CONCEPT FORMATION IN AUTISTIC SPECTRUM DISORDERS

MACKENZIE STABILE

A thesis submitted to
The Department of Psychology
Rutgers University

Written under the direction of
Professor Karin Stromswold
Of the Department of Psychology
Rutgers University
New Brunswick, New Jersey

April 23rd, 2014
Abstract

A large body of research supports the theory that typically developing (TD) children undergo a shift from thematic to taxonomic conceptual organization around the age of 7. Despite ample support for this shift in TD children, the thematic-to-taxonomic shift has never been explored in individuals with autistic spectrum disorders (ASD). In order to investigate how TD and ASD individuals categorize information, as well as the mechanisms that underlie the shift from thematic-to-taxonomic categories, 24 ASD participants were matched by mental age (MA) to 26 TD participants. Participants completed a battery of grouping tasks that assessed categorization and justification ability across basic, intermediate, and superordinate level categories. In order to examine independent predictors of taxonomic justification ability, participants were also tested on tasks measuring cognitive flexibility, metamemorial awareness, linguistic sophistication, and declarative knowledge. A series of multiple logistic regression analyses revealed that MA was the best predictor of taxonomic categorization on all tasks. Although the ASD group demonstrated significant impairments in metamemory, declarative knowledge, and cognitive flexibility, there was no difference between ASD and TD performance on any of the categorization tasks. These findings indicate that individuals with ASD may not form concepts differently than TD children.
The Thematic-to-Taxonomic Shift in Categorization

A large body of research supports the theory that typically developing children may have fundamentally different types of categories than adults. This research suggests that, prior to the age of 7, children tend to categorize information in terms of thematic events (e.g. PRESENT and EGGNOG are both Christmas items), whereas adults tend to categorize information in terms of abstract concepts or “taxonomies” (e.g. EGGNOG and HOT CHOCOLATE are both warm beverages). Despite a considerable amount of research, the mechanisms that underlie children’s shift from thematic-to-taxonomic categories are not entirely understood.

This thesis explores the nature of categories in typically developing individuals and in individuals with autistic spectrum disorders (ASD), as well as the mechanisms that underlie the shift from thematic-to-taxonomic categories. The organization of the thesis is as follows. In the first section, we begin with a brief review of existing research on the thematic-to-taxonomic shift in typically developing children, highlighting key proposals on why the shift occurs. We then review the literature on categorization development in ASD. Because no previous research has investigated the thematic-to-taxonomic shift in autistic spectrum disorders, we focus our review on those factors that researchers have proposed underlie the thematic-to-taxonomic shift in typically developing children. In the second section, we present the experiment we conducted on typically developing and ASD individuals’ categories. In the final section, we discuss how the results of this experiment elucidate the factors that underlie the thematic-to-taxonomic shift in typically developing children, and the nature of categories in typically developing and ASD individuals.
Thematic Categories: A matter of Preference?

Why do children’s categories shift from being event-based thematic concepts to taxonomic concepts around the age of 7? Because even infants appear to have taxonomic categories like LIVING and NON-LIVING (e.g., Gelman & Opfer, 2002; Rostad, Yott, & Poulin-Dubois, 2012), and taxonomic information is an essential component for word learning (Nelson, 1977), it is highly unlikely that children lack taxonomic categories entirely.

If very young children have some taxonomic categories, one possible explanation for the thematic-to-taxonomic shift is that it results from a change in children’s categorization preference rather than a change in ability. Researchers have argued that young children’s preference for thematic categories may be due to their relevance in everyday life (Inhelder & Piaget, 1964), and because children practice thematic relations in their spontaneous play (Nelson & Seidman, 1984).

One study that supports the possibility that the thematic-to-taxonomic shift reflects different preferences depending on situational demands is a study by Smiley and Brown (1979) that shows that children appear to have different types of categories depending on the task they are perform. For example, their match-to-sample triad task, in which children were asked to match a standard object (e.g., BEACH BALL) to their choice of a taxonomically related target (e.g., BASEBALL) or a thematically related target (e.g., SAND), Smiley and Brown (1979) found that five- and six-year-old children tended to choose the thematic target. However, when asked to justify matching the

---

1 Word learning requires an individual to practice a certain degree of abstraction and generalization. For instance, in order to learn the word TABLE, a child must integrate multiple visually distinct exemplars (e.g., a wooden table with short legs, a coffee table, a glass table) into his/her conception of TABLE. Thus, word learning requires taxonomic categorization.
standard with a taxonomic target, the children were able to do so. This suggests that children are capable of understanding and producing taxonomic categories, but they either prefer thematic categories, or thematic categories are more salient to them.

**Thematic-to-Taxonomic Shift and Cognitive Development**

Rather than the thematic-to-taxonomic shift resulting from changes in children’s preference, another possibility is that the thematic-to-taxonomic shift is driven by children’s emerging abilities in one or more specific areas, such as declarative knowledge, language, cognitive flexibility or metamemory.

**Declarative knowledge.** Some researchers have proposed that children’s thematic-to-taxonomic shift results from their acquiring more world knowledge (Sheng, McGregor, & Marion, 2006). Logically, a certain degree of knowledge about a given set of objects is essential for taxonomic categorization. For instance, in order to create the category ANIMAL, one must know that animals share certain properties (e.g., that all things that are animals are alive, must eat and breathe, and that when they procreate, their offspring are the same “type” as they are). The more specific the category, the more knowledge is necessary. For example, in order to create a REPTILE category, one must first be able to identify what animals are considered reptiles and then infer that “cold-blooded” is a feature of the REPTILE category. By this reasoning, the more knowledge a child possesses about a given set of objects, the more likely he or she will be to create taxonomic categories.

---

2 For the purposes of this paper, we assume that for all but the simplest categories like ALIVE or NOT ALIVE (which may be innate), children start by identifying one or two exemplars of a category, infer the features shared by the members of the category, and then are able to use these features to identify new members of a category.
**Cognitive Flexibility.** The term cognitive flexibility refers, broadly speaking, to the ability to shift thoughts or actions depending on situational demands. Categorizations are inherently contingent upon cognitive flexibility. As such, the thematic-to-taxonomic shift could be driven by children’s developing capacity to manipulate information mentally. If this possibility is correct, we might expect that cognitive flexibility would develop in concert with children’s shift from thematic-to-taxonomic categories.

Researchers have proposed various definitions of cognitive flexibility pertaining to multiple aspects of categorization. Blaye, Bernard-Peyron, Paour, and Bonthoux (2006) propose two types of categorical flexibility: response flexibility (i.e., the ability to create multiple sorts in a given task) and conceptual flexibility (i.e., the ability to switch between multiple types of conceptual organizations, such as thematic and taxonomic, in the same set of stimuli). The ability to switch between multiple types of conceptual organizations is also referred to as cross-classification by Nguyen and Murphy (2003). For the purposes of this paper, we are primarily interested in Blaye et. al’s (2006) notion of “conceptual flexibility,” which we will henceforth refer to as cognitive flexibility.

**Metamemory.** Another cognitive ability that could contribute to children’s thematic-to-taxonomic shift is metamemory. Metamemory, a type of metacognition, is one’s knowledge and awareness of one’s own memory, including information storage and retrieval (Flavell & Wellman, 1977). Metamemory is thought to aid in information retrieval by enabling individuals to utilize strategic organizational strategies.

Following their discovery that children, unlike adults, did not semantically cluster on rapid-naming tasks (i.e., name multiple objects from one category before switching to another), Weinert and Schneider (1995) proposed that children’s lack of clustering might
result from poor metamorial awareness. According to Weinert and Schneider’s (1995) “strategy emergence theory,” children between the ages of 8 and 10 have acquired enough metamemorial awareness to implement organizational strategies on free recall and categorization tasks. Increased metamemorial awareness should therefore result in better clustering and switching strategies. For example, if a child is asked to name animals, he or she might begin by naming CAT. A child with high metamemorial awareness might be aware that recalling multiple animals that belong to the same category (e.g., PETS) is an efficient strategy to assist in recall. A younger child, on the other hand, might not possess the metamemorial awareness to utilize this specific strategy. If the “strategy emergence theory” is correct, increases in metamemorial awareness might contribute to the shift from thematic-to-taxonomic categorization in children. Consistent with this, some previous research has shown that metacognitive abilities such as metamemory play a crucial role in concept formation and learning. For example, one study found that metamemory has a strong impact on an individual’s ability to form concepts (Anderson & Nashon, 2007).

**Linguistic Development.** Another possibility is that the thematic-to-taxonomic shift is facilitated by general linguistic development. One aspect of linguistic development that appears to be related to the thematic-to-taxonomic shift is the syntagmatic-paradigmatic shift. Woodworth and Schlosberg (1954) found that when adults are given a word association task, adults often respond to a stimulus word (e.g., PENCIL) with a parallel and related word (e.g., PEN), whereas children’s responses tend to be thematically related (e.g., PAPER). The term “syntactic-paradigmatic shift,” coined by Brown and Berko (1960), refers to the finding that children tend to respond to this task
with words found in syntactic contiguity, (i.e., with words that co-occur with a stimulus word). For example, if given a stimulus word COLD, a child might respond by naming OUTSIDE. On the other hand, adults tend to respond to such tasks with words found in paradigmatic contiguity (i.e., words of the same “type”). An example of a paradigmatic response to the stimulus word COLD would be HOT. Incidentally, syntagmatic responses typically involve event-based words (thematic), and paradigmatic responses tend to reflect similar “kinds” (taxonomic).

The syntactic-paradigmatic shift is quite similar to the thematic-to-taxonomic shift. Given the similarity between the syntagmatic-paradigmatic shift and the thematic-to-taxonomic shift, it is possible that the two shifts are driven by similar developments. Brown and Berko (1960) argued that the syntagmatic-paradigmatic results from increased syntactic sophistication. Syntax, therefore, might play an important role in the thematic-to-taxonomic shift as well. Because syntactic sophistication entails knowledge that words fall into specific categories (e.g., NOUNS, VERBS), increases in syntactic sophistication might bolster a child’s awareness that objects fall into specific semantic categories, or taxonomies (e.g., FURNITURE, or ANIMAL).

Another possibility is that the syntactic-paradigmatic shift is facilitated through children’s broadening vocabulary with age. Although Brown and Berko (1960) found age-related trends in paradigmatic responses, with first graders giving mainly syntagmatic responses and adults giving mainly paradigmatic responses, Ervin (1989) discovered that the type of response children gave on word association tasks was related to the type of stimulus word given. He found that even most kindergarteners often
responded paradigmatically to mass or count nouns. Thus, he concluded that vocabulary size plays a fundamental role in the syntagmatic-paradigmatic shift.

