2013-09-13

Children’s Understanding of Antonymy

Phillips, Catherine

UNIVERSITY OF CALGARY

Children's Understanding of Antonymy

by

Catherine I. Phillips

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF PSYCHOLOGY

CALGARY, ALBERTA

AUGUST, 2013

Abstract

Understanding antonymy (the opposite meaning relationship between certain words) is an important conceptual development that likely supports, and is itself supported by, the acquisition of other skills. The purpose of the present research was three-fold: to determine the development of understanding for the concept of antonymy with a novel, less verbal task than used in previous research; to determine whether there are other cognitive developments related to the development of the concept of antonymy; and to evaluate whether there was evidence of a sensitivity to, or an implicit understanding of, the concept of antonymy prior to behavioural evidence of understanding.

Results showed that children's accuracy on the opposite appreciation task improved with age. Furthermore, 4- and 5-year-old children performed significantly above chance, whereas 3-year-old children did not. Children's performance was correlated with the context subscale of the CCC-2; children who performed better on the Opposite task also tended to be perceived as making use of context to understand language. Children whose parents reported having books or games that are focused on opposites were more likely to perform above chance than were those who did not have access to these materials in the home. Children's performance on the opposite task was not correlated with receptive vocabulary or working memory. Eye gaze, response latency and item differences on the opposite task did not provide evidence for latent understanding of the concept of antonymy. Overall, children were significantly more accurate for the word pair *big – small* than the word pair *awake – asleep*. However, 3-year-old children did not perform above chance on the word pair for which they

displayed the best performance. Eye gaze analyses did not provide any evidence for latent understanding of the concept of opposites in 3-year-old children.

These results suggest that 4- and 5-year-old children, but not 3-year-old children, show an appreciation for the antonymy relationship, and that this appreciation generalizes to a number of different antonym pairs. Furthermore, children demonstrate this appreciation only when the label “opposite” is used in the task, suggesting that antonymy is not a relationship made salient by stimulus properties alone.

Acknowledgements

In every job that must be done there is an element of fun. You find the fun, and – SNAP – the job's a game!
- Mary Poppins

This dissertation is the culmination of years of hard work. Without the wonderful support structure I have, I'm not confident that I would have been able to accomplish all that I have. There isn't sufficient space to individually thank all of the friends, family and colleagues who have helped me along the way and made the process much more fun.

There are, however, a few are worthy of particular note. I would like to thank my husband, Cory Phillips; my parents, Henry and Claudia Gross; and my supervisor, Dr. Penny Pexman.

Dedication

This work is dedicated to my Nonna Lucia Castellani.

Table of Contents

Abstract ii
Acknowledgements iv
Dedication v
Table of Contents vi
List of Tables vii
List of Figures ix
Children's Understanding of Antonymy 1
Learning the Concept of Opposite 4
Subcategorization of Antonyms 6
Acquisition of Antonymy 13
Present research 26
Method 32
Participants 32
Procedure 32
Opposite Task 33
Additional Measures 40
Results 43
Behaviour 43
Correlations & Demographics 46
Processing Measures: Response Latency and Eye Gaze 52
Discussion 64
Development of the Concept 64
Implicit Understanding 70
Limitations 75
Further Future Directions 79
Conclusions 83
References 85
Appendix A 95
Appendix B 96
Appendix C 99
Appendix D 102
Appendix E 103

List of Tables

Table 1. Children's Mean Proportions of Target Item Selections in Each Condition 43

Table 2. Children's Mean Proportions of Target Item Selections for Each Word Pair 47

Table 3. Children's Mean Scores on Additional Measures and Demographic Characteristics in the Opposite Label Condition 48

Table 4. Children's Mean Scores on Additional Measures and Demographic Characteristics in the Another One Label Condition 49

Table 5. Children's Mean Scores on Additional Measures and Demographic Characteristics in the No Label Condition 50

Table 6. Children's Mean Response Latencies (in ms) in the Opposite Label Condition on Trials Ending with a Correct Selection 54

Table 7. Children's Mean Response Latencies (in ms) in the Opposite Label Condition on Trials Ending with an Incorrect Selection 56

Table 8. Mean Number of Fixations and Mean number of Correct First Fixations in the Opposite Label Condition 57

Table 9. Mean Proportion of Time Spent Looking at the Target Image in the Opposite Label Condition on Trials Ending with a Correct Selection..58

Table 10. Mean Proportion of Time Spent Looking at the Target Image in the Opposite Label Condition on Trials Ending with an Incorrect Selection 59

Table 11. Mean Proportion of Time Spent Looking at the Distractor Image in the Opposite Label Condition on Trials Ending with a Correct Selection 60

Table 12. Mean Proportion of Time Spent Looking at the Distractor Image in the Opposite Label Condition on Trials Ending with an Incorrect Selection.........................61

Table 13. Mean Proportion of Time Spent Looking at the Experimenter in the Opposite Label Condition on Trials Ending with a Correct Selection..62

Table 14. Mean Proportion of Time Spent Looking at the Experimenter in the Opposite Label Condition on Trials Ending with an Incorrect Selection 63

Table 15. Mean Proportion of Time Spent Looking at the Image Selected by the Experimenter in the Opposite Label Condition on Trials Ending with a Correct Selection 63

Table 16. Mean Proportion of Time Spent Looking at the Image Selected by the Experimenter in the Opposite Label Condition on Trials Ending with an Incorrect Selection..64

List of Figures

Figure 1. Schematic of the testing situation..36

Children's Understanding of Antonymy

Understanding antonymy, or lexical opposition, is an important conceptual development that likely supports, and is itself supported by, the acquisition of other skills. Antonyms are pairs of words that are both minimally and maximally different, they typically differ maximally but only on a single dimension (e.g., length or age); therefore, their meanings are simultaneously similar and different (H. H. Clark, 1970; Jones, 2002; Murphy, 2003; Owens, 2008). Antonyms are also an important component of adult- and child-directed speech, as well as adult- and child-produced speech and adult-written language (Jones, 2002; 2006; 2007; Jones & Murphy, 2005; Murphy & Jones, 2008). For example, Murphy and Jones (2008) found that antonyms co-occurred once every 135 turns of speech produced by 4-year-old children. Given that their study evaluated a limited list of antonymic pairs, they suggest that this is likely a low estimate of the true frequency of antonym co-occurrence. In a large corpora composed of newspaper text Jones (2002) found that as many as one in 50 written sentences contained an antonym pair. Given the pervasiveness of antonyms in everyday communication it is important to understand children's appreciation of antonymy.

As will be reviewed, children's use of antonymy in discourse may precede their understanding of antonymy. In order to understand antonymy a child must learn that the meanings of some words are related, and that a pair of words that form an antonymic pair have a binary relationship with each other (Murphy, 2003). Children must also learn the specific relationship that the words have with each other in order to distinguish antonymy from other lexical relations, such as synonymy. Some researchers consider children's use of negation to be equivalent with their use of antonymy (e.g., Kreezer & Dallenbach, 1929). I, however, take the view that although an adult's use of negation may be conceptually similar to his use of antonymy, a stricter definition of understanding is necessary with children. As others have noted, there is a level of insight into the mind of an adult that allows for some things to be taken for granted or assumed, that require further evidence when children are being evaluated and described (Close, Hahn, Hodgetts, & Pothos, 2010). For children, the use of negation could simply indicate that they recognize a difference in the meaning of two words; it does not allow us to infer that this difference is specifically one of opposition. Prior to accepting negation as an antonymic response, the child must demonstrate an understanding of antonymy by producing canonical antonymic responses.

Children might use their understanding of antonymy to organize their lexical and conceptual understanding of the world (E. V. Clark, 2004; Murphy & Jones, 2008). Antonymy could help children understand and strengthen the association between pairs of words and help them notice the similarities and differences between pairs of words. This, in turn, could help them establish the meaning of the words. For example, antonymy is likely related to children’s understanding of dimensional adjectives (e.g., to understand that more is greater in quantity than less; Palermo, 1973). Antonymy could also help children organize information they have about the world into binary categories (e.g., this is a good cat and that is a bad cat). Finally, young children might use antonymy verbally to reduce the chance of being misunderstood (e.g., I want the big cookie, not the little cookie; Murphy & Jones, 2008).

Understanding antonymy is also likely related to the development of a variety of cognitive skills. For example, the resolution of conversational ambiguity that is required for appreciation of counterfactual irony is likely related to the concept of antonymy because, in this form of communication, the literal meaning of a remark is often the opposite of the intended meaning (Climie & Pexman, 2008; Pexman, 2008). If someone says "Nice shot!" to a person who has just missed making a winning goal, the speaker intends a negative meaning with a literally positive statement. It seems probable that an understanding of antonymy supports, and perhaps is critical to, the understanding of this form of verbal irony. Arithmetic skills may also depend, to some extent, upon the antonymy since conceptual knowledge of addition and subtraction requires comprehension of the concept of inversion (where $a + b - b$ must equal a, there is no calculation required; Bryant, 1992; Vilette, 2002). To fully appreciate addition and subtraction you need to understand the relationship between them, that is, that they have opposite effects.

