonomically but did not justify taxonomically.

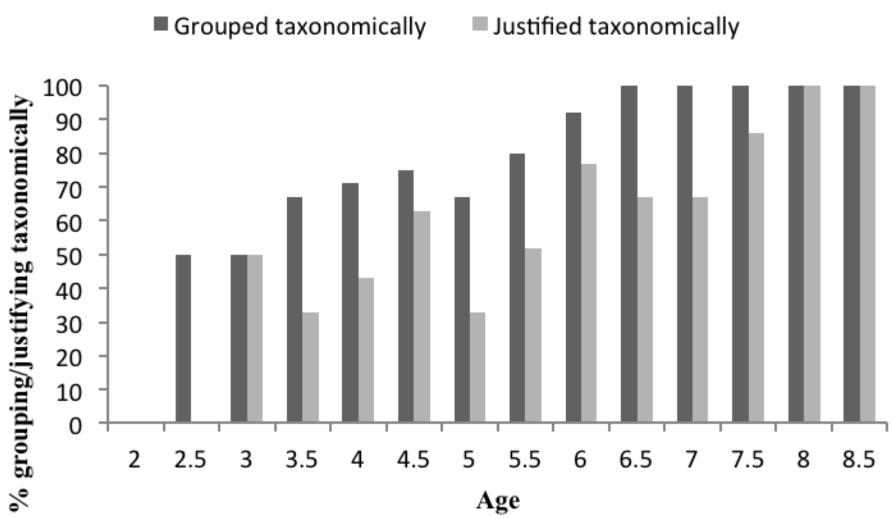


Fig 2.17. Taxonomic grouping vs. justification on the Superordinate level task

As shown in Figure 2.18, on the Basic level task, 12 children (13%) grouped taxonomically but did not justify taxonomically. Note that even on this task in which justifying taxonomically was equivalent to naming animals at the basic level, there was still a sizeable number of children who were able to group but not justify.

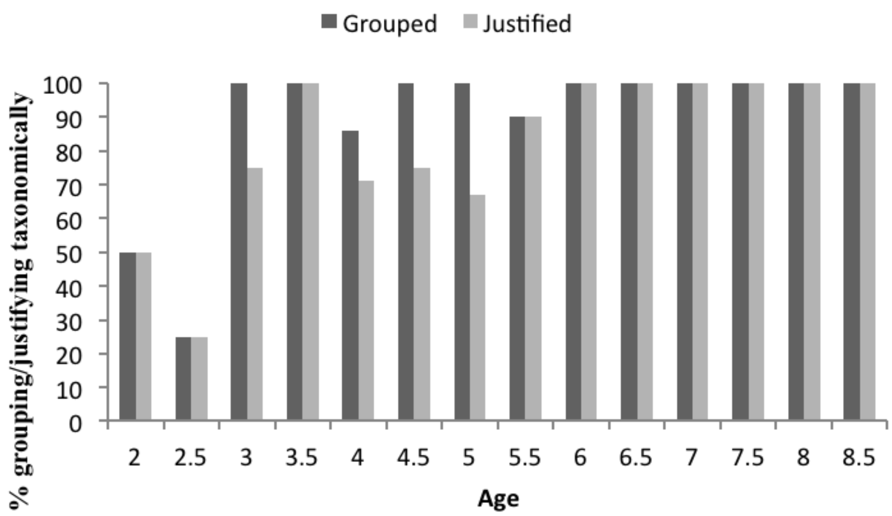


Fig 2.18. Taxonomic grouping vs. justification on the Basic level task

As shown in Figure 2.19, in contrast to the substantial grouping-justification differences on the abstract categorization tasks, on the Perceptual Grouping task, there were only five children (5%) who grouped but did not justify.

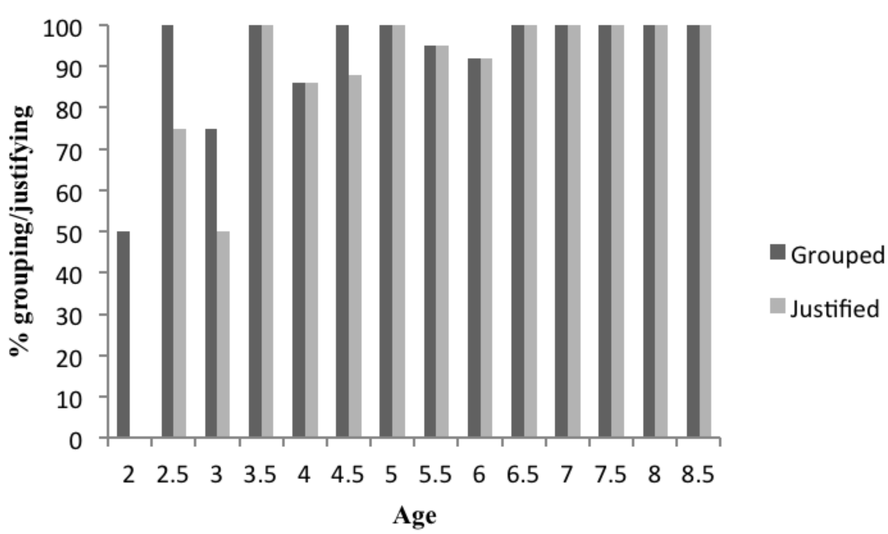


Fig 2.19. Grouping vs. justification on the Perceptual Grouping task

Declarative knowledge vs. taxonomic justification

As we reported earlier, 19 out of 20 of our Animal Declarative Knowledge questions probed taxonomic categories (e.g., which ones are mammals?). As shown in Figure 2.20, 81 children (89%) selected appropriate animals on at least 10 questions on this task. Of these 81 children, 32 (41%) did *not* justify taxonomically even *once* on the Intermediate level task. These results demonstrate that declarative knowledge is not sufficient for taxonomic justification. Furthermore, eight children correctly identified animals on at least 10 questions but did not justify in *any* way (i.e., not even thematically or perceptually) on the Intermediate level task. For these children in particular, the mismatch between declarative knowledge and justification cannot be driven merely by a preference for other ways of justification but rather by a lack of ability to justify taxonomically despite a wealth of relevant declarative knowledge.

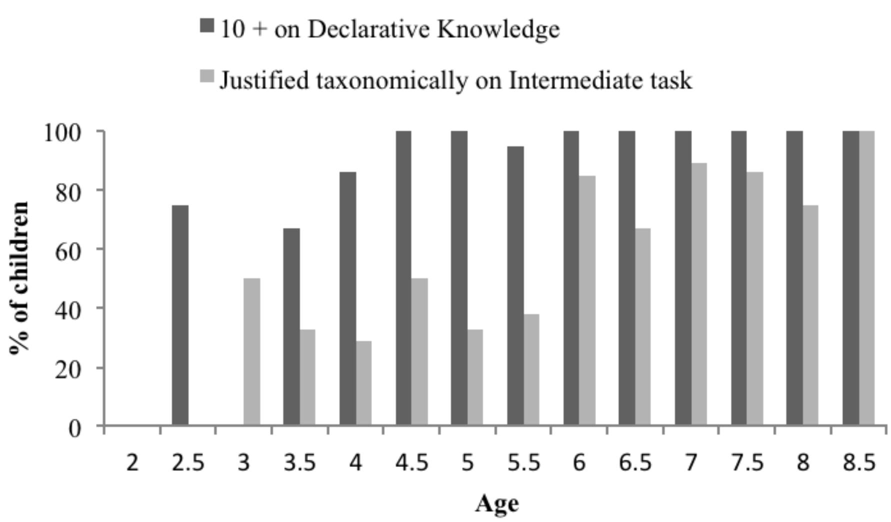


Fig 2.20. Percent of children with 10+ on the Animal Declarative Knowledge task and percent justifying taxonomically on the Intermediate level task

Consistency between levels

To investigate whether children were consistent in how they justified across levels, we conducted a series of omnibus multiple logistic regression analyses (i.e., both independent variables were entered in one step in each analysis).

Taxonomic justification. In the first analysis, we entered two independent variables—justifying taxonomically on the Superordinate level task and age—and one dependent varaible—justifying taxonomically on the Intermediate level task—in order to evaluate whether justifying taxonomically on one task predicted doing so on the other task independent of age. As shown in Table 2.1, the overall model was a good fit for data, with both Superordinate level taxonomic justification and Age independently predicting Intermediate level taxonomic justification.

Table 2.1. Logistic Regression Analysis of Intermediate Level Taxonomic Justification with Superordinate Level Taxonomic Justification and Age as Predictors

Predictor	**B**	**Wald's χ^2**	***df***	***p***	**Odds Ratio**
Constant	-1.95	2.50	1	.11	.14
Age	.55	7.19	1	.007	1.74
Superordinate taxonomic justification (1 = yes, 0 = No)	-2.28	17.26	1	< .0005	.10
Overall Model Evaluation	**Log Likelihood**	**χ^2**	***df***	***p***	**Nagelkerke R^2**
	-42.43	39.96	2	< .0005	.48

Another set of logistic regression analyses with Basic level taxonomic justification and Age as independent variables revealed that Basic level justification predicted neither Superordinate level taxonomic justification nor Intermediate level taxonomic justification. Basic level taxonomic justification did not independently predict either Superordinate or Intermediate level taxonomic justification (both *p*s > .05).

Thematic justification. Next, we investigated whether children who justify thematically do so consistently between levels. Given that we found a U-shaped effect of Age on thematic justification, we conducted a logistic regression analysis with Superordinate thematic justification and the negative absolute value of the *z* score of age as independent variables and Intermediate thematic justification as the dependent variable. As shown in Table 2.2, although the overall model was a good fit for the Intermediate thematic justification data, with Age being a significant independent predictor. However, Superordinate thematic justification was not an independent predictor.[11] Given that so few children justified thematically on the Basic level task, we did not further analyze these data.

Table 2.2. Logistic Regression Analysis of Intermediate Level Thematic Justification with Superordinate Level Thematic Justification and Age as Predictors

Predictor	**B**	**Wald's χ^2**	***df***	***p***	**Odds Ratio**
Constant	.58	1.20	1	.27	1.78
- |z score (age)|	2.78	9.71	1	.002	16.11

[11] When age was not included as a predictor, Superordinate thematic justification significantly predicted Intermediate thematic justification (Log likelihood = -41.91, Nagelkerke R^2 = .11, Wald $\chi 2$ (1, *N* = 91) = 6.64, *p* = .01, Odds ratio = .24).

Superordinate thematic justification (1 = yes, 0 = No)	-.95	2.42	1	.12	.39
Overall Model Evaluation	**Log Likelihood**	χ^2	***df***	***p***	**Nagelkerke *R*²**
	-33.67	23.18	2	< .0005	.36

Perceptual Justification. A logistic regression analysis with Superordinate perceptual justification as the Independent variable and Intermediate perceptual justification as the dependent variable revealed that this model was not a good fit for the data (Log likelihood = -58.87, Nagelkerke R^2 = .11, Wald χ2 (1, N = 91) < .0005, p = .999, Odds ratio < .0005). Therefore, we did not further analyze these data. However, see the Discussion of this chapter for an account of why children were more likely to justify perceptually on the Intermediate level (with more than half of children doing so) than on the other two levels.

Order of category formation

There were no children who justified taxonomically on the Intermediate level task but did *not* justify taxonomically on the Superordinate level task; however, there were children who did the reverse. Additionally, there were no children who justified perceptually on the Superordinate level task but did *not* justify perceptually on the Intermediate level task; however there were children who did the reverse. There were no other classification rules. For example, some children justified taxonomically on the Superordinate or Intermediate level tasks but not on the Basic level task, and vice versa, and some children justified thematically on the Superordinate level task but not on the Intermediate level task and vice versa.

Clustering results

Recall that one possible explanation for why the 375 children in our Verbal Fluency task did not cluster is that we used "adult" features in our cluster analyses. If children cluster using different features than adults, this would account for our failure to find evidence of verbal fluency clustering. In order to address this possibility, we reanalyzed our verbal fluency data using the justifications that children generated on the Intermediate level task. These justifications added 55 new features to our feature set and included taxonomic, thematic, and perceptual justifications.

We analyzed the verbal fluency data both collapsing across age groups and separately for three-, four-, and five-year-olds. Under both methods of analysis, there were no features for which children clustered significantly more in their actual lists than in the randomzied versions of their lists (all *p*s > .05). At age five only, clustering approached signficance at the .01 level for three features: BIRD, FEATHERS, and WINGS (all *p*s = .03).

Discussion

I began this chapter by describing three possible accounts for why children did not semantically cluster in the Verbal Fluency task. I can now evaluate these accounts.

Account I: Within the animal domain, preschool-age children lack the declarative knowledge to access animal names at an Intermediate level. In other words, children do *not access multiple consecutive pets because they don't know what pets are.* This account is *not* consistent with the data. In the Animal Declarative Knowledge task, all but one child (99%) were able to select animals correctly on *at least four* questions, and 81 children (89%) were able to select animals correctly on *at least ten* questions. More specifically, of the 46 three- through five-year-olds in the current study (the ages of the children in the verbal fluency study), 45 (98%) correctly selected animals on at least four questions, and 39 (87%) correctly selected animals on at least ten questions. Preschoolers have the relevant declarative knowledge but still do not cluster during a verbal fluency task.

Next I will consider Account III.

Account III: Children do cluster, but they do so by different features than adults. In particular, one possibility is that children clustered by thematic rather than taxonomic features. This account is also *not* consistent with the data. When I used any justification—taxonomic, thematic, or perceptual—provided by any child in the Intermediate level task as features in the Verbal Fluency data, I still did not find any significant clusters. Although the possibility remains that children clustered by features that I did not test, I ruled out 62 features (55 generated by children in the Intermediate Grouping task and an additional seven generated *only* by adults in the Explicit Grouping task), that encompassed all three types of semantic features.

One limitation to my clustering methodology is that it assumes participants share features. For example, if only one participant clustered by pets and only one participant clustered by mammals, I would not find evidence of clustering by either feature. This is potentially more problematic for thematic categories, which are likely more idiosyncratic (i.e., they could be based on children's idiosyncratic life experiences) than taxonomic or perceptual properties. However, many of the children's thematic justifications in the Intermediate level task did not appear to be particularly idiosyncratic (e.g., "Cats chase mice;" "Dogs and cats are enemies;" "Birds eat worms"), and therefore, if children clustered by thematic features, one would have expected these more common thematic features to be significant. Thus, it is unlikely that children semantically clustered.

I now return to Account II.

Account II. Preschool-aged children do have intermediate level declarative knowledge about animals but this information is not stored in their lexicons and/or they do not access words via intermediate level categories. Having ruled out the other two accounts, Account II is the best explanation of the data. Analysis of verbal fluency data from children older than five is necessarily to investigate when and how children's lexicons and/or lexical access becomes adult-like. Given that three clustering features approached significance in the five-year-olds' data, it is likely that this shift happens gradually.

Disentangling lexical organization from lexical access is extremely difficult. Psycholinguistic studies investigating lexical organization necessarily investigate lexical access, and in both adult and developmental studies, lexical access results are often taken as evidence about the structure of the lexicon. My results suggest that preschoolers' lack of semantic clustering is due to differences in either the structure of their lexicons or in their lexical access, but it is not currently possible to spell out these differences further.

Is perceptual part of taxonomic?

As I reported in the introduction, in some studies toddlers appear to have exhibited a *taxonomic* preference. However, one problem with interpreting these results is that the taxonomic matches were also more perceptually-similar to the targets than the thematic matches were to the targets. Thus, it is difficult to discern whether toddlers had an abstract taxonomic preference or a perceptual preference. My study can inform this debate in terms of how children's perceptual and (abstract) taxonomic *justifications* develop. Interestingly, in my study, two- and three-year-olds who did justify tended to do so either taxonomically or thematically, whereas children ages four through six were much more likely to justify perceptually. These results suggest that at least in some cases, young children may not be as perceptually-bound as has previously been suggested (see Springer, 2001, for discussion).

Representations vs. *ad hoc* categories vs. *post hoc* explanations

It was often unclear whether children's justifications were reflections of their represented categories, whether they were constructing *ad hoc* categories when grouping, or whether they were constructing *post hoc* justifications. This distinction is clearest for perceptual groups/justifications. Some children appeared to be identifying perceptual features of the *stimuli* themselves, either during grouping or during justification. For example, a child who said, "This has a lump and this has a lump" while pointing to little (seemingly non-meaningful) indentations in the drawings of two animals appeared to be comparing the *drawings* rather than the things which they were intended to represent. In these cases, children were not accessing categories stored in memory but rather, were solving a problem based on novel information; that is, children were following the rules of the task by finding similarities among the stimuli, rather than by retrieving categorical information. This behavior is consistent with evidence that even older children and preteenagers struggle with dual representation, that is, with seeing something as both a symbol and a referent (e.g., Brown, McNeil, & Glenberg, 2009; Uttal, O'Doherty, Newland, Hand, & DeLoache, 2009).

In other cases, children gave perceptual justifications that could not have been based on the stimuli themselves. For example, a child who said that both a hippo and an elephant were big—despite the fact that the drawings of these animals were not larger than the drawings of the mouse or the worm—must have been accessing stored information. Further analyses are necessary to determine whether the developmental trajectories differ for represented perceptual categories, *ad hoc* perceptual categories, and *post hoc* perceptual justifications.

Although it is less clear which thematic justifications were reflections of represented categories and which were *ad hoc* or *post hoc*, it is likely that children provided each of these types of thematic justifications. For example, a child who grouped a cat with a mouse because "cats chase mice" either could have been retrieving this thematic category from semantic memory or could have been retrieving an episodic memory of a cat chasing a mouse and constructing an *ad hoc* or *post hoc* category. Less clear still is whether children could retrieve thematic knowledge (e.g., cats chase mice) or taxonomic knowledge (e.g., cats and mice are mammals) from semantic memory without having this information represented as a category and then construct *ad hoc* or *post hoc* categories based on this knowledge. Again, future work is necessary to disentangle the processes of justifying in these two ways.

Cause of the shift

No measure surpassed age in predicting taxonomic grouping or justification. One possibility is that developments in language, ability to group and justify, flexible thinking, declarative knowledge, and metamemory do *not* play independent roles in the thematic-taxonomic shift. A second possibility is that these developments do play independent roles in the thematic-taxonomic shift but that my tests were not sensitive measures of these abilities. A third possibility is that the interaction between age and some cognitive ability I did not test would have surpassed age in predicting taxonomic grouping and/or justification. Further investigation is necessary to determine which of these possibilities is right and whether the thematic-taxonomic shift occurs as a result of general cognitive development or whether development in one or more specific cognitive abilities causes this shift.

Grouping vs. justification

Throughout the thematic-taxonomic shift literature, researchers use children's justifications as a proxy for the nature of their categories. There is a practical reason for this. Without a justification, it is impossible to discern whether a child is grouping SHIRT and PANTS because they are both "clothing" or because they are both "things [the child] use[s] in the morning time" (both justifications for this grouping given by children on the Superordinate level task). A review of stimuli from experiments that do *not* evaluate children's justifications suggests that experimenters may have misinterpreted children's matches. For example, Scott, Serchuk, & Mundy (1982) characterized SOCK and SHOE as thematically related, even though these items also have a clear abstract taxonomic relationship (i.e., clothing worn on the feet) as well as salient perceptual similarities.

Evaluating children's matches or groups is particularly problematic given Nelson's (1986) argument that during the process of switching from thematic to taxonomic thinking, children develop "slot-filler" categories—taxonomic-like categories grounded in schemas (e.g., LION and TIGER are animals seen at the zoo). It is also problematic given that many members of the same taxonomic category also share perceptual characteristics, especially at the basic and subordinate levels.

However, my data suggest that many children are able to group but not justify taxonomically and therefore that it is problematic to rely solely on justifications for insight into category representations. On the other hand, justifications but not groupings provide direct insight into children's understanding of their category structure.

Individual differences

There is some evidence for persistent individual categorization preferences. For example, in a longitudinal study, Dunham and Dunham (1995) found that children maintained taxonomic or thematic preferences at 13- , 24- , and 36-months. In another study, Kalenine and Bonthoux (2006) tested three- and four-year-olds in two sessions separated by a month and found that they maintained their categorization preferences for living but not nonliving things. In my study, children who justified thematically or taxonomically on the Superordinate level task also tended to do so on the Intermediate level task. This was true for taxonomic justification even when controlling for age. Therefore, it is possible that taxonomic preference is somewhat attributable to individual preference.

However, justifying perceptually on the Superordinate level task did not predict doing so on the Intermediate level task. One possibility is that when taxonomic or thematic categories are not salient, children construct perceptually-based *ad hoc* groupings or perceptually-based *post hoc* justifications, which are easier to construct on the fly than taxonomic or thematic groupings/justifications. Therefore, a child could provide taxonomic or thematic justifications on one task in which these categories were more salient but perceptual justifications on another task in which they were less salient. Consistent with this account, 2/91 children (2%) justified perceptually on the Basic level task, and 6/91 children (7%) did so on the Superordinate level task, but 48/91 children (53%) did so on the Intermediate level task. These results are in keeping with the intuition that taxonomic categories were most salient on the basic level (where justifications are equivalent to naming items), followed by the Superordinate level (where there were *a priori* categories), but that taxonomic categories were not as salient on the Intermediate level.

Qualitative effects of age

Finally, there were apparent qualitative differences in older and younger children's performance on the Animal Declarative Knowledge task. When an experimenter asked a question (e.g., "Which ones live on a farm?"), older children tended to answer immediately and rapidly, often without consulting the picture cards. When older children did refer to the cards, they appeared to be scouring the cards quickly to see if they had missed any animals. Conversely, younger children often looked deliberately at each individual card as if the cards were novel stimuli about which the children were expected to solve a novel a problem. Younger children's responses were often more hesitant (e.g., "I think maybe a chicken lives on a farm.")