**Effect of Level.** As discussed above, some researchers have argued that even infants have primitive taxonomic categories like LIVING and NON-LIVING (e.g., Gelman & Opfer, 2002; Rostad, Yott, & Poulin-Dubois, 2012). Perhaps the thematic-to-taxonomic shift is not a global phenomenon, but rather a shift that occurs earlier for superordinate categories (like ANIMAL and FOOD) and basic-object categories (like DOG and CAT) than for intermediate level categories (like PET and ZOO ANIMAL or VEGETABLE or DESSERT) that require domain-specific knowledge.³

Although research suggests that taxonomic categorization may develop sooner in some levels than others, there is disagreement about which levels develop before others. Mandler and Bauer (1998) found that 16 to 20 month old children are better able to differentiate between objects that belong to different superordinate level categories (e.g. CAT vs. HOUSE) than they are able to differentiate between basic level categories that belong to the same category (e.g. DOG vs. CAT). This suggests that taxonomic categorization develops first at the superordinate level. In contrast, some studies have shown that it is easier for two year olds to group items that share the same basic object level (e.g., Daehler, Lonardo, & Bukakto, 1979; Rosch, et al., 1976). Interestingly, relatively little research has investigated the development of “intermediate” level categories.

³ We use the term “intermediate level” to refer to a level of categorization between the superordinate and basic-object levels. For instance, CAT and DOG are both members of the superordinate level category ANIMAL, but CAT and DOG are distinct basic-object level categories (e.g. DOGS vs. CATS). CAT and DOG both belong to the intermediate level category PET category, in addition to both belonging to the superordinate level category ANIMAL.
**Concept Formation in Autism**

Previous research suggests that individuals with autism form concepts differently than typically developing individuals. However, the results of studies on concept formation in autism are varied and inconsistent. Some studies have shown that autistic children have particular difficulty forming concepts related to size and spatial relationships (Ohta, 1987) or to temporal relationships and artifacts (Perkins, Dobinson, Boucher, Bols & Bloom, 2006). Minshew, Meyer and Goldstein (2002) found that autistic children had irregular concept formation, but they had no problems with concept identification. From this, they hypothesized that children with autism display category formation irregularities because they encode concepts at lower levels of abstraction than typically developing children. Klinger and Dawson (2001) found that children with autism performed just as well as typically developing children on tasks requiring them to learn new concepts based on rules. However, they performed worse on tasks requiring them to learn new concepts based on prototype information.

In contrast to the studies above, some studies suggest that autistic children may not have categorization deficits after all. For example, in a study that required participants to categorize randomized dot patterns using operational concept definitions, Froehlich et al. (2012) found that autistic individuals used prototype information the same way that typically developing children who were “matched” with autistic children for mental age. Similarly, in a matching-to-sample study on basic and superordinate level concept formation in typically developing, autistic, and non-autistic mentally retarded children who were matched on the basis of mental age, Tager-Flusberg (1985) found no
apparent differences between autistic and typically-developing children’s ability to categorize information and form abstract concepts.

**Abstraction and Taxonomic Categorization in Autism.** Although the thematic-to-taxonomic shift has been extensively studied in typically developing children, we are not aware of previous research on the thematic-to-taxonomic shift in children with autism spectrum disorders. As mentioned above, Minshew, et al. (2002) posit that autistic individuals' impairments in categorization are due to autistic individuals’ limited capacity for abstraction. Indeed, impaired abstraction ability is a widely reported and accepted feature of autism (Minshew & Goldstein, 1998; Meyer & Minshew, 2002; Minshew & Goldstein, 2002). The literature suggests that individuals with autism encounter difficulty synthesizing information, and focus on “parts” rather than the “whole” (Meyer & Goldstein, 2002). Because taxonomic categorization is inherently abstract, we would expect autistic individuals to develop taxonomic categories later than typically developing individuals.

**Cognitive Flexibility in Autism.** If cognitive flexibility plays a key role in the development of taxonomical organization of categories, we would predict that the shift from thematic-to-taxonomic categorization might occur at a later age for people who are less cognitively flexible such as children with autism (Hammes & Langdell, 1981) than for typically developing children.

One hypothesis is that autistic children’s difficulty with categorization reflects their compulsively concrete thinking patterns (Noache, 1974). For example, Hammes and Langdell (1981) found that autistic individuals lack flexibility in their formation and use of symbols. Impairments in cognitive flexibility would explain why concept formation
abnormalities appear present in most individuals with autism, including high functioning autistic (HFA) individuals with normal to near-normal IQs. In support of this theory, when Yang (1998) examined concept formation as a function of input modality in children with autism and non-autistic children with mental retardation, he found that ASD children had categorization deficits across three modes of input, whereas non-autistic children with mental retardation demonstrated categorization deficits only in one type of input modality.

**Metamemory in Autism.** Surprisingly, the metamemory capabilities of individuals with autism have received very little attention in research. The few studies that have been conducted on metamemory in autism have reported mixed findings. Some research suggests that autistic children do not differ from typically developing children in their metamemorial awareness (Wojcik, Allen, Brown, & Souchay, 2011; Wojcik, Waterman, Lestie, Moulin, & Souchay, 2013), while other research indicates that autistic children have selective metamemory impairments (Wojcik, 2011; Farrant, Blades, & Boucher, 1999). Due to the paucity of research on the subject, we cannot say with certainty whether children with autism do or do not suffer metamemory deficits.

However, given that individuals with autism typically have broadly impaired meta-cognitive functioning (e.g., Wilson, Best, Minshew & Strauss, 2010), it is plausible that metamemory, a subset of metacognition, is impaired in autism. Flavel et al. (1993) identified three variables that affect metamemorial awareness in typically developing children. These variables are knowledge about metamemory tasks (i.e. what makes one memory task more difficult than another), knowledge of one’s own and others’ memory abilities, and knowledge of mnemonic strategies. If autistic individuals lack Theory of
Mind, (e.g., Happe, 1995), we would predict that they would be impaired in at least one of the three components of metamemory awareness (namely, knowledge of one’s own and others’ memory abilities).

In keeping with this theory, one study that investigated autistic children’s metacognitive monitoring (i.e. estimation of future performance) and metacognitive control (i.e. the ability to utilize memory strategies) revealed that children with autism show selective difficulties with metacognitive control (Wojcik, 2011). Therefore, we might expect autistic individuals to perform worse on categorization tasks. In fact, there is some evidence that the implementation of metacognitive training in children with learning disabilities has a significant effect on improving concept formation (Lauth, Husein & Spies, 2006).

**Category levels.** The effect of level on categorization ability in autism is relatively understudied, although Tager-Flusberg (1985) did report that autistic children had more difficulty grouping objects from the same superordinate level than objects from the same basic level. At the intermediate level, objects can be grouped in many ways. For example, if shown the animals CAT, DOG, LION, WOLF, MOUSE, and SQUIRREL, a child could conceivably group CAT, DOG, and MOUSE because they are all PETS (a taxonomic grouping). Alternatively, a child could group CAT, MOUSE, and SQUIRREL because cats chase mice and squirrels (a thematic grouping), or they could group CAT and together because they are both FELINES (a taxonomic grouping).

Because autistic children have particular difficulty performing tasks with poorly defined rules (Cielieski & Harris, 1997) and the “rules” of grouping on the intermediate level are
fluid, we were particularly interested in how our autistic participants performed on intermediate level grouping tasks.

In the study that follows, we investigated the nature of categories in autistic individuals and mental-age matched typically developing children in order to elucidate the nature of young children’s categories and to investigate the factors that underlie the shift to adult-like categories. In addition to helping to elucidate the nature of categories in typically developing children and the factors underlying the apparent thematic-to-taxonomic shift, understanding the processes by which people with autism use categories has important educational and clinical implications.\(^4\)

If cognitive flexibility plays a key role in the development of taxonomical organization of categories, we would predict that the shift from thematic-to-taxonomic categorization would occur at a later age for populations that demonstrate less cognitive flexibility such as autism (Hammes & Langdell, 1981). Similarly, if metamemory skills are critical, we would expect autistic individuals to have a delayed thematic-to-taxonomic shifts because autistic individuals may not have as well-developed metamemories as typically developing children.

Whereas autistic individuals often have difficulties with metacognitive abilities and flexible thinking, high functioning autistic individuals often have comparatively intact declarative knowledge, vocabulary and grammar. Thus, if these skills are critical in

\(^4\) For example, should cognitive flexibility play an important role in concept formation, adapting teaching methods to be more rule-based and straightforward could help to facilitate learning in a school setting. On the other hand, an increased emphasis on promoting flexible thinking in therapeutic settings could help to remediate impairments in conceptual processes. Should metamemory play a significant role in concept formation, an increased emphasis on metamemory training in the classroom might benefit autistic children.
driving the thematic-to-taxonomic shift, we might expect that the thematic-to-taxonomic shift would be relatively similar in autistic and mental age-matched typically developing children.

Given that taxonomic categorization ability may develop more rapidly at some levels than others and autistic individuals might be expected to have more difficulty with abstract categories, we investigated concept formation at three levels: the superordinate level (e.g. ANIMAL vs. FOOD), the basic level (CAT vs. DOG), and the intermediate level (PET vs. ZOO ANIMAL). If cognitive flexibility plays a critical role in the thematic-to-taxonomic shift, we would predict that autistic and age-matched participants would be more similar to one another on superordinate and basic level categories than on intermediate level categories.
Methods

Participants

Participants were 24 autistic, monolingual, English-speaking individuals who ranged in age from 3.92 to 19.96 years (mean = 11.92 years, SE = .91) and 26 typically-developing, monolingual English-speaking children who ranged in age from 2.39 years to 8.83 years (mean = 5.08 years, SE = .45). Children were recruited from schools and camps in New Jersey. The protocol was approved by the Rutgers University Institutional Review Board for Human Subjects.

Testing Procedure

Children were tested in a quiet room with no other children present in two half-hour sessions. During the first session, each participant completed the entire battery of experimental tasks. Approximately one week later, each participant was administered the Kaufmann Brief Intelligence Test-2 (Kaufmann, 2005). All the sessions were video-recorded for later coding.

Standardized Assessment

Kaufmann Brief Intelligence Test-2 (K-BIT-2, Kaufman, 2005). Both typically developing and autistic participants were given the KBIT-2. The KBIT-2 is a commonly used and well-normed IQ test for individuals between the ages of 4 and 90. The test has high internal-consistency reliability and the correlation between KBIT-2 Verbal scores and corresponding portions of the WISC-III are .83 and .79, indicating high external validity. The test has also been proven to 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.

**Experimental Battery**

**General grouping procedure.** In the grouping tasks in our study, we showed participants pictures of cards and asked them to “put the ones that are kind of the same together” in trading card holders. We carefully designed the task and the instructions this way because a review of categorization studies revealed that the task and instructions used to assess children’s categories can profoundly affect the type of groups that children form.

For example, many researchers investigate the thematic-to-taxonomic shift by using a match-to-sample task. In the match-to-sample task, children are shown a target item (e.g., boat), a taxonomic match (e.g., car), and a thematic match (e.g., ocean) and asked which one is “the same as” or which “goes with” the target. However, as several researchers have noted, compared to a free sort task in which children are free to group items however they want, 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).

Details of the matching or sorting task can also affect how children group objects. For example, Markman, Cox, and Machida (1981) found that using a “spatially-extended surface” such as a table with no compartments biased preschoolers towards thinking thematically. Having children put items in piles is problematic because then they cannot see what items are already in a group. We chose to have the participants put cards in a
card holder because this enabled them to see what the other members of the group were, while providing an unambiguous sense of items belonging to a set.