The purpose of the proposed research is to investigate the lexical development of antonymy in children, with a focus on antonymous adjectives. Though antonymous nouns (e.g., *life – death*) and verbs (e.g., *increase - decrease*) also exist, I focused this investigation on adjectives because adjectives lent themselves best to the format of the novel task I developed. In my task children were presented with images of animals that depicted the adjective without additional objects in the image. Limiting myself to a single word class also resulted in less variability between trials in the wording of the task. Research investigating children's acquisition of antonymy is relatively sparse. As a result, in the following sections I will set the stage by reviewing children's acquisition of adjectives in general, providing an overview of how antonyms have been subcategorized (gradability and markedness/polarity), and describing extant theories on how children acquire the concept of antonymy. I will end with a review how children's understanding of antonymy has been studied to date and how my research builds on what has been done by addressing a major limitation in the metalinguistic studies of children's understanding of antonymy that have so far been conducted.

Learning the Concept of Opposite

Adjectives. The acquisition of adjectives is difficult for young children for a number of reasons (Graham, Cameron, & Welder, 2005; Graham, Welder, & McCrimmon, 2003; Waxman & Markow, 1998). First, adjectives can describe a wide range of properties. Adjectives can refer to transient physiological states (e.g., *hungry*), transient emotional states (e.g., *angry*), stable traits (e.g., *friendly*), observable perceptual properties (e.g., *green*), and unobservable properties (e.g., *smart*). Second, the precise meaning of an adjective is influenced by the noun that it modifies. For example, the adjective *bright* conveys one meaning when used in the phrase *the bright girl* (smart) and a different meaning when used in the phrase *the bright sun* (luminous). The meaning of dimensional adjectives also depends upon the context of comparison, a *big hippopotamus* and a *big mouse* are not expected to be the same size (Justeson & Katz, 1991; L. B. Smith, Cooney, & McCord, 1986; Syrett, Kennedy, & Lidz, 2010). And finally, adjectives have different ranges of application. While some adjectives can be applied across members of a category (e.g., *wooden* tables), and across members of different categories (e.g., *wooden* tables and *wooden* floors), other adjectives are more restricted in their application (e.g., *sleepy* cat, but not *sleepy* table). Given the challenges that children face when acquiring adjectival meanings, it is not surprising that adjectives are relatively rare in the early receptive and productive vocabularies of young children, especially in comparison to nouns (Caselli et al., 1995; Waxman & Markow, 1998).

Recent evidence suggests that children are able to use semantic information to distinguish adjectives from other lexical categories and that children appreciate that adjectives refer to object properties (e.g., Booth & Waxman, 2003; Gelman & Taylor, 1984; Hall, 1994; Hall & Graham, 1999; Hall & Moore, 1997; Waxman & Markow, 1998). When an object is labelled with a novel word modelled as an adjective, as in "*this is an X one*" children will extend the novel word to other objects that share similar properties. In contrast, when an object is labelled with a novel word modelled as a count noun, as in "*this is an X*" children will extend the novel word to other objects that are of the same category. Finally, when an object is labelled with a novel word modelled as a proper name, as in "*this is X*" children restrict their extension of the novel word. Children are also able to use semantic information to determine whether an adjective should be interpreted restrictively (Diesendruck, Hall, & Graham, 2006). Children were more likely to choose an unlabelled object when a novel adjective had been used in the prenominal position (e.g., *"this is a very BLICKY cat"*) than when a novel adjective had been used in the predicate position (e.g., *"this cat is very BLICKY"*). Children are also able to use information from the syntactic frame in which adjectives are embedded to help them establish the range of extension for a novel adjective. For example, Graham and colleagues (2003) found that children used information from the syntactic frame to determine whether a novel adjective could be extended to other members of the same basic-level category (e.g., dog). Children were more willing to extend the novel label to another German Shepherd when they were told that a German Shepherd "*is very DAXY*" versus when they were told that a German Shepherd "*feels very DAXY*".

In addition to semantic and syntactic information, there is a large body of evidence indicating that basic-level categories play a critical role in children's acquisition of novel adjectives (e.g., Graham et al., 2003; 2005; Klibanoff & Waxman, 2000; Waxman & Markow, 1998). For example, Graham and colleagues (2005) found that preschoolers were more likely to extend trait adjectives (e.g., *friendly*) to other members of the same basic-level category than to members of the same superordinate-level category or inanimate objects. Similarly, Waxman and Markow (1998) found that 21-month-old infants were more likely to extend a novel adjective to other objects with shared properties (e.g., colour) when all the objects were members of the same basic-level category than when they were members of different basic-level categories. In conclusion, although adjectives are difficult to acquire, preschool aged children comprehend adjectives and are able to use a variety of cues to identify and acquire novel adjectives. It is plausible that preschool aged children are also beginning to acquire the concept of antonymy when conveyed by adjectives.

Subcategorization of Antonyms

Antonyms have typically been subcategorized in terms of gradability or polarity. When categorized based on gradability, antonym pairs are labelled as either gradable (i.e., the application of one antonym does not automatically preclude the application of the other; e.g., *big – little*) or non-gradable (i.e., the antonyms are mutually exclusive, therefore the application of one precludes the other; e.g., *alive – dead*; Jones, 2002; Kennedy & McNally, 2005; Syrett, Bradley, Kennedy, & Lidz, 2006). Because gradable antonyms are composed of gradable adjectives (which describe properties on which objects can take different values on a scale), they are modifiable (e.g., *very big*), and can be used comparatively (e.g., *bigger*). A limitation of this categorization structure is that the distinction between non-gradable and gradable antonyms is not as transparent in everyday language use as it is in theory and definition. Non-gradable antonyms are frequently modified in a manner similar to gradable antonyms in everyday discourse, for example writers use the phrase "*very pregnant*" to describe someone in the later stages of a pregnancy (Jones, 2002).

The other traditional categorization structure of antonyms is based on polarity (also referred to as markedness). By this categorization structure, antonym pairs are divided into an unmarked member (positive pole) and a marked member (negative pole) using several criteria to determine which member is associated with which pole (Bracken, 1988; H. H. Clark, 1970; Murphy, 2003). The most frequently cited criteria are that the unmarked member generally represents a greater amount on the dimension of interest, typically fosters the dimensional name, and can be used neutrally (H. H. Clark, 1970; but see also Jones, 2002 for examples indicating that writers do not always feel that the unmarked term is sufficiently neutral). The member of an antonym pair that meets the most of these criteria is typically deemed the unmarked member. For example, *long* is the unmarked member of the antonym pair *long* – *short* as it represents a greater size, fosters the dimensional name (length) and can be used in a neutral sense (e.g., how *long* is that board?). However, recently Murphy (2003) has argued that a categorization structure based upon gradability or polarity does not account for all instances of antonymy. Instead Murphy suggests that antonymy should be considered as a subcategory of lexical relations that can be described as being the same in all "contextually relevant properties but one" (Murphy, 2003, p. 44). Within this categorization structure Murphy also includes other lexical relations, such as synonymy.

Gradability. Some researchers propose a further subcategorization of gradable adjectives into relative gradable adjectives (sometimes referred to as dimensional adjectives) whose interpretation is context dependent (e.g., *tall – short*) and absolute gradable adjectives whose interpretation is not context dependent (Kennedy & McNally, 2005; Syrett et al., 2006; 2010). Furthermore, absolute gradable adjectives can have scales with minimal (e.g., *spotted*) or maximal endpoints (e.g., *full*) that provide default standards of comparison. Research suggests that children as young as 3 years old distinguish between these subcategories of gradable adjectives (Bartlett, 1976; Eilers, Oller, & Ellington, 1974; Heidenheimer, 1975; O'Dowd, 1980; Syrett et al., 2006; 2010; Townsend, 1976).

Children will shift their standard of comparison with relative gradable adjectives but not with absolute gradable adjectives, demonstrating that they are sensitive to the fact that the former, but not the latter, are context-dependent. Children were asked to determine whether they could fulfil a puppet's request for one of two items based on the information in the request (1970; Syrett et al., 2010). Infelicitous requests were either not specific enough (e.g., "*the full one*" when neither container was completely full) or did not identify either object (e.g., "*the red one*" when the items were white and yellow). Like adults, children shifted their standard of comparison when the request included a relative gradable adjective (e.g., "*the long one*"). Also like adults, children did not shift their standard of comparison when the request included an absolute gradable adjective with a minimal standard (e.g., "*the spotted one*"). The only difference between children and adults was that children were more tolerant of some imprecision with absolute gradable adjectives with a maximal standard (e.g., "*the full one*"). The authors argue that children were not simply treating absolute gradable adjectives with maximal standards as if they were relative gradable adjectives; their reaction times to the former were longer than their reaction times to the latter, suggesting that there was an additional level of processing invoked and that their willingness to allow for the imprecision was not an automatic response.