This contrast between older and younger children's approaches was especially striking in response to the question, "Which ones have four legs?" As usual, older children immediately and rapidly named four-legged animals and then glanced at the cards to complete their responses. However, younger children often counted the legs on each animal. Sometimes the younger children expressed uncertainty about whether something in a picture was a leg. Other times, when a leg was partially occluded by another body part, younger children failed to count that leg and pronounced that an animal had three legs. These differences are consistent with older but not younger children having intermediate-level taxonomic categories (e.g., FARM ANIMALS, FOUR-LEGGED ANIMALS) from which they can efficiently retrieve members.

Our coding scheme did not capture these qualitative differences. That is, a child received credit for naming “cow” as a four-legged animal whether he did so by retrieving “cow” from his FOUR-LEGGED ANIMALS category or whether he counted the legs on the picture of the cow. There are several ways that this difference could be investigated in the future. First, if I measured the amount of time each child spent answering a given question on the Animal Declarative Knowledge task, I might find that older children spent significantly less time than younger children. Second, if I compared older and younger children’s eye movements, I might find that older children spent proportionately less time looking at the stimuli. Either of these differences would suggest that older and younger children answer the questions using different strategies. Finally, I could investigate children’s performance on this task in a new condition in which picture cards are *not* provided for reference. If older children’s performance in this condition did not decline as steeply as younger children’s performance, this would suggest that older children tend to retrieve answers from memory whereas younger children rely more heavily on the stimuli.

The results in this chapter on typically-developing children suggest that lexical representation and access is qualitatively different in young children and that intermediate level taxonomic develop categories after basic and superordinate level taxonomic categories. In the next chapter, I investigate the lexicon and categorization in children with autism.

III. Lexical representation, access, and categorization in children with autism

In Chapter II, I investigated categorization development in typically-developing children. In Chapter III, I investigate parallel questions in children with autism spectrum disorders (ASD).[12] There were two primary motivations for the current study.

First, investigating these questions in autistic children, whose cognitive strengths and weaknesses are often more pronounced than those of typically developing (TD) children, could elucidate the cognitive abilities instrumental in the thematic-taxonomic shift. For example, ASD children often have impaired language and are often less cognitively flexible than TD children of the same mental age. Therefore, if sophisticated language and/or cognitive flexibility is necessary for creating taxonomic groups and/or for justifying taxonomically, I would expect ASD children to group and/or justify taxonomically less than TD children. However, if categorization development occurs as a result of increases in mental age more generally, I would expect ASD children to group and/or justify taxonomically to the same extent as TD children of a similar mental age.

Second, although there has been extensive previous research on categorization development in ASD, much remains unknown, and the literature is currently divided on whether categorization development differs in ASD children. Furthermore, little is known about the trajectory of categorization development in ASD. In particular, the current study is the first to investigate the thematic-taxonomic shift in children with ASD.

---

[12] Throughout this chapter, I restrict my review and investigation to *high functioning* autistic individuals.

Some studies suggest that ASD children are impaired in generating categories (e.g., in grouping tasks such as those in the previous chapter) but not in following explicit categorization rules. For example, Minshew, Meyer, & Goldstein (2002) gave a battery of tests to ASD children and TD children matched for mental age. The ASD children performed worse than the TD children on all tasks that required them to generate categories (i.e., the Goldstein-Scheer Object Sorting Task, the Picture Absurdities subtest, and the 20 Questions task). However, the ASD and typical children performed equally well on tests that required them to follow predetermined rules (i.e., the Trail Making Test Part B and the Halstead Category Test).

In another study, Klinger and Dawson (2001) taught novel animal-like categories to three groups of children matched for mental age: TD children, children with ASD, and children with Down Syndrome. The children were taught in one of three ways: by seeing prototypes, by seeing pictures from which a rule could be deduced and being told that a rule could be deduced, or by seeing pictures from which a rule could be deduced but *not* being told so. All three groups successfully learned the categories via rules—both when the children were and were not told to learn from a rule—but only the typical group learned the categories via prototypes. Taken together, the studies by Minshew et al. and by Klinger suggest that ASD children have difficulty generating categories but not identifying or adhering to category rules.

However, in another study in which participants were asked to categorize patterns of random dots, Froehlich et al. (2012) found that ASD children were as good as TD mental age matches at learning from prototypes. One possibility is that ASD children have the capacity to learn from prototypes when the stimuli are more straightforward and/or perceptual (e.g., with patterns of dots) but not when they are more abstract (e.g., with animal-like categories). This possibility is consistent with Frith and Happe's (1994) claim that ASD children lack *central coherence,* that is, the ability to draw together diverse information to construct high-level meaning in context. Frith and Happe gave ASD and mental age-matched TD children the Children's Embedded Figure Test, in which they were asked to locate a part (e.g., a triangle) in a whole object (e.g., a baby carriage). The ASD children were faster than the TD children at identifying the part, suggesting that ASD children tend to focus on details rather than on an abstract whole (i.e., they lack central coherence).

One could argue that if ASD children lack central coherence, they should not be able to develop taxonomic categories. In order to develop the categories "animal," "pet," or even "dog," one must abstract away from the features of particular animals, pets, or dogs and construct an abstract whole. Note that central coherence is *not* necessary for developing thematic categories, whose members do not share abstract features. Nor is central coherence necessary for developing perceptual categories. In fact, if autistic children lack central coherence, one might then expect them to excel at developing perceptual categories, which require ignoring the abstract whole and focusing on perceptual details.

However, the claim that ASD children lack central coherence is inconsistent with Tager-Flusberg's (1985) finding that ASD children have both superordinate and basic level categories. Tager-Flusberg tested ASD and mental age-matched TD children on two versions of a match-to-sample task: a picture-matching version, and a word-matching version. In the picture version, Tager-Flusberg showed the child a target (e.g., a kitchen chair), a match (e.g., a rocking chair), and a distractor (e.g., a dog) and asked the child to point to the one that was like the target. Some children were given basic level matches (e.g., the example given), and some children were given superordinate level matches (e.g., matching an apple to a banana rather than to an elephant). The word matching version was identical except the experimenter said the target item, puppets said the match and distractor items, and the child pointed to the puppet who correctly matched the experimenter. In both conditions and on both levels, ASD children performed as well as the typical matches.

An alternative to the Central Coherence hypothesis is that ASD individuals' categorization difficulties result from impairments in processing. Gastgeb, Strauss, and Minshew (2006) gave TD and ASD children and adolescents (matched by four criteria: chronological age, IQ, verbal IQ, and performance IQ) a category verification measuring prototypicality effects. In the task, participants saw visual stimuli and heard a word from four basic level categories (dog, cat, couch, and chair) and were asked whether the visual and auditory stimuli matched. For both ASD and TD individuals, there was a significant effect of prototypicality on both speed and accuracy, and both groups improved with age. ASD and TD children were equally accurate on both prototypical and non-prototypical items. However ASD individuals were significantly slower than TD individuals, and this difference was greater for non-prototypical items. Thus, it is possible that categories are qualitatively similar in ASD and TD children but that ASD individuals' processing impairments impede their categorization abilities. ASD children's processing impairment is consistent with the argument that ASD children's range of cognitive and linguistic deficits stems from their underconnectivity (i.e., lower degree of information integration and synchronization between cortical areas (Just, Cherkassky, Keller, and Minshew, 2004; Nielson et al., 2014; Uden, Supekar, & Menon, 2013).

Cognitive predictors in ASD

In this section, I review the literature on language, declarative knowledge, cognitive flexibility, and metamemory in ASD. Next, I investigate the relationships between each of these abilities and the thematic-taxonomic shift in ASD. It is important to note that there is considerable variability on all aspects of cognition within ASD. For example, twenty percent of ASD individuals are completely nonverbal, whereas most high functioning autistic individuals (HFA) individuals have fully intact language. Therefore, the literature reviewed in this section reflects general findings about relatively HFA individuals (i.e., those with language) rather than diagnostic criteria. Further, in many of the studies reported in this section, ASD children were matched to TD children of the same mental age but had a higher chronological age. Therefore, even when HFA children are reported to have performed normally on a task, the HFA children may still be impaired compared to children of the same chronological age. In these cases, the findings reveal that a particular ability is intact compared to other abilities rather than that the HFA children perform normally for their age.

Language. Compared with mentally retarded or TD children matched for nonverbal mental age, verbal ASD individuals often present with greater impairments in pragmatics and morphosyntax than in vocabulary or other forms of lexical knowledge.

Pragmatics. In terms of pragmatics, even the most HFA individuals frequently present with deficits in nonverbal communicative gesture (Charman, Drew, Baird, & Baird, 2003); speech acts (Wetherby, 1986); discourse pragmatics (Capps, Kehres, &

Sigman, 1998); nonliteral language (MacKay & Shaw, 2004; Williams et al., 2013); and pragmatic functions of prosody (Shriberg et al., 2001). Deficits in pragmatics are not surprising given that ASD individuals are thought to lack a theory of mind (Baron-Cohen, Leslie, & Frith, 1985), which is integral to pragmatics (see Tager-Flusberg, 2000, for discussion).

Morphosyntax. ASD children tend to produce less syntactically-complex utterances (Scarborough, Rescorla, Tager-Flusberg, Fowler, & Sudhalter, 1991) and tend to perform more poorly on the expressive syntax portion of the Clinical Evaluation of Language Fundamentals (Botting & Conti-Ramsden, 2003) than other children of the same nonverbal mental age. Further, Eigsti, Bennetto, & Dadlani (2007) found that ASD children were impaired on morphosyntax (including lower MLUs) in comparison to TD and developmentally-delayed children *matched by score on the Peabody Picture Vocabulary Test* (PPVT), suggesting that ASD children have a deficit specific to some aspects of language (including grammar) rather than in language more generally.

Some of ASD individuals' impairments in syntax may be related to their theory of mind deficit. For example, both Pinker and Bloom (1990) and Tager-Flusberg (2000) have argued that facility with complementizer phrases is related to understanding propositional attitudes, which is contingent on having a theory of mind.

However, ASD individuals also demonstrate deficits in aspects of grammar that have not been explained by a theory of mind deficit. For example, even HFA individuals often omit requisite closed-class items, such as "the" at the start of a phrase (Bartolucci, Pierce, & Streiner, 1980) and often omit inflectional morphology, such as "ing" and "ed" (Botting & Conti-Ramsden, 2003). Interestingly, in a study in which regular and irregular past tense forms were elicited from HFA and chronological age-matched TD children, the HFA children outperformed the typical children in the speed with which they named the *irregular* forms (Walenski, Mostofsky, Gidley, & Ullman, 2005). Given that the irregular forms of words are stored in the lexicon, this finding is consistent with evidence that HFA children have intact lexical knowledge (see Vocabulary and the lexicon, below).

Finally, ASD and TD individuals have been shown to have different patterns of neural activation when processing syntactic information (Just et al., 2004). High-functioning ASD adults and verbal mental age-matched adults were given active sentences (e.g., "The cook thanked the father") or passive sentences (e.g., "The father was thanked by the cook") and were given a forced choice question (e.g., "Who was thanked?"). Surprisingly, the groups did not differ in terms of accuracy, and the ASD group outperformed the TD group in terms of speed. However, the groups showed activation differences, with the ASD group having more activation than the TD group in Wernicke's area (associated with lexical processing), and the TD group having more activation the ASD group in Broca's area (associated with syntactic processing). Further, the TD group had greater functional connectivity between the cortical areas, which the authors argue is necessary not only for language but also for the other abilities in which ASD individuals are often impaired (e.g., motor function, memory, abstraction, and social aptitude).

Vocabulary and the lexicon. In a large-scale study of ASD children and adolescents, some participants performed extremely poorly on a test of expressive vocabulary whereas others were in the normal range for their age, demonstrating the heterogeneity of ASD (Kjelgaard & Tager-Flusberg, 2001).

Some HFA children and adults have relatively similar lexical and conceptual representations as TD individuals of the same mental age. For example, in a study in which preschoolers were taught novel words, ASD children performed similarly to TD children matched by verbal mental age (Eigsti, 2001). Thus, in ASD children, the ability to learn words may be intact relative to other aspects of language. Additionally, ASD children (of normal to low-normal mental age) have been found to increase their lexical diversity (i.e., use of words with different roots) and distribution of form class (i.e., use of nouns, verbs, etc.) at the same rate as had previously been reported for TD children of the same chronological age (Tager-Flusberg et al.,1990).

In addition, HFA individuals have been shown to have qualitatively normal lexical and conceptual representations. For example, HFA children give the same prototypicality ratings as TD children on both the basic and superordinate levels (Tager-Flusberg, 1985), and HFA adults perform normally on priming tasks (e.g., when they hear “doctor,” they respond “nurse,” Toichi & Kamio, 2001).

Lexical access also appears to be relatively intact in HFA individuals. For example, Minshew et al. (1992) found that HFA adolescents and young adults of normal intelligence performed as well as IQ- and age-matched TD individuals on a verbal fluency task, and Minshew, Goldstein, & Siegel (1997) found that HFA adults of normal intelligence performed as well as IQ- and age-matched TD adults on the verbal fluency portion of the Wechsler Adult Intelligence Scale-Revised. Muller et al. (1999) found that HFA adults of normal intelligence performed as well as TD adults of normal intelligence on both the associative fluency and the rapid picture naming portions of the Clinical Evaluation of Language Fundamentals-Revised: Screening Test. Finally, Boucher (1988) found that HFA children performed as well as age- and vocabulary-matched TD children on verbal fluency tasks when given a category (e.g., “name animals”), although they performed poorly when not given a category (“Say as many words as you can think of”). Boucher’s results are consistent with the categorization results (above) in which HFA individuals were able to adhere to category rules but not to generate their own categories.

Declarative knowledge. Declarative memory is composed of semantic memory (i.e., memory for facts) and episodic memory (i.e., memory for events). HFA individuals often have intact semantic memory relative to their other cognitive functions. For example, HFA children outperformed chronological- and verbal IQ- matched, learning disabled children on tests of rote memory, such as memorization of phone numbers (Bennetto, Pennington, & Rogers, 1996), and HFA children performed as well as chronological- and IQ-matched TD children at memorizing arbitrary word pairings (Minshew & Goldstein, 2001). Given that semantic memory often remains intact in HFA and given that the lexicon is part of semantic memory, it is not surprising that HFA children can have intact lexicons. Conversely, ASD children often have impaired episodic memory. For example, Bennetto et al., (1996) found that ASD children and adolescents have impaired temporal order memory (i.e., memory for the order in which things happened) and impaired source memory (i.e., memory for when and where things happened) compared to chronological- and IQ-matched TD children. Similarly, in a review paper, Shalom (2003) concluded that HFA individuals generally have impaired episodic memory but often have intact semantic memory (and an intact perceptual representation system).

Cognitive flexibility. Even HFA children are often less cognitively flexible than TD children of the same mental age. For example, two review papers (Pennington & Ozonoff, 1996; Seargent, Geurts, & Oosterlaan, 2002) investigating executive function in disordered populations both reported that HFA individuals had poorer cognitive flexibility than any of the clinical populations they studied (attention deficit hyperactivity disorder, conduct disorder, and Tourette's syndrome) had on *any* executive function measure (e.g., working memory, inhibition, and several others).

Furthermore, several studies have demonstrated that ASD children perform more poorly than TD children on the DCCS, a standard test of cognitive flexibility, when matched by chronological age, verbal IQ, and nonverbal IQ (Yasumura et al., 2012), just by chronological age (Dichter et al., 2010), or just by PPVT score (Frye, Zelazo, & Palfai, 1995), although they performed similarly to TD children when matched by nonverbal IQ (Frye et al., 1995). Finally, in a recent verbal fluency study, Begeer et al. (2013) found that HFA children and adolescents of normal verbal intelligence switched between clusters less frequently than chronological- and verbal age-matched TD children and adolescents, suggesting that HFA children are less cognitively flexible[13].

Metamemory. It is currently unclear whether metamemory is generally impaired in ASD individuals or whether ASD individuals' poor performance on some metamemory tasks is due to their impairment in other aspects of cognition. For example, a common task for measuring metamemorial awareness is a Judgment-of-confidence task. In this task, participants estimate how confident they are that they will perform well

---

[13] Begeer et al. used Troyer et al's (1997) methodology. See Chapter One for a discussion of problems with this methodology.

on a memory task. In one study, ASD and verbal mental age-matched TD children first saw a series of photographs and then, in a second round, were shown some of the same and some new photographs (Rouse, Donnelly, Hadwin, & Brown, 2004). ASD children's confidence in whether they had previously seen a photograph was not correlated with their accuracy. However, in another study, ASD and TD children matched for age and IQ were given a series of verbal instructions (e.g., "Pick up the red ruler and put it in the yellow box, then touch the blue pencil") and were then asked to follow the instruction and rate their confidence that they carried out the instruction properly (Wojcik, Allen, Brown, & Souchey, 2011). In this study, ASD and TD children were equally accurate at estimating how well they remembered the instructions. It is currently unclear why the results of the two studies differed, but given the difference in task demands, it is possible that ASD children's abilities in other areas (e.g., their perceptual representation system, language) affected their performance on the tasks. Another possibility is that the different matching criteria (age and verbal IQ vs. age and total IQ) resulted in different findings.

Interestingly, ASD individuals appear to have worse metamemorial awareness of their episodic memory than of their semantic memory. For example, Wojcik et al. (2013) conducted a study with two conditions: episodic and semantic. In the episodic condition, HFA children were first asked to memorize a list of unrelated word pairs and then were given the first word of each pair and asked for the second. If a child could not remember a word, the child was asked how likely he would be to recognize the word if he was given choices. In the semantic condition, HFA children were given words and asked to provide definitions. If a child could not provide a definition, the child was asked how likely he would be to recognize the definition if given choices. The HFA children were significantly more accurate at predicting their ability to recognize the word/definition in the semantic condition than in the episodic condition (Wojcik et al., 2013). These results suggest that not only do HFA children often have impaired episodic memory and intact semantic memory—as discussed earlier in this chapter—but that this discrepancy follows for HFA children's metamemorial awareness.

In this study, I first investigate the developmental trajectories of several cognitive abilities (language, cognitive flexibility, declarative knowledge, and metamemorial awareness) in ASD children and TD children matched in two ways: by mental age and (in a separate set of analyses) by chronological age. Next, I investigate the developmental trajectories of categorization in these two populations. My goals are to determine the relationships between developments in cognition and categorization in both populations and to explore whether ASD children follow typical patterns of categorization development.

Methods

Participants

The ASD participants were 24 children between 3.92 and 19.96 years (mean = 11.92 years, *SE* = .91) who had received a formal diagnosis of ASD. The TD participants were 26 children between 2.39 and 8.83 years (mean = 5.08 years, *SE* = .45) who had not been diagnosed with any language or learning disability. All but one of the TD children were also included in the analyses reported in Chapter II[14]. The children in both groups were monolingual speakers of English and were recruited from New Jersey schools. The

[14] This child was excluded from the previous study for being several standard deviations older than the mean age.

protocol was approved by the Rutgers University Institutional Review Board for Human Subjects.

Matching criteria

A longstanding issue in research on impaired populations is whether to match impaired to typical participants by mental age (i.e., intelligence) or by chronological age. On the one hand, controlling for general intelligence reduces the likelihood that differences obtained between groups merely reflect known differences in intellectual abilities between the populations. Controlling for general intelligence allows the researcher to uncover any selective deficits that might exist in the atypical population.

On the other hand, controlling for mental age has several limitations. First, there is no standard definition of general intelligence. Therefore, matching criteria are unclear. Second, it is generally accepted that intelligence encompasses a wide variety of cognitive abilities, which could conceal impairments or aptitudes in the selective abilities a study was designed to identify. Given that the criteria for intelligence are so varied *and* that controlling for performance on a particular intelligence test could obscure selective abilities, results on experimental measures could vary considerably by which intelligence test is selected as the matching criterion. For example, if impaired and typical participants are matched by a test emphasizing cognitive flexibility, this could obscure the importance of cognitive flexibility on the experimental measures. However, if participants are matched by a test emphasizing vocabulary, this could obscure the importance of vocabulary on the experimental measures. And so on.

To maximize the benefits and minimize the drawbacks of matching by an intelligence test, I matched ASD to TD participants in two ways 1. Mental age (MA) and 2. Chronological age (CA), and conducted all analyses twice.

MA was calculated based on scores on the Kaufman Brief Intelligence Test-2 (KBIT-2), a commonly-used and well-normed IQ test for individuals between four and 90 years (Kaufman & Kaufman, 2004).[15] The test has high internal-consistency reliability and high external validity. For example, the correlation between the KBIT-2 Verbal scores and corresponding portions of the WISC-III are .83 and .79. The test can also effectively assess individuals with lower cognitive abilities (Homack & Reynolds, 2007).

The K-BIT-2 has three subsections: Verbal Knowledge, Matrices, and Riddles. Scores on the Verbal Knowledge subtest and Riddles subtest are used to calculate a verbal IQ score, and scores on the Matrices subtest are used to calculate a nonverbal IQ score. MA was calculated by averaging the verbal MA and nonverbal MA scores.

Procedure

Children were tested in two sessions separated by approximately one week. In the first session, the child completed the battery of tasks described in Chapter II. In the second session, the child completed the KBIT-2. All sessions were video recorded for later coding.

Results

Mental age

[15] Because the KBIT-2 is designed for children ages four and older, any score with a corresponding mental age of less than four was coded as 3.5.

The ASD group had a mean MA of 6.06 (*SE* = .52, range = 3.50 - 11.33), and the TD group had a mean MA of 5.18 (*SE* = .40, range = 3.5-12.33). This difference was not significant (unpaired $t(48) = .45$, $p = .65$). In the TD group, MA and CA were very highly correlated ($r = .93$, $p < .0005$), suggesting that my TD participants were of normal intelligence. However, not surprisingly, in the ASD group, MA and CA were not correlated ($r = .24$, $p = .27$). Visual inspection of the data revealed that there were not outliers in either the TD or ASD group in terms of the relationship between MA and CA.