The instructions children are given also affects their grouping behavior. Researchers who employ either a match-to-sample task or a free sorting task tend to either ask children to group items that are “the same” or to group items that “go together.” The problem with asking children to group items that “go together” is that doing so tends to bias children to think thematically (see Denney & Moulton, 1976; Waxman & Namy, 1997). The problem with asking children to group items that are “the same” is that, during pilot testing for the current study, children seemed to interpret “the same” as meaning “exactly the same.” In order to avoid these two problems, we asked participants to “put the ones that are kind of the same together.”

**Task 1: Familiarization task.** The first task in the experimental battery was designed to familiarize participants with the grouping procedure. Each participant was shown two trading card holders, one filled with Disney cards and the other with Sesame Street cards (see Fig 1).
Fig 1. Familiarization task materials

The experimenter pointed at the Disney holder and asked, “Do you know who any of these characters are?” If the participant named at least one of the characters, the experimenter responded, “Good!” If not, the experimenter said, “That’s okay!” The experimenter then asked the participant where the characters were from. If the participant did not know the answer, the experimenter explained that they were all Disney characters. Next, the experimenter pointed at Sesame Street card holder and asked the same set of questions. Finally, the experimenter revealed a Mickey Mouse card (see Figure 2) and asked the participant to name the character on the card. If the participant did not say Mickey Mouse, the experimenter said, “This is Mickey Mouse!”
Next, 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 participant responded “Disney,” the experimenter said, “That’s great! Now you know how to play the game. Let’s try another one.” If the participant did not answer correctly, 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.”

**Task 2: Superordinate level task.** In the second task, participants named, grouped, and then justified their groupings of eight cards (see Figure 3). These eight cards had 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).
Two pseudorandomized orders of cards were created, and the cards were presented in one of the two orders to each participant. This task contained three subparts: vocabulary, grouping, and justification.

I. Procedure.

a. Vocabulary. At the beginning of the superordinate task, a stack of blank cardholders was placed in front of the participant. The experimenter then put down one of the 8 cards and asked the participant, “Do you know what this is?” If the participant correctly identified the item, the experimenter said “Good” and put down the next card. If the participant answered incorrectly or did not respond, the experimenter said “It’s a (banana)” and went to the next card. After all eight cards were laid out in front of the participant, the experimenter asked the participant to name any of the items that the child did not initially name correctly. Again, the experimenter responded with “Good” or “It’s a (banana)” after each item.

---

5 One reason we had participants label the items was because Deak & Bauer (1996) found that preschool children often label lined drawings at the superordinate level (e.g., labeling a banana as “food”), and we wanted to be sure that the participants 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.
b. **Grouping.** After the participant had named all 8 food and clothing cards, the experimenter introduced the grouping task by saying, “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 participant “Which ones should we put in here?” and allowed the participant to place cards into the holder, assisting if necessary. When a participant stopped putting cards into the holder for three seconds, the experimenter asked “Any more or just those?” and gave the participant the opportunity to add more cards. This question was repeated until the participant did not want to add more cards.

c. **Justification.** Once the participant had finished adding cards to a cardholder, the experimenter pointed to the cardholder with cards in it and said, “Great, how are these kind of the same?” to elicit the participant’s justification. The experimenter then picked another empty cardholder 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 participant’s attention to the remaining cards. The above procedure was repeated until all cards were used or until the participant 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 participants named correctly on the first try. For some items, several answers were considered correct (e.g., pants or jeans, shirt or t-shirt, etc.)

b. **Grouping.** For the purposes of scoring for this task and all subsequent tasks, a participant was said to form a group if he or she deliberately put two or more cards in a cardholder. A grouping was not counted if the participant appeared to be arbitrarily
grouping cards. Specifically, groupings were excluded if a participant chose consecutive cards in a row or column without justifying this grouping; if a participant grabbed several cards at once without justifying this grouping; or if a participant included all items in one grouping.

In this task, food and clothing were identified \textit{a priori} as the two taxonomic features by which participants could group. A participant 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. In the superordinate level task, no thematic or perceptual features were identified \textit{a priori} and, thus, it was not possible to code whether participants formed thematic or perceptual groups.

\textbf{c. Justification.} A participant’s justification for a group 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’’); \textsuperscript{6} or none (if a participant gave no response or gave a non-justification such as “They’re just the same’’). All of the justifications from twenty randomly selected participants were independently coded by two coders. Coders agreed on all but one justification (Cohen’s kappa = .95).

\textbf{Task 3: Perceptual Grouping task.}

In this task, participants grouped twenty “Set” cards that had objects on them that varied along four dimensions (see Figure 3). The four dimensions were shape of objects (diamond, oval, squiggly); color of objects (red, green, purple); texture of objects (solid, lined, blank); and number of objects (one, two, three).

\textsuperscript{6} Researchers disagree about whether to count perceptual justifications as taxonomic. We separated them in order to explore the development of abstract (i.e., non-perceptual) categories.
Two randomized orders were created and participants were randomly assigned to an order.

**I. Procedure.** 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 participant the opportunity to add more cards until the participant was finished and then asking for the participant’s justification. However, after a participant’s justification, the experimenter praised the participant 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 participant said that there were no more groupings.

**II. Coding.** Responses were coded in several ways.
1. **Grouping.** A participant was classified as having grouped if all cards in a least one grouping of two or more cards contained only cards that shared the same value for one or more of the 4 dimensions (e.g. a grouping with only red cards, a grouping with only diamond cards).

2. **Justifying.** A participant was classified as having justified if he or she accurately justified at least one group. He or she could do this by saying that all of the cards had the same value for a dimension (e.g., “They’re all the same color”) or by saying the value that all of the cards shared (e.g., “They’re all red”).

3. **Switching Grouping Dimensions.** Notice that one can use any of the four dimensions to group the Set cards. Being able to switch grouping dimensions or to justify using different dimensions is often taken as a measure of cognitive flexibility (Zelazo, 2006). A participant was classified as having switched grouped dimensions if he or she grouped using different dimensions in different groupings. For example, a participant received credit for switching dimensions if he or she grouped cards that had objects of the same color (e.g., cards with red objects) in one grouping and then grouped cards that had objects of the same shape (e.g., cards with ovals). A participant also received credit for switching grouping dimension if he or she grouped by shape and color in one grouping (e.g., objects that were diamonds and red) and by shape and texture in another grouping (e.g., objects that were diamond and striped). Note that grouping first by “red” and then by “green” was not considered switching dimensions because both red and green are values of the dimension “color.” Additionally, grouping by shape and color (e.g., red diamonds) in one grouping and again by shape and color (e.g., green ovals) in another grouping was not considered switching.
4. Switching Justification Dimensions. A participant received credit for switching justification dimensions if he or she justified by more than one dimension either within or across groupings. For example, if a participant grouped red objects together and said, “They’re red” and then grouped diamonds together and said, “They’re diamonds” for the second grouping, this was considered justifying by more than one dimension. Additionally, if a participant formed a single group of cards that had objects that were both red and diamonds said “They’re red diamonds,” this was considered justifying by more than one dimension because the justification involved two dimensions (color and shape). In contrast, if a participant grouped red objects together and said, “They’re red” and then grouped green objects together and said, “They’re green” for the second grouping, this was not considered switching because red and green are values of the same dimension (color).

Task 4: Intermediate level Task. In this task, participants named and grouped twenty cards that depicted black and white line drawings of 20 distinct animals (see Figure 4). The 20 animals on the cards were the twenty animals that were most frequently named by a group of 275 three- to five-year olds in a separate verbal fluency task (see Isacoff & Stromswold, 2014). Two randomized orders of the cards
were created, and participants were randomly assigned to an order. As was the case with the superordinate level task, the intermediate level task had three components: vocabulary, grouping, and justification.

I. Procedure. The experimenter had the participant name each animal using the same elicitation procedure that was used in the superordinate level task. Next, the experimenter used the procedures outlined in the perceptual level task to elicit animal groupings and justifications.

II. Coding. A participant’s animal vocabulary score was the number of animals (out of 20) that the participant correctly named on the first try. For some animals, several answers were considered correct (e.g., kitty, cat, or kitten; mouse or rat). Given that we did not posit any grouping features \textit{a priori}, we did not code participant’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). All justifications from twenty randomly selected participants were independently coded by two coders. Coders agreed on all justifications.

Task 5: Basic level task.

I. Procedure. In this task, participants grouped black and white line drawings depicting multiple visually distinct exemplars of the same traditional basic level categories (two cats, two dogs, two giraffes, three elephants, four fish, see Figure 5).

![Figure 5](image_url)  
**Fig 5.** Basic Level Task cards.

Two random orders of the cards were created, and participants were randomly assigned to an order. We purposefully did not have participants name the animals in this task because we did not want to increase 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 those used in the other grouping tasks.

II. Coding. Five groupings (cat, dog, elephant, fish, and giraffe) were determined a priori to be the possible taxonomic groupings. Just as in the superordinate level task, participants were given credit for taxonomically grouping if at least one of their groupings only contained members from one of these categories. Participant’s justifications were coded as taxonomic (e.g., “They’re dogs”). All of the participants’ justifications were basic object level taxonomic justifications.

Task 6: Animal declarative knowledge task. The animal declarative knowledge task was, in some ways, the reverse of the intermediate-grouping task in that children were asked to identify animals that had certain properties. For example, participants were asked “Which ones are pets? Which ones are scary?” and so on. Participants were asked a total of twenty questions that correspond to the 20 features that correspond to the 20 features that adults used most often to group animals (see appendix A for list of features) (Isacoff, Cohen, Liu, Hou-Imerman & Stromswold, EPA, March, 2013)

I. Procedure. The experimenter placed the cards in front of the participant in the same layout as in the intermediate level task. 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.” Participants were encouraged to point to the cards that best answered the questions. Verbal responses were accepted as well. The same exclusion criteria applied as in the grouping tasks (i.e., if participants 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 participants were randomly assigned to an order.

II. Coding. A participant’s animal declarative knowledge score was the number of questions (out of twenty) for which the participant correctly identified at least one animal and did not misidentify any animals. For example, in response to “Which ones are mammals?” a participant would receive credit for selecting the lion but not for selecting both the lion and the bird. In order to verify which animals counted for each question, we used adult subjects’ responses from the verification task described in the previous chapter.

Task 7: Metamemory task. This task was adapted from Flavell’s (1976) test of metamemory. In this task, the participant 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 participant then heard five sets of facts about Megan and Henry. For example, in one question, the participant was told “Megan (pointing) is trying to learn 18 new words. Henry (pointing) is trying to learn three new words. The facts were randomly assigned to each puppet. (See appendix B for list of Metamemory items).

Task 8: Spontaneous speech. In this task, the participant saw a colored drawing of a playground, and were told, “Here is a picture of a playground. Can you tell me what you see?” After the participant described the picture, or if the participant did not provide a description, the participant was asked “Can you tell me a story about the picture?” The mean number of morphemes per utterance (i.e., mean length of utterance, or MLU) was calculated using the procedure outlined in Brown (1973).
Results

Because of the number of statistical comparisons that were made in this study, we set the critical value for significance at $p = .01$, rather than $p = .05$.