In another task children were asked to determine the appropriateness of an adjective when judging sets of seven objects that varied on a gradable adjective dimension (E. V. Clark, 1973; Donaldson & Balfour, 1968; Holland & Palermo, 1975; Palermo, 1973; 1974; Syrett et al., 2006). Children judged the extreme poles (positive and negative) of relative and absolute gradable adjectives the same as adults, but their treatment of the objects that fell between the extreme poles differed. Children judged absolute gradable adjectives with a maximal standard (e.g., *full*) as meaning something like possessing the property (e.g., something like "*is full*") and therefore judged some objects near, but not at, the positive pole as possessing the relevant property (e.g., *full*). As with requests, children appeared to be more willing to accept imprecision than adults. Children's judgments of absolute gradable adjectives with minimal standards (e.g., *spotted*) did not differ from adults.

With relative gradable adjectives the determination of the standard of comparison is context-dependent (what counts as tall may vary from one context to another) and therefore there is an added layer of complexity to their interpretation (Barner & Snedeker, 2008; Ebeling & Gelman, 1994; , but see also Eilers et al., 1974; Syrett et al., 2006; 2010;

Townsend, 1976; Weiner, 1974). In addition to being sensitive to the fact that only relative gradable adjectives are context dependent, children are able to use contextual information to generate a standard of comparison that is sensitive to changes in the composition of the set of items being compared and are sensitive to different types of context (Barner & Snedeker, 2008; Palermo, 1973).

A perceptual judgment of the applicability of a relative gradable adjective (e.g., *tall*) made on a group of items requires that an abstract standard of comparison be established (Barner & Snedeker, 2008; 1970). Typically the standard is computed based on real world knowledge but it can also be spontaneously generated for a set of novel objects. Barner and Snedeker (2008) found that information about contextual factors (e.g., perceptual similarity and explicit mention of set membership) plays a role in children's grouping of items for comparison. Four-year-old children spontaneously considered novel objects that were perceptually similar as members of a set and generated a *tall – short* standard of comparison for the novel objects. Children were also sensitive to the statistical properties of an array of objects. The standard of comparison was affected by the addition of tall or short objects that were either perceptually similar or labelled as the same kind, but was not affected by the addition of tall or short objects that differed perceptually and were labelled as a different kind.

Not only is a relative gradable adjective evaluated in the relation to other objects of the same kind, the comparison that is made can be normative (compared to a stored mental representation), perceptual (compared to something else of the same kind that is physically present), or functional (judged with regards to its intended use; e.g., Ebeling & Gelman, 1994). Ebeling and Gelman (1994) found that children as young as 2-years-old are sensitive to how different contexts affect the interpretation of a familiar relative gradable adjective (*big – little*) and can switch between contexts with varying ease. Switching to a normative context was much more difficult than switching from a normative context even for 4-year-old children. In conclusion, by at least 4 years of age, children are not only sensitive to the influence of context on the interpretation of relative gradable adjectives (Syrett et al., 2006; 2010), they are also able to generate standards of comparison with familiar adjectives for novel objects, can shift this standard of comparison when new members of the set are introduced, and finally, children are sensitive to the different kinds of contexts that can be used to make comparisons (Barner & Snedeker, 2008; Ebeling & Gelman, 1994).

Polarity. The other way that antonyms have been subcategorized in the literature is in terms of polarity. This subcategorization structure has been used by some researchers to predict three stages in acquisition of polar terms (H. H. Clark, 1970). In the first stage of acquisition, the nominal meaning (i.e., the unmarked or positive pole term) is learned as the name of the dimension but is not associated with either end of the dimension. In the second stage of acquisition, both the marked and unmarked terms are associated with the positive end of the dimension. In other words, in this stage of acquisition the antonymous terms are treated as synonyms. In the final stage of acquisition, the unmarked term is correctly associated with the positive end of the dimension (i.e., the extended end) and the marked term is correctly associated with the negative end of the dimension (i.e., the relative lack of extent). In this final stage of acquisition the child's understanding of an antonym pair resembles an adult's understanding of an antonym pair.

As predicted by the first stage of acquisition, an asymmetric order of acquisition of polar antonyms has been found, with the unmarked (positive pole) member of a pair being acquired before the marked (negative pole) member of the pairs (Bracken, 1988; Brewer & Stone, 1975; E. V. Clark, 1972; 1973; Donaldson & Balfour, 1968; Klatzky, Clark, & Macken, 1973; Palermo, 1973; M. Smith, Johnston, & Coop, 1979). This asymmetry could be the result of differences in the frequency of usage in adult discourse since the unmarked member often fosters the name of the dimension and can be used neutrally (H. H. Clark, 1970). However, it has been argued that this is unlikely to be the sole contributor to the asymmetry (Klatzky et al., 1973). When Klatzky and colleagues (1973) replaced the unmarked and marked English words in four adjective pairs with nonsense syllables, children took longer to learn, and made more errors with, the syllables associated with the negative end of a dimension. Since the results didn't match what would be predicted if children were translating the syllables into their English equivalents or if children were using the antonym pair *big – small* as a mediator for learning, it was concluded that the results supported the hypothesis that the asymmetry in the acquisition of polar adjectives is due to an underlying cognitive asymmetry. What this asymmetry might entail was not described and because the asymmetry of acquisition is not consistently observed (Bartlett, 1976; Eilers et al., 1974; Heidenheimer, 1975; O'Dowd, 1980; Townsend, 1976), I am reluctant to draw strong conclusions about the order of acquisition of members of antonym pairs from this body of research.

In line with the second stage of acquisition proposed by Clark (1970), in which children associate both the marked and unmarked members terms with the positive end of the dimension, it has been found that young children who do not know the meaning of the term *less* tend to respond as if it was the synonym of *more* (E. V. Clark, 1973; Donaldson & Balfour, 1968; Holland & Palermo, 1975; Palermo, 1973; 1974 but see also Eilers et al., 1974; Townsend, 1976; Weiner, 1974). Performance improved with age, with children eventually correctly associating the terms *less* and *more* with the appropriate scales of the dimension (Palermo, 1973), as predicted by the third stage of acquisition proposed by Clark (1970).

Acquisition of Antonymy

There is some disagreement on the issue of how children learn the concept of antonymy. Previous research provides evidence that frequency of co-occurrence, the use of common or prototypical sentence frames, discourse goals, and prosody may play a role in how children learn the concept of antonymy and/or individual antonym pairs. Some researchers postulate that antonym word pairs are associated with each other because they occur together within a sentence more often than would be predicted by chance (co-occurrence hypothesis; Charles & Miller, 1989; Fellbaum, 1995; Justeson & Katz, 1991). For example, Justeson and Katz (1991) found that in English text antonyms co-occur in more than 8.6 times as many sentences as expected by chance, and found a co-occurrence rate of one in every 14.7 sentences. A few studies have also found that antonymous adjectives that co-occur in a sentence are often substitutable for each other, and that this may be one of the mechanisms by which the antonymous pair of words is identified (Justeson & Katz, 1991). However, Fellbaum (1995) found that antonyms can co-occur across word classes and suggested that this would limit the use of co-occurrence and substitutability in terms of a mechanism for children's learning of the concept of antonymy by making it a less reliable cue. These results suggest that co-occurrence and substitutability are likely one of several mechanisms that children use to learn antonymy.

The major limitation of the aforementioned research is that it uses adult corpora and data to explain how a child learns the concept of antonymy. This is problematic because discourse analyses have found some differences between the functions of antonymy in child- and adult-directed speech, as well as between child-directed and child-produced speech (Jones & Murphy, 2005; Murphy & Jones, 2008). Research has also found that there are changes in the discourse functions used by children at different ages that are not accompanied by similar changes in child-directed speech (Jones, 2007; Murphy & Jones, 2008). In a corpus study of antonymy, Jones (2002) found that antonyms served specific discourse goals (he originally identified two major discourse goals and six minor discourse goals; another minor discourse goal was added after evaluating child-directed speech; Jones & Murphy, 2005). These goals were often associated with (but not limited to) specific sentence frames. Together discourse analyses suggest that children's acquisition of the concept of antonymy is not simply the result of exposure to antonyms co-occurring in sentences and may instead also be related to the sentence frames in which the antonyms co-occur and the discourse goals for which antonyms are used (Jones, 2002; Jones & Murphy, 2005; Justeson & Katz, 1991; Murphy & Jones, 2008). Jones and Murphy have postulated that co-occurrence of word pairs in these sentence frames might highlight their contrastive nature and help the language learner appreciate the antonymous semantic relation between the words (Jones, 2002; 2007; Murphy & Jones, 2008).

Prosody may also be used to help children recognize the contrastive nature of antonym word pairs, and the use of prosody as a cue to recognize antonyms is not mutually exclusive with the aforementioned modes of antonym acquisition. Some of the ways that prosody could act as a cue to the link between antonymous word pairs are through verbal emphasis or acoustical features (Murphy, 2003). When English-speaking adults are asked to produce phrases containing novel adjectives that are intended to convey the meaning of known antonym word pairs they produced reliable acoustical (prosodic) features (e.g., fundamental frequency (F_0)) that were correlated with the overall word valence (positivity or negativity) and also each novel adjective pair had unique acoustical features (Nygaard, Herold, & Namy, 2009). This was referred to as generating meaningful prosody. These acoustical features were absent when a meaning was not specified to the speakers. When adults and children were asked to choose between two pictures they were able to select the intended picture when provided with meaningful prosody but not when provided with neutral prosody (Herold, Nygaard, Chicos, & Namy, 2011; Nygaard et al., 2009). Although not yet evaluated, children may be able to use these reliable acoustical features to learn the meaning of the words that form an antonymic pair, and/or help children establish a link between the words that form an antonymic pair. The prosodic emphasis on the antonymic words may help draw attention to them and may highlight their contrastive nature.