Development of cognitive predictors in ASD

As in the previous study, one goal of this study was to determine the relationship between developments in children's categories and in children's other cognitive abilities (vocabulary, ability to group and justify, flexible thinking, declarative knowledge, metamemorial awareness, and productive syntax). I first investigated the relationships between mental and chronological age and each of these cognitive abilities in both ASD and TD children so that I could later see whether any of these abilities predict categorization behavior beyond what can be predicted by mental and chronological age.

Language.

Superordinate level vocabulary. On their first attempt, participants correctly identified 7.28 out of 8 possible items (*SE* = .16), with ASD participants correctly identifying 7.29 items (*SE* = .23), and TD participants correctly identifying 7.27 items (*SE* = .23). As shown in Figure 3.1, Pearson's correlation analyses revealed that MA and Vocabulary scores were not correlated in the TD group ($r = .40$, $p = .052$) or in the TD group ($r = .29$, $p = .151$). CA and Vocabulary scores were also not correlated in the ASD group ($r = .40$, $p = .054$), although they were marginally correlated in the TD group ($r =$

.40, p = .043). An ANCOVA with MA as a covariate and Group as a between subjects variable revealed that older children had marginally better Vocabulary scores than younger children (p = .015), although Group (p = .926) was not significant predictor. However, an ANCOVA with CA as a covariate and Group as a between subjects variable revealed CA was a significant predictor of Vocabulary score ($F(1,47) = 7.37$, p = .009), although Group was not (p = .071).

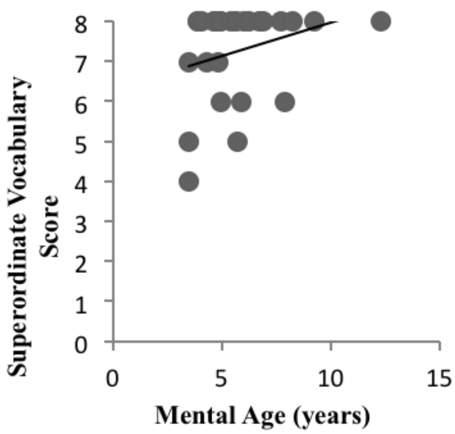


Fig 3.1A. TD participants

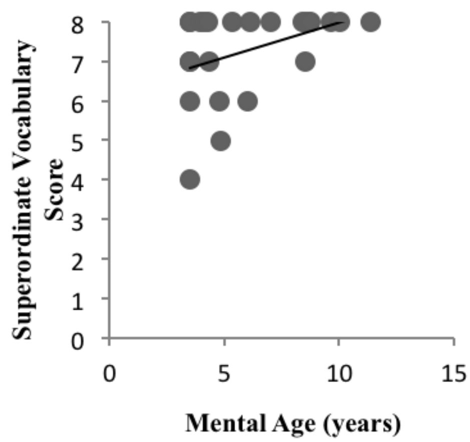


Fig 3.1B. ASD participants

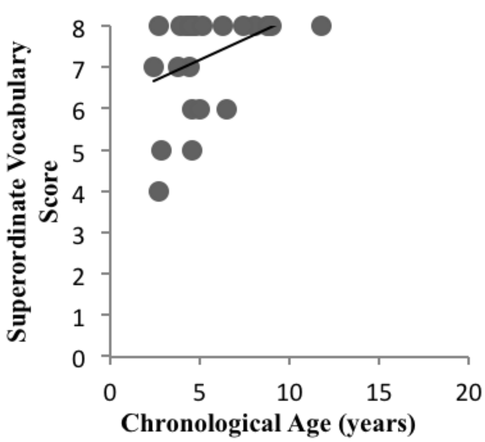


Fig 3.1C. TD participants

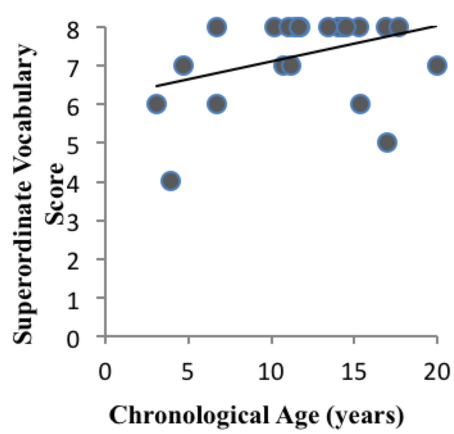


Fig 3.1D. ASD participants

Fig 3.1. Age and Superordinate Vocabulary in TD and ASD

Intermediate level vocabulary. Participants identified an average of 18.70/20 animals (range = 11 – 20, SE = .31) on their first attempt, with the ASD group correctly identifying an average of 18.13 animals (range = 11 – 20, *SE* = .59) and the TD group correctly identifying 19.23 animals (range = 17 – 20, *SE* = .20). As shown in Figure 3.2, Pearson's correlation analyses revealed that MA and vocabulary scores were correlated in the ASD group (r = .50, p = .01) but were only marginally correlated in the TD group (r = .50, p = .03). Conversely, CA and vocabulary score were correlated in the TD group (r = .52, p = .006), but not in the ASD group (r = .19, p = .38). An ANCOVA with MA as a covariate and Group as a between subjects variable revealed that MA was a significant predictor of Vocabulary score (F(1,47) = 12.12, p =.001), with higher MA predicting higher Vocabulary scores. However, TD children had only marginally better Vocabulary scores than ASD children (F(1,47) = 5.04, p = .03). An ANCOVA with CA as a covariate and Group as a between subjects variable revealed that CA did not predict Vocabulary score (p = .093), although again, TD children had marginally better Vocabulary scores than ASD children (p = .015).

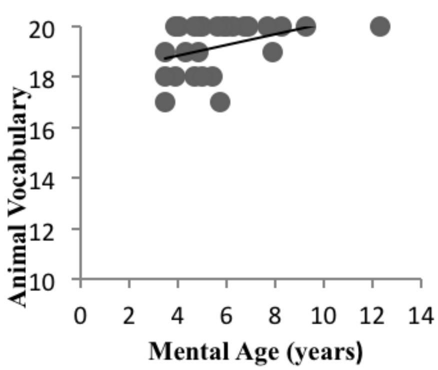


Fig 3.2A. TD participants

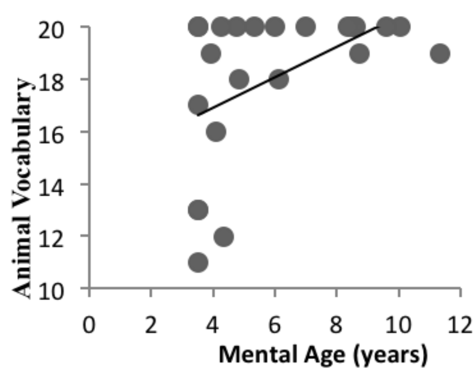


Fig 3.2B. ASD participants

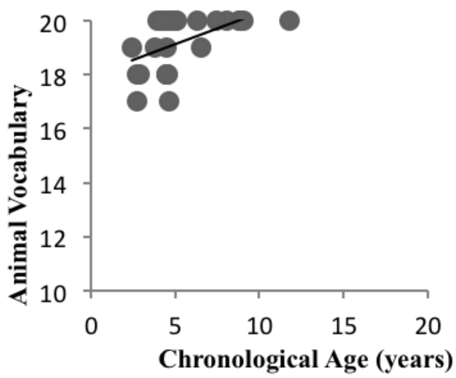


Fig 3.2C. TD participants

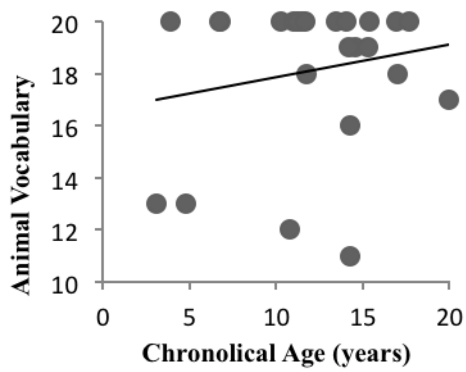
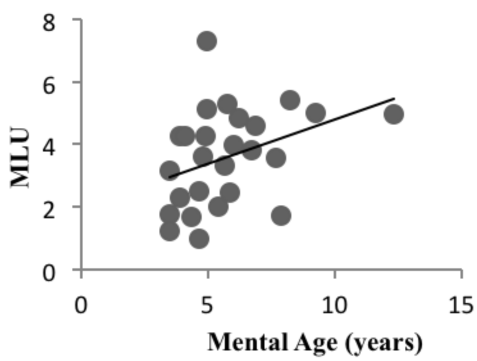


Fig 3.2D. ASD participants

Figure 3.2. Age and Intermediate Vocabulary in TD and ASD

Productive syntax. Across groups, average MLU was 3.18 (range = 0 – 7.30, *SE* = .27), with the ASD group having an average MLU of 2.74 (range = 0 – 7.20, *SE* = .44) and the TD group having an average MLU of 3.59 (range = 1.00 – 7.30, *SE* = .30). As shown in Figure 3.3, Pearson's correlation analyses revealed that MA was correlated with MLU in the ASD group (r's = .62, p = .001) but not in the TD group (p = .06). However, CA was not correlated with MLU in the ASD group (r = .12, p = .58) but was marginally correlated with MLU in the TD group (r = .40, p = .044).

An ANCOVA with MA as a covariate and Group as a between subjects variable revealed a significant effect of MA on MLU (F(1,47) = 17.42, p < .0005), with higher MA predicting longer MLU. However, there was not a significant effect of Group on MLU (F(1,47) = 4.56, p = .038). An ANCOVA with CA as a covariate rather than MA revealed that neither CA did not significantly predict MLU (F(1,47) = 1.91, p = .174) but that TD children had marginally greater MLUs than ASD children (F(1,47) = 4.52, p = .039).



Fix 3.3A. TD participants

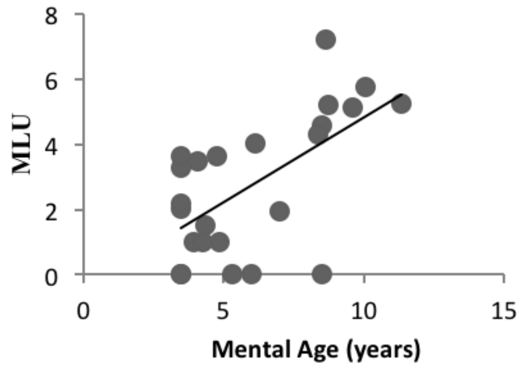


Fig 3.3B. ASD participants

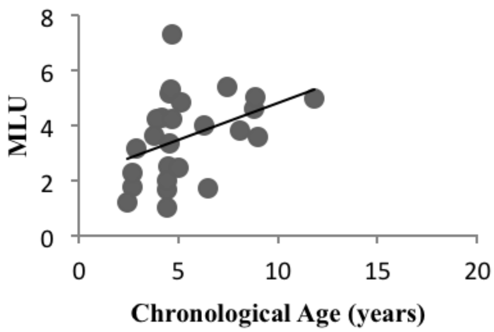


Fig 3.3C. TD participants

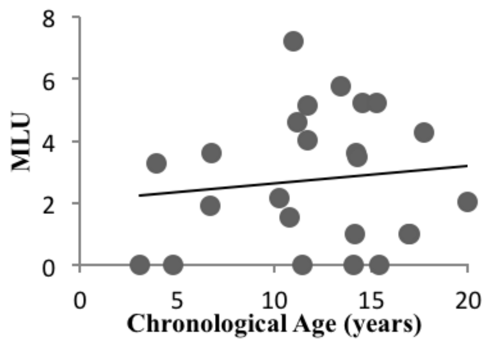


Fig 3.3D. ASD participants

Fig 3.3. Age and MLU in TD and ASD

Perceptual Grouping.

Grouping. Overall, 49/50 participants (98%) grouped on the Perceptual Grouping task, including all 26 TD children and all but one ASD child (95.83%). On this task, 21/24 ASD children (88%) and 21/26 TD children (81%) grouped by shape; 17/24 ASD children (71%) and 17/26 TD children (65%) grouped by color; 17/24 ASD children (71%), and 18/26 TD children (69%) grouped by texture; and 21/24 ASD children (88%) and 20/26 TD children (77%) grouped by number. A set of multiple logistic regression analyses with MA and Group as predictors and each dimension as a dependent variable revealed that the overall fit of the models were not good, with neither MA nor Group being a significant independent predictor of any of the grouping dimensions. Similarly, a set of multiple logistic regression analyses with CA rather than MA as a predictor revealed that the overall fit of the models were not good, with neither CA nor Group being a significant independent predictor of any of the grouping dimensions. These results are consistent with those in the previous chapter in which there was no effect of (chronological) age on grouping dimensions.

Justification. Overall, 39/50 participants (78%) justified, including 18/24 ASD participants (75%) and 21/26 TD participants (80.80%). As shown in Table 3.1, a multiple logistic regression analysis of the perceptual justification data with MA and Group as predictors revealed that the overall fit of the model was good (Log Likelihood = -16.80, $\chi^2(2)= 19.11$, $p < .0005$, Nagelkerke $R^2 = .49$), with MA but not Group being a good independent predictor of perceptual justification.

Table 3.1. Logistic Regression Analysis of Perceptual Justification with MA and Group as Predictors

Predictor	**B**	**Wald's χ^2**	***df***	***P***	**Odds Ratio**
Constant	-6.22	5.04	1	.025	.00
MA	1.61	6.61	1	.010	4.98
Group (1 = ASD, 0 = TD)	-.01	.00	1	.995	.99
Overall Model Evaluation	**Log Likelihood**	**χ^2**	***df***	***p***	**Nagelkerke R^2**
	-16.68	19.11	2	< .0005	.49

As shown in Table 3.2, a logistic regression analysis with CA instead of MA revealed that the overall fit of the model was only fair (Log Likelihood = 22.27, $\chi^2(2)$ = 8.15, p = .017, Nagelkerke R^2 = .23), that older children justified marginally more than younger children, and that TD children justified marginally more than ASD children.

Table 3.2. Logistic Regression Analysis of Perceptual Justification with CA and Group as Predictors

Predictor	**B**	**Wald's χ^2**	***df***	***p***	**Odds Ratio**
Constant	-.10	.02	1	.895	.11
CA	.00	5.90	1	.015	1.00
Group (1 = ASD, 0 = TD)	-2.20	4.43	1	.035	.11
Overall Model Evaluation	**Log Likelihood**	**χ^2**	***df***	***P***	**Nagelkerke R^2**
	22.27	8.15	2	.017	.23

On the Perceptual Grouping task, 12/24 ASD children (50%) and 17/26 TD children (65%) justified by shape; 10/24 ASD children (42%) and 14/26 TD children (54%) justified by color; 4/24 ASD children (17%) and 7/26 TD children (27%) justified by texture; and 4/24 ASD children (17%) and 9/26 TD children (35%) justified by number. A set of multiple logistic regression analyses with MA and Group as predictors and each justification dimension as a dependent variable revealed that the overall fit of the model was only good when Shape was entered as the dependent variable. As shown in Table 3.3, in this model MA but not Group was a significant independent predictor of justifying by shape, with older children justifying more than younger children.

Table 3.3 Logistic Regression Analysis of Shape Justification on the Perceptual Grouping Task with MA and Group as Predictors

Predictor	**B**	**Wald's χ^2**	***df***	***p***	**Odds Ratio**
Constant	-3.81	7.96	1	.005	.02
MA	.84	10.84	1	.001	2.33
Group (1 = ASD, 0 = TD)	-1.05	2.00	1	.16	.35
Overall Model Evaluation	**Log Likelihood**	**χ^2**	***df***	***p***	**Nagelkerke R^2**
	-23.44	21.16	2	< .0005	.46

Similarly, as shown in Table 3.4, when the multiple logistic regression analysis was repeated with CA rather than MA as a predictor, the overall fit of the model was only good when Shape was entered as a dependent variable, and both CA and Group were significant independent predictors of justifying by shape, with older children justifying more than younger children and TD children justifying more than ASD children. These results are consistent with those in the previous chapter in which Shape was the only justification dimension for which there was an effect of age.

Table 3.4 Logistic Regression Analysis of Shape Justification on the Perceptual Task with CA and Group as Predictors

Predictor	**B**	**Wald's χ^2**	***df***	***p***	**Odds Ratio**
Constant	-1.01	1.82	1	.18	.36
CA	.32	6.44	1	.01	1.38

Group (1 = ASD, 0 = TD)	-2.98	6.26	1	.01	.05
Overall Model Evaluation	**Log Likelihood**	***χ²***	***df***	***p***	**Nagelkerke *R*²**
	-28.59	10.86	2	.004	.26

Switching grouping dimensions. As shown in Figure 3.4, 36/50 participants (72%) grouped by two or more dimensions, including 14/24 ASD participants (58.33%), and 22/26 TD participants (84.62%).[16] As shown in Table 3.5, a multiple logistic regression analysis with MA and Group as predictors revealed that the overall model was only a fair for the data (Log Likelihood = -25.99, $\chi^2(2) = 7.33$, $p = .026$, Nagelkerke $R^2 = .20$) and that MA did not predict switching grouping dimensions, although TD children switched grouping dimensions marginally more than ASD children.

[16] Given that group sizes were smaller in this study than in Chapter II, in the graphs I divided the children into one-year age brackets (as opposed to half-year age brackets). As in the previous study, use of these brackets and of percentages is for the purpose of depiction only.

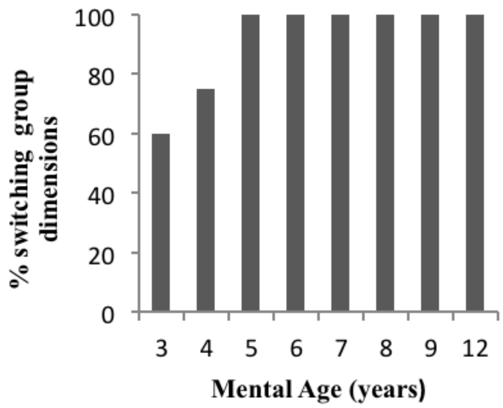


Fig 3.4A. TD participants

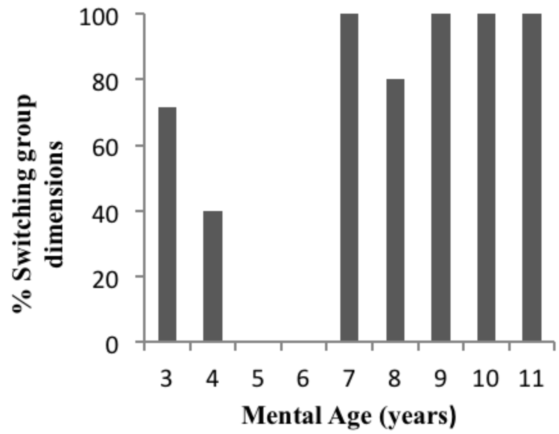


Fig 3.4B. ASD participants

Figure 3.4. Age and dimension-switching in Perceptual task groups in TD and ASD

Table 3.5. Logistic Regression Analysis of Switching Perceptual Grouping Dimensions with MA and Group as Predictors

Predictor	**B**	**Wald's χ^2**	***df***	***p***	**Odds Ratio**
Constant	.23	1.03	1	.825	1.26
MA	.27	2.56	1	.109	1.31
Group (1 = ASD, 0 = TD)	-1.49	4.45	1	.035	.23
Overall Model Evaluation	**Log Likelihood**	**χ^2**	***df***	***p***	**Nagelkerke R^2**
	25.99	7.33	2	.026	.20

As shown in Table 3.6, when CA was used rather than MA, the overall fit of the model was also was only fair, and CA was not a significant predictor, but Group was a significant independent predictor of switching grouping dimensions (p = .008), with TD children switching grouping dimensions more than ASD children.

Table 3.6. Logistic Regression Analysis of Switching Perceptual Grouping dimensions with CA and Group as Predictors

Predictor	**B**	**Wald's χ^2**	***df***	***p***	**Odds Ratio**
Constant	.70	.87	1	.350	2.01
CA	.00	3.52	1	.061	1.00
Group (1 = ASD, 0 = TD)	-2.73	6.97	1	.008	.07
Overall Model Evaluation	**Log Likelihood**	χ^2	***df***	***P***	**Nagelkerke R^2**
	-25.36	8.59	2	.014	.23

Switching justification dimensions. As shown in Figure 3.5, 22/50 participants (44%) switched justification dimensions, including 7/24 ASD participants (29.17%) and 15/ 26 TD participants (57.69%). As shown in Table 3.7, a multiple logistic regression analysis with MA and Group as independent variables revealed that the overall model was a good fit for the data, with there being a positive linear effect of MA and with TD children switching justification dimensions marginally more than ASD children.