Mental Age Matching

Autistic participants were matched to typically developing participants on the basis of their average mental ages (henceforth “MA”), which were calculated by averaging the verbal MA and nonverbal MAs obtained from the KBIT-2. Because the KBIT-2 is designed for children ages 4 and up, any score with a corresponding mental age of “less than 4 was coded as a 3.5 MA. The MA for the ASD group was 6.06 ($SE = .52$, range = 3.50 - 11.33), and the MA for the TD group was 5.18 ($SE = .40$, range = 3.5-12.33). This difference was not significant (unpaired $t(48) = .45$, $p = .65$).

Analyses

A long-standing issue in research involving participants who are members of atypical (impaired) populations is whether or not to match participants to typically developing participants on the basis of IQ. On the one hand, controlling for general intelligence reduces the likelihood that results obtained merely reflect the known differences in intellectual abilities of the atypical and typical group, and allows one to uncover any selective deficits that might exist in the atypical population. On the other hand, matching or controlling for mental age has several limitations.

First, the nature of “intelligence” is hotly debated, and encompasses a wide variety of cognitive abilities, from working-memory to processing speed. Second, depending on what tasks are included in the general intelligence test, one may inadvertently obscure atypical participants’ selective deficits. For example, if the IQ test
that is used has many items that require cognitive flexibility, matching typical and atypical participants for mental age may undermine or eliminate the importance of cognitive flexibility on the experimental task.

For this reason, analyses were conducted in two ways. In order to determine whether specific features of autism contribute to atypical concept formation above and beyond autistic participants’ cognitive impairments, all analyses were performed twice: once using MA and once using chronological age (CA).

**Animal Declarative Knowledge Task**

Overall, the participants correctly answered 13.92 of the 20 animal knowledge questions ($SE = .69$), with the ASD participants correctly answering an average of 12.46 questions ($SE = 1.18$) and the TD children correctly answering an average of 15.27 questions ($SE = .67$). As shown in Figure 6, Pearson’s correlation analyses revealed that both ASD and typically developing participants’ animal knowledge scores were highly correlated with both MA and CA (all $r’s > .55$, all $p’s < .005$).
Participants’ animal knowledge scores were analyzed using an analysis of covariance (ANCOVA) with MA as a covariate and Group (ASD vs. TD) as a between subjects variable. Using this significance level, the ANCOVA revealed that there was a significant difference in the ASD and TD participant’s scores ($F(1,47) = 9.67, p = .003$) and MA had a significant effect, with participants with higher MAs correctly answering more questions than participants with lower MAs ($F(1,47) = 33.11, p < .0005$). When animal knowledge scores were analyzed using an ANCOVA with chronological age (CA) as the covariate and Group (ASD vs. TD) as the between subjects variable, typically developing children correctly answered more questions than ASD children ($F(1,47) = 9.67, p = .003$).
27.16, \( p < .0005 \) and chronologically older participants answered more questions correctly than younger participants (\( F(1,47) = 23.25, \ p < .0005 \)).

**Animal Vocabulary**

Overall, the participants knew the names of 18.70 of the 20 animals (\( SE = .31 \)), with the ASD group knowing the names of 18.13 animals (\( SE = .59 \)) and the TD group knowing the names of 19.23 animals (\( SE = .20 \)). As shown in Figure 7, simple linear regression analyses revealed that MA was a marginally significant predictor of animal vocabulary in both ASD and TD children (both \( p \)’s < .05). CA was a significant predictor of TD participant’s animal vocabulary (\( r = .52, \ p = .006 \)), but CA was not a marginally significant predictor of ASD participants’ Animal Vocabulary (\( r = .19, \ p = .375 \)).

**Fig 7A.** Typically developing participants

**Fig 7B.** ASD participants

**Fig 7C.** Typically developing participants

**Fig 7D.** ASD participants
Figure 7. Age and Animal Vocabulary scores in typically developing and Autistic Spectrum Disorder (ASD) participants.

Participants’ Animal Vocabulary scores were analyzed using an ANCOVA with MA as a covariate and Group (ASD vs. TD) as a between subjects variable. This analysis revealed that participants with higher MAs knew the names of more animals than participants with lower MAs ($F(1,47) = 12.12, p = .001$), but there was only a marginally significant difference in the number of animals known by ASD and TD participants ($F(1,47) = 5.04, p = .03$). Participants’ animal vocabulary scores were then analyzed in an ANCOVA with CA as a covariate and Group (ASD vs TD) as a between subjects variable. The analysis revealed that neither CA ($p = .093$) nor Group ($p = .015$) played a significant role in the number of animals

**Metamemory**

Overall, the participants correctly answered 2.70 out of 5 metamemory questions ($SE = .17$), with the ASD group correctly answering 2.27 questions ($SE = .27$) and the TD group correctly answering 3.08 questions ($SE = .21$). As shown in Figure 8, linear regression 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 participant’s metamemory scores ($r = .60, p = .001$), but CA was not even a marginally significant predictor of ASD participants’ animal vocabulary ($r = .34, p = .09$).
Participants’ metamemory scores were analyzed using an ANCOVA with MA as a covariate and Group (ASD vs. TD) as a between subjects variable. This analysis revealed that TD participants correctly answered significantly more questions than ASD participants ($F(1,47) = 12.55$, $p = .001$) and that participants with higher MAs correctly answered significantly more questions than participants with lower MAs ($F(1,47) = 38.60$, $p < .0005$). When this analysis was performed with CA (rather than MA) as a covariate, the results were virtually the same, with significant effects of both CA ($F(1,47) = 10.36$, $p = .002$) and Group ($F(1,47) = 16.68$, $p < .0005$).

**Spontaneous Speech Task**
Overall, participants had an average Mean Length of Utterance (MLU) of 3.18 ($SE = .27$), with the ASD group having an MLU of 2.74 ($SE = .44$) and the TD group having an MLU of 3.59 ($SE = .30$). As shown in Figure 9, linear regression analyses revealed that MA was a significant predictor of ASD participants’ MLUs ($r's = .62, p = .001$), but not TD participants’ MLUs ($p > .05$). In contrast, CA was a marginally significant predictor of TD participant’s MLUs ($r = .40, p = .044$), but CA was not even a marginally significant predictor of ASD participants’ MLUs ($r = .12, p = .58$).

![Fig 9A. Typically developing participants](image1)

![Fig 9B. ASD participants](image2)

![Fig 9C. Typically developing participants](image3)

![Fig 9D. ASD participants](image4)

**Fig 9.** Age and Mean Length of Utterance in typically developing and Autistic Spectrum Disorder (ASD) participants.

When participants’ MLUs were analyzed using an ANCOVA with MA as a covariate and Group (ASD vs. TD) as a between subjects variable, participants with
higher MAs had larger MLUs than participants with lower MAs \(F(1,47) = 17.42, p = .000\), but the difference between ASD and TD participants was only marginally significant \(F(1,47) = 4.56, p = .038\). When the same analysis was performed with CA as a covariate rather than MA, the difference between ASD and TD MLU’s was still marginally significant \(F(1,47) = 4.52, p = .039\), but CA did not play a significant role on MLUs \(F(1,47) = 1.91, p = .174\).

**Perceptual Task**

**Grouping.** Overall, 49 out of 50 participants (98%) grouped on the perceptual task, with all of the 26 TD children, and all but one of the ASD participants (95.83%) forming at least one perceptual group. Because all but one participant perceptually, we did not conduct further analyses of the presence or absence of grouping on the perceptual task.

**Justification.** Overall, 39 out of 50 participants (78%) justified at least one of their perceptual groups, with 18 of the 24 ASD participants (75%) and 21 of the 26 TD participants (80.80%) doing so. A multiple logistic regression analysis of the perceptual justification data with the variables MA and Group (ASD/TD) was performed. As shown in Table 1, the overall fit of the model was good (Log Likelihood = -16.80, \(\chi^2(2) = 19.11, p < .0005\), Nagelkerke’s \(R^2 = .49\)), with MA, but not Group, being a good independent predictor of participants’ perceptual justification data.
As shown in Table 2, when the perceptual justification data were analyzed using logistic regression with the variables CA and Group (ASD/TD), the overall fit of the model was only fair (Log Likelihood = 22.27, $\chi^2(2) = 8.15, p = .017$, Nagelkerke’s $R^2 = .23$), with both CA and group approaching significance as independent predictors of perceptual justification.
Number of groupings. Although there was a ceiling effect on whether or not participants grouped perceptually, there was variance in how many perceptual groups participants formed. Overall, the participants formed an average of 3.34 groups ($SE = .26$), with the ASD group forming 3.29 groups ($SE = .35$) and the TD group forming 3.38 groups ($SE = .38$). As depicted in Figure 10, linear regression analyses revealed that neither MA nor CA was a significant predictor of the number of perceptual groups formed by either the ASD or TD participants.

Figure 10. Age and number of perceptual groups in typically developing and Autistic Spectrum Disorder (ASD) participants.
When the number of perceptual groups was analyzed using an ANCOVA with MA as a covariate and Group (ASD vs. TD) as a between subjects variable, neither MA nor Group played a significant role in the number of perceptual groups participants formed (MA $p = .202$, Group $p = .792$). When the ANCOVA was repeated with CA as a covariate, again, neither group nor CA played a significant role in the number of groups formed (CA $p = .205$, Group $p = .309$).

**Number of justifications.** Overall, the participants gave an average of 3.02 justifications for their perceptual groups ($SE = .47$), with the ASD group giving 2.69 justifications ($SE = .63$) and the TD participants giving 3.38 justifications ($SE = .69$). As depicted in Figure 11, linear regression analyses revealed that MA was a significant predictor of the number of perceptual justification made by ASD participants ($r = .64, p = .001$), but not the number formed by TD participants. In contrast, CA was not a significant predictor for either ASD or TD participants.
Figure 11. Age and number of perceptual justifications in typically developing and Autistic Spectrum Disorder (ASD) participants.

An ANCOVA with MA as a covariate and Group (ASD vs. TD) as a between subjects variable revealed that neither MA nor Group played a significant role in the number of perceptual justifications participants provided (MA $p = .071$, Group $p = .348$). When the ANCOVA was repeated with CA as a covariate, again, neither CA nor group played a significant role in the number of perceptual justifications provided (CA $p = .501$, Group $p = .300$).

**Switching grouping dimensions.** We were particularly interested in the relationship between the ability to switch grouping dimensions (e.g., grouping first by color, and then by shape) because some researchers have argued that cognitive flexibility
plays a key role in the thematic-to-taxonomic shift (see introduction section), and the ability to switch the dimensions by which one groups can be taken as a sign of cognitive flexibility (see Methods section).

Overall, 36 out of 50 participants (72%) grouped by two or more dimensions with 14 out of 24 ASD participants (58.33%), and 22 out of 26 TD participants (84.62%) switching grouping dimensions. As shown in Figure 12, for the TD children, there was a linear relationship between MA and switching grouping dimensions, with all of the TD children with MAs above 4 switching perceptual grouping dimension. In contrast, for the ASD participants, the relationship between MA and switching grouping dimensions appeared non-monotonic.