Approaches to the study of antonym acquisition. Empirical research systematically assessing the development of the concept of antonymy generally concludes that appreciation for the concept of antonymy begins to develop around the age of four years (E. V. Clark, 1972; Heidenheimer, 1975; Morris, 2003). However, there is also some evidence of later (Kreezer & Dallenbach, 1929) and younger ages of acquisition and usage (Jones, 2007; Jones & Murphy, 2005; Murphy & Jones, 2008). This is most likely the result of a wide range of tasks and behaviours having been used to evaluate children's knowledge of antonyms and the concept of antonymy. There are also differences in the criteria used to establish whether a child has acquired the concept of antonymy, with only some researchers considering negation as positive evidence. Broadly speaking, these studies fall into one of four categories: studies of the acquisition of meaning, studies of the expression of contrast, metalinguistic studies that focus on the acquisition of the lexical relation of antonymy, and discourse analysis studies. Knowing and using antonyms and appreciating the concept of antonymy involves a variety of linguistic and communicative skills, and together these studies reveal the complexity that is the development of children's appreciation of antonymy.

Acquisition of meaning. This group of studies evaluates children's knowledge about the meaning of the words that form an antonym pair. One of the major themes of this research is to determine how children come to differentiate the meanings of the terms, and is closely tied to the polarity research reviewed earlier. A number of these studies look at children's acquisition of the meaning of *more* and *less* (e.g., Donaldson & Balfour, 1968; Eilers et al., 1974; Holland & Palermo, 1975; Palermo, 1973; 1974; Townsend, 1976; Weiner, 1974), but studies have also considered other antonym pairs, such as *high – low*, *tall –short*, *long – short*, *wide – narrow*, and *before – after* (E. V. Clark, 1971; Eilers et al., 1974; L. B. Smith et al., 1986; Townsend, 1976). Some researchers conclude that children initially treat *more* and *less* as synonyms and only later differentiate the terms and associate them with the opposite poles of the dimension to which they refer (Bracken, 1988; Brewer & Stone, 1975; E. V. Clark, 1972; H. H. Clark, 1970; Holland & Palermo, 1975; Klatzky et al., 1973; Palermo, 1973; 1974). However, as mentioned previously the results of the studies evaluating children's understanding of the words *more* and *less* do not allow for strong conclusions to be drawn because the evidence is mixed (Bartlett, 1976; Heidenheimer, 1975; Kuczaj & Maratsos, 1975; O'Dowd, 1980; Townsend, 1976).

An alternative explanation on the acquisition of the meaning of the words that form an antonym pair is that they are learned individually (e.g., *hot – not hot* and *cold – not cold*) and only later associated with each other (e.g., *hot – cold*). Word association is one way that children's understanding of the antonymic relationship between word pairs has been evaluated and the results may shed light on how the meanings of words in an antonymic pair become associated with each other (Heidenheimer, 1975; 1978). Heidenheimer (1975) found that there was a significant increase in antonym production in a word association task between the ages of 5;3 and 5;9, but that children as young as 4;0 would spontaneously produce antonyms in this word association task, suggesting that even these young children have a nascent understanding of antonymy, or at the very least the words that form an antonymic pair are beginning to be linked to each other by these young children. During this developmental time period children's percentage of antonym responses was found to increase by almost 50% and prior to this increase in antonymic responding there was an increase in the number of negation responses produced (e.g., *big – not big*).

It was postulated that the increase in negation responses prior to the increase in antonym responses was indicative of an intermediate developmental stage that precedes the comprehension of the concept of antonymy, suggesting that children's acquisition of the concept of antonymy is gradual and not characterized by sudden insight. The increase in negation responses followed by the increase in antonymic responses suggests that children may be learning the meanings of the antonymic words individually and then forming the antonymic relation between them at a later stage of development rather than initially considering them to have the same meaning. Heidenheimer (1978) evaluated children in grades 1, 3 and 5 and found that children's production of synonyms lagged behind their production of antonyms in a word association task, and that the production of synonyms increased significantly with age. If children are initially conceptualizing both members of an antonym pair as having the same meaning, as theorized by Clark (1970), they would initially be conceptualizing the members of the antonym pair as synonyms and one might expect a different pattern of antonym and synonym responses in a word association task. More specifically, one might expect synonym responses to precede antonym responses, or at the very least for them to emerge at the same time.

The other major theme of this type of research involves evaluating whether some antonym pairs are acquired before others. The general consensus is that there is an order of acquisition when antonym pairs are from the same conceptual domain (e.g., size), with more general antonym pairs (e.g., *big – small*) being acquired before more specific antonym pairs (e.g., *tall – short*; Bartlett, 1976; Brewer & Stone, 1975; E. V. Clark, 1972; Eilers et al., 1974; L. B. Smith et al., 1986; M. Smith et al., 1979). For example, *big – small* is considered more general than *tall – short* because the latter requires the more extended dimension be vertical whereas the former has no such requirement. A *big* cat can be big on any or all of three dimensions, whereas a *tall* cat must be tall in the vertical dimension. It is typically found that children make fewer errors with the more general antonym pairs. Clark (1972) further found that young children substituted better-known words for lesser-known words in a semantically appropriate manner (e.g., replacing *tall* with *big* in the *tall – short* word pair) and that older children produced more correct answers (Morris, 2003, did not find a similar increase in performance with age when evaluating similarly aged children). However, Ravn and Gelman (1984) found conflicting evidence, reporting that for children between 3- and 5-years-old the terms *big* and *little* are typically used to refer to height instead of the more general meaning of greatest extent. They postulated that the correct meaning of *big* and *little* may, in fact, be more difficult for preschool children to acquire precisely because the terms are more general and therefore used in a less consistent manner.

Expression of contrast. This type of study evaluates children's use of antonyms to express contrast. Similar to discourse analysis studies, these studies examine the function of antonymy in language. For example, Akiyama (1985) focused on antonymous adjectives and asked English- and Japanese-speaking 4- and 5-year-old children to generate a denial of a sentence by saying the opposite (e.g., *The ship is large*). Of interest was whether children would prefer to use a semantic denial (usually an antonym; e.g., *The ship is small*) or a syntactic denial (e.g., *The ship is not large*). Most of the children did not know the meaning of the word *opposite* so they were given several example sentences and it was explained to them that there were two ways of saying the opposite (semantic and syntactic denials). English-speaking 4-year-old children used a semantic denial over 50% of the time while Japanese-speaking 4-year-old children used this strategy less than 25% of the time. The difference in preference was less dramatic, but still present, in 5-year-old children. Because both groups of children displayed a similar proportion of incorrect antonym use Akiyama concluded that the difference in production was not due to a lack of antonym knowledge in the Japanese-speaking 4-year-old children but rather due to cultural and language differences. Kim, Shatz, and Akiyama (1990) replicated this study with Korean-speaking children and found that Korean-speaking 4-year-old children used semantic denials on 43% of the trials. Over all age groups evaluated, they found that children produced semantic denials on 65% of the trials. Clark, Carpenter and Deutsch (1995) focused instead on verbs and evaluated how 2-, 3- and 4-year-old children used the productive option for creating the reversal of a verb with the prefix un- (e.g., shoes that have been *tied* can be *untied*). Children of all ages produced semantically appropriate reversals, including the use of antonyms (e.g., *off* elicited the response *on*).

Metalinguistic studies. Another group of studies evaluates children's understanding of the lexical relation of antonymy, versus children's use of antonyms as a strategy to create a contrast or a denial. To date, metalinguistic studies have been restricted to verbal games wherein the child responds to a question along the lines of "*what is the opposite of X?*" The present research study adopts a related approach, therefore a closer examination of this body of research will be provided.

The initial study to investigate children's appreciation for the concept of antonymy was conducted by Kreezer and Dallenbach (1929). Their task required that the child give verbal answers in response to a series of questions (e.g., *"What is the opposite of slow?"*). Each child was provided with example trials, and test trials with and without feedback. A child was considered to know the concept of antonymy if he produced the corresponding antonym (e.g., *fast*) or if he utilized the strategy of negation (e.g., *not fast*). The authors stated that they did not distinguish between children's use of antonyms and negation because they were interested in children's understanding of the concept of antonymy and not children's vocabulary skills. Furthermore, a child was considered to have learned the concept of antonymy during the course of the study if he began responding correctly during the test trials that contained feedback.

Several children exhibited learning of the concept of antonymy within the course of the study, and when they did so learning did not appear to be gradual. That is, after a child produced a correct answer, he tended to continue doing so until the completion of the study. The researchers described this learning as if it occurred with sudden insight, which is in contrast to the learning process suggested by Heidenheimer's (1975) research. Although children were not retested, the researchers believed that this learning was permanent.