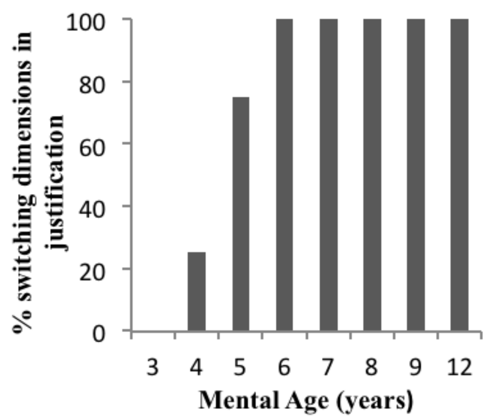


Fig 3.5A. TD participants

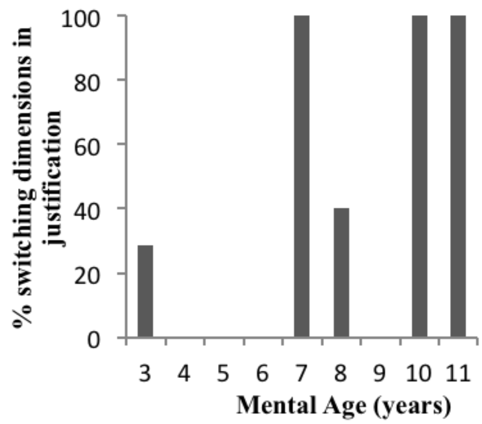


Fig 3.5B. ASD participants

Figure 3.5. Age and dimension-switching in perceptual task justifications in TD and ASD

Table 3.7. Logistic Regression Analysis of Switching Perceptual Grouping Justification Dimensions with MA and Group as Predictors

Predictor	B	Wald's χ^2	*df*	*p*	Odds Ratio
Constant	-2.94	.05	1	.825	1.26
MA	.59	9.10	1	.003	1.31
Group (1 = ASD, 0 = TD)	-1.89	5.69	1	.017	.23
Overall Model Evaluation	**Log Likelihood**	χ^2	*df*	*p*	**Nagelkerke R^2**
	-25.64	17.31	2	< .0005	.39

Similarly, as shown in Table 3.8, when CA rather than MA was included as a predictor, the overall model was a good for the data, and CA was a marginally significant independent predictor, with older children switching dimensions marginally more than younger children. However, in this model, Group was also a significant independent predictor.

Table 3.8. Logistic Regression Analysis of Switching Perceptual Grouping Justification Dimensions with CA and Group as Predictors

Predictor	**B**	**Wald's χ^2**	***df***	***p***	**Odds Ratio**
Constant	-1.54	3.46	1	.006	.22
CA	.001	6.13	1	.013	1.00
Group (1 = ASD, 0 = TD)	-4.09	7.68	1	.006	.02
Overall Model Evaluation	**Log Likelihood**	**χ^2**	***df***	***p***	**Nagelkerke R^2**
	-27.56	13.47	2	.001	.32

Animal Declarative Knowledge. Overall, participants correctly answered an average of 13.92/20 questions on the Animal Declarative Knowledge task (range = 1 – 20, *SE* = .69), with the ASD participants correctly answering an average of 12.46 questions (range = 1 – 20, *SE* = 1.18) and the TD children correctly answering an average of 15.27 questions (range = 7 – 20, *SE* = .67). As shown in Figure 3.6, Pearson's correlation analyses revealed that both ASD and TD participants' Declarative Animal Knowledge scores were highly correlated with both MA and CA (all *r*s > .55, all *p*s < .005).

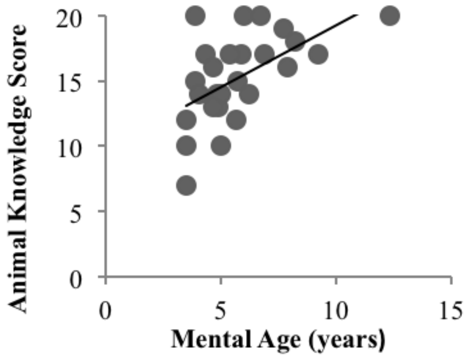


Fig 3.6A. TD participants

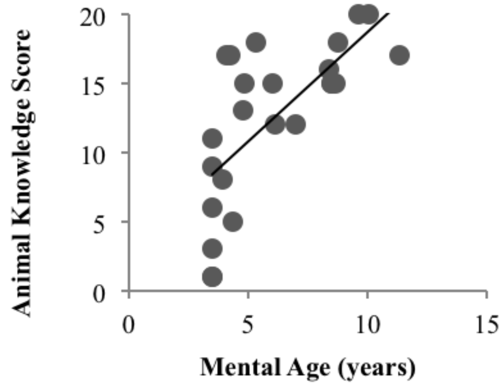


Fig 3.6B. ASD participants

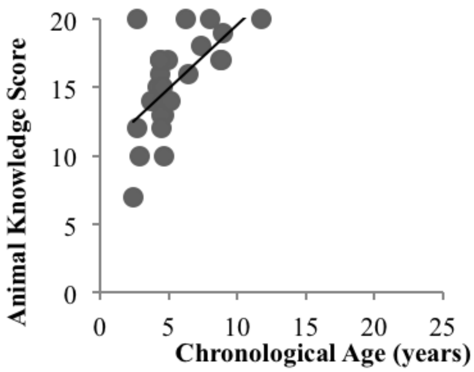


Fig 3.6C. TD participants

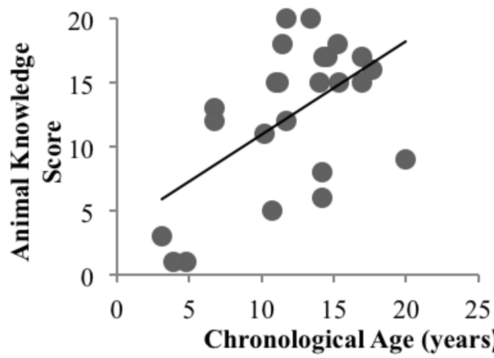


Fig 3.6D. ASD participants

Fig 3.6. Age and Animal Declarative Knowledge in TD and ASD

An ANCOVA with MA as a covariate and Group as a between subjects variable. revealed a significant effect of MA ($F(1,47) = 33.11$, $p < .0005$), with higher MA predicting higher Animal Declarative Knowledge score, and a significant effect of Group ($F(1,47) = 9.67$, $p = .003$), with TD children outperforming ASD children.

Similarly, when CA was used as a covariate, there was a significant effect of CA ($F(1,47) = 23.25$, $p < .0005$), with chronologically older children correctly answering more questions than chronologically younger children, and a significant effect of Group ($F(1,47) = 27.16$, $p < .0005$), with TD children outperforming ASD children.

Metamemory. Participants correctly answered an average of 2.70/ 5 Metamemory questions (range = 0 – 5, *SE* = .17), with the ASD group correctly answering an average of 2.27 questions (range = 0 – 5, *SE* = .27) and the TD group correctly answering an average of 3.08 questions (range = 1 – 5, *SE* =.21). As shown in Figure 3.7, Pearson's correlation analyses revealed that MA was a significant predictor of both ASD and TD children's Metamemory scores (both *r*s > .55, both *p*s < .005), and CA was a significant predictor of TD participants' Metamemory scores ($r = .60$, $p = .001$), but CA was not a significant predictor of ASD participants' Metamemory scores ($r = .34$, $p = .09$).

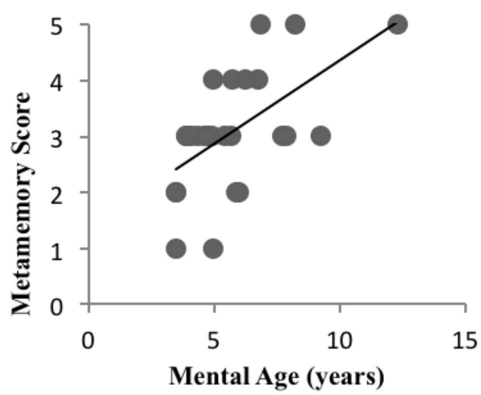


Fig 3.7A. TD participants

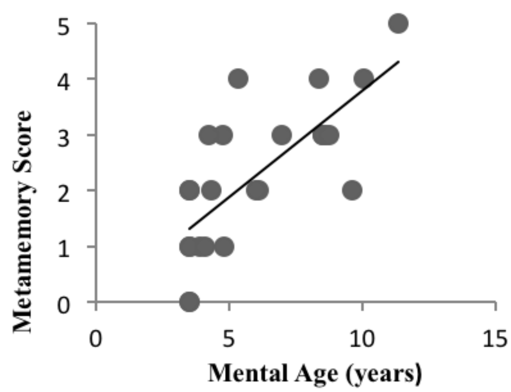


Fig 3.7B. ASD participants

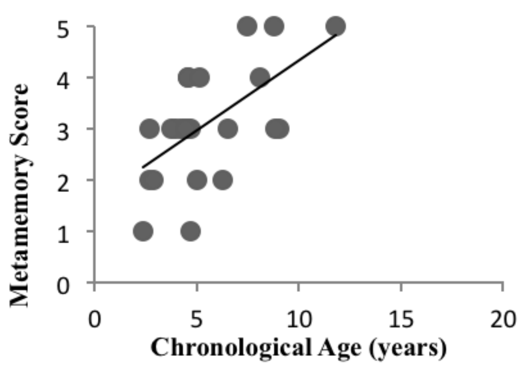


Fig 3.7C. TD participants

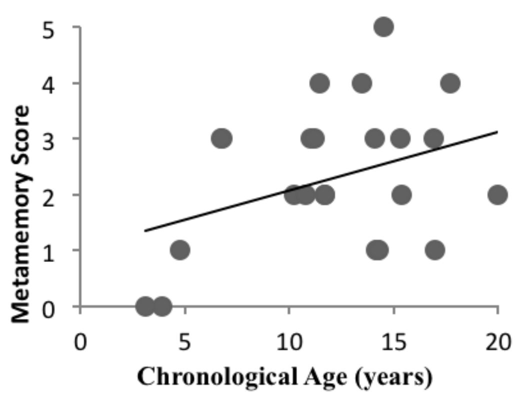


Fig 3.7D. ASD participants

Fig 3.7. Age and metamemory in TD and ASD

An ANCOVA with MA as a covariate and Group as a between subjects variable revealed that participants with higher MAs correctly answered significantly more questions than participants with lower MAs ($F(1,47) = 38.60$, $p < .0005$) and that TD participants correctly answered significantly more Metamemory questions than ASD participants ($F(1,47) = 12.55$, $p = .001$). When this analysis was performed with CA rather than MA as a covariate, the results were similar, with significant effects of both CA ($F(1,47) = 10.36$, $p = .002$) and Group ($F(1,47) = 16.68$, $p < .0005$).

Effects of MA, CA, and Group on categorization

In the previous chapter, I investigated the effect of (chronological) age on children's grouping and justification on the Superordinate level, Intermediate level, and Basic level tasks. Similarly, in this chapter, I use logistic regression analyses to investigate the effects of mental/chronological age and group on categorization behavior on these tasks.

Superordinate level task.

Grouping. As shown in Figures 3.8A and 3.8B, 24/50 participants (48%) grouped taxonomically (i.e., created a food or clothing group), including 10/24 ASD participants (41.67%) and 21/26 TD participants (80.80%) doing so. As shown in Table 3.9, a multiple logistic regression analysis with MA and Group as independent variables revealed that the overall model was a good fit for the data, with MA but not Group being a significant predictor of taxonomic grouping.

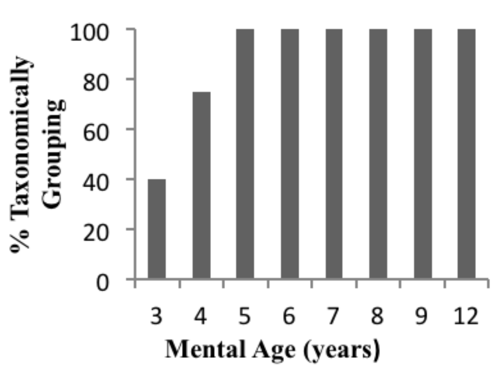


Fig 3.8A. TD participants

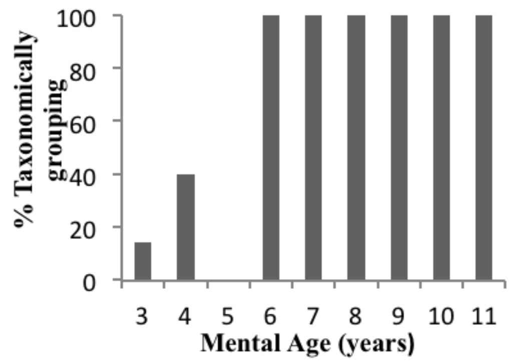


Fig 3.8B. ASD participants

Fig 3.8. Age and taxonomic grouping on the Superordinate level task in TD and ASD

Table 3.9. Logistic Regression Analysis of Taxonomic Grouping on the Superordinate Level Task with MA and Group as Predictors

Predictor	**B**	**Wald's χ^2**	***df***	***p***	**Odds Ratio**
Constant	-7.41	6.90	1	.009	.001
MA	1.88	8.56	1	.003	6.56
Group (1 = ASD, 0 = TD)	-1.59	2.82	1	.093	.21
Overall Model Evaluation	**Log Likelihood**	**χ^2**	***df***	***p***	**Nagelkerke R^2**
	-14.52	32.05	2	< .0005	.67

As shown in Table 3.10, when the logistic regression analysis was repeated using CA rather than MA, the model was a good fit for the data, but neither CA nor Group was a significant independent predictor of taxonomic grouping.

Table 3.10. Logistic Regression Analysis of Taxonomic Grouping on the Superordinate Level Task with CA and Group as Predictors

Predictor	B	Wald's χ^2	*df*	*p*	Odds Ratio
Constant	.984	2.08	1	.149	.2.67
CA	< .0005	.90	1	.343	1.00
Group (1 = ASD, 0 = TD)	-1.68	3.50	1	.061	.19
Overall Model Evaluation	**Log Likelihood**	χ^2	*df*	*p*	**Nagelkerke R^2**
	-14.05	8.61	2	.01	.30

Taxonomic justification. As shown in Figure 3.9, 28/50 participants (56%) justified taxonomically on the Superordinate level task, including 12/24 ASD participants (50%) and 16/26 TD participants (51.54%). As shown in Table 3.11, a multiple logistic regression analysis with MA and Group as independent variables revealed that the overall model was a good fit for the data, with MA but not Group being a significant independent predictor.

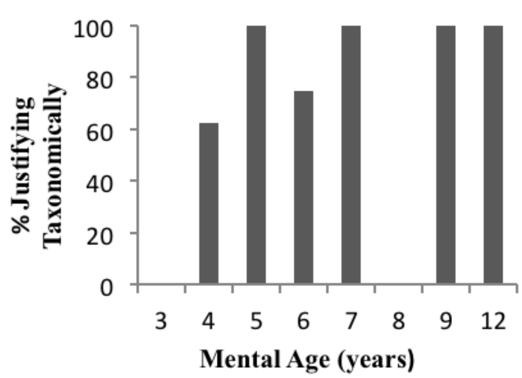


Fig 3.9A. TD participants

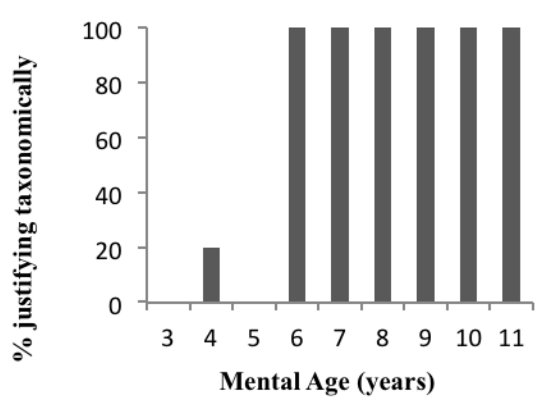


Fig 3.9B. ASD participants

Fig 3.9. Age and taxonomic justification on the Superordinate level task in TD and ASD

Table 3.11. Logistic Regression Analysis of Superordinate Level Taxonomic Justification with MA and Group as Predictors

Predictor	**B**	**Wald's χ^2**	***df***	***p***	**Odds Ratio**
Constant	-6.59	10.40	1	.001	.001
MA	1.36	11.41	1	.001	3.90
Group (1 = ASD, 0 = TD)	-.94	1.17	1	.280	.39
Overall Model Evaluation	**Log Likelihood**	χ^2	***df***	***p***	**Nagelkerke R^2**
	-18.24	32.14	2	< .0005	.64

As shown in Table 3.12, when the analysis was repeated with CA rather than MA, the overall model was not a good fit, and neither average CA nor Group was an independent predictor of taxonomic justification.

Table 3.12. Logistic Regression Analysis of Taxonomic Justification on the Superordinate Level Task with CA and Group as Predictors

Predictor	**B**	**Wald's χ^2**	***Df***	***p***	**Odds Ratio**
Constant	-.724	1.15	1	.284	.49
CA	.00	4.66	1	.031	1.00
Group (1 = ASD, 0 = TD)	-2.07	4.33	1	.037	.13
Overall Model Evaluation	**Log Likelihood**	χ^2	***df***	***p***	**Nagelkerke R^2**

-31.02	6.56	2	.038	.17

Thematic justification. As shown in Figure 3.10, 6/50 participants (12%) justified thematically on the Superordinate level task, including 0/26 ASD participants (0%) and 6/26 TD participants (23%). As shown in Table 3.13, a multiple logistic regression analysis with MA and Group revealed that the model was not a good fit for the data and that neither MA nor Group was a significant independent predictor. In the previous chapter I found an inverted U-shaped effect of age on thematic justification. In this study, given that only six TD children justified taxonomically, when I conducted a logistic regression analysis with the negative absolute value of the *z* score of TD children's MA as the independent variable and TD children's thematic justification as the dependent variable, the model was only a fair fit for the data (Log likelihood = -10.77, Nagelkerke R^2 = .34, Wald $\chi 2$ (1, N = 91) = 3.90, p = .05, Odds ratio = 36.02). More specifically, all six children who justified thematically fell between the mental ages of four and six years; thus, the pattern is the same for TD children in both studies. Conversely, no ASD children justified thematically on the Superordinate level.

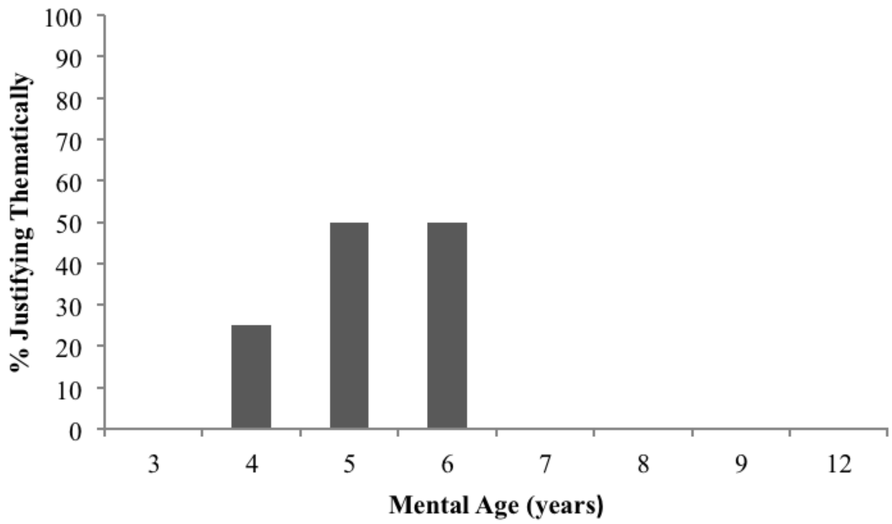


Fig 3.10. Age and thematic justification on the Superordinate level task in TD

Table 3.13. Logistic Regression Analysis of Superordinate Level Thematic Justification with MA and Group as Predictors

Predictor	**B**	**Wald's χ^2**	***df***	***p***	**Odds Ratio**
Constant	-1.13	.631	1	.428	.32
MA	-.01	.003	1	.955	.99
Group (1 = ASD, 0 = TD)	-20.00	.00	1	.998	.00
Overall Model Evaluation	**Log Likelihood**	**χ^2**	***df***	***p***	**Nagelkerke R^2**
	-14.05	8.61	2	.014	.30

Similarly, as shown in Table 3.14, when the analysis was repeated with CA rather than MA, the model was a good fit for the data, but neither CA nor Group was a good independent predictor of thematic justification. Additionally, there was not a U-shaped effect of CA on TD children thematic justification (Log likelihood = -12.55, Nagelkerke

R^2 = .17, Wald $\chi 2$ (1, N = 91) = 1.95, p = .16, Odds ratio = 7.44). However, all thematically-justifying children fell between the chronological ages of 4.5 and 6.5, again following the pattern observed in the previous chapter.

Table 3.14. Logistic Regression Analysis of Superordinate Level Thematic Justification with CA and Group as Predictors

Predictor	**B**	**Wald's χ^2**	***Df***	***p***	**Odds Ratio**
Constant	-1.26	1.11	1	.29	.28
CA	.00	.003	1	.96	1.00
Group (1 = ASD, 0 = TD)	-20.07	.00	1	.99	.00
Overall Model Evaluation	**Log Likelihood**	**χ^2**	***df***	***p***	**Nagelkerke R^2**
	-14.05	8.61	2	.01	.30

Perceptual Justification. No ASD children and only one TD child justified perceptually on the Superordinate level task. Therefore, I did not analyze this data further.

Intermediate level Task.