A multiple logistic regression analysis of the perceptual group switching data with the independent variables MA and Group (ASD/TD) was performed. The overall model was only a fair fit to the data (Log Likelihood = -25.99, $\chi^2(2) = 7.33$, $p = .026$, Nagelkerke’s $R^2 = .20$), and neither Group nor MA was a significant independent
predictor of switching group dimensions in the Perceptual task (See Table 3).

Furthermore, the null model (with no independent variables) correctly predicted 72% of cases and the model with 2 variables (MA and Group) predicted 68% of cases correctly. A log ratio test confirmed that the null model was as good a fit to the data as the more complex 2 variable model.

Table 3. Logistic Regression Analysis of Perceptual Group Switching Data with MA and Group as Independent Variables.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>β</th>
<th>Wald’s $\chi^2$</th>
<th>df</th>
<th>p</th>
<th>Odds Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>.23</td>
<td>1.03</td>
<td>1</td>
<td>.825</td>
<td>1.26</td>
</tr>
<tr>
<td>MA</td>
<td>.27</td>
<td>2.56</td>
<td>1</td>
<td>.109</td>
<td>1.31</td>
</tr>
<tr>
<td>Group (1 = ASD, 0 = TD)</td>
<td>-1.49</td>
<td>4.45</td>
<td>1</td>
<td>.035</td>
<td>.23</td>
</tr>
</tbody>
</table>

Overall Model Evaluation

<table>
<thead>
<tr>
<th>Log Likelihood</th>
<th>$\chi^2$</th>
<th>df</th>
<th>p</th>
<th>Nagelkerke’s $R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>25.99</td>
<td>7.33</td>
<td>.026</td>
<td>.20</td>
</tr>
</tbody>
</table>

As shown in Table 4, when CA was used rather than MA, the overall fit of the model to the data was only fair, with Group being a significant independent predictor of switching dimensions in perceptual groups ($p = .008$), but not CA (See Table 4). However, a log ratio test revealed that the model with CA and Group was not a significantly better model than the null model.
Table 4. Logistic Regression Analysis of Perceptual Group Switching Data with CA and Group as Independent Variables.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>β</th>
<th>Wald’s $\chi^2$</th>
<th>df</th>
<th>$p$</th>
<th>Odds Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>.70</td>
<td>.87</td>
<td>1</td>
<td>.350</td>
<td>2.01</td>
</tr>
<tr>
<td>CA</td>
<td>.00</td>
<td>3.52</td>
<td>1</td>
<td>.061</td>
<td>1.00</td>
</tr>
<tr>
<td>Group (1 = ASD, 0 = TD)</td>
<td>-2.73</td>
<td>6.97</td>
<td>1</td>
<td>.008</td>
<td>.07</td>
</tr>
</tbody>
</table>

Overall Model Evaluation

<table>
<thead>
<tr>
<th>Log Likelihood</th>
<th>$\chi^2$</th>
<th>df</th>
<th>$p$</th>
<th>Nagelkerke’s $R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-25.36</td>
<td>8.59</td>
<td>2</td>
<td>.014</td>
<td>.23</td>
</tr>
</tbody>
</table>

**Switching justification dimensions.** We also investigated whether Group or age predicted whether participants switched the dimensions by which they justified their groups. Overall, 22 out of the 50 participants (44%) switched justification dimensions, with 7 out of 24 ASD participants (29.17%) and 15 out of 26 TD participants (57.69%) doing so. As depicted in Fig 13, as TD participants’ MAs increased, there was a linear increase in switching justification dimensions, with all of the TD participants with MAs of 6 or above switching dimensions. In contrast, for the ASD participants, the relationship between MA and switching justification dimensions appeared non-monotonic.
Figure 13. Age and justification-switching on the perceptual task in typically developing and Autistic Spectrum Disorder (ASD) participants.

A multiple logistic regression analysis of the perceptual task justification-switching data with MA and Group (ASD/TD) as variables was conducted. As shown in Table 5, the overall model was a very good fit to the data, with MA being a significant independent predictor of justification-switching on the perceptual task and Group being a marginally significant independent predictor.

Table 5. Logistic Regression Analysis of Perceptual Justification-Switching Data with MA and Group as Independent Variables.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>β</th>
<th>Wald’s χ²</th>
<th>df</th>
<th>p</th>
<th>Odds Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-2.94</td>
<td>.05</td>
<td>1</td>
<td>.825</td>
<td>1.26</td>
</tr>
<tr>
<td>MA</td>
<td>.59</td>
<td>9.10</td>
<td>1</td>
<td>.003</td>
<td>1.31</td>
</tr>
<tr>
<td>Group (1 = ASD, 0 = TD)</td>
<td>-1.89</td>
<td>5.69</td>
<td>1</td>
<td>.017</td>
<td>.23</td>
</tr>
</tbody>
</table>

Overall Model Evaluation

<table>
<thead>
<tr>
<th>Log Likelihood</th>
<th>χ²</th>
<th>df</th>
<th>p</th>
<th>Nagelkerke’s R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>-25.64</td>
<td>17.31</td>
<td>2</td>
<td>.000</td>
<td>.39</td>
</tr>
</tbody>
</table>
A multiple logistic regression analysis of the perceptual task justification-switching data with CA and Group (ASD/TD) as variables was then conducted. As shown in Table 6, the overall model was again a good fit to the data, with both ASD and CA being significant independent predictors of justification-switching on the perceptual task.

**Table 6.** Logistic Regression Analysis of Perceptual Justification Switching Data with CA and Group as Independent Variables.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>β</th>
<th>Wald’s $\chi^2$</th>
<th>df</th>
<th>$p$</th>
<th>Odds Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-1.54</td>
<td>3.46</td>
<td>1</td>
<td>.006</td>
<td>.22</td>
</tr>
<tr>
<td>CA</td>
<td>.001</td>
<td>6.13</td>
<td>1</td>
<td>.013</td>
<td>1.00</td>
</tr>
<tr>
<td>Group (1 = ASD, 0 = TD)</td>
<td>-4.09</td>
<td>7.68</td>
<td>1</td>
<td>.006</td>
<td>.02</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Overall Model Evaluation</th>
<th>Log Likelihood</th>
<th>$\chi^2$</th>
<th>df</th>
<th>$p$</th>
<th>Nagelkerke’s $R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-27.56</td>
<td>13.47</td>
<td>2</td>
<td>.001</td>
<td>.32</td>
</tr>
</tbody>
</table>

**Basic Level Task**

**Basic level grouping.** Overall, participants formed an average of 3.66 taxonomic groups out of a possible 5 ($SE = .29$), with ASD participants forming 3.42 taxonomic groups ($SE = .45$) and TD participants forming 3.88 taxonomic groups ($SE = .38$). As depicted in Figure 14, neither MA nor CA was a significant predictor of the number of basic object level taxonomic groups formed by either ASD or TD children (all $p$’s > .05).
Figure 14. Age and number of taxonomic sorts on the basic level task in typically developing and Autistic Spectrum Disorder (ASD) participants.

The number of correct taxonomic groupings was analyzed using an ANCOVA with MA as a covariate and Group (ASD vs. TD) as a between subjects variable. This analysis revealed that neither MA nor Group played a significant role in the number of taxonomic groups participants formed (MA $p = .275$, Group $p = .391$). Basic level taxonomic groupings were then analyzed using an ANCOVA with CA as a covariate and Group (ASD vs. TD) as a between subjects variable. This analysis revealed marginally
significant effects of both CA ($F(1,47) = 5.64, p = .022$) and Group ($F(1,47) = 5.09, p = .029$) on the number of taxonomic groups formed.

**Basic level justifications.** Overall, participants gave an average of 3.25 out of 5 possible taxonomic justifications ($SE = .32$), with ASD participants giving 3.25 justifications ($SE = .46$) and TD participants giving 3.23 taxonomic justifications ($SE = .46$). As depicted in Figure 15, MA was not a significant predictor of the number of basic object level taxonomic justifications for either the ASD or TD participants (both $p$’s $>.05$). CA was a marginally significant predictor of justifications for the ASD participants ($r = .43, p = .037$), but not for the TD participants ($p > .10$).

![Fig 15A. Typically developing participants](image1)

![Fig 15B. ASD participants](image2)

![Fig 15C. Typically developing participants](image3)

![Fig 15D. ASD participants](image4)
**Figure 15.** Age and taxonomic justification on the basic level task in typically developing and Autistic Spectrum Disorder (ASD) participants.

The number of taxonomic justifications was analyzed using an ANCOVA with MA as a covariate and Group (ASD vs. TD) as a between subjects variable. This analysis revealed that participants with higher MAs tended to give more correct justifications ($F(1,47) = 5.42, p = .024$), but Group did not play a significant role ($F(1,47) = .02, p = .90$). When the ANCOVA was repeated with CA as the covariate, there was a significant effect of CA, with older participants correctly justifying more than younger participants ($F(1,47) = 7.14, p = .010$), but Group did not play a significant role ($F(1,47) = 3.41, p = .071$).

**Superordinate Level Task**

**Grouping.** Overall, 24 out of 50 participants (48%) separated clothing and food into separate groups, with 10 of the 24 ASD participants (41.67%) doing so and 14 of the 26 TD participants (53.85%) doing so. As depicted in Figure 16, as TD’s participants’ MAs increased, there was a linear increase in superordinate taxonomic grouping, with all of the TD participants with MAs of 5 or above grouping taxonomically. In contrast, for the ASD participants, the relationship between MA and superordinate taxonomic grouping appeared non-monotonic prior to MA 6, after which all of the ASD participants grouped taxonomically.
Fig16A. Typically developing participants  
Fig16B. ASD participants

Figure 16. MA and taxonomic grouping on the superordinate level task in typically developing and Autistic Spectrum Disorder (ASD) participants.

A multiple logistic regression analysis of the taxonomic grouping data with MA and Group (ASD/TD) as independent variables was performed. The overall model was a very good fit for the data, with MA being a significant independent predictor of superordinate taxonomic grouping, whereas Group was not (see Table 7).

Table 7. Logistic Regression Analysis of Taxonomic Sorting on the Superordinate Level Task with MA and Group as variables

<table>
<thead>
<tr>
<th>Predictor</th>
<th>β</th>
<th>Wald’s χ²</th>
<th>df</th>
<th>p</th>
<th>Odds Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-7.41</td>
<td>6.90</td>
<td>1</td>
<td>.009</td>
<td>.001</td>
</tr>
<tr>
<td>MA</td>
<td>1.88</td>
<td>8.56</td>
<td>1</td>
<td>.003</td>
<td>6.56</td>
</tr>
<tr>
<td>Group (1 = ASD, 0 = TD)</td>
<td>-1.59</td>
<td>2.82</td>
<td>1</td>
<td>.093</td>
<td>.21</td>
</tr>
</tbody>
</table>

Overall Model Evaluation

<table>
<thead>
<tr>
<th>Log Likelihood</th>
<th>χ²</th>
<th>df</th>
<th>p</th>
<th>Nagelkerke’s R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>-14.52</td>
<td>32.05</td>
<td>2</td>
<td>.00</td>
<td>.67</td>
</tr>
</tbody>
</table>
When the logistic regression analysis was repeated using CA and Group (ASD/TD) as independent variables, the log ratio test revealed that the model with CA and Group was no better than the null model, and neither CA nor Group was a significant independent predictor of taxonomic grouping (see Table 8).