It was concluded that 50% of children can be taught the antonymic relation between the ages of 5;3 and 5;9 if provided with instruction and feedback. Based on the trend of increasing appreciation for the concept of antonymy as the age of the children increased, the authors suggested that a child could learn the concept of antonymy on their own (i.e., without instruction) between the ages of 8 and 9. Interestingly, and of consequence for all metalinguistic studies of antonymy, it was found that some children who stated that they did not know the concept of antonymy actually did know the relation but not the name for it, and vice versa. This suggests that studies that rely extensively on children's understanding of the word *opposite*, especially those that do not provide any feedback during the course of the study, may be unnecessarily difficult for children.

Further research revealed that the age of acquisition of the concept of opposite is likely younger than was found by Kreezer and Dallenbach (E. V. Clark, 1972; Heidenheimer, 1975; Kreezer & Dallenbach, 1929; Morris, 2003). With a similar verbal task (e.g., "*What is the opposite of X?*"), Clark (1972) found that children as young as 4 could provide a correct opposite response to some of the stimulus words. Children younger than 4 could not understand the task and therefore were not evaluated beyond the pilot testing phase. Children participated in two experiments separated by approximately two weeks; in the first experiment the stimulus words were presented in isolation (e.g., *big*), and in the second experiment the stimulus words were presented in a three-word phrase (e.g., *the big horse*). Children heard an example trial and completed a practice trial before being administered the remainder of the task, but no feedback was provided in any phase of the study. There were no differences between the results of these two experiments so the data were combined.

Most notably, in contrast with the sudden learning found in the study conducted by Kreezer and Dallenbach (1929), a distinct order of acquisition was found; when making a semantically appropriate substitution, children always replaced a difficult word with an easier word. For example, a child might replace tall with big in the *tall – short* word pair but never the reverse (i.e., replacing *big* with *tall* in the *big – small* pair). There was also a correlation between age and the number of semantically appropriate responses (including but not limited to adult correct responses), which Clark (1972) suggested was the result of children in the younger age groups (4;0-4;5 and 4;6-4;11) knowing the easier, less complex, word pairs (e.g., *big – small*) and none of the more difficult word pairs (e.g., *deep – shallow*). These results suggest that the younger children tested knew the concept of opposite but did not have a sufficiently developed vocabulary to successfully perform the task with all word pairs. Perhaps vocabulary was also a limiting factor for the children who were unable to understand the task.

Utilizing a verbal task similar to the previous two studies reviewed, Morris (2003) evaluated whether children between the ages of 4 and 5 years understand the concept of opposite. Before being administered the test trials, children were provided with warm up trials and a definition of opposites (“An opposite is something unlike something else. So, the opposite of tiny is big.” p. 426), but no feedback was provided on individual trials. Children were able to consistently generate opposites, however, in contrast to the results obtained by Clark (1972), no effect of age was found; 4- and 5-year-old children appeared to understand the concept of opposite equally well. These results again suggest that children’s verbal behaviour reflects an understanding of the concept of opposite at a much younger age than found by Kreezer and Dallenbach (1929), which raises the question of whether changes in the education and/or socialization of children over the past 80 years may have lowered the age of acquisition.

Discourse Analysis. It is also possible that the concept of antonymy begins to develop before the age of 4, especially given that not all children understand the instructions provided in the metalinguistic studies (i.e., not all children understand the meaning of the word *opposite* at the start of the study). If children don’t fully understand the instructions provided then the studies are liable to conclude that the concept of antonymy develops at a later age than actually occurs, especially if there is no feedback or training provided, as was the case in several of the aforementioned metalinguistic studies.

This suggestion is supported by several recent discourse analysis studies that identify the ways antonyms are used in discourse by children and adults. Whereas the previous three areas of research used elicitation techniques, this final type of study uses a corpus-based approach and looks at the co-occurrence of antonym pairs in turns of speech (spoken language) or sentences (written language). In contrast to the previous studies that looked at a child's metalinguistic awareness of the concept of antonymy, these studies look at a child's usage of antonyms. The studies in this group had three major goals: to identify when children begin using antonyms in everyday speech, to determine whether the discourse functions of antonymy varied as a function of the age of the user, and to determine whether the discourse functions used by children paralleled those used by adults when speaking to children (Jones, 2002; 2007; Jones & Murphy, 2005; Murphy & Jones, 2008). The studies evaluated the types of antonyms used in up to four different language corpora: adult-produced writing, adult-produced speech, child-produced speech, and child-directed speech (Jones, 2007; Jones & Murphy, 2005; Murphy & Jones, 2008). Both of the child corpora came from longitudinal recordings of family interactions with either three (Jones, 2007; Jones & Murphy, 2005) or five (Murphy & Jones, 2008) children between the ages of 2;4 and 4;11 years. The adult corpora came from either a newspaper corpus or the spoken component of the British National Corpus (Jones, 2007; Murphy & Jones, 2008).

Spontaneous usage of antonyms was found in children as young as two years old (e.g., "He's a girl and you're a boy"; Jones & Murphy, 2005, p. 408), and children of all ages were found to use antonyms in most of the discourse functions that have been identified in the three corpora of adult usage. An additional discourse function, interrogative antonymy (in which antonyms are used in the context of a question or choice; "Is she a good mommy or a bad mommy?" Jones & Murphy, 2005, p. 413), that was not designated as an independent discourse function in previous analyses of adult-produced writing (Jones, 2002) was found to be sufficiently prevalent in child-directed and child-produced speech as to warrant its addition (Jones & Murphy, 2005). A subsequent analysis of a British spoken word corpora similarly found that interrogative antonymy was more common in adult-directed speech than in adult-produced writing and therefore its prevalence was not unique to child-produced and child-directed language (Jones, 2006).

Overall it was found that the usage of the discourse functions of antonymy was both similar and different in the four corpora evaluated. The two most common discourse functions were Ancillary Antonymy (when the contrast between an antonym pair is used to highlight, affirm, or create a contrast between another pair of words; "Milk is good for you but gum is bad for you." Jones & Murphy, 2005, p. 409) and Coordinated Antonymy (when antonyms are used to signal the inclusiveness or exhaustiveness of a scale, in other words, to signal the unimportance of the scale; "I wanna play with my big cars and my little cars." Jones & Murphy, 2005, p. 409). The frequency with which the minor classes of antonymy (so called because they are used much less frequently than Ancillary and Coordinated Antonymy) were used was found to differ more between the corpora; the pattern of usage in child-directed speech differed from that found in child-produced speech. In addition, distinguished antonymy (when the writer or speaker explicitly refers to the distinction between a pair of words; "I think sometimes, and I'm not saying there's any difference between male and female as far as that's concerned, but I think it's a way of your, getting your er aggression out." Jones, 2007, p. 1114) was not found at all in either child-produced or child-directed speech.

Antonym use was always proportionally greater in child-produced speech than in child-directed speech and was also found to significantly increase in child-produced speech with age, a change that was not paralleled in the child-directed speech (Jones, 2007; Jones & Murphy, 2005; Murphy & Jones, 2008). The frequencies of the minor classes of antonymy were also found to differ within child-produced speech according to the age of the child, however there was not a corresponding change in child-directed speech. Based on these differences in the patterns of discourse function usage, the authors concluded that children are active, not passive, users of antonyms in discourse and tentatively concluded that children are not merely reproducing the discourse functions that they hear most often. If children's usage can be taken as a measure of their understanding, or at least the initial stages of their development of the concept of antonymy, it is possible that children may display an understanding of the concept of antonymy in a metalinguistic task before the age of 4 if the task demands are either changed or decreased (Jones, 2007; Murphy & Jones, 2008). However, it is important to remember that distinguished antonymy, the discourse function that relies most heavily on metalinguistic awareness, was not found at all in child-produced or child-directed speech (Jones, 2007). Young children's use of antonymy in discourse may simply be a developmental precursor to children's metalinguistic understanding of antonymy.

Present research

The goals of the present study were to determine the extent to which young children have developed a lexical concept of antonymy when it is evaluated with a less verbal metalinguistic task, whether there is evidence of conceptual understanding before children are able to demonstrate this understanding behaviourally, and whether other cognitive developments are related to the development of the concept of antonymy. Murphy and Jones (2008) suggested that the metalinguistic studies that have been conducted thus far may have failed to find evidence for children's understanding of the concept of opposite before the age of 4 because the response tasks used have been too difficult for young children. Verbal reports are difficult for young children, so children's performance in previous metalinguistic studies could have been hindered by the requirement of a verbal response (Vilette, 2002). Furthermore, children's verbal recall of an event can lag behind nonverbal measures of memory for the same event and can also lag behind general verbal skill (Simcock & Hayne, 2003). Thus there is the possibility that a more engaging and less demanding task, for example a task requiring a nonverbal response, would reveal the development of the concept of antonymy in children under the age of 4 years. This possibility was tested in the present study.