Taxonomic Justification. As shown in Figure 3.11, 24/50 participants (48%) provided taxonomic justifications, including 10/24 ASD participants (41.66%) and 14/26 TD participants (53.85%). As shown in Table 3.15, a multiple logistic regression analysis with MA and Group as independent variables revealed that the overall model was a good fit for the data, but only MA was a good independent predictor of taxonomic justification ($p < .0005$)

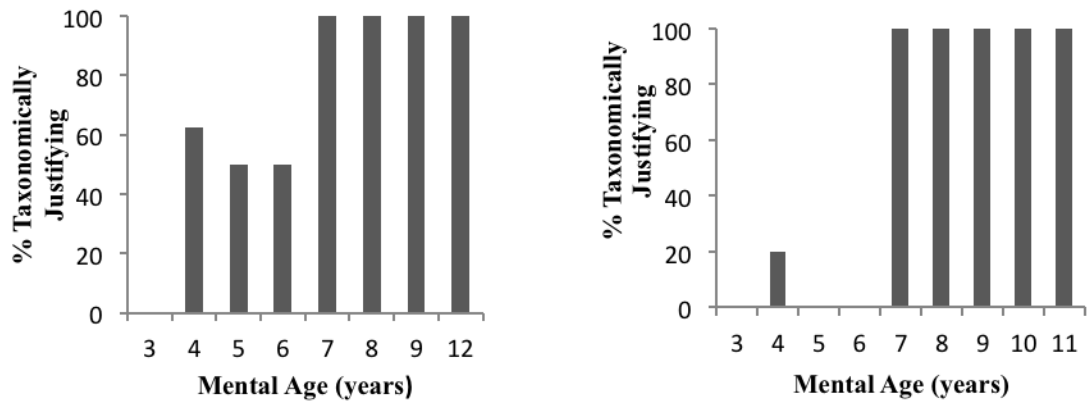


Fig 3.11A. TD participants **Fig 3.11B.** ASD participants

Figure 3.11. Age and taxonomic justification on the Intermediate level task in TD and ASD

Table 3.15. Logistic Regression Analysis of Intermediate Level Taxonomic Justification with MA and Group as Predictors

Predictor	**B**	**Wald's χ^2**	***df***	***p***	**Odds Ratio**
Constant	-6.21	12.54	1	< .0005	.002
MA	1.17	13.18	1	< .0005	3.21
Group (1 = ASD, 0 = TD)	-1.35	2.19	1	.139	.26
Overall Model Evaluation	**Log Likelihood**	**χ^2**	***df***	***p***	**Nagelkerke R^2**
	-18.92	31.39	2	< .0005	.62

As shown in Table 3.16, when the analysis was repeated with CA rather than MA as an independent variable, the overall model was only a fair fit for the data, older children taxonomically justified marginally more than younger children, and TD children taxonomically justified marginally more than ASD children.

Table 3.16. Logistic Regression Analysis of Intermediate Level Taxonomic Justification with CA and Group as Predictors

Predictor	**B**	**Wald's χ^2**	***df***	***p***	**Odds Ratio**
Constant	-1.02	2.21	1	.137	.36
CA	.001	4.31	1	.038	1.00
Group (1 = ASD, 0 = TD)	-2.08	4.15	1	.042	.125
Overall Model Evaluation	**Log Likelihood**	**χ^2**	***df***	***p***	**Nagelkerke R^2**
	-31.55	6.13	2	.047	.15

Thematic justification. Overall, 5/50 participants (10%) justified thematically, including 3/24 ASD participants (12.5%) and 2/26 TD participants (7.69%). Given that so few participants justified thematically, I did not further analyze these data.

Perceptual justification. As shown in Figure 3.12, 19/50 participants (38%) justified perceptually, including 5/24 ASD participants (20.83%) and 14/26 TD participants (53.85%). As shown in Table 3.17, a multiple logistic regression analysis with MA and Group as independent variables revealed that the overall model was only a fair fit for the data, and that MA was not a significant predictor, but that TD children perceptually justified marginally more than ASD children.

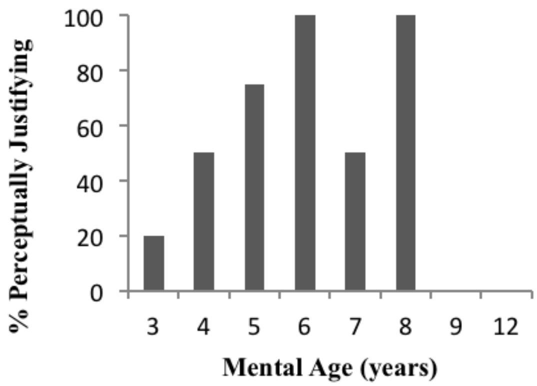


Fig 3.12A. TD participants

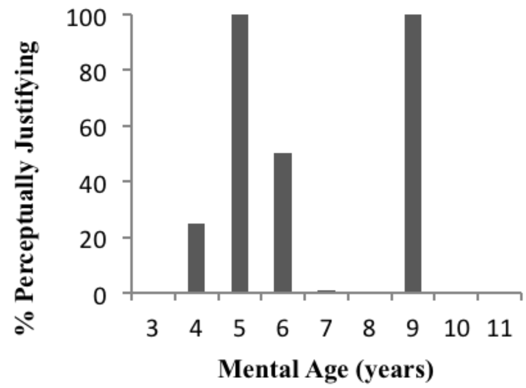


Fig 3.12B. ASD participants

Fig 3.12. Age and perceptual justification on the Intermediate level task in TD and ASD

Table 3.17. Logistic Regression Analysis of Intermediate Level Perceptual Justification with MA and Group as Predictors

Predictor	**B**	**Wald's χ^2**	***df***	***p***	**Odds Ratio**
Constant	-.19	.05	1	.831	.82
MA	.06	.18	1	.669	1.06
Group (1 = ASD, 0 = TD)	-1.51	5.53	1	.019	.22
Overall Model Evaluation	**Log Likelihood**	**χ^2**	***df***	***p***	**Nagelkerke R^2**
	-30.14	6.14	2	.047	.16

As shown in Table 3.18, when the analysis was repeated with CA rather than MA, the overall model was only a fair fit for the data, and neither Group nor CA was a significant independent predictor.

Table 3.18. Logistic Regression Analysis of Intermediate Level Perceptual Justification with CA and Group as Predictors

Predictor	**B**	**Wald's χ^2**	***df***	***p***	**Odds Ratio**
Constant	.47	.53	1	.47	1.60
CA	.00	.38	1	.54	1.00
Group (1 = ASD, 0 = TD)	-1.11	1.67	1	.20	.33
Overall Model Evaluation	**Log Likelihood**	**χ^2**	***df***	***p***	**Nagelkerke R^2**
	-30.04	6.34	2	.04	.16

Notably, among participants with lower mental ages, there appears to be more variability in the TD than the ASD group. As shown in Figure 3.13, about half of TD participants with an MA of less than seven justified taxonomically. In stark contrast, as shown in Figure 3.14, almost none of the ASD participants under this mental age justified taxonomically.

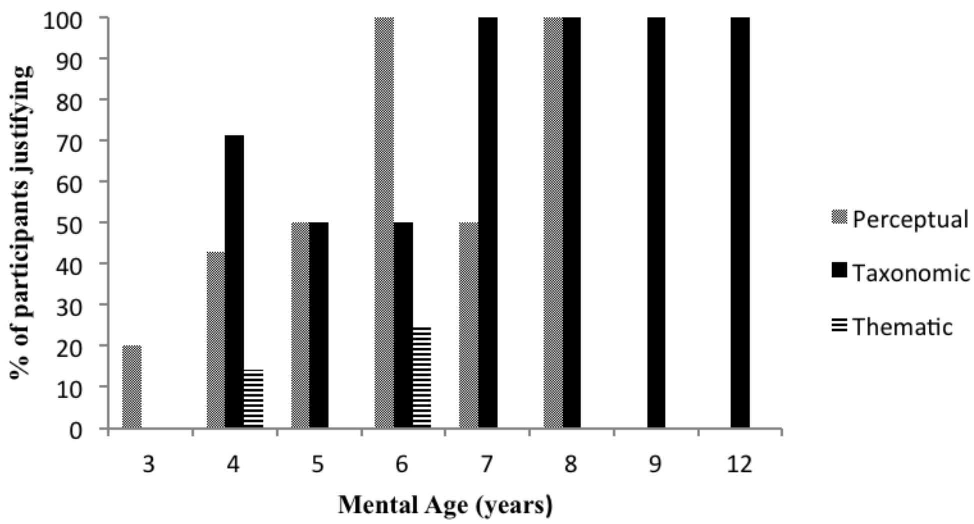


Fig 3.13. Age and justification type on the Intermediate level task in TD

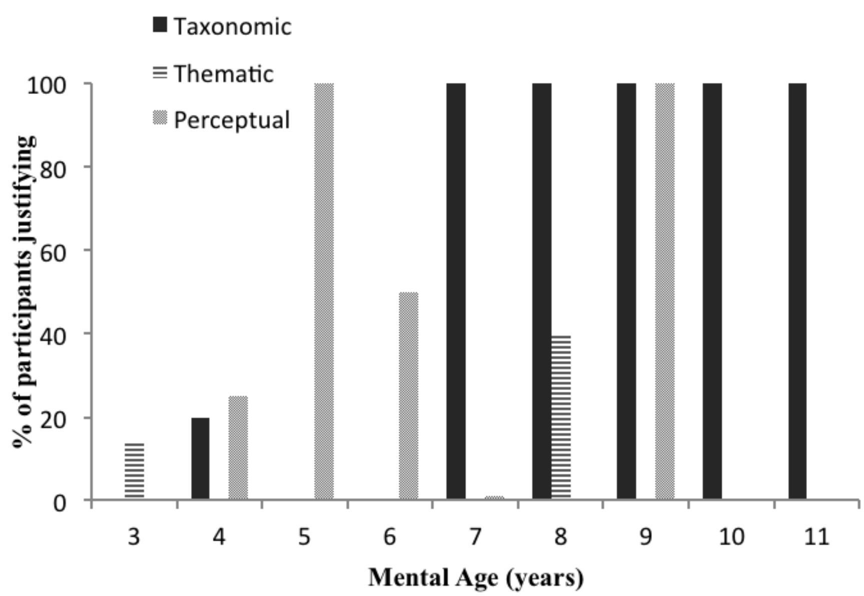


Fig 3.14. Age and justification type on the Intermediate level task in ASD

Basic level task.

Grouping. As shown in Figure 3.15, 41/50 participants grouped taxonomically, including 19/24 ASD participants and 22/26 TD participants. As shown in Table 3.19, a multiple logistic regression analysis with MA and Group as independent variables revealed that the overall model was not a good fit for the data, with neither MA nor Group being a significant independent predictor.

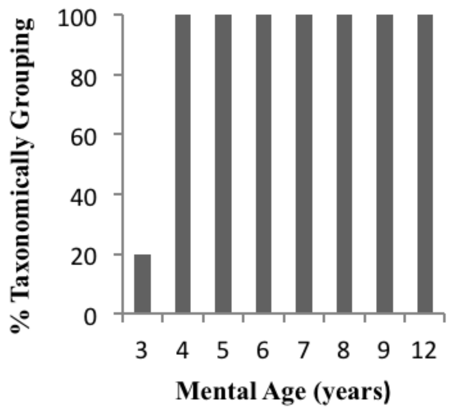


Fig 3.15A. TD participants

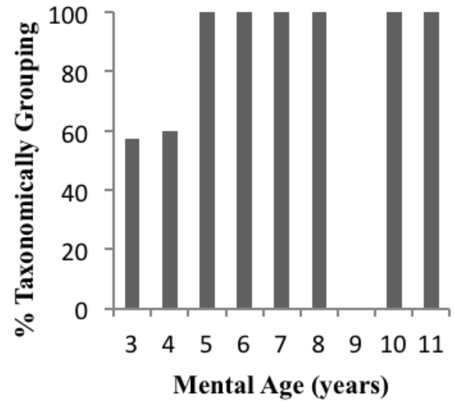


Fig 3.15B. ASD participants

Fig 3.15. Age and taxonomic grouping on the Basic level task

Table 3.19. Logistic Regression Analysis of Basic Level Grouping with MA and Group as Predictors

Predictor	B	Wald's χ^2	*df*	*p*	Odds Ratio
Constant	-1.08	.53	1	.466	.34
MA	.54	3.52	1	.061	1.71
Group (1 = ASD, 0 = TD)	-.38	.24	1	.627	.68
Overall Model Evaluation	**Log Likelihood**	χ^2	***df***	***p***	**Nagelkerke** R^2
	-20.65	5.84	2	.054	.16

Similarly, as shown in Table 3.2, when using CA rather than MA as an independent variable, the analysis revealed that the overall model was not a good fit for the data, and that CA was not a predictor, although TD children grouped marginally more than ASD children.

Table 3.20. Logistic Regression Analysis of Basic level Grouping with CA and Group as Predictors

Predictor	**B**	**Wald's χ^2**	***df***	***p***	**Odds Ratio**
Constant	.44	.30	1	.583	1.55
CA	.001	4.30	1	.038	1.00
Group (1 = ASD, 0 = TD)	-1.87	3.36	1	.067	.16
Overall Model Evaluation	**Log Likelihood**	**χ^2**	***df***	***p***	**Nagelkerke R^2**
	-20.85	5.45	2	.066	.17

Justification. As shown in Figure 3.16, 36/50 participants (72%) justified taxonomically on the Basic level task, including 18/24 ASD participants (75%) and 18/26 TD participants (69%). As shown in Table 3.21, a logistic regression analysis with MA and Group as independent variables revealed that the overall model was a good fit for the data, and that older children justified marginally more than younger children, although Group was not a significant independent predictor.

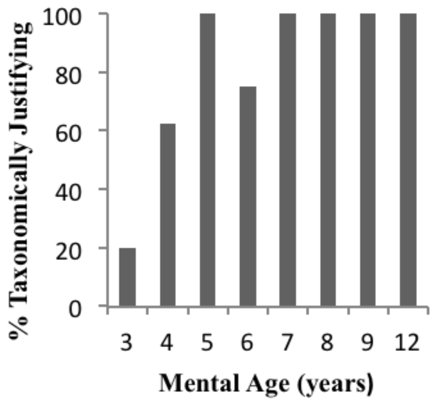


Fig 3.16A. TD participants

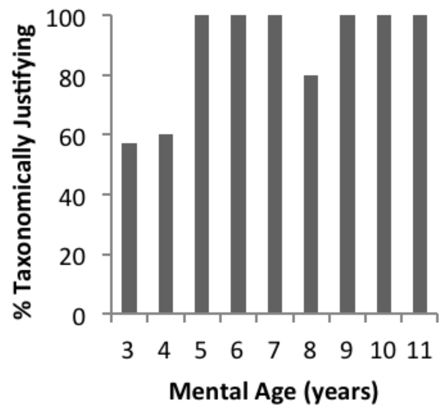


Fig 3.16B. ASD participants

Fig 3.16. Age and taxonomic justification on the Basic level task in TD and ASD

Table 3.21. Logistic Regression Analysis of Basic Level Justification with MA and Group as Predictors

Predictor	**B**	**Wald's χ^2**	***df***	***p***	**Odds Ratio**
Constant	-2.36	3.04	1	.081	.095
MA	.59	5.47	1	.019	1.81
Group (1 = ASD, 0 = TD)	.35	.25	1	.62	1.41
Overall Model Evaluation	**Log Likelihood**	**χ^2**	***df***	***p***	**Nagelkerke R^2**
	-25.09	9.11	2	.01	.24

As shown in Table 3.22, when CA rather than MA was entered as an independent variable, the model was also a good fit for the data, and CA, but not Group, was an independent predictor of taxonomic justification.

Table 3.22. Logistic Regression Analysis of Basic Level Justification with CA and Group as Predictors

Predictor	B	Wald's χ^2	*df*	*p*	Odds Ratio
Constant	-1.06	1.88	1	.170	.345
CA	.001	7.56	1	.006	1.00
Group (1 = ASD, 0 = TD)	-1.89	3.23	1	.072	.151
Overall Model Evaluation	**Log Likelihood**	χ^2	*df*	*p*	**Nagelkerke R^2**
	-23.97	11.36	2	.003	.29

In the previous chapter, a series of multiple logistic regression analyses revealed that none of the cognitive measures exceeded (chronological) *age* in predicting categorization behavior. Similarly, to determine whether any of my measures were significant predictors of categorization behavior independent of *MA* and *Group*, I conducted a series of multiple logistic regression analyses with MA, Group, and one other variable entered as independent variables and Intermediate taxonomic justification as the dependent variable. Just as chronological age was the best predictor in the previous chapter, in each of the current analyses, MA predicted Intermediate taxonomic justification better than any of my cognitive measures.

Grouping vs. Justification

To assess the relationship between taxonomic grouping and taxonomic justification, I compared the grouping and justification results on the three tasks for which these data were available: the Superordinate level task, the Basic level task, and the Perceptual Grouping task.

Superordinate level task. Overall, 35/50 participants (70%) taxonomically grouped and 28/50 participants (56%) taxonomically justified. More specifically, as shown in Figure 3.17, 21/26 TD children (80.80%) grouped taxonomically, whereas only 16/26 (61.50%) justified taxonomically. The difference between grouping and justification was less pronounced in the ASD children. As shown in Figure 3.18, 14/24 ASD children (58.30%) grouped taxonomically, and 12/24 (50%) justified taxonomically. Visual inspection of the data reveals different patterns in the two groups. In the TD group, although there is somewhat of a positive linear effect of age, there is considerable variability. In contrast, in the ASD group, there is a ceiling effect beginning at age six for both grouping and justification. This is similar to the results for the Intermediate level task in which TD children demonstrated greater variability than ASD children.

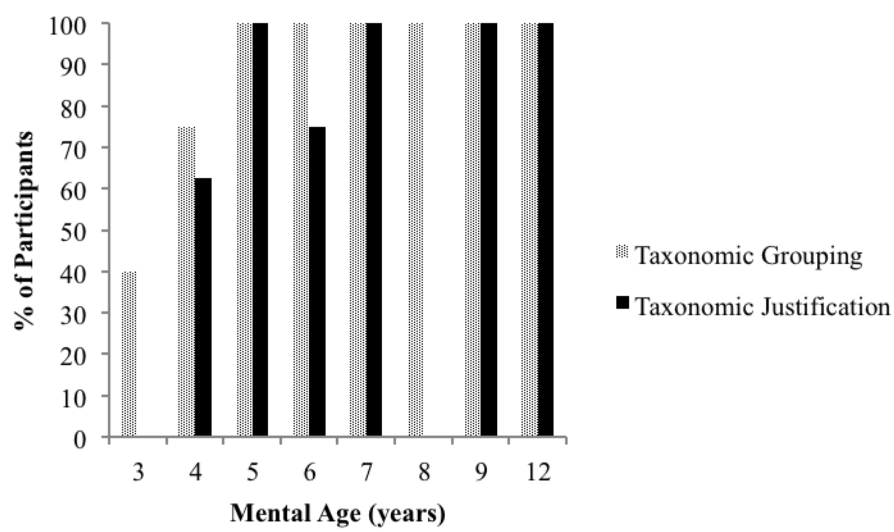


Fig 3.17. Age and taxonomic grouping and justification on the Superordinate level task in TD

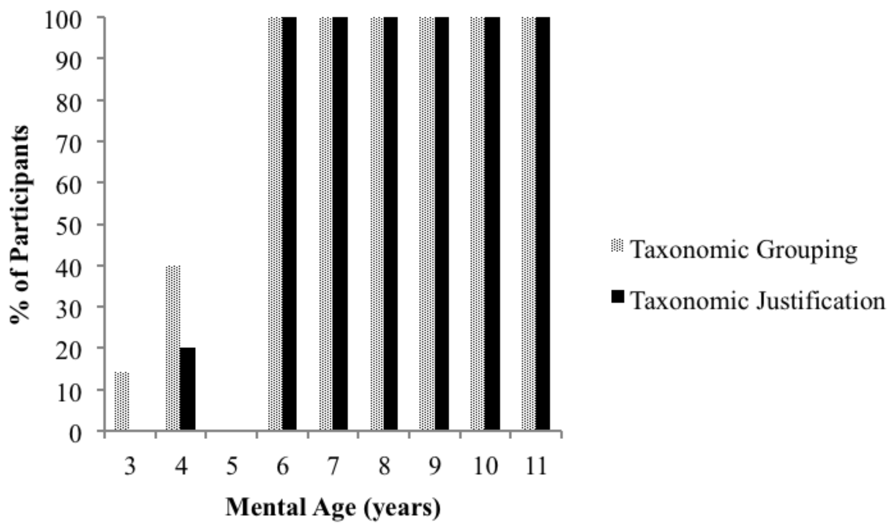


Fig 3.18. Age and taxonomic grouping and justification on the Superordinate level task in ASD

Basic level task. Overall, 41/50 participants (82%) grouped and 36/50 participants (72%) justified on the basic level. More specifically, as shown in Figure 3.19, 22/26 TD children (84.60 %) grouped and 18/26 TD children (69.20%) justified at the basic level. As shown in Figure 3.2, consistent with the results of the Superordinate level task, ASD children exhibited less difference between grouping and justification than TD children did, with 19/24 ASD children (79.20%) grouping and 18/24 ASD children (75.00%) justifying at the basic level.