**Table 8.** Logistic Regression Analysis of Taxonomic Grouping on the Superordinate Level Task with CA and Group as Independent Variables.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>β</th>
<th>Wald’s χ²</th>
<th>df</th>
<th>p</th>
<th>Odds Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>.984</td>
<td>2.08</td>
<td>1</td>
<td>.149</td>
<td>2.67</td>
</tr>
<tr>
<td>CA</td>
<td>.000</td>
<td>.90</td>
<td>1</td>
<td>.343</td>
<td>1.00</td>
</tr>
<tr>
<td>Group (1 = ASD, 0 = TD)</td>
<td>-1.68</td>
<td>3.50</td>
<td>1</td>
<td>.061</td>
<td>.19</td>
</tr>
</tbody>
</table>

**Overall Model Evaluation**

<table>
<thead>
<tr>
<th>Log Likelihood</th>
<th>χ²</th>
<th>df</th>
<th>p</th>
<th>Nagelkerke’s R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>-14.05</td>
<td>8.61</td>
<td>2</td>
<td>.01</td>
<td>.30</td>
</tr>
</tbody>
</table>

**Justifications.** Overall, 6 out of 50 participants (12%) justified thematically in the superordinate level task, with 6 out of 26 of TD participants (23.08%) doing so and none of the ASD participants doing so. A multiple logistic regression analysis of the superordinate thematic justification data was conducted with MA and Group (ASD/TD) as IVs. A log ratio test revealed that the model with MA and Group did not fit the data better than the null model, and neither MA nor Group was a significant independent predictor of thematic justification on the superordinate-level task (see Table 9).
Table 9. Logistic Regression Analysis of Superordinate Level Thematic Justification data with MA and Group as Independent Variables.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>β</th>
<th>Wald’s $\chi^2$</th>
<th>df</th>
<th>p</th>
<th>Odds Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-1.13</td>
<td>.631</td>
<td>1</td>
<td>.428</td>
<td>.32</td>
</tr>
<tr>
<td>MA</td>
<td>- .01</td>
<td>.003</td>
<td>1</td>
<td>.955</td>
<td>.99</td>
</tr>
<tr>
<td>Group (1 = ASD, 0 = TD)</td>
<td>-20.00</td>
<td>.00</td>
<td>1</td>
<td>.998</td>
<td>.00</td>
</tr>
</tbody>
</table>

Overall Model Evaluation

<table>
<thead>
<tr>
<th>Log Likelihood</th>
<th>$\chi^2$</th>
<th>df</th>
<th>p</th>
<th>Nagelkerke’s $R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-14.05</td>
<td>8.61</td>
<td>2</td>
<td>.014</td>
<td>.30</td>
</tr>
</tbody>
</table>

The multiple logistic regression analysis was repeated with CA and Group as variables. Again, a log ratio test revealed that the model with CA and Group was not a better fit of the superordinate thematic justification data than the null model was, and neither Group nor CA was a good independent predictor of thematic justification on the superordinate level task (See Table 10).

Table 10. Logistic Regression Analysis of Superordinate Level Thematic Justification data with CA and Group as variables.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>β</th>
<th>Wald’s $\chi^2$</th>
<th>df</th>
<th>p</th>
<th>Odds Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-1.26</td>
<td>1.11</td>
<td>1</td>
<td>.29</td>
<td>.28</td>
</tr>
<tr>
<td>CA</td>
<td>.00</td>
<td>.003</td>
<td>1</td>
<td>.96</td>
<td>1.00</td>
</tr>
<tr>
<td>Group (1 = ASD, 0 = TD)</td>
<td>-20.07</td>
<td>.00</td>
<td>1</td>
<td>.99</td>
<td>.00</td>
</tr>
</tbody>
</table>

Overall Model Evaluation

<table>
<thead>
<tr>
<th>Log Likelihood</th>
<th>$\chi^2$</th>
<th>df</th>
<th>p</th>
<th>Nagelkerke’s $R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-14.05</td>
<td>8.61</td>
<td>2</td>
<td>.01</td>
<td>.30</td>
</tr>
</tbody>
</table>
Overall, 28 out of 50 participants (56%) justified taxonomically on the superordinate level task, with 12 out of 24 ASD participants (50%) and 16 out of 26 TD participants (51.54%) doing so. As depicted in Figure 17, for the TD participants, there was a non-monotonic relationship between MA and superordinate taxonomic justifications, whereas essentially no ASD participants with MAs of less than 6 provided superordinate taxonomic justification and all of the ASD participants with MAs of 6 or greater provided superordinate taxonomic justifications.

**Fig 17A.** Typically developing participants  
**Fig 17B.** ASD participants

**Fig 17.** Age and taxonomic justification on the superordinate level task in typically developing and Autistic Spectrum Disorder (ASD) participants.

A multiple logistic regression analysis of taxonomic justification was conducted with MA and Group (ASD/TD) as independent variables. As shown in Table 11, the overall model was an excellent fit to the data, with only MA being a good independent predictor of taxonomic justification.
**Table 11.** Logistic regression analysis of superordinate level taxonomic justification data with MA and Group as independent variables.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>β</th>
<th>Wald’s χ²</th>
<th>df</th>
<th>p</th>
<th>Odds Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-6.59</td>
<td>10.40</td>
<td>1</td>
<td>.001</td>
<td>.001</td>
</tr>
<tr>
<td>MA</td>
<td>1.36</td>
<td>11.41</td>
<td>1</td>
<td>.001</td>
<td>3.90</td>
</tr>
<tr>
<td>Group (1 = ASD, 0 = TD)</td>
<td>-.94</td>
<td>1.17</td>
<td>1</td>
<td>.280</td>
<td>.39</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Overall Model Evaluation</th>
<th>Log Likelihood</th>
<th>χ²</th>
<th>df</th>
<th>p</th>
<th>Nagelkerke’s R²</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-18.24</td>
<td>32.14</td>
<td>2</td>
<td>.000</td>
<td>.64</td>
</tr>
</tbody>
</table>

When the analysis was repeated with CA rather than MA, the overall model was not as good a fit, with neither average CA nor Group being good independent predictors of taxonomic justification on the superordinate level task (see Table 12).

**Table 12.** Logistic Regression Analysis of taxonomic justification on the superordinate level task with CA and Group as independent variables.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>β</th>
<th>Wald’s χ²</th>
<th>df</th>
<th>p</th>
<th>Odds Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-.724</td>
<td>1.15</td>
<td>1</td>
<td>.284</td>
<td>.49</td>
</tr>
<tr>
<td>CA</td>
<td>.00</td>
<td>4.66</td>
<td>1</td>
<td>.031</td>
<td>1.00</td>
</tr>
<tr>
<td>Group (1 = ASD, 0 = TD)</td>
<td>-2.07</td>
<td>4.33</td>
<td>1</td>
<td>.037</td>
<td>.13</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Overall Model Evaluation</th>
<th>Log Likelihood</th>
<th>χ²</th>
<th>df</th>
<th>p</th>
<th>Nagelkerke’s R²</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-31.02</td>
<td>6.56</td>
<td>2</td>
<td>.038</td>
<td>.17</td>
</tr>
</tbody>
</table>
Intermediate Level Task (animal groups).

**Thematic justifications.** Only 3 out of 24 ASD participants (12.5%) provided thematic justifications of animal groups, and only 2 out of 26 TD participants (7.69%) did so. Inspection of Figure 17 does not reveal any clear pattern between MA and thematic justifications by either ASD or TD participants.

![Fig 17A. Typically developing participants](image1)

![Fig 17B. ASD Participants](image2)

**Figure 17.** Mental age and thematic justification on the intermediate level task in typically developing and Autistic Spectrum Disorder (ASD) participants.

A multiple logistic regression analysis of the thematic justification data was conducted with MA and Group (ASD/TD) as independent variables. A Log ratio test revealed that the overall model was not a better fit to the data than the null model, and neither Group nor MA was a good independent predictor of thematic justification on the intermediate-level task (see Table 13).
Table 13. Logistic Regression Analysis of Thematic Justification data on the Intermediate Level task with MA and Group as Independent Variables.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>β</th>
<th>Wald’s $\chi^2$</th>
<th>df</th>
<th>p</th>
<th>Odds Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-2.93</td>
<td>4.38</td>
<td>1</td>
<td>.036</td>
<td>.05</td>
</tr>
<tr>
<td>MA</td>
<td>.08</td>
<td>.15</td>
<td>1</td>
<td>.703</td>
<td>1.08</td>
</tr>
<tr>
<td>Group (1 = ASD, 0 = TD)</td>
<td>.51</td>
<td>.28</td>
<td>1</td>
<td>.594</td>
<td>1.67</td>
</tr>
</tbody>
</table>

Overall Model Evaluation

<table>
<thead>
<tr>
<th>Log Likelihood</th>
<th>$\chi^2$</th>
<th>df</th>
<th>p</th>
<th>Nagelkerke’s $R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-16.03</td>
<td>.46</td>
<td>2</td>
<td>.793</td>
<td>.02</td>
</tr>
</tbody>
</table>

Similar results were obtained when the multiple logistic regression analysis was repeated with CA rather than MA. Again, a log ratio test revealed that the overall model was not a better fit for the data than the null model, and neither Group nor CA was a good independent predictor of thematic justification (see Table 14).

Table 14. Logistic regression analysis of thematic justification data on the intermediate level task with CA and Group as independent variables.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>β</th>
<th>Wald’s $\chi^2$</th>
<th>df</th>
<th>p</th>
<th>Odds Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-3.9</td>
<td>8.98</td>
<td>1</td>
<td>.003</td>
<td>.02</td>
</tr>
<tr>
<td>CA</td>
<td>.001</td>
<td>2.14</td>
<td>1</td>
<td>.144</td>
<td>1.00</td>
</tr>
<tr>
<td>Group (1 = ASD, 0 = TD)</td>
<td>-1.26</td>
<td>.58</td>
<td>1</td>
<td>.445</td>
<td>.29</td>
</tr>
</tbody>
</table>

Overall Model Evaluation

<table>
<thead>
<tr>
<th>Log Likelihood</th>
<th>$\chi^2$</th>
<th>df</th>
<th>p</th>
<th>Nagelkerke’s $R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-14.85</td>
<td>2.85</td>
<td>2</td>
<td>.240</td>
<td>.12</td>
</tr>
</tbody>
</table>
**Perceptual justifications.** Overall, 19 out of 50 participants provided perceptual justifications for their intermediate animal groupings, with 5 out of 24 ASD participants (20.83%) doing so and 14 out of 26 TD participants (53.85%) doing so. As depicted in Figure 18, among the TD participants, there was a linear increase in perceptual justifications up to MA of 6, followed by a dip and then a return to 100%. For the ASD participants, there wasn’t a readily apparent relationship between MA and providing perceptual justifications.

![Figure 18A](image1.png)  ![Figure 18B](image2.png)

**Figure 18.** Age and perceptual justification on the intermediate level task in typically developing and Autistic Spectrum Disorder (ASD) participants.