Although my goal was to make the task more engaging and potentially less demanding, I also wanted to ensure that I was assessing genuine comprehension. That is, I did not want any features of the task or stimuli to make that concept self-evident. I also wanted to assess whether children were performing above chance in my task by understanding that there was a relationship between the words without understanding that the nature of the relationship was one of antonymy. Specifically, the present study evaluated the development of the concept of antonymy in 3- 4- and 5-year-old children, using adjectives and a task that assessed comprehension but did not require the child to make a verbal response. By eliminating the need for a verbal response I address the major limitation of previous metalinguistic studies of children's understanding of the concept of antonymy. For this purpose I devised a novel task, hereafter referred to as the Opposite Task. In brief, the child is presented with three images. The experimenter selects one image, and then asks the child to choose an image. The experiment was a between subjects design, children were video recorded while they participated in one of three conditions. In the Opposite Label Condition the images were labelled with an adjective (e.g., *big dog – small dog – happy dog*) and the child was asked to choose the *opposite one*. In the Another One Label Condition the images were similarly labelled but the child was asked to choose *another one*. Finally, in No Label Condition the images were labelled with basic-level category information only (e.g., *dog – dog – dog*) and the child was asked to choose *another one*. Video recordings were later used to measure response latency and to code eye gaze behaviours. In each condition, children's working memory and receptive vocabulary were also evaluated. All parents were asked to complete the Child Communication Checklist Revised (CCC-2; Norbury, Nash, Baird, & Bishop, 2004) and a demographic questionnaire.

Working memory may be positively correlated with performance on the Opposite Task since the child must keep in mind the labels provided for each image, the image the experimenter selected, and that the goal is to select the opposite image. The demands on working memory were minimized, but not eliminated, in the design of the Opposite Task by ensuring that the child could see the image that the experimenter selected while making his decision about which image to choose. Repeating the task instructions on each trial also reduced the demands placed on the child's working memory.

Receptive vocabulary may be positively correlated with performance on the Opposite Task for at least two reasons. First, if the development of the concept of antonymy requires that the child know both words in an opposite pair before being able to understand the oppositional relationship between the words in the pair, a child who has a larger receptive vocabulary could learn the concept of antonymy before a child who has a smaller receptive vocabulary. Second, a child may need to know several pairs of opposites before being able to develop the concept; a child with a larger vocabulary likely knows more pairs of words that have an oppositional relationship than a child with a smaller vocabulary.

If one of the mechanisms by which children learn antonymy is through the word pairs co-occurring in sentence frames, as suggested by the discourse analysis research and the co-occurrence hypothesis, then overall communicative ability or performance on either the Context and/or Syntax subscales of the CCC-2 might be positively correlated with performance on the Opposite Task. On a related note, children who are more skilled language users, as measured by the CCC-2, might be expected to be more knowledgeable about antonymy.

Hypotheses. Based upon prior research and the arguably reduced task demands in my novel task relative to the tasks used in previous studies, it was hypothesized that 5- and 4-year-old children, and perhaps 3-year-old children, would demonstrate knowledge of the concept of antonymy with above chance performance in the Opposite Label Condition of the Opposite Task (E. V. Clark, 1972; Heidenheimer, 1975; Morris, 2003), and that accuracy on the Opposite Task would improve with age. In the Another One Label Condition it is possible that children who are aware of the antonymy relation between the adjective pair would respond by selecting the opposite image even though the experimenter does not explicitly ask for it, in which case it was expected that they would perform above chance. On the other hand, it is possible that even if children know the concept of antonymy, they may need the concept label in order to perform above chance on the Opposite Task, in which case it was expected that they would perform at chance. Finally, if, as intended, the images did not themselves convey the adjectives children were expected to perform at chance in No Label Condition of the Opposite Task. It was further expected that older children would make accurate judgements faster than younger children in the conditions in which they have above chance performance.

An order of acquisition was expected, with some opposite word pairs being acquired before others. As mentioned, several researchers have found a distinct order of acquisition of dimensional adjectives in which more general antonym pairs were acquired before more specific antonym pairs (Bartlett, 1976; Brewer & Stone, 1975; E. V. Clark, 1972; Eilers et al., 1974; L. B. Smith et al., 1986). In addition, the words that compose the opposite pairs are acquired at different ages which could cause different performance on the word pairs in the present research (Dale & Fenson, 1996). Based on previous research, and the age of acquisition for specific words, it was expected that the best performance would be observed in the Opposite Label Condition for the following word pairs: *big – small*, *clean – dirty*, and *happy – sad*.

In terms of the additional measures collected, it was hypothesized that children's accuracy in the Opposite Label Condition of the Opposite Task would be positively correlated with their working memory, receptive vocabulary and parental responses on the CCC-2 even when the effect of age was controlled for with partial correlations. No specific associations or relationships were hypothesized with regards to the demographic data; rather this information was collected in an exploratory manner to further shed light on the development of the concept of opposites.

One area of disagreement in the literature is whether children learn the concept of antonymy gradually or with sudden insight. The following eye gaze and latency measures were included to investigate whether children show a sensitivity to, or an implicit understanding of, the concept of antonymy prior to demonstrating their knowledge behaviourally. If they do not show any sensitivity it would suggest that children acquire the concept of antonymy rapidly, as if with sudden insight. However, it is important to note that this is not in conflict with my previous hypothesis about the order of acquisition of word pairs because acquiring the concept of antonymy is not considered to be equivalent to understanding the concept across all antonym pairs. The following eye gaze measures were used to evaluate whether the acquisition of antonymy could be considered gradual or with sudden insight: proportion of time looking at the target image (i.e., the correct response in the Opposite Label Condition), proportion of time looking at the distractor image (i.e., the incorrect response in the Opposite Label Condition), proportion of time looking at the experimenter, proportion of time looking at the image selected by the experimenter, the overall number of fixations, the proportion of first fixations on the target, and the response latency.

Age differences were expected in the Opposite Label Condition on all eye gaze measures. It was expected that older children would spend a greater proportion of time looking at the target image relative to younger children, would make fewer fixations during the course of a trial, and would show shorter overall response latencies, all of which would suggest greater levels of efficiency and confidence in their responses. By analysing the trials ending with the selection of the distractor (i.e., incorrect trials in the Opposite Label Condition) separately from the trials ending with the selection of the target (i.e., correct trials in the Opposite Label Condition) I was able to examine whether, when children select the distractor image, their eye gaze behaviour showed any evidence that they were sensitive to the opposite relationship, and whether this changed as a function of age. Age differences could occur if older children have a better implicit appreciation that the target image is potentially the desired answer versus younger children who might make select the distractor image because they do not even consider the alternative choice.

Method

Participants

Participants were 60 three-year-old children (*M* = 39.07, *SD* = 2.07, 30 girls and 30 boys), 84 four-year-old children (*M* = 50.45, *SD* = 1.87, 42 girls and 42 boys), and 60 five-year-old children (*M* = 62.10, *SD* = 2.02, 30 girls and 30 boys), recruited from the University of Calgary ChILD database. Across ages, 68 children completed each condition of the Opposite Task (equal numbers of males and females). An additional eight children participated but were excluded due to experimenter error (*n* = 3), placing two images in the box on four or more trials (*n* = 3), unwillingness to participate (*n* = 1), or a parent-reported speech delay (*n* = 1). All participants were from English speaking homes, and in all but three homes, English was the children's first language.

Procedure

Participants were tested individually in a quiet room at the University of Calgary. Each child first completed the Opposite Task (one of three conditions) followed by a receptive vocabulary task (TELD3) and then a working memory task (Memory for Objects). Parents were asked to complete a brief demographic questionnaire and the CCC-2, United States Edition (Norbury et al., 2004) while their child participated.

All of the tasks were completed in a single session lasting between 20 and 30 minutes. Children were video recorded while completing the Opposite Task so that responses and eye gaze behaviour could be later analyzed in a detailed fashion.

Opposite Task

Materials. The stimuli were colour images of animals that a preschooler is likely to be familiar with and that can vary in several ways (e.g., cat, dog, dinosaur). The images were presented in sets of three (e.g., *dirty pig – clean pig – small pig*) and within a set all of the images were from the same basic level category (i.e., the same type of animal). The images measured a maximum of 8 cm on the largest dimension, with the exception of the images depicting *big*, *small*, *tall* or *short* animals. Images labelled *big* or *tall* measured just under 10 cm on the largest dimension and images labelled *small* or *short* measured 5 cm on the largest dimension. All images were printed on white paper that was then cut to measure 10 cm x 10 cm so that the image was centred. The images were laminated and the adjective depicted in the image was printed on the back of the card (this label was not shown to the child). Since preschool aged children may use visual similarity (e.g., shape, texture, colour) to guide categorization of objects (Baldwin, 1992; Bonthoux & Kalenine, 2007; Morgan & Greene, 1994), care was taken to control the visual similarity (e.g., colour, posture, size) of the images within sets of images so that no one image looked particularly different than the other two (e.g., one black cat, one white cat, and one orange cat).

The images depicted eight pairs (16 words) of opposite adjectives; please refer to Appendix A for a list of the word pairs used. Two of the eight word pairs were used for training and the remaining six pairs were used for the test trials; each word pair was presented once. The word pairs were selected based upon prior research on the concept of opposite or antonymy (Bracken, 1988; E. V. Clark, 1972; Heidenheimer, 1975; Jones, 2007; Kreezer & Dallenbach, 1929). That is, previous research suggested that these pairs are among the antonym pairs that children understand first.