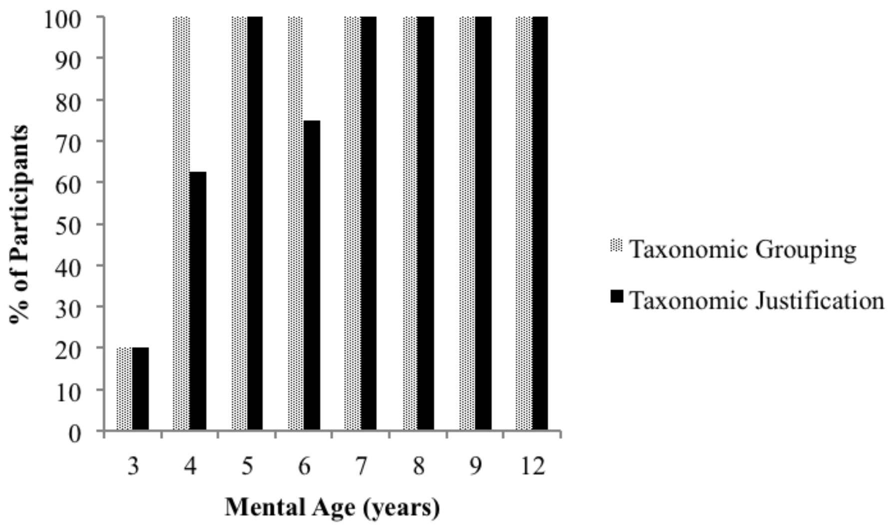


Fig 3.19. Age and taxonomic grouping and justification on the Basic level task in TD

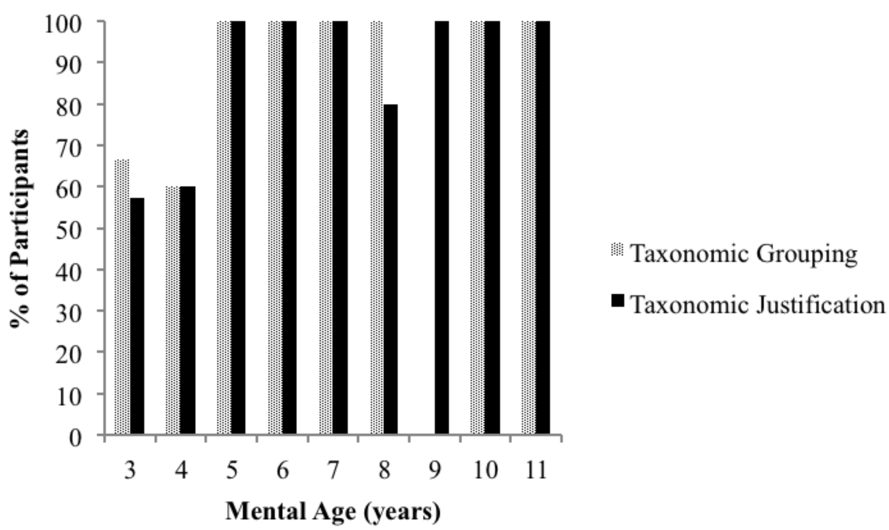


Fig 3.20. Age and taxonomic grouping and justification on the Basic level task in ASD

Perceptual Grouping. Overall, 49/50 participants (98%) grouped and 39/50 participants (78%) justified on the Perceptual Grouping task. In particular, as shown in Figure 3.21, 26/26 TD participants (100%) grouped and 21/26 participants (80.8%)

justified. Following a similar pattern, as shown in Figure 3.22, 23/24 ASD participants (95.8%) grouped and 18/24 ASD participants (75%). Visual inspection of the data revealed that the two groups followed a similar developmental trajectory.

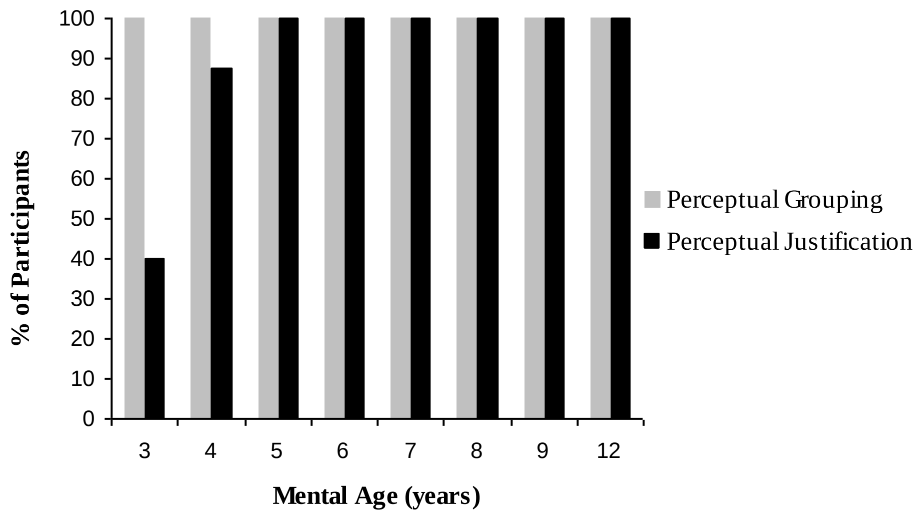


Fig 3.21. Age and grouping and justification on the Perceptual Grouping task in TD

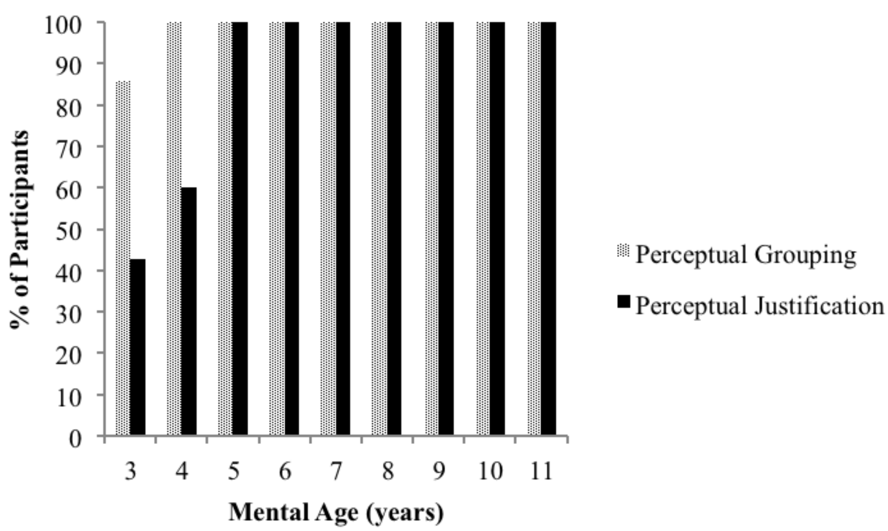


Fig 3.22. Age and grouping and justification on the Perceptual Grouping task in ASD

Discussion

This study investigated categorization development in ASD children and TD children matched by mental age and chronological age. The goals of this study were to determine whether ASD and TD children's categories follow similar patterns of development and to explore whether specific cognitive abilities predict the thematic-taxonomic shift in each of these populations.

Effects of mental age

Children with greater MAs were more likely to group and justify taxonomically on the Superordinate and Intermediate level tasks. Additionally, children were greater MAs knew the names of more animals, knew more facts about animals, had larger MLUs, had greater metamemorial awareness, and were more likely to justify and to switch justification dimensions on the Perceptual Grouping task (i.e., they were better justifiers and had greater cognitive flexibility).

However, there were a few measures that MA did not predict. First, MA did not predict thematic or perceptual grouping on the any of the categorization tasks. This is likely because so few of the children in either group justified thematically or perceptually. It is unclear why proportionally fewer TD children justified thematically or perceptually in this study than in the previous study. However, on the Superordinate level task, there were non-significant U-shaped effects of both MA and CA on thematic justification in the TD group. Thus, the data followed a similar pattern as in the previous study. Conversely, no ASD children justified thematically on the Superordinate level task. Further investigation is necessary to determine whether ASD children really have less thematic categorization than TD children.

Second, MA did not predict performance on the Basic level task. This is not surprising given that there was a ceiling effect in both groups. Finally, MA did not predict switching grouping dimensions on the Perceptual Grouping task. Visual inspection of the data reveals that all TD children with a mental age of five and above switched grouping dimensions; thus, there was close to a ceiling effect for the TD children. In the ASD group, there was somewhat of a bimodal distribution, with only children of low and high mental ages switching grouping dimensions but not children with more moderate mental ages doing so. My intuition is that the ASD results are idiosyncratic and would not replicate, although further investigation is necessary to rule out the possibility that this distribution is real.

Level in TD and ASD children

In both the TD and ASD groups, more children justified taxonomically on the Basic level task (69% and 75%, respectively) than on either the Superordinate level task (52% and 50%, respectively) or the Intermediate level task (54% and 42%, respectively). These findings are consistent with previous research demonstrating that autistic children have more difficulty grouping objects that share the same superordinate level than objects sharing the same basic level (Tager-Flusberg, 1985) and with evidence that basic level categories are the first to develop in typical children (see “Effect of Level” in Chapter II).

Knowledge vs categorization in ASD

In the previous chapter, I demonstrated that relevant declarative knowledge was not sufficient for categorization. Similarly, in this study, three-quarters of the ASD participants correctly answered at least half of the Animal Declarative Knowledge questions but only about a third provided taxonomic justifications on the Intermediate level task. Furthermore, Animal Declarative Knowledge scores were not a significant predictor of taxonomic justification independent of mental age.

Categorization in ASD

Despite the fact that the ASD participants were less cognitively flexible, had less declarative knowledge about animals and less metamemorial awareness, and tended to be less linguistically advanced (with marginally smaller MLUs and animal vocabularies) than the MA-matched typically developing participants, there were no significant differences between the ASD group and the TD group on any of the categorization tasks. There are several possible explanations for these findings.

One possibility is that categories are the same in ASD and TD children with similar mental ages. Although previous studies have suggested that ASD children are impaired in constructing *new* categories (e.g., Klinger & Dawson, 2001; Minshew et al, 2002), it is less clear whether ASD children eventually acquire TD-like categories. My results are consistent with previous studies suggesting that categories are qualitatively similar in ASD and TD individuals (e.g., Gasteb et al., 2006; Tager-Flusberg, 1985).

Metamemory

Although the ASD children performed significantly worse than the TD children on the Metamemory task, they performed similarly on the categorization tasks. Therefore, one possibility is that strong metamemorial awareness is not necessary for taxonomic categorization. Another possibility is that the ASD children performed poorly on the metamemory task for non-metamemory-related reasons. Recall that Flavell's task measures a child's knowledge of his or her own memory by asking about the memories of two puppets. Thus, it requires the child to engage in pretend play (treating puppets as if they were people) and to attribute mental states to other (i.e., to have a theory of mind). Given that autism is marked by deficits in pretend play (e.g., Jarrold, 2003) and in Theory of Mind (e.g., Happe, 1995), the ASD children's metamemory scores may have been affected by impairments in these other areas. This possibility seems especially likely given previous studies suggesting relatively intact metamemory in autism (e.g., Wojcik et al., 2011; Wojcik, Moulin, & Souchey, 2013).

Matching by mental age

In the Methods section, I discussed problems with matching impaired and typical children by mental age. I now consider the possibility that matching ASD and TD children by mental age obscured differences in how the two populations categorize. In the Nonverbal portion of the KBIT-2, an experimenter showed children a target picture (e.g., a car) and asked them which of several other pictures goes with the target (e.g., a truck, a frying pan, a sun, an apple, or a zipper; see Figure 3.23). This task is similar to my categorization tasks (although, see the introduction of Chapter II for a discussion of how the KBIT-2's use of "goes with" is more thematic-y, whereas my use of "kind of the same as" is more taxonomic-y).

Although the KBIT-2 instruction manual does not discuss the types of relationships between targets and correct matches on their task, 5/20 correct matches (25%) are taxonomically related to the target (e.g., a car and a truck are both members of the superordinate level category VEHICLES), whereas 15/20 correct matches (75%) are thematically related (e.g., a ring and a hand, see Figure 3.24). Crucially, in this example, a child's selection of "sock" would be marked incorrect even though RING and SOCK are both (non-prototypical) members of the superordinate category CLOTHING. To the extent that the KBIT-2 nonverbal IQ scores reflect categorization preferences, it is not surprising that the ASD participants behaved like typically developing children who had similar scores on the KBIT-2.

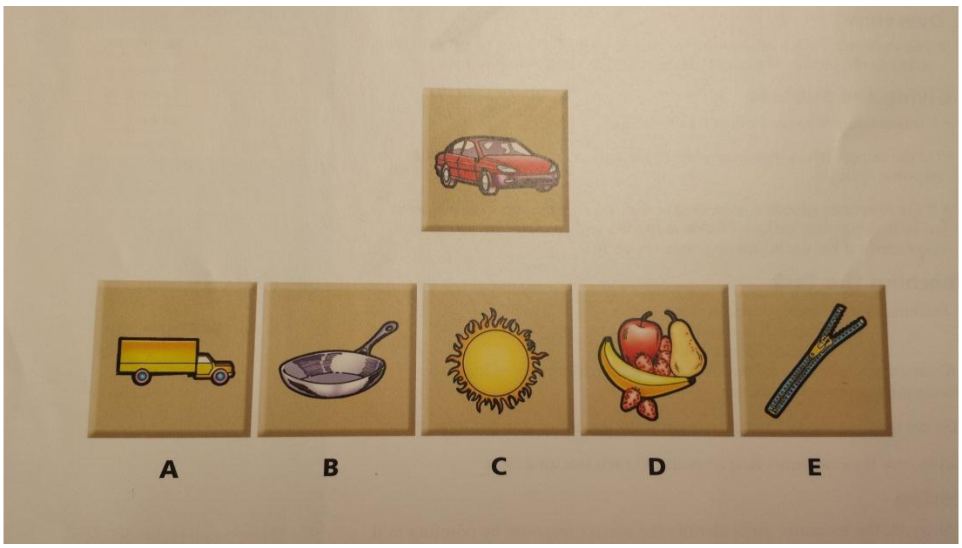


Fig 3.23. KBIT-2 Matrices Sample A: Taxonomic association

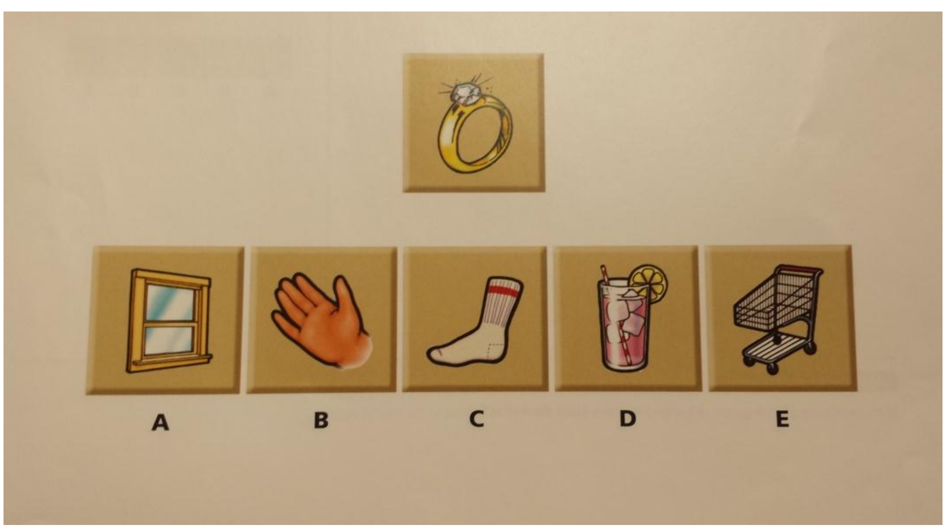


Fig 3.24. KBIT-2 Matrices Sample B: Thematic association

The KBIT-2 also has some similarities to my other cognitive measures. For example, it is likely that the Nonverbal task (described above) also measures cognitive flexibility. Recall that Blaye et. al (2001) measured cognitive flexibility as a child's ability to switch between multiple types of conceptual organizations. To succeed on the KBIT-2 task, the child was required to switch between taxonomic (e.g., Figure 3.23) and thematic (e.g., Figure 3.24) categorization.

Other Nonverbal KBIT-2 items are similar to my Perceptual Grouping task. As shown in Figures 3.25, 3.26, and 3.27, children are shown incomplete matrices of abstract stimuli and must select the items that complete the matrices. In both this task and in my Perceptual Grouping task, the child must isolate perceptual features of a group of objects.

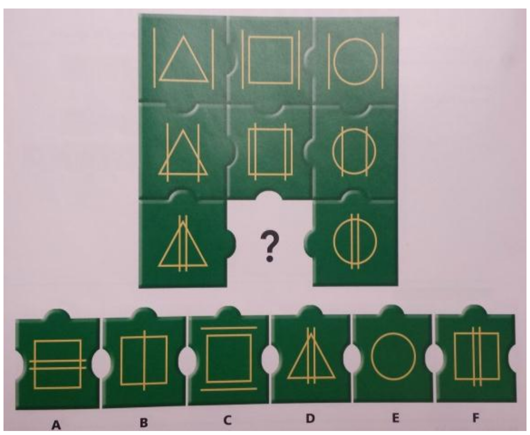


Fig 3.25. KBIT-2 Matrices Item 24

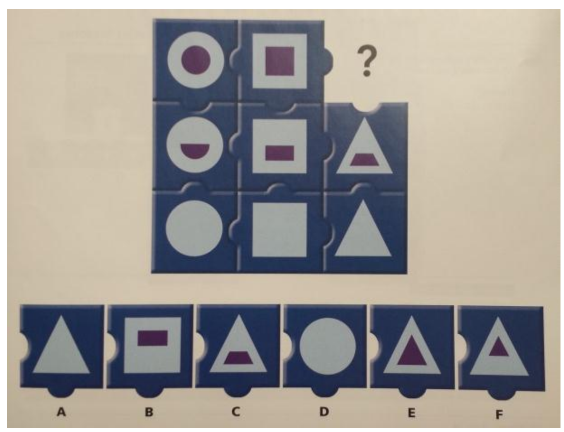


Fig 3.26. KBIT-2 Matrices Item 25

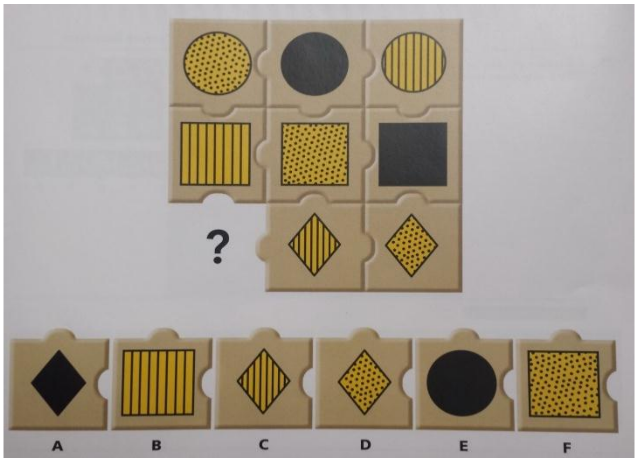


Fig 3.27. KBIT-2 Matrices Item 26

Finally, on the Verbal Knowledge subtest, an experimenter asked children a question (e.g., “Which one tells you how much something weighs?”) and asked the child to point to the relevant picture (e.g., a scale). This is almost identical to my Animal Declarative Knowledge task. Given the similarity between portions of the KBIT-2 and my tasks, it is possible that matching by MA obscured group differences. This is especially worrisome given that when I matched my groups by CA rather than by MA, the TD group taxonomically justified marginally more than the ASD group on the Superordinate and Intermediate level tasks. However, the fact that the difference between groups was greater on most cognitive measures than it was on the categorization tasks suggests ASD children may follow relatively typical patterns of categorization development. A final possibility is that ASD children perform similarly to TD children on categorization tasks but that they do so by different processes (i.e., they have different patterns of neuronal activation). To address this possibility, I designed an ERP experiment.

ERP experiment proposal

Previous ERP research has reinforced the behavioral finding that related words prime each other. For example, several studies have demonstrated that in unmasked lexical decision tasks (LTDs), the N400 component is attenuated when the prime and target words are related compared to when they are unrelated (e.g., Franklin, Dien, Neely, Huber, & Watson, 2007; Ruz, Madrid, Lupianez, & Tudela, 2003). This finding has been replicated in studies in which direction was attended away from the task (e.g. Relander, Rama, & Kujala, 2008), suggesting that activation of related words is automatic. However, with only two notable exceptions, ERP studies of typical populations have not distinguished between taxonomically and thematically related words, or have specifically only used taxonomically related words (e.g. Rossell, Price, & Nobre, 2003). Neither of these exceptions (Khateb et al, 2003, Maguire et al., 2010), which utilized different experimental paradigms, found a significant effect of type of relationship (taxonomic/thematic) on the N400, although both found an attenuated N400 for unrelated compared to related pairs, as predicted. However, a study of adults with right hemisphere damage found an attenuated N400 for taxonomic compared to thematic pairs (Hagoort et al., 1996), suggesting that N400 amplitude is a reasonable measure of category salience. Finally, as previously discussed, no studies have investigated the thematic-taxonomic shift in ASD children.

The study I have designed aims to modify Maguire et al.'s study to investigate 1. Whether ERP data supports the finding that typical preschoolers find thematic relationships more salient than taxonomic ones 2. Whether similar findings obtain in TD and ASD children. Additionally, comparing behavioral results (RT and accuracy in an

LDT) with ERP results would speak to the degree to which each of these methodologies captures lexical categorization and the degree to which these types of results can predict each other.

I hypothesize that TD preschoolers would be fastest and most accurate for thematic targets, followed by taxonomic targets, followed by unrelated targets. Similarly, for TD preschoolers, I hypothesize that N400's would be attenuated for thematic compared to taxonomic targets, and attenuated for taxonomic targets compared to unrelated targets. Conversely, I hypothesize that older children show patterns similar to the adults in Maguire et al's study (i.e., faster, more accurate, and more activation for taxonomic targets than for thematic targets). If ASD children's results from similar to those of TD children of the same MA and/or CA, this would suggest that ASD children follow typical patterns of category development. However, if I found group differences on any measure, this would highlight differences between the way categories develop in TD and ASD children.

General Discussion

This dissertation explored the development of lexical organization, lexical access, and categorization. In Chapter I, I investigated lexical representation and access in adults by comparing the way adult speakers of English and Mandarin access words during a verbal fluency task with the way they explicitly categorize these words.

First, the results from this study suggest that lexical items are organized around semantic features by which adults can access these words. Second, the results suggest that words are organized around multiple features and that adults access words by different features when automatically accessing words than when explicating categorizing words. In particular, adults tend to use personally-relevant features (e.g., SCARY) when automatically accessing words, whereas, they tend to use learned features (e.g., MAMMAL) when explicitly categorizing.

Third, although this automatic/explicit distinction obtains across languages, some features are linguistically-relative. For example, Mandarin speakers used the feature CHINESE ZODIAC for both automatic access and explicit categorization, whereas, not surprisingly, English speakers did not use this feature in either task. An unexpected finding is that some of the English speaker's lexical access features were only explicit grouping features for Mandarin speakers (e.g., LIVESTOCK), and some of the Mandarin speaker's lexical access features were only explicit grouping features for English speakers (e.g., TRANSPORTATION). These results in particular demonstrate the need for empirical investigations of lexical organization and access. That is, there are limitations to speaker judgments (expressed through explicit categorization) about the organization of the lexicon (i.e., how speakers automatically access words).