A multiple logistic regression analysis of the perceptual justification data was conducted with MA and Group (ASD/TD) as independent variables. The overall model was only a fair fit to the data, with group approaching significance as an independent predictor of perceptual justification ($p = .02$), and MA having no effect (See Table 15).

**Table 15.** Logistic Regression Analysis of Intermediate Level Perceptual Justification data with MA and Group as Independent Variables.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>$\beta$</th>
<th>Wald’s $\chi^2$</th>
<th>$df$</th>
<th>$p$</th>
<th>Odds Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-.19</td>
<td>.05</td>
<td>1</td>
<td>.831</td>
<td>.82</td>
</tr>
</tbody>
</table>
When the analysis was repeated with CA rather than MA, the overall model was again only a fair fit to the data, but now neither Group nor CA was a good independent predictor of perceptual justification (see Table 16).

### Table 16. Logistic Regression Analysis of Intermediate Level Perceptual Justification data with CA and Group as Independent Variables.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>β</th>
<th>Wald’s χ²</th>
<th>df</th>
<th>p</th>
<th>Odds Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>.47</td>
<td>.53</td>
<td>1</td>
<td>.47</td>
<td>1.60</td>
</tr>
<tr>
<td>CA</td>
<td>.00</td>
<td>.38</td>
<td>1</td>
<td>.54</td>
<td>1.00</td>
</tr>
<tr>
<td>Group (1 = ASD, 0 = TD)</td>
<td>-1.11</td>
<td>1.67</td>
<td>1</td>
<td>.20</td>
<td>.33</td>
</tr>
</tbody>
</table>

When the analysis was repeated with CA rather than MA, the overall model was again only a fair fit to the data, but now neither Group nor CA was a good independent predictor of perceptual justification (see Table 16).

### Taxonomic Justifications. Overall, 24 out of 50 participants (48%) provided taxonomic justifications for groups, with 10 out of 24 ASD participants (41.66%) doing so, and 14 out of 26 TD participants (53.85%) doing so. As depicted in Figure 19, about half of the TD participants with MAs of between 4 and 6 taxonomically justified, whereas virtually none of the ASD participants with MAs of between 4 and 6 did. All of
the participants with MAs of 7 or above taxonomically justified, regardless of whether they were TD or ASD.

**Figure 19.** Age and taxonomic justification on the intermediate level task in typically developing and Autistic Spectrum Disorder (ASD) participants.

A multiple logistic regression analysis of the intermediate taxonomic justification data was conducted with MA and Group (ASD/TD) as independent variables. As shown in Table 17, the overall model was a very good fit to the data, but only MA was a good independent predictor of taxonomic justification ($p = .000$)

**Table 17.** Logistic regression analysis of intermediate level taxonomic justification data with MA and Group as independent variables.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>$\beta$</th>
<th>Wald’s $\chi^2$</th>
<th>$df$</th>
<th>$p$</th>
<th>Odds Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-6.21</td>
<td>12.54</td>
<td>1</td>
<td>.000</td>
<td>.002</td>
</tr>
<tr>
<td>MA</td>
<td>1.17</td>
<td>13.18</td>
<td>1</td>
<td>.000</td>
<td>3.21</td>
</tr>
<tr>
<td>Group (1 = ASD, 0 = TD)</td>
<td>-1.35</td>
<td>2.19</td>
<td>1</td>
<td>.139</td>
<td>.26</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Overall Model Evaluation</th>
<th>Log Likelihood $\chi^2$</th>
<th>$df$</th>
<th>$p$</th>
<th>Nagelkerke’s $R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-18.92</td>
<td>2</td>
<td>.000</td>
<td>.62</td>
</tr>
</tbody>
</table>

**Fig 19A.** Typically developing participants

**Fig 19B.** ASD participants
When the analysis was repeated with CA rather than MA as an independent variable, the overall model was only a fair fit to the data, with both Group and CA being marginally good independent predictors of thematic justification (see Table 18).

**Table 18.** Logistic Regression Analysis of Intermediate Level Taxonomic Justification data with CA and Group as Variables.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>β</th>
<th>Wald’s $\chi^2$</th>
<th>df</th>
<th>p</th>
<th>Odds Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-1.02</td>
<td>2.21</td>
<td>1</td>
<td>.137</td>
<td>.36</td>
</tr>
<tr>
<td>CA</td>
<td>.001</td>
<td>4.31</td>
<td>1</td>
<td>.038</td>
<td>1.00</td>
</tr>
<tr>
<td>Group (1 = ASD, 0 = TD)</td>
<td>-2.08</td>
<td>4.15</td>
<td>1</td>
<td>.042</td>
<td>.125</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Overall Model Evaluation</th>
<th>Log Likelihood</th>
<th>$\chi^2$</th>
<th>df</th>
<th>p</th>
<th>Nagelkerke’s $R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-31.55</td>
<td>6.13</td>
<td>2</td>
<td>.047</td>
<td>.15</td>
</tr>
</tbody>
</table>

As demonstrated in Figs 20 and 21, ASD participants exhibited a rather abrupt onset of taxonomic categorization ability on the intermediate level task, whereas typically developing appeared to gradually demonstrate taxonomic justification with age.

![Fig 20. MA and justification type on the intermediate level task in ASD participants.](image-url)
Finally, we attempted to uncover specific predictors of taxonomic justification on the intermediate level task that were significant predictors independent of MA and Group. We did this by running a series of multiple logistic regressions in which MA, Group, and one other variable were entered as independent variables, with intermediate taxonomic justification as the dependent variable. In this way, we tested a total of 14 variables: animal declarative knowledge scores, animal vocabulary scores, metamemory scores, MLU, justification on the perceptual task, # of perceptual groups created, # of perceptual justifications, taxonomic grouping on the superordinate task, and taxonomic justification on the superordinate task. In all cases, MA was the only significant predictor.

**Grouping vs. Justification**

In order to assess the relationship between taxonomic grouping and taxonomic justification, we compared the grouping and justification results from the Basic level task.
and the Superordinate level task. Note: the intermediate level task was excluded from this section because there were no taxonomic groups determined a priori.

**Basic level task.** Overall, 41 out of 50 participants (82%) were able to form at least one taxonomic group on the basic level task, and 36 out of 50 participants (72%) provided at least one taxonomic justification on the basic level task.

**Superordinate level task.** Overall, 35 out of 50 participants (70%) taxonomically grouped the superordinate level stimuli and 28 out of 50 participants (56%) provided taxonomic justifications for their superordinate groupings.
Recall that the primary goals of this study were to investigate:

1) The nature of categories in autistic and typically developing children.

2) The development of categories in autistic and typically developing children; specifically, whether autistic children undergo a shift from event-based thematic categories to adult-like taxonomic categories.

3) Whether and how characteristic features of autism contribute to any differences that are found in the nature or development of categorization in autism.

We begin by discussing how the results of this study shed light on proposals about what drives the developmental shift from thematic to taxonomic categories. Specifically, we will address how age, level of categorization, declarative knowledge, cognitive flexibility and issues such as grouping vs. justification affect the thematic-to-taxonomic shift. We then shift our focus to categorization ability in autism and consider several interpretations of our findings. Finally, we discuss the implications of our ASD data with respect to the mechanisms underlying the thematic-to-taxonomic shift.

**What Drives the Thematic-to-Taxonomic Shift?**

**Age.** To date, the majority of research conducted on the thematic-to-taxonomic shift suggests that children’s categories become increasingly taxonomic as they get older. The period from infancy to early elementary school is a period during which typically developing children undergo profound development in many cognitive and non-cognitive domains. One of the goals of this study was to tease apart the specific factors associated with getting older that drive the thematic-to-taxonomic shift. Consistent with the results of previous studies that have shown that advancing chronological age is highly associated
with the shift from thematic to taxonomic categorization, we found that our participants’ mental ages were strong predictors of whether they grouped and justified these groupings taxonomically. That said, we failed to uncover any specific aspect(s) of linguistic or cognitive development that were significant predictors of taxonomic categorization independent of mental age.

To review, MA had a significant effect on most measures of performance in each of our experimental tasks except for the number of groups created on the perceptual task, the number of justifications provided on the perceptual task, switching-dimensions in groups on the perceptual task, the number of basic level taxonomic groups created, perceptual justification on the intermediate level, and thematic justification on the superordinate and intermediate levels.

Specifically, participants with higher MAs knew the names of more animals, knew more facts about animals, had larger MLUs, and had greater metamemory awareness than children with lower MAs. In the grouping and justification tasks, children with greater MAs also formed more taxonomic groups in the superordinate level task, justified more of their groupings in the perceptual task, were more likely to switch justification-dimensions on the perceptual task, and provided more taxonomic justifications for groups on the basic, superordinate, and intermediate level tasks. Of particular interest, participants with greater MAs exhibited greater cognitive flexibility as measured by the ability to switch dimensions that they used to justify in the perceptual task.

**Thematic Justification.** Most studies suggest that the shift from thematic to taxonomic categories happens at or around the age of 6 or 7. Surprisingly, despite the fact
that almost three fourths of the participants in our study were under the age of 6, very few of our participants justified thematically on any of the categorization tasks. Although the failure of most of our subjects to justify thematically is intriguing, the paucity of thematic justification data in our study limits our ability to draw conclusions about the effect of MA on thematic categorization.

Age, Taxonomic Justification, and Level of Categorization. Participants with greater MAs were more likely to group taxonomically on the superordinate level (i.e., grouping items of food and clothing separately), and to justify taxonomically their groupings on the basic object level (e.g., CATS vs. DOGS), intermediate level (e.g., PETS, ZOO ANIMALS) and superordinate level (CLOTHING vs. FOOD). All but nine of the participants formed taxonomic groups at the basic object level, and MA had no effect on the number of taxonomic groups or justifications at the basic level in either the ASD or typically developing group. This suggests that taxonomic categories develop at a younger age at the basic-object level than at the superordinate or intermediate levels. This finding is consistent with past research demonstrating that autistic children have more difficulty grouping objects that share the same superordinate level than objects that share the same basic-object level (Tager-Flusberg, 1985) and other studies that have shown that it is easier for typically developing infants to group items from the same basic-object level than items from the same superordinate level (e.g., Daehler et al., 1979; Rosch, et al., 1976).

Knowledge is not enough. Some researchers have suggested that as children learn more about a topic, their categories shift from being thematic to taxonomic (see

7 One possibility is that the ability to justify one’s actions requires fairly advanced metacognitive skills and, thus, by the time that children are old enough to justify why they grouped in certain ways, they have already undergone the thematic-to-taxonomic shift.
Introduction). Our results suggest that, at least for intermediate level categories like FARM ANIMAL, knowledge is not sufficient for taxonomic categorization. In our study, all but one typically developing participant correctly answered at least half of the questions on the Animal Knowledge Task, yet only half gave taxonomic justifications for their animal groups. Similarly, three-quarters of our ASD participants correctly answered at least half of the Animal Knowledge Task questions, and yet only about a third gave taxonomic justifications on the intermediate level task. Furthermore, animal knowledge scores were not a significant predictor of taxonomic justification independent of mental age.