In each set of images presented to the child, two of the images depicted a pair of antonymous adjectives (e.g., *dirty – clean*), and one image depicted an unrelated distractor (e.g., *small*; please refer to Appendix B for the images used). The distractors were generated from the eight pairs of adjectives. One member of each opposite pair was used once as a distractor, with the limitation that the words *short* and *tall* were not used as the distractor for the *big – small* word pair and vice versa. To offset concerns about priming, the word pair *big – small* and the word pair *short – tall* were never presented consecutively. Furthermore, the distractor was never part of the word pair used in the set of images immediately preceding or following it (e.g., *dirty* was never be used as a distractor in a set of images when the previous or following set contained the word pair *clean – dirty*).

For some of the word pairs the target and the distractor images were quite similar. For example, the *small pig* did not have any dirt on it, and could, therefore, also be considered a *clean pig*. If the experimenter placed the *dirty pig* in the box the child would have had to select between two images that had no dirt on them (the *clean pig* and the *small pig*). To help mitigate this problem, the experimenter always placed the image that is confusable with the distractor image (in this example, the *clean pig*) into the box.

Procedure. On each of the eight trials, a child was presented with a set of three images (e.g., *dirty pig – clean pig – small pig*). On the table was a brightly coloured box (29 cm x 22 cm x 2 cm) tilted slightly towards the child and positioned directly in front of the child. Also on the table, and placed between the box and the child, were placeholders to ensure that the spacing of the images was consistent across trials and between children. The images were placed on the placeholders during the course of a trial. Refer to Figure 1 for a schematic of the testing situation.

At the start of the session the child was brought into the testing room and seated at a small table across from both the experimenter and a video camera. All sessions were recorded as digital video with the camera angled to record the child's eye gaze. The assumption, adopted by many language processing researchers, is that "the mind is going where the eye is going" (Trueswell & Gleitman, 2004, p. 320). Therefore, by monitoring children's eye gaze as they performed the task it was possible to assess what they are "considering" and when they begin "considering" the alternative responses.

The child was told that he or she would be playing an animal game with the experimenter. The experimenter then introduced the task to the child through a training trial. The experimenter told the child the following: "The first game we're going to play today is an animal game. This game has just two rules. The first rule is that you sit on your chair. The second rule is that you keep your hands in your lap until it is your turn to play. Can you do that for me? Good job. Now I'm going to show you how to play the rest of the game." The experimenter then placed a set of images (e.g., *dirty pig – clean pig –*

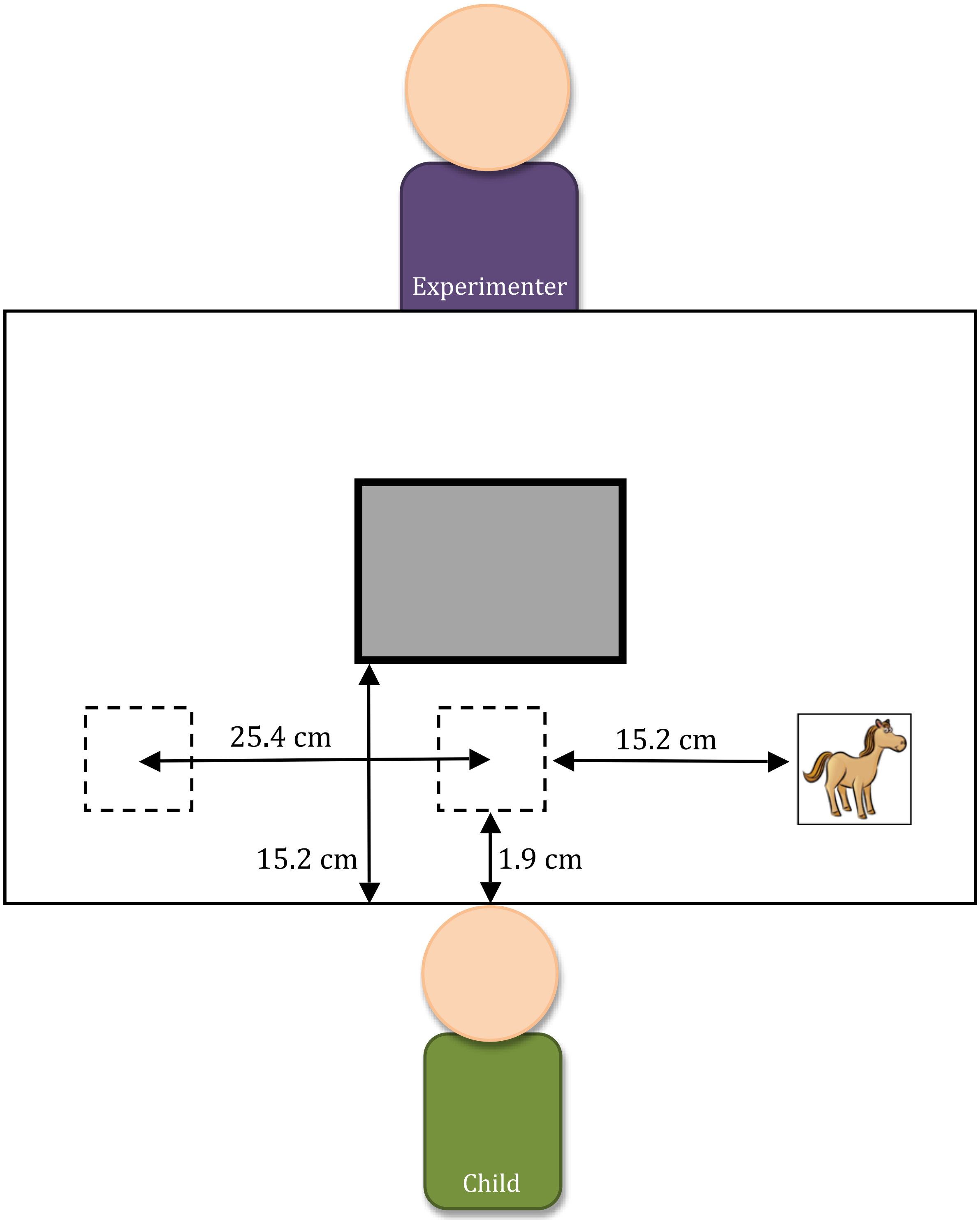


Figure 1. Schematic of the testing situation.

small pig) on the table in front of the child and labelled each image while pointing to it. The experimenter always placed the first image directly in front of the child. The placement of the second and third image was counterbalanced, such that on half of the trials the second image was placed to the left of the first image. This balancing ensured that children who always chose the image placed last (or second) performed at chance. The location of the other two images was counterbalanced, such that in half of the trials the target was placed on the right hand side of the child. This balancing ensured that children who always chose the image on the right (or on the left) performed at chance. The placement of the other two images was also partially counterbalanced, such that four trials had the target placed before the distractor image (or vice versa). If the child was not paying attention to the task the experimenter re-labelled each of the images while pointing to each a second time. Following the labelling of the images, the experimenter put the middle image (e.g., *clean pig*) into the box, and while doing so labelled the image again. The experimenter always placed the middle image into the box in order to maximize the distance between the target image and the distractor image so as to facilitate coding of eye gaze behaviours from the video.

Each child saw two training trials (presented in a fixed order) followed by one of four possible orders of presentation of the six test trials (the presentation orders are hereafter referred to as versions). In the Opposite Label Condition the experimenter then asked the child to place "the opposite one in the box" (e.g., *dirty pig*). This was the signal to the child that it is his turn to place an image into the box. In the Another One Label Condition and the No Label Condition children were asked to place "another one in the box". No Label Condition differed from the Opposite Label Condition and the Another

One Label Condition in that the images were not described with adjectives (e.g., *pig* – *pig* – *pig*). Please refer to Appendix C for more details on the wording of the task. In the two training trials presented in the Opposite Label Condition (but not in the Another One Label Condition or the No Label Condition), the child received feedback about their image selection. If the child selected the distractor image the experimenter presented the correct image with the image that the experimenter had selected and verbally told the child the correct response. If the child selected the target image the experimenter provided positive feedback. No feedback was provided in the six test trials. In the Another One Label Condition and the No Label Condition the child was provided with feedback about their performance in putting one picture into the box.

Coding system. Participant responses were coded as either a target selection (i.e., correct response on the Opposite Label Condition) or a distractor selection (i.e., incorrect response on the Opposite Label Condition). In other words, did the child place the opposite image in the box? Trials were then divided on response accuracy such that trials ending with the selection of the target were analyzed separately from trials ending with the selection of the distractor. Measures coded from the video included response latency, the proportion of first fixations on the target, the total number of fixations, and the durations of looks to each response image (target and distractor), the experimenter and the image selected by the experimenter (each expressed as the proportion of looking time relative to total looking time). Looks to other locations were coded as "other" and were not analyzed further.