Chapter II investigated the nature and development of lexical organization, access, and categorization in typically-developing children ages two through nine. The results from this chapter reveal important differences between the way adults and the way young children organize, access, and categorize words

First, when adults access words, they tend to cluster by taxonomic features. However, when preschoolers access words, they do not cluster by semantic features of any type: taxonomic, thematic, or perceptual. This finding is particularly striking because preschoolers do have declarative knowledge about these features (e.g., they know that cats are pets).

Second, Chapter II explored developments in children's categories by investigating how they group items and how they justify these groupings. First, consistent with previous findings, taxonomic grouping and taxonomic justification increased with age. Interestingly, many children grouped taxonomically but did not justify taxonomically. This grouping-justification incongruity is analogous to the finding (in Chapter I) that adults do not have conscious access to their lexical access features.

Chapter II also revealed the different rates at which different levels of taxonomic categorization develop. Consistent with the majority of previous studies, my results suggest that basic level categories develop before superordinate level categories. This dissertation made a novel contribution to our understanding of category levels by being the first to investigate the development of intermediate level categories. My data suggest that intermediate level categories develop after superordinate level categories. Together, these findings demonstrate that taxonomic categorical structure develops gradually and becomes more fine-grained with development. The findings that preschoolers neither semantically cluster on a verbally fluency task nor justify taxonomically on an explicit categorization task are converging evidence that fine-grained taxonomic categories develop later in childhood.

An additional goal of Chapter II was to investigate whether developments in specific cognitive abilities cause the thematic-taxonomic shift (and perhaps in tandem, the advance of adult-like lexical representation and access). I hypothesized that developments in language, knowledge, cognitive flexibility, and/or metamemorial awareness could play a role in this shift. However, I found that age was a better predictor of taxonomic grouping and justification than any of these cognitive abilities. Further research will be necessary to determine whether different measures of these cognitive abilities would better predict this shift; whether cognitive abilities I did not test would predict this shift; or whether this shift is the result of general maturation. One difficulty in deciding among these possibilities is that age is collinear with the development of every cognitive ability I predicted could play a role in the thematic-taxonomic shift. A possible solution to the collinearity problem would be to use more fine grained measures of the cognitive predictors and/or of categorization behavior (e.g., reaction time, eye gaze, or electrophysiology).

Finally, in Chapter III, I investigated the nature and development of categorization in autism. One reason for this study was that the literature is currently divided about the nature of categorization in autism, with evidence for and against autistic children having impaired categorization. Furthermore, very little is known about the developmental trajectory of category development in autistic children. For example, my study is the first to investigate the thematic-taxonomic shift in autism.

The premise of this study was the following: If autistic children are impaired in a cognitive ability that is central to developing taxonomic categories, one would expect autistic children to be less likely to group and justify than children of the same mental age. Alternatively, if specific cognitive abilities are not central to developing taxonomic categories, one would expect autistic children to develop taxonomic categories at the same rate as mental-age-matched typical children.

Relative to typical children of the same mental age, the autistic children in this study demonstrated impairments in every cognitive ability that I hypothesized could play a role in the thematic-taxonomic shift. They were less cognitive flexibility, had less declarative knowledge, and had less metamemorial awareness, and they had marginally poorer scores on tests of vocabulary and productive syntax. However, the autistic children did not perform significantly differently than the mental-aged-matched typical children on any categorization task.

These results are important for two reasons. First, they strengthen the finding (in Chapter II) that none of the cognitive abilities I measured contribute independently to the thematic-taxonomic shift. Second, they suggest that although autistic children often present with numerous cognitive impairments, they may follow typical patterns of category development. However, just as in Chapter II, the collinearity between chronological age my cognitive measures may have obscured the importance of these cognitive abilities for categorization development, in this chapter, controlling for mental age may have obscured differences the autistic and typical groups. In particular, the overlap between the KBIT-2 (my measure of mental age) and both my cognitive and categorization measures could have masked differences in the autistic and typical children's patterns of category development. Further investigation using different matching criteria (i.e., a different measure of mental age) could help to clarify whether autistic children follow typical patterns of categorization development. Additionally, studies use electrophysiological or neuroimaging methodology could pinpoint differences in ASD children's underlying categorization processes. For this reason, I designed an Event Related Potential (ERP) study to investigate the thematic-taxonomic shift. One benefit of ERP experiments is that they minimize performance demands (i.e., subjects are not asked to complete a task) and thus, they can provide more direct evidence about representation (rather than evidence about participants' metacogntive awareness of their representations).

Finally, I began this dissertation having operationalized *feature* but being agnostic as to what kinds of features link words in the lexicon. For example, I stated that these features could be semantic, associative, or phonological; they could be taxonomic, thematic, or perceptual; they could include personally-relevant features and explicitly taught features. My data can narrow these possibilities.

First, adults access words by different features depending on the goal of the task. That is, words are organized in multiple ways, and lexical access is context-dependent. More specifically, the features by which adults explicitly categorize are almost entirely semantic and tend to be more explicitly taught. Although I did not report on adults phonological clustering in this dissertation, the semantic features by which adults cluster (i.e., more automatically access words) tend to be more personally relevant.

Second, typically-developing preschoolers have a qualitatively poorer lexical organization. They do not have intermediate level features by which words are linked, or at least, they do not access words via these features. Third, as both typically developing children and children with autism spectrum disorders develop, the features by which they categorize words become increasingly taxonomic.

There are still many open questions. First, throughout my dissertation, I used the same terms to refer to words (e.g., “mammal”) and to features (e.g., MAMMAL). One question is the extent to which acquiring a word is necessary for or aids in acquiring a feature. My data cannot speak to whether a person can have the feature MAMMAL—that is, whether a person accesses multiple consecutive mammals via this feature—without having acquired the word. My hypothesis is that it is possible to have a feature without having a word but that having the word strengthens this feature.

Second, although my data give some insights about the development of features, there is more to be learned about how they acquired. One possibility is that some features are innate. For example, ANIMATE and INANIMATE could be innate features that exist prior to word learning. If this is the case, as words are learned, they could be attached directly to the features. Conversely, features that are acquired after relevant words are learned—e.g., PET is likely acquired after “dog” and “cat”—could be attached to words as they are acquired.

Third, I have not solved the problem of what a feature is. Although I have assumed that words are atomic units, I have not described whether a features is also atomic, whether it is a bundle of other features, or whether it is something else. Given the problems with decomposition, it seems safest to say that a feature, like a word, is an atomic unit but that a set of words could be activated by multiple features simultaneously. For example “dog” and “cat” could share the features PET, MAMMAL, ANIMAL, and so on. Thus, one place for future research is to determine whether words sharing multiple features are activated more quickly than those sharing only one feature and whether the effects of sharing multiple features is additive. Despite these open questions, this dissertation provides insights into the nature of lexical organization and its development.

Appendix A. Animal Declarative Knowledge questions

Which ones are wild?

Which ones do people ride on?

Which ones are big?

Which ones are disgusting?

Which ones are pets?

Which ones are mammals?

Which ones live in Africa?

Which ones are cats?

Which ones live in the water?

Which ones eat other animals?

Which ones live in the circus?

Which ones fly?

Which ones are birds?

Which ones are rodents?

Which ones are apes?

Which ones live in the backyard?

Which ones are scary?

Which ones do people eat?

Which ones live on a farm?

Which ones have four legs?

Appendix B. Metamemory Task Questions

Megan and Henry are trying to learn some new words.

1. Megan is in a very noisy room. Henry is in a very quiet room. Is it harder for Megan or Henry?

2. Megan has a long time to learn the words. Henry has a short time to learn the words. Is it harder for Megan or Henry?

3. Megan is learning the words all by herself. Henry has help from his teacher. Is it harder for Megan or Henry?

4. Megan is trying to learn eighteen new words. Henry is trying to learn three new words. Is it harder for Megan or Henry?

5. Megan is drawing pictures to help her learn the words. Henry is not drawing pictures to help him learn the words. Is it harder for Megan or Henry?

References

Anglin, J. M. (1977). *Word, object, and conceptual development*. New York: Norton.

Armstrong, S. L., Gleitman, L. R., & Gleitman, H. (1983). What some concepts might not be. *Cognition, 13*(3), 263-308.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition*, *21*(1), 37-46. Barsalou, L. W. (1983). Ad hoc categories. *Memory & Cognition, 11*(3), 211-227.

Barsalou, L. W., & Hale, C. R. (1993). Components of conceptual representation: From feature lists to recursive frames. In I. Van Mechelen, J. Hampton, R. Michalski & P. Theuns (Eds.), *Categories and concepts: Theoretical views and inductive data analysis* (pp. 97-144). San Diego, CA: Academic Press.

Bartolucci, G., Pierce, S. J., & Streiner, D. (1980). Cross-sectional studies of grammatical morphemes in autistic and mentally retarded children. *Journal of Autism and Developmental Disorders, 10*(1), 39-50.

Begeer, S., Wierda, M., Scheeren, A. M., Teunisse, J. P., Koot, H. M., & Geurts, H. M. (2013). Verbal fluency in children with autism spectrum disorders: Clustering and switching strategies. *Autism: The International Journal of Research and Practice.*

Bennetto, L., Pennington, B. F., & Rogers, S. J. (1996). Intact and impaired memory functions in autism. *Child Development, 67*(4), 1816-1835.

Blaye, A., & Bonthoux, F. (2001). Thematic and taxonomic relations in preschoolers: The development of flexibility in categorization choices. *British Journal of Developmental Psychology, 19*, 395-412.

Borge-Holthoefer, J., & Arenas, A. (2009). Navigating word association norms to extract semantic information. *Proceedings of the 31st Annual Conference of the Cognitive Science Society*.

Boroditsky, L. (2001). Does language shape thought? Mandarin and English speakers' conceptions of time. *Cognitive Psychology, 43*(1), 1-22.

Botting, N., & Conti-Ramsden, G. (2003). Autism, primary pragmatic difficulties, and specific language impairment: Can we distinguish them using psycholinguistic markers? *Developmental Medicine and Child Neurology, 45*(8), 515-524.

Boucher, J. (1988). Word fluency in high-functioning autistic children. *Journal of Autism and Developmental Disorders, 18*(4), 637-645.

Brown, M. C., McNeil, N. M., & Glenberg, A. M. (2009). Using concreteness in education: Real problems, potential solutions. *Child Development Perspectives, 3*(3), 160-164.

Brown, R., & McNeil, D. (1966). The "tip of the tongue" phenomenon. *Journal of Verbal Learning and Verbal Behavior, 5*, 325-337.

Brown, R., & Berko, J. (1960). Word association and the acquisition of grammar. *Child Development, 31*, 1-14.

Buchanan, L., Brown, N. R., Cabeza, R., & Maitson, C. (1999). False memories and semantic lexicon arrangement. *Brain and Language, 68*(1-2), 172-177.

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodal inference: A practical information-theoretic approach* (2nd ed.). Fort Collins, CO: Springer-Verlag.

Capps, L., Kehres, J., & Sigman, M. (1998). Conversational abilities among children with autism and children with developmental delays. *Autism, 2*, 325-344.

Carey, S. (1999). Knowledge acquisition: Enrichment or coneptual change? In E. Margolis, & S. Laurence (Eds.), *Concept: Core readings* (pp. 459-488). Cambridge, MA: Massachusetts Institute of Technology.

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

Carey, S. (1985). In Gleitman L., Carey S., Newport E. and Spelke E. (Eds.), *Conceptual change in childhood*. Cambridge, MA: The MIT Press.

Cartwright, K. B. (2002). Cognitive development and reading: The relation of reading-specific multiple classification skill to reading comprehension in elementary school children. *Journal of Educational Psychology, 94*, 56-63.

Casasanto, D. (2008). Who's afraid of the big bag Whorf? cross-linguistic differences in temporal language and thought. *Language Learning, 58*(1), 63-79.

Charman, T., Drew, A., Baird, C., & Baird, G. (2003). Measuring early language development in preschool children with autism spectrum disorder using the MacArthur communicative development inventory (infant form). *Journal of Child Language, 30*(1), 213-236.

Chen, J. Y. (2007). Do Chinese and English speakers think about time differently? Failure of replicating Boroditsky (2001). *Cognition, 104*(2), 427-436.

Chereau, C., Gaskell, M. G., & Dumay, N. (2007). Reading spoken words: Orthographic effects in auditory priming *Cognition, 102*(3), 341-360.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge: M.I.T. Press.

Clark, E. V. (1983). Meanings and concepts. In J. H. Flavell, & E. M. Markman (Eds.), *Handbook of child psychology cognitive development* (4th ed., pp. 787-840). New York: John Wiley & Sons.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82*(6), 407-428.

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior, 8*, 240-247.

Cronin, V. S. (2002). The syntagmatic-paradigmatic shift and reading development. *Journal of Child Language*, *29*(1), 189-204.

Cronin, V. S., Pratt, M., Abraham, J., Howell, D., Bishop, S., & Manning, A. (1986). Word association and the acquisition of reading. *Journal of Psycholinguistic Research*, *15*, 1-11.

Daehler, M. W., Lonardo, R., & Bukatko, D. (1979). Matching and equivalence judgments in very young children. *Child Development, 50*(1), 170-179.

Deak, G. O., & Bauer, P. J. (1996). The dynamics of preschoolers' categorization choices. *Child Development, 67*(3), 740-767.

Denney, D. R., & Moulton, P. A. (1976). Conceptual preferences among preschool children. *Developmental Psychology*, *12*(6), 509-513.

DePaul, M. R., & Ramsey, W. (1997). *Rethinking intuition: The psychology of intuition and its role in philosophical inquiry*. Lanham, Md.: Rowman & Littlefield.

Dichter, G. S., Radonovich, K. J., Turner-Brown, L. M., Lam, K. S., Holtzclaw, T. N., & Bodfish, J. W. (2010). Performance of children with autism spectrum disorders on the dimension-change card sort task. *Journal of Autism and Developmental Disorders*, *40*(4), 448-456.

Dunham, P., & Dunham, F. (1995). Developmental antecedents of taxonomic and thematic strategies at three years of age. *Developmental Psychology, 31*(3), 483-493.

Eigsti, I. M. (2001). *Word learning and memory functions in young children with autism.* (Unpublished PhD). University of Rochester, Rochester, NY.

Eigsti, I. M., Bennetto, L., & Dadlani, M. B. (2007). Beyond pragmatics: Morphosyntactic development in autism. *Journal of Autism and Developmental Disorders, 37*(6), 1007-1023.

Federmeier, K. D., McLennan, D. B., De Ochoa, E., & Kutas, M. (2002). The impact of semantic memory organization and sentence context information on spoken language processing by younger and older adults: An ERP study. *Psychophysiology, 39*(2), 133-146.

Fenson, L., Vella, D., & Kennedy, M. (1989). Children's knowledge of thematic and taxonomic relations at two years of age. *Child Development, 60*(4), 911-919.

Ferrand, L., & New, B. (2003). Semantic and associative priming in the mental lexicon. In P. Bonin (Ed.), *The mental lexicon* (pp. 25-43). New York: Nova Science Publishers.

Flavell, J. H., & Wellman, H. M. (1977). Metamemory. In R. Kail, & J. Hagen (Eds.), *Perspectives on the development of memory and cognition* (pp. 3-33). Hillsdale, NJ: Erlbaum.

Fodor, J. A. (1970). Three reasons for not deriving 'kill' from 'cause to die'. *Linguistic Inquiry, 1*, 429-438.

Forster, K. I., Mohan, K., & Hector, J. (2003). The mechanics of masked priming. In S. Kinoshita, & S. J. Lupker (Eds.), *Masked priming: State of the art* (pp. 3-37). Hove, UK: Psychology Press.

Franklin, M. S., Dien, J., Neely, J. H., Huber, E., & Waterson, L. D. (2007). Semantic priming modulates the N400, N300, and N400RP. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology, 118*(5), 1053-1068.

Frege, G. (1884/1974). *The foundations of arithmetic: A logico-mathematical enquiry into the concept of number* [Die Grundlagen der Arithmetik: eine logisch-mathematische Untersuchung über den Begriff der Zahl] (J. L. Austin Trans.). (2nd Revised Edition ed.). Evanston, Il.: Northwestern University Press.

Frenck-Mestre, C., & Bueno, S. (1999). Semantic features and semantic categories: Differences in rapid activation of the lexicon. *Brain and Language, 68*(1-2), 199-204.

Frith, U., & Happe, F. (1994). Autism: Beyond "theory of mind". *Cognition, 50*, 115-132.

Frith, C. D., Friston, K. J., Liddle, P. F., & Frackowiak, R. S. (1991). A PET study of word finding. *Neuropsychologia, 29*(12), 1137-1148.

Froehlich, A. L., Anderson, J. S., Bigler, E. D., Miller, J. S., Lange, N. T., Dubray, M. B., . . . Lainhart, J. E. (2012). Intact prototype formation but impaired generalization in autism. *Research in Autism Spectrum Disorders, 6*(2), 921-930.

Fromkin, V. (1980). *Errors in linguistic performance: Slips of the tongue, ear, pen, and hand*. San Francisco: Academic Press.

Gallese, V., & Lakoff, G. (2005). The brain's concepts: The role of the sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology, 22*(3), 455-479.

Gastgeb, H. Z., Strauss, M. S., & Minshew, N. J. (2006). Do individuals with autism process categories differently? the effect of typicality and development. *Child Development*, *77*(6), 1717-1729.

Gelman, S. A. (2004). Psychological essentialism in children. *Trends in Cognitive Sciences, 8*(9), 404-409.

Gelman, S. A., & Coley, J. D. (1990). The importance of knowing a dodo is a bird: Categories and inferences in two-year-old children. *Developmental Psychology, 26*, 796-804.

Gelman, S. A., & Markman, E. M. (1987). Young children's inductions from natural kinds: The role of categories and appearances. *Child Development*, *58*(6), 1532-1541.

Gennari, S. P., Sloman, S. A., Malt, B. C., & Fitch, W. T. (2002). Motion events in language and cognition. *Cognition*, *83*(1), 49-79.

Genome, J., & Lombrozo, T. (2012). Concept possession, experimental semantics, and hybrid theories of reference. *Philosophical Psychology*, *25*, 1-26.

Gentner, D., & Namy, L. L. (2006). Analogical processes in language learning. *Current Directions in Psychological Science, 15*(6), 297-301.

Goldman, A. (2007). Philosophical intuitions: Their target, their source, and their epistemic status. *Grazer Philosophische Studien, 74,* 1-26.

Goni, J., Arrondo, G., Sepulcre, J., Martincorena, I., Velez de Mendizabal, N., Corominas-Murtra, B., Villoslada, P. (2011). The semantic organization of the animal category: Evidence from semantic verbal fluency and network theory. *Cognitive Processing, 12*(2), 183-196.

Griffiths, T. L., Steyvers, M., & Tenenbaum, J. B. (2007). Topics in semantic representation. *Psychological Review, 114*(2), 211-244.

Guastavino, C. (2007). Categorization of environmental sounds. *Canadian Journal of Experimental Psychology = Revue Canadienne De Psychologie Experimentale, 61*(1), 54-63.

Hagoort, P., Brown, C. M., & Swaab, T. Y. (1996). Lexical-semantic event-related potential effects in patients with left hemisphere lesions and aphasia, and patients with right hemisphere lesions without aphasia. *Brain: A Journal of Neurology, 119(2),* 627-649

Hajibayova, L. (2013). Basic-level categories: A review. *Journal of Information Science, 39*(5), 676-687.

Happe, F. G. (1995). The role of age and verbal ability in the theory of mind task performance of subjects with autism. *Child Development*, *66*(3), 843-855.

Hauk, O., & Tschentscher, N. (2013). The body of evidence: What can neuroscience tell us about embodied semantics? *Frontiers in Psychology, 4*, 50.

Henley, N. M. (1969). A psychological study of the semantics of animal terms. *Journal of Verbal Learning and Verbal Behavior*, *8*, 176-184.

Hirschfeld, L. (1996). *Race in the making: Cognition, culture, and the child's construction of human kinds*. Cambridge, MA: MIT Press.

Homack, S. R., & Reynolds, C. R. (2007). *Essentials of assessment with brief intelligence tests*. Hoboken, NJ: John Wiley & Sons, Inc.

Imai, M., Gentner, D., & Uchida, N. (1994). Children's theories of word meaning: The role of shape similarity in early acquisition. *Cognitive Development*, *9*, 45-75.

Inhelder, B., & Piaget, J. (1964 (originally written in 1959)). *The early growth of logic in the child, classification and seriation* (E. A. Lunzar, D. Papert Trans.). New York: Harper & Row.

Isacoff, N., Liu, D., Hou-Imerman, T., & Stromswold, K. (2013). Is lexical organization linguistically relative: Evidence from English & Mandarin. *The 84th Annual Meeting of the Eastern Psychological Association,* New York, NY.

Isacoff, N., & Stromswold, K. (2011). Semantic organization and its development. *82nd Annual Meeting of the Eastern Psychological Association*, Cambridge, M.A.

Isacoff, N. M., & Stromswold, K. (2014). Not all lexical access tasks are created equal: Lexical development between three and five. *First Language, 34*(1), 43-57.

Iwasaki, N., Vinson, D., & Vigliocco, G. (2010). Does the grammatical count/mass distinction affect semantic representations? evidence from experiments in English and Japanese. *Language and Cognitive Processes, 25*(2), 189-223.

Jackendoff, R. (1983). *Semantics and cognition*. Cambridge, MA: MIT Press.

Jackendoff, R. (2010). *Meaning and the lexicon: The parallel architecture 1975-2010*. Oxford: Oxford University Press.