**Grouping Versus justification.** One possible concern is that our taxonomic justification measure may underestimate whether children have taxonomic categories. In the two tasks where we had *a priori* taxonomic categories (FOOD vs CLOTHING in the superordinate task and DOG vs CAT, etc. in the basic level task), more participants grouped taxonomically than justified taxonomically. Because around three quarters of participants were able to taxonomically group the superordinate level stimuli and only around half provided taxonomic justifications for the superordinate level stimuli, we can infer that participants’ justifications may underestimate their ability to categorize taxonomically. This gap between children’s taxonomic groupings and justifications has important implications for research on categorization because it is often not possible to define *a priori* what the possible thematic or taxonomic categories are, especially at the intermediate level. Without *a priori* categories, researchers are forced to rely on children’s justifications to determine whether children’s groupings are thematic or taxonomic.
Is Categorization Atypical in Autism?

Despite the fact that the ASD participants were less cognitively flexible, had less declarative knowledge about animals and less metamemory skills, and tended to be less linguistically advanced (with marginally smaller MLUs and animal vocabularies) than our 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.

Perhaps we didn’t find differences between our ASD and MA-matched typically developing children’s categorization abilities because ASD individuals’ categories are the same as typically developing children with similar IQs. On its face, this explanation seems unlikely because many categorization studies have demonstrated clear differences between ASD and typically developing children’s categorization ability. However, a more careful read of the literature reveals that many of the categorization studies which have demonstrated differences between typically developing and ASD children have examined ASD children’s ability to acquire new categories (e.g., Ohta, 1987; Minshew et al, 2002; Klinger & Dawson, 2001). In contrast to these studies, in our study, the ASD children were tested on familiar categories (food, clothing, animals, colors, shape, etc.) and familiar objects. Thus, the results of our study may highlight an important distinction between the acquisition of categories and their eventual structure. In our literature review, we only found one other study that investigated categorization structure with familiar objects in ASD children (Tager-Flusberg, 1985). Consistent with the results of our study, Tager-Flusberg (1985) failed to find differences between ASD participants and MA matched typically developing children (Tager-Flusberg, 1985).
As discussed in the introduction, children’s developing metamemory abilities have been proposed as one possible reason for the developmental shift from thematic-to-taxonomic categories. However, in our study, we found very few differences in our ASD and mental age-matched TD participants’ categorization, yet our ASD participants did significantly worse than our TD participants on Flavell’s 1977 classic metamemory task (see Appendix B).

Before concluding that metamemory plays no role in category formation, let us examine more closely Flavell’s task. Recall that Flavell’s task measures a child’s knowledge of his or her own memory by asking about the memory abilities of two puppets (“Henry” and “Megan”). Thus, it requires that the child engage in pretend play (treating puppets as if they were people). Flavell’s task also requires that the child assume that other people have minds that are similar to the child’s. In other words, doing well on Flavell’s metamemory task requires having a well-developed ability to attribute mental states to one’s self and to others, (i.e., Theory of Mind). Thus, Flavell’s task may have underestimated our ASD participants’ metamemory skills because deficits in pretend play (e.g., Jarrold, 2003) and deficits in Theory of Mind (e.g., Happe, 1995) are hallmarks of autism. The few studies that have been conducted on ASD and metamemory (e.g., Wojcik et al., 2011; Wojcik et al., 2013) suggest that ASD individuals may not have impairments in metamemory.

A third factor that may have contributed to our failure to find differences between ASD and typically developing children’s categorization is that, by matching our ASD and typically developing participants on the basis of mental age, we were matching them for the critical skills needed for categorization. Indeed, the KBIT-2 tests some of the same
cognitive abilities as the categorization tasks in this study. For example, in the KBIT-2’s Verbal Knowledge subtest, the examiner places 6 pictures in front of the client and asks a question or says a word, and the client responds by pointing to the picture that best depicts the word or answers the question. Thus, the KBIT-2’s Verbal Knowledge subtest is similar to our Animal Vocabulary and Animal Knowledge tests. The KBIT-2’s Riddles subtest, which is intended to measure verbal comprehension, reasoning, and vocabulary knowledge, requires the examinee to answer riddles. One could argue that this draws on many of the same skills as the test we used to calculate our participants’ MLUs.

The Nonverbal portion of the KBIT-2 consists of matrices in which the examiner asks the client to select which of five pictures goes best with a target picture. For some of the items, the target is a taxonomically- or thematically-related item. For example, in one item, the examinee is shown a picture of a car and must choose between a truck, a frying pan, a sun, fruit items, or a zipper (See Figure 22). The correct answer is truck, which is taxonomically related to “car” in that they are both vehicles. In another KBIT-2 item, the client is shown a ring, and must choose between a window, a hand, a sock, a glass of lemonade, or a shopping cart (See Fig. 23). Here, the correct answer, “hand”, is thematically related to “ring,” because a ring goes on a hand. Note that selecting the sock might be an example of a taxonomic categorization because both “ring” and “socks” are items that are worn. To the extent that the KBIT-2 nonverbal IQ scores reflect performance on taxonomic and thematically related items, it is not surprising that our ASD participants behaved like typically developing children who had similar scores on the KBIT-2.
In addition to measuring both taxonomic and thematic categorization ability, the matrices items described above likely test cognitive flexibility. Recall Blaye et al’s notion of “conceptual flexibility” (i.e., the ability to switch between multiple types of conceptual organizations). Items on the first two portions of the KBIT-2 such as those described above indirectly tap cognitive flexibility because in order to answer the questions correctly, an examinee must adjust his or her categorization depending on
whether a test item depicts a taxonomic (e.g., Fig 22) or thematic (e.g., Fig 23) relationship between the stimulus picture and the correct response.

The remaining nonverbal KBIT-2 subtest items directly assess cognitive flexibility. In these items, the client is shown incomplete matrices of abstract stimuli and must select the items that complete the matrices. According to the KBIT-2’s designers, each abstract item requires nonverbal reasoning and “flexibility in applying a problem-solving strategy” (Kaufman & Kaufman, 2004, p. 30). These KBIT-2 items are similar to the items in our “Perceptual Task” in that the examinee must isolate perceptual features of a given group of objects in order to solve the problem (See Figs 24, 25 and 26).

**Fig 24.** KBIT-2 Matrices Item 24.  
**Fig 25.** KBIT-2 Matrices Item 25.  
**Fig 26.** KBIT-2 Matrices Item 26.
Given the similarity between portions of the KBIT-2 and our tasks, it is possible we would have found greater differences between typical and ASD participants if we had used a different measure of mental age, or if we had matched ASD and typically developing participants by chronological age. Although we did analyze our data using both chronological age and mental age, our typically developing participants were purposefully selected to be good mental age matches for our ASD participants and, thus, they were chronologically much younger than our ASD participants.

A third factor that may have contributed to our failure to find differences in categorization is that, perhaps our ASD participants were not autistic or were too high functioning, and thus did not have deficits in knowledge, language, cognitive flexibility and metamemory that have been argued to play an important role in categorization. However, the fact that our ASD participants performed worse than typically developing children on our knowledge, language, cognitive flexibility and metamemory tasks, but performed the same as typically developing children on our categorization tasks suggests that this is not the case.

Despite the fact that our ASD participants generally did not perform quantitatively differently than our MA-matched TD participants on our categorization tasks, close examination of the data suggest there may be qualitative differences between groups in the developmentally trajectory of taxonomic categorization development. For example, typically developing children demonstrate a gradual increase in taxonomic grouping behavior on the superordinate level task until the mental age of five, at which point all participants were able to group food and clothing items separately (see Figure 16A). In contrast, there appears to be no trend toward taxonomic grouping behavior in
the ASD group prior to the mental age of six, at which point all participants formed taxonomic groups (Figure 16B).

A large proportion of the typically developing individuals with mental ages of less than 6 justified taxonomically at the superordinate level (Figure 17A), whereas ASD participants demonstrated little-to-no taxonomic justification behavior until the mental age of six, at which point all participants justified taxonomically (Figure 17B). On the intermediate level task, around half of the typically developing children justified taxonomically prior to mental age six, at which point all participants justified taxonomically (Figure 19A). In contrast, ASD participants provided comparatively fewer taxonomic justifications between the mental ages of three and six, but justified just as much as typically developing children starting at the mental age of seven (Figure 19B).

These results are consistent with the development of taxonomic justification and categorization following a relatively gradual trajectory in typically developing children. For the ASD participants, the onset of taxonomic justification appears to be much more abrupt, with little-to-no taxonomic grouping or justification by the younger mental age ASD participants, and virtually 100% taxonomic categorization and justification by the older mental age ASD participants. If this characterization is accurate, it might suggest that ASD children suffer from impairments in the use or acquisition of taxonomic categories until they reach a certain mental age, at which point their categorization is similar to that of typically developing children. Further research is necessary to determine whether there are indeed early differences in the developmental trajectories of taxonomic categories ASD and typically developing children.

**Future Directions**
Justification tasks (and to a lesser extent grouping tasks) require advanced meta-cognitive awareness and verbal skills. This is particularly problematic for children with autism, who often are impaired in both of these areas. Future research on categorization development using more implicit measures of categorization such as eye-tracking, event related potential and neuroimaging techniques may help to elucidate whether there are quantitative or qualitative differences between ASD and typically developing children’s categories, and what drives the shift from thematic to taxonomic categories in ASD and typically developing children.
References


Nguyen, S.P. & Murphy, G.L. (2003) An apple is more than just fruit: cross-classification


CONCEPT FORMATION IN AUTISM

Psychology. 40(3), 450-469.


**Appendix A. 20 Features for Animal Declarative Knowledge Task**

<table>
<thead>
<tr>
<th>Order 1</th>
<th>Order 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild</td>
<td>Cat</td>
</tr>
<tr>
<td>People ride on</td>
<td>Water</td>
</tr>
<tr>
<td>Big</td>
<td>Large</td>
</tr>
<tr>
<td>Disgusting</td>
<td>Scary</td>
</tr>
<tr>
<td>Pet</td>
<td>Have 4 legs</td>
</tr>
<tr>
<td>Mammal</td>
<td>Mammal</td>
</tr>
<tr>
<td>Africa</td>
<td>Ape</td>
</tr>
<tr>
<td>Cat</td>
<td>Africa</td>
</tr>
<tr>
<td>Water</td>
<td>Circus</td>
</tr>
<tr>
<td>Eat other animals</td>
<td>People ride on</td>
</tr>
<tr>
<td>Circus</td>
<td>Bird</td>
</tr>
<tr>
<td>Fly</td>
<td>Rodent</td>
</tr>
<tr>
<td>Bird</td>
<td>People eat</td>
</tr>
<tr>
<td>Rodent</td>
<td>Wild</td>
</tr>
<tr>
<td>Ape</td>
<td>Disgusting</td>
</tr>
<tr>
<td>Live in backyard</td>
<td>Fly</td>
</tr>
<tr>
<td>Scary</td>
<td>Farm</td>
</tr>
<tr>
<td>People eat</td>
<td>Pet</td>
</tr>
<tr>
<td>Farm</td>
<td>Live in backyard</td>
</tr>
<tr>
<td>Have 4 legs</td>
<td>Eat other animals</td>
</tr>
</tbody>
</table>
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 for 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 for Henry?

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

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