January, D., & Kako, E. (2007). Re-evaluating evidence for linguistic relativity: Reply to boroditsky (2001). *Cognition, 104*(2), 417-426.

Jarrold, C. (2003). A review of research into pretend play in autism. *Autism : The International Journal of Research and Practice, 7*(4), 379-390.

Johnson-Laird, P. N., Herrmann, D. J., & Chaffin, R. (1984). Only connections: A critique of semantic networks. *Psychol.Bull., 96*(2), 292-315.

Just, M. A., Cherkassky, V. L., Keller, T. A., Kana, R. K., & Minshew, N. J. (2007). Functional and anatomical cortical underconnectivity in autism: Evidence from an

FMRI study of an executive function task and corpus callosum morphometry. *Cerebral Cortex (New York, N.Y.: 1991), 17*(4), 951-961.

Kalenine, S., & Bonthoux, F. (2006). The formation of living and non-living superordinate concepts as a function of individual differences. *Current Psychology Letters, 19*(2), 443-458.

Kalenine, S., Peyrin, C., Pichat, C., Segebarth, C., Bonthoux, F., & Baciu, M. (2009). The sensory-motor specificity of taxonomic and thematic conceptual relations: A behavioral and fMRI study. *Neuroimage, 44*(3), 1152-1162.

Katz, J., & Fodor, J. A. (1963). The structure of a semantic theory. *Language, 39*, 170-210.

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman brief intelligence test* (2nd ed.). Bloomington, MN: NCS Pearson.

Keil, F. C. (1986). The acquisition of natural kind and artifact terms. In W. Demopoulos, & A. Marras (Eds.), *Language learning and concept acquisition* (pp. 133-153). Norwood, NJ: Ablex.

Khateb, A., Michel, C. M., Pegna, A. J., O'Dochartaigh, S. D., Landis, T., & Annoni, J. M. (2003). Processing of semantic categorical and associative relations: An ERP mapping study. *International Journal of Psychophysiology : Official Journal of the International Organization of Psychophysiology, 49*(1), 41-55.

Kjelgaard, M. M., & Tager-Flusberg, H. (2001). An investigation of language impairment in autism: Implications for genetic subgroups. *Language and Cognitive Processes, 16*(2-3), 287-308.

Klinger, L. G., & Dawson, G. (2001). Prototype formation in autism. *Development and Psychopathology, 13*(1), 111-124.

Kripke, S. (1980). *Naming and necessity*. Cambridge, MA: Harvard University Press.

Labov, W. (1973). The boundaries of words and their meanings. In C. J. N. Bailey, & R. W. Shuy (Eds.), *New ways of analyzing variation in english* (pp. 340-373). Georgetown: Georgetown University Press.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago.

Landau, B., Dessalegn, B., & Goldberg, A. M. (2010). Language and space: Momentary interactions. In P. Chilton, & V. Evans (Eds.), *Language, cognition and space: The state of the art and new directions.* London: Equinox Publishing.

Landauer, T. K., & Dumais, S. T. (1997). A solution to plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review, 104*(2), 211-240.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). Introduction to latent semantic analysis. *Discourse Processes, 25*, 259-284.

Leslie, S. J., Khemlani, S., & Glucksberg, S. (2011). All ducks lay eggs: The generic overgeneralization effect. *Journal of Memory and Language, 65*, 15-31.

Li, P., Abarbanell, L., Gleitman, L., & Papafragou, A. (2011). Spatial reasoning in tenejapan mayans. *Cognition, 120*(1), 33-53.

Low, A., Bentin, S., Rockstroh, B., Silberman, Y., Gomolla, A., Cohen, R., & Elbert, T. (2003). Semantic categorization in the human brain: Spatiotemporal dynamics revealed by magnetoencephalography. *Psychological Science, 14*(4), 367-372.

Lucy, J. A., & Gaskins, S. (2001). Grammatical categories and the development of classification preferences: A comparative approach. In S. C. Levinson, & M. Bowerman (Eds.), *Language acquisition and conceptual development* (pp. 257-283). Cambridge: Cambridge University Press.

MacKay, G., & Shaw, A. (2004). A comparative study of figurative language in children with autism spectrum disorders. *Child Language Teaching and Therapy, 20*, 13-32.

Maguire, M. J., Brier, M. R., & Ferree, T. C. (2010). EEG theta and alpha responses reveal qualitative differences in processing taxonomic versus thematic semantic relationships. *Brain and Language, 114*(1), 16-25.

Malt, B. (1994). Water is not H20. *Cognitive Psychology, 27*, 41-70.

Malt, B., Sloman, S., & Gennari, S. (2003). Speaking versus thinking about objects and actions. In D. Gentner, & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp. 81-112)

Mandler, J. M., & Bauer, P. J. (1988). The cradle of categorization: Is the basic level basic? *Cognitive Development*, *3*, 247-264.

Mandler, J. M., Bauer, P. J., & McDonough, L. (1991). Separating the sheep from the goats: Differentiating global categories. *Cognitive Psychology, 23,* 263-298.

Mandler, J. M. (1979). Language comprehension. *Science (New York, N.Y.), 203*(4377), 259-260. doi:10.1126/science.203.4377.259

Mari-Beffa, P., Valdes, B., Cullen, D. J., Catena, A., & Houghton, G. (2005). ERP analyses of task effects on semantic processing from words. *Brain Research.Cognitive Brain Research, 23*(2-3), 293-305.

Markman, E. M., Cox, B., & Machida, S. (1981). The standard object-sorting task as a measure of conceptual organization. *Developmental Psychology*, *17*(1), 115-117.

Markman, E. M. (1989). *Categorization and naming in children: Problems of induction*. Cambridge: MIT.

Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, *20*(2), 121-157.

Medin, D. L., Wattenmaker, W. D., & Hampson, S. E. (1987). Family resemblance, conceptual cohesiveness, and category construction. *Cognitive Psychology, 19*(2), 242-279.

Melkman, R., Tversky, B., & Baratz, D. (1981). Developmental trends in the use of perceptual and conceptual attributes in grouping, clustering, and retrieval. *Journal of Experimental Child Psychology, 31*(3), 470-486.

Mervis, C. B., & Crisafi, M. A. (1982). Order of acquisition of subordinates-, basic-, and superordinate-level categories. *Child Development, 53*, 258-266.

Mervis, C. B., & Rosch, E. (1981). Categorization of natural objects. *Ann. Rev. Psychol., 32*, 89-115.

Meyer, D. E., & Schvanevelt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operators. *Journal of Experimental Psychology, 90*(2), 227-234.

Minshew, N. J., & Goldstein, G. (2001). The pattern of intact and impaired memory functions in autism. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 42*(8), 1095-1101.

Minshew, N. J., Goldstein, G., Muenz, L. R., & Payton, J. B. (1992). Neuropsychological functioning in nonmentally retarded autistic individuals. *Journal of Clinical and Experimental Neuropsychology, 14*(5), 749-761.

Minshew, N. J., Goldstein, G., & Siegel, D. J. (1997). Neuropsychologic functioning in autism: Profile of a complex information processing disorder. *Journal of the International Neuropsychological Society : JINS, 3*(4), 303-316.

Minshew, N. J., Meyer, J., & Goldstein, G. (2002). Abstract reasoning in autism: A dissociation between concept formation and concept identification. *Neuropsychology, 16*(3), 327-334.

Mo, L., Xu, G., Kay, P., & Tan, L. H.Electrophysiological evidence for the left-lateralized effect of language on preattentive categorical perception of color. *Proceedings of the National Academy of Sciences, 108*(34) 14026-14030.

Moller, J., Jansma, B. M., Rodriguez-Fornells, A., & Munte, T. F. (2007). What the brain does before the tongue slips. *Cerebral Cortex (New York, N.Y.: 1991), 17*(5), 1173-1178.

Muller, R. A., Behen, M. E., Rothermel, R. D., Chugani, D. C., Muzik, O., Mangner, T. J., & Chugani, H. T. (1999). Brain mapping of language and auditory perception in high-functioning autistic adults: A PET study. *Journal of Autism and Developmental Disorders, 29*(1), 19-31.

Murphy, G. L. (2001). Causes of taxonomic sorting by adults: A test of the thematic-to-taxonomic shift. *Psychonomic Bulletin & Review, 8*(4), 834-839.

Murphy, G. L., & Allopenna, P. D. (1994). The locus of knowledge effects in concept learning. *Journal of Experimental Psychology.Learning, Memory, and Cognition, 20*(4), 904-919.

Nation, K., & Snowling, M. J. (1999). Developmental differences in sensitivity to semantic relations among good and poor comprehenders: Evidence from semantic priming. *Cognition, 70*(1), B1-13.

Nelson, K. (1986). *Event knowledge: Structure and function in development*. Hillsdale, NJ: Erlbaum.

Nelson, K. (1977). The syntagmatic-paradigmatic shift revisited: A review of research and theory. *Psychological Bulletin*, *84*(1), 93-116.

Nelson, K., & Seidman, S. (1984). Playing with scripts. In I. Bretherton (Ed.), *Symbolic play: The development of social understanding* (pp. 45-71). Orlando, Fl.: Academic Press.

Nielsen, J. A., Zielinski, B. A., Fletcher, P. T., Alexander, A. L., Lange, N., Bigler, E. D., Anderson, J. S. (2014). Abnormal lateralization of functional connectivity between language and default mode regions in autism. *Molecular Autism, 5*(1), 8-2392-5-8.

Pennington, B. F., & Ozonoff, S. (1996). Executive functions and developmental psychopathology. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *37*(1), 51-87.

Piaget, J. (1962 (originally written in 1923)). *The language and thought of the child*. London: Routledge & Kegan Paul.

Pinker, S., & Bloom, P. (1990). Natural language and natural selection. *Behav..Brain Sci.*, *13*(4), 707-784.

Quinn, P. C., Eimas, P. D., & Rosenkrantz, S. L. (1993). Evidence for representations of perceptually similar natural categories by 3-month-old and 4-month-old infants. *Perception, 22*(4), 463-475.

Relander, K., Rama, P., & Kujala, T. (2009). Word semantics is processed even without attentional effort. *Journal of Cognitive Neuroscience, 21*(8), 1511-1522.

Rips, L. J., Shoben, E. J., & Smith, E. E. (1973). Semantic distance and the verification of semantic relations. *Journal of Verbal Learning and Verbal Behavior, 12,* 1-20.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8*, 382-452.

Rosch, E., Simpson, C., & Miller, R. S. (1976). Structural bases of typicality effects. *Journal of Experimental Psychology: Human Perception and Performance, 2*(491), 502.

Rosch, E. (1973). Natural categories. *Cognitive Psychology, 4*, 328-350.

Ross, B. H., & Murphy, G. L. (1999). Food for thought: Cross-classification and category organization in a complex real-world domain. *Cognitive Psychology, 38*(4), 495-553.

Rossell, S. L., Price, C. J., & Nobre, A. C. (2003). The anatomy and time course of semantic priming investigated by fMRI and ERPs. *Neuropsychologia, 41*(5), 550-564.

Rostad, K., Yott, J., & Poulin-Dubois, D. (2012). Development of categorization in infancy: Advancing forward to the animate/inanimate level. *Infant Behavior & Development, 35*(3), 584-595.

Rouse, H., Donnelly, N., Hadwin, J. A., & Brown, T. (2004). Do children with autism perceive second-order relational features? The case of the Thatcher illusion. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 45*(7), 1246-1257.

Rubin, D. C., & Olson, M. J. (1980). Recall of semantic domains. *Memory & Cognition, 8*(4), 354-356.

Ruz, M., Madrid, E., Lupianez, J., & Tudela, P. (2003). High density ERP indices of conscious and unconscious semantic priming. *Brain Research: Cognitive Brain Research, 17*(3), 719-731.

Sauzeon, H., Lestage, P., Raboutet, C., N'Kaoua, B., & Claverie, B. (2004). Verbal fluency output in children aged 7-16 as a function of the production criterion: Qualitative analysis of clustering, switching processes, and semantic network exploitation. *Brain and Language, 89*(1), 192-202.

Scarborough, H. S., Rescorla, L., Tager-Flusberg, H., Fowler, A. E., & Sudhalter, V. (1991). The relation of utterance length to grammatical complexity in normal and language-disordered groups. *Applied Psycholinguistics, 12*, 23-45.

Schmitz, M., & Wentura, D. (2012). Evaluative priming of naming and semantic categorization responses revisited: A mutual facilitation explanation. *Journal of Experimental Psychology.Learning, Memory, and Cognition.*

Schwartz, B. L., & Metcalfe, J. (2011). Tip-of-the-tongue (TOT) states: Retrieval, behavior, and experience. *Memory & Cognition, 39*(5), 737-749.

Schwartz, M. F., Kimberg, D. Y., Walker, G. M., Brecher, A., Faseyitan, O. K., Dell, G. S., Coslett, H. B. (2011). Neuroanatomical dissociation for taxonomic and thematic knowledge in the human brain. *Proceedings of the National Academy of Sciences of the United States of America, 108*(20), 8520-8524.

Schwartz, S., Baldo, J., Graves, R. E., & Brugger, P. (2003). Pervasive influence of semantics in letter and category fluency: A multidimensional approach. *Brain and Language, 87*(3), 400-411.

Scott, W. A. (1962). Cognitive complexity and cognitive flexibility. *American Sociological Association, 25*, 405-414.

Scott, M. S., Serchuk, R., & Mundy, P. (1982). Taxonomic and complementary picture pairs: Ability in two- to five-year-olds. *International Journal of Behavioral Development, 5*, 243-256.

Segui, J., Mehler, J., Frauenfelder, U., & Morton, J. (1982). The word frequency effect and lexical access. *Neuropsychologia, 20*(6), 615-627.

Sergeant, J. A., Geurts, H., & Oosterlaan, J. (2002). How specific is a deficit of executive functioning for attention-deficit/hyperactivity disorder**?**. *Behav. Brain Res., 130*(1-2), 3-28.

Sergent, J., Zuck, E., Levesque, M., & MacDonald, B. (1992). Positron emission tomography study of letter and object processing: Empirical findings and methodological considerations. *Cerebral Cortex (New York, N.Y.: 1991), 2*(1), 68-80.

Setoh, P., Wu, D., Baillargeon, R., & Gelman, R. (2013). Young infants have biological expectations about animals. *Proceedings of the National Academy of Sciences of the United States of America,110*(40), 15937-15942

Sheng, L., McGregor, K. K., & Marian, V. (2006). Lexical-semantic organization in bilingual children: Evidence from a repeated word association task. *Journal of Speech, Language, and Hearing Research : JSLHR, 49*(3), 572-587.

Shriberg, L. D., Paul, R., McSweeny, J. L., Klin, A. M., Cohen, D. J., & Volkmar, F. R. (2001). Speech and prosody characteristics of adolescents and adults with high-functioning autism and asperger syndrome. *Journal of Speech, Language, and Hearing Research : JSLHR, 44*(5), 1097-1115.

Slobin, D. (1987). Thinking for speaking. *Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society*, 435-444.

Slowiaczek, L. M., & Hamburger, M. (1992). Prelexical facilitation and lexical interference in auditory word recognition. *Journal of Experimental Psychology.Learning, Memory, and Cognition*, *18*(6), 1239-1250.

Smith, L. B., & Heise, D. (1992). Perceptual similarity and conceptual structure. In B. Burns (Ed.), *Percepts, concepts, and categories* (pp. 233-272).

Spelke, E. (1999). Physical knowledge in infancy: Reflections on piaget's theory. In S. Carey, & R. Gelman (Eds.), *The epigenesis of mind: Essays on biology and cognition* (pp. 133-169). Hillsdale, NJ: Laurence Erlbaum Associates.

Spiro, R. J., & Jehng, J. C. (1990). Cognitive flexibility and hypertext: Theory and technlogy for the nonlinear and multidimensional traversal of complex subject matter. In D. Nix, & R. J. Spiro (Eds.), *Cognition, education, and multimedia: Exploring ideas in high technology* (). Hillsdale, NJ: Lawrence Erlbaum Associates.

Springer, K. (2001). Perceptual boundedness and perceptual support in conceptual development. *Psychological Review*, *108*(4), 691-708.

Strauss, E., Sherman, E., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary*. Oxford: Oxford University Press.

Strevens, M. (2000). The essentialist aspect of naive theories. *Cognition*, *74*, 149-175.

Stromswold, K. (2006). Why aren't identical twins linguistically identical? genetic, prenatal and postnatal factors. *Cognition, 101*(2), 333-384.

Tager-Flusberg, H. (1985). Constraints on the representation of word meaning: Evidence from autistic and mentally retarded children. In S. A. Kuczaj, & M. Barrett (Eds.), *The development of word meaning* (pp. 139-166). New York, NY: Springer-Verlag.

Tager-Flusberg, H. (2000). Language and understanding minds: Connections in autism. In S. Baron-Cohen, H. Tager-Flusberg & D. J. Cohen (Eds.), *Understanding other minds: Perspectives from autism and developmental cognitive neuroscience* (2nd ed., pp. 1-45). Oxford: Oxford University Press.

Takahashi, T., Takikawa, Y., Kawagoe, R., Shibuya, S., Iwano, T., & Kitazawa, S. (2011). Influence of skin blood flow on near-infrared spectroscopy signals measured on the forehead during a verbal fluency task. *Neuroimage, 57*(3), 991-1002.

Toichi, M., & Kamio, Y. (2001). Verbal association for simple common words in high-functioning autism. *Journal of Autism and Developmental Disorders, 31*(5), 483-490.

Troyer, A. K. (2000). Normative data for clustering and switching on verbal fluency tasks. *Journal of Clinical and Experimental Neuropsychology, 22*(3), 370-378.

Troyer, A. K., Moscovitch, M., & Winocur, G. (1997). Clustering and switching as two components of verbal fluency: Evidence from younger and older healthy adults. *Neuropsychology, 11*(1), 138-146.

Tulving, E. (1962). Subjective organization in free recall of "unrelated" words. *Psychological Review*, *69*(4), 344-354.

Tversky, B. (1985). Development of taxonomic organization of named and pictured categories. *Developmental Psychology, 21*(6), 1111-1119.

Uddin, L. Q., Supekar, K., & Menon, V. (2013). Reconceptualizing functional brain connectivity in autism from a developmental perspective. *Frontiers in Human Neuroscience, 7*(458), 1-11.

Unsworth, N., Spillers, G. J., & Brewer, G. A. (2011). Variation in verbal fluency: A latent variable analysis of clustering, switching, and overall performance. *Quarterly Journal of Experimental Psychology (2006), 64*(3), 447-466.

Uttal, D. H., O'Doherty, K., Newland, R., Hand, L. L., & Deloache, J. S. (2009). Dual representation and the linking of concrete and symbolic representations. *Child Development Perspectives, 3*(3), 156-159.

Vigliocco, G., Vinson, D. P., Paganelli, F., & Dworzynski, K. (2005). Grammatical gender effects on cognition: Implications for language learning and language use. *Journal of Experimental Psychology.General*, *134*(4), 501-520.

Vygotsky, L. S. (1962 (originally written in 1934)). *Thought and language*. Cambridge, MA: MIT Press.

Walenski, M., Mostofsky, S. H., Gidley Larson, J., & Ullman, M. T. (2005). Fast grammatical processing in autism. *Twelfth Annual Meeting of the Cognitive Neuroscience Society,* New York, NY.

Waxman, S. R. (1991). Covergences between semantic and conceptual organization in the preschool years. In S. A. Gelman, & J. P. Byrnes (Eds.), *Perspectives on language and thought* (pp. 107-144)

Waxman, S. R., & Namy, L. L. (1997). Challenging the notion of a thematic preference in young children. *Developmental Psychology, 33*(3), 555-567.

Wetherby, A. M. (1986). Ontogeny of communicative functions in autism. *Journal of Autism and Developmental Disorders, 16*(3), 295-316.

Wierzbicka, A. (1996). *Semantics: Primes and universals*. Oxford: Oxford University Press.

Williams, D. L., Cherkassky, V. L., Mason, R. A., Keller, T. A., Minshew, N. J., & Just, M. A. (2013). Brain function differences in language processing in children and adults with autism. *Autism Research : Official Journal of the International Society for Autism Research, 6*(4), 288-302.

Wittgenstein, L. (1921/1922/2007). *Tractatus logico-philosophicus* [Logisch-Philosophische abhandlung, annalen der naturphilosophie] (C. K. Ogden Trans.). Bibliobazaar.

Wojcik, D. Z., Allen, R. J., Brown, C., & Souchay, C. (2011). Memory for actions in autism spectrum disorder. *Memory, 19*(6), 549-558.

Wojcik, D. Z., Moulin, C. J., & Souchay, C. (2013). Metamemory in children with autism: Exploring "feeling-of-knowing" in episodic and semantic memory. *Neuropsychology, 27*(1), 19-27.

Woodcock, R. W. (1973). *Woodcock reading mastery test*. Circle Pines, MN: American Guidance Service.

Woodworth, R. S., & Schlosberg, H. (1954). In Holt R. &. W. (Ed.), *Experimental psychology*. New York: Holt, Rinehart & Winston.

Yasumura, A., Kokubo, N., Yamamoto, H., Yasumura, Y., Moriguchi, Y., Nakagawa, E., . . . Hiraki, K. (2012). Neurobehavioral and hemodynamic evaluation of cognitive shifting in children with autism spectrum disorder. *Journal of Behavioral and Brain Science, 2*(4), 463-470.

Yee, E., & Sedivy, J. C. (2006). Eye movements to pictures reveal transient semantic activation during spoken word recognition. *Journal of Experimental Psychology, Learning, Memory, and Cognition, 32*(1), 1-14.

Zelazo, P. D. (2006). The dimensional change card sort (DCCS): A method of assessing executive function in children. *Nature Protocols, 1*(1), 297-301.