The video record was examined on a frame-by-frame basis (each frame = 33 ms) using FinalCut Pro 5.0.4, with audio and video signals synchronized. In addition to overall response latencies (total time children took to respond from the experimenter's cue) I also wanted to measure response latencies in a somewhat more fine-grained way; that is, distinguishing early from later processing. In order to do so, I divided total response latencies as a function of the child's actions. Thus, response latencies were measured from the point in time at which the experimenter began voicing "can you put the opposite/another one in the box" to the point when the child initiated movement of his/her arm to pick up an image (Early phase) then from the initiation of movement to the point when the child touched the image they would eventually put into the box (Middle phase), and finally from the first touch to the point when the child placed the image in the box (Late phase). Response latencies were measured from the initiation of movement when it preceded the voicing of the response request because although the child was instructed to keep their hands in their lap they were not instructed regarding where to look. Thus, a child might look at the images on the table as soon as the experimenter finished labelling the images. There is a certain level of arbitrariness in parsing the response latency into these phases on the basis of the child's action however there is a large body of literature linking action to cognition (e.g., Huettig, Rommers, & Meyer, 2011). The alternative would be to parse the response latency based on an absolute time point. By linking it to a child's action I am able to couple the measure to each individual child. Individual trials were excluded from the eye gaze analysis if the child was not facing in the direction of the camera at the time that the experimenter began voicing the response request, if the child knocked an image off the table, if the child moved the image selected by the experimenter, if the child asked a question of the experimenter, if the child needed to be encouraged to respond, if the experimenter asked the child to sit down or keep their hands in their lap after the trial had started, or if the eye gaze or movement of the child was obscured in the video (e.g., by the experimenter's arm). A total of 79 trials (6.45%) were excluded from the eye gaze analyses for these reasons.

A second coder coded data for 51 of the participants (25% of the data). Neither coder was fully blind to the hypotheses because audio information about which condition the child was in was linked to audio information that was essential to the coding process. Differences of greater than two video frames were reviewed and discussed by the coders. The final agreement between first and second coders (in terms of number of video frames elapsed) was acceptable, and inter-rater reliability was assessed using a two-way mixed-model absolute agreement single-measures intraclass correlation coefficient (ICC; Hallgren, 2012; McGraw & Wong, 1996). The resulting ICC was in the excellent range for all variables evaluated (Cicchetti, 1994): $ICC = 0.99$ for the Early phase, $ICC = 0.99$ for the Middle phase, and $ICC = 0.99$ for the Late phase. The numbers of times the child looked at the target, the distractor, the experimenter, and the box/image selected by the experimenter on each trial were also coded, as well as the duration of looks (in terms of number of video frames elapsed) to each of these locations. Agreement between the first and second coders was $ICC = 1.00$ for number of fixations to each location. Agreement was $ICC = 0.99$ for the duration of looks to the target, $ICC = .99$ for the duration of looks to the distractor, $ICC = 0.99$ for the duration of looks to the experimenter, and $ICC = .98$ for the duration of looks to the box/image selected by the experimenter.

Additional Measures

Memory for Objects task. The Memory for Objects task evaluates children's working memory capacity (Nilsen & Graham, 2009). In this task, the child was shown between one and three sets of three pictures, was asked to name the pictures (from left to right), and asked to remember the last picture (item) in each row. Pictures were viewed through a card so that only one row was visible at a time, and the final row of pictures was covered before the child made a response. Refer to Appendix D for more details on the wording of the Memory for Objects task.

Increasing the number of rows presented in a trial from one to three increases the task difficulty by increasing the number of items to recall. The child was presented with two practice trials (1 single row trial and 1 two row trial) and up to 18 test trials (4 single row trials, 10 two row trials, and 4 three row trials). Testing continued until the child failed to recall a single item from four consecutive trials. The measure of interest was the number of items correctly recalled, regardless of order (scored out of 36).

Test of Early Language Development 3. The Test of Early Language Development 3 (TELD3) is a copyrighted, standardized testing tool used to assess children's receptive, expressive and overall spoken language with children aged 2;0 to 7;11 years. In this study only the receptive vocabulary component of this tool was administered.

In this task, the child was asked to respond to three types of questions. On the first type of question, the child was instructed to look at a page that has printed images (e.g., a boy inside a house and a boy standing beside a house) on it and asked to point to a specific image on the page (e.g., the boy standing next to the house). The second type of question consists of the child being asked verbal questions and the child responding to each question verbally (e.g., Do you eat breakfast in the morning or the afternoon?). The final type of question involves the child being given "toys" (e.g., blocks and coins) and instructions on how to interact with the toy (e.g., place one coin on each block). If a child did not respond after 15 seconds, the experimenter encouraged a response, but if no response was given afterwards, it was scored as incorrect.

A child was tested until the basal questions and the ceiling questions were established. The basal questions are the highest three consecutive questions answered correctly. The ceiling question is the last correctly completed question prior to the child answering three consecutive questions incorrectly. The age of the child determined the starting item (age 3 = item 10, age 4 = item 15, age 5 = item 20) and testing continued until the ceiling item was reached. Raw scores were compared to norms to generate percentile rank scores.

Parental demographic questionnaire and CCC-2. In addition to working memory and receptive vocabulary, other factors may be related to the development of the concept of opposite. Parents were asked to complete a brief demographic questionnaire while their child participated in the tasks described above. The questionnaire was composed of nine questions which asked about the age of the child, the presence of siblings in the household, the birth order of children in the household, the languages spoken in the household, the child's childcare experience, and the child's access to games and books that deal with the concept of opposite. See Appendix E for a copy of the parental questionnaire.

Parents were also asked to complete the CCC-2 while their child participated. The CCC-2 is a copyrighted, standardized parental checklist that is composed of 70 items that parents rate on a scale from 0 to 3 (20 of the items are reverse coded). Population norms for 10 subscales and a general communication composite (GCC) score are available for children between the ages of 4;0 and 16;11. Parents of 3-year-old children were also asked to complete the CCC-2 to ensure uniformity across participants but these responses were not included in any of the analyses reported. Scores on the general communication composite measure and the syntax and context subscales were of interest. These subscales are of interest because of research suggesting that children are able to use semantic information to distinguish adjectives from count nouns and proper names, as well as to determine whether an adjective should be interpreted restrictively (Diesendruck et al., 2006; Gelman & Taylor, 1984; Hall & Moore, 1997; Waxman & Markow, 1998).

Results

Behaviour

Children's proportions of target selections (number of times the target image was selected divided by the number of valid trials, hereafter referred to as accuracy) in each of the conditions, as a function of age, are presented in Table 1.

Table 1

Children's Mean Proportions of Target Item Selections in Each Condition

	Opposite Label Condition		Another One Label Condition		No Label Condition	
Age	*M*	*SD*	*M*	*SD*	*M*	*SD*
3	0.50	0.16	0.44	0.15	0.46	0.13
4	0.64	0.22	0.54	0.19	0.48	0.21
5	0.79	0.23	0.55	0.22	0.51	0.20

Accuracy was first evaluated with a 2 (gender: male, female) x 3 (age: 3-, 4-, 5-years-old) x 3 (condition: Opposite Label Condition, Another One Label Condition, No

Label Condition) x 4 (version: 1, 2, 3, 4) between subjects ANOVA. There was a significant main effect of age, $F(2, 132) = 8.46$, $p < .001$, a significant main effect of condition, $F(2, 132) = 13.03$, $p < .001$ and a significant main effect of gender, $F(1, 132) = 4.76$, $p = .031$ (males ($M = 0.57$, $SD = .22$) outperformed females ($M = 0.52$, $SD = 0.21$)). There was no main effect of version, $F(3, 132) = 1.00$, $p = .397$, nor were any interactions significant (all *p*s > .10).

Because gender and version effects were not of interest in the current study, and did not interact with the variables of interest, these variables were dropped from the analysis and children's accuracy was re-evaluated with a 3 (age: 3-, 4-, 5-years-old) x 3 (condition: Opposite Label Condition, Another One Label Condition, No Label Condition) between subjects ANOVA. There was, again, a main effect of age, $F(2,195) = 8.87$, $p < .001$, and a main effect of condition $F(2, 195) = 13.15$, $p < .001$. Most importantly, there was a marginally significant Age x Condition interaction, $F(4, 195) = 2.27$, $p = .063$. Follow up simple main effect tests for each condition revealed the nature of this interaction. In the Opposite Label Condition there was a significant main effect of age, $F(2, 68) = 9.88$, $p < .001$. Five-year-old children outperformed 4-year-old children ($t(65) = 2.47$, $p = .016$) and 3-year-old children ($t(65) = 4.45$, $p < .001$). Four-year-old children also outperformed 3-year-old children, $t(65) = 2.33$, $p = .023$. Fisher's LSD correction was applied to these t-tests and all subsequent follow-up t-tests reported unless otherwise noted. There was no significant main effect of age in the Another One Label Condition or the No Label Condition (all *p*s > .141).

Children's accuracy relative to chance performance (i.e., 50% target image selection, or three trials out of six) in each condition was evaluated with a two-tailed single sample t-test for each of the conditions. Accuracy in the Opposite Label Condition was significantly above chance, $t(67) = 5.02$, $p <. 001$. This effect was driven by both 5- and 4-year-old children performing above chance ($t(19) = 5.71$, $p < .001$ and $t(27) = 3.31$, $p = .003$ respectively). Three-year-old children performed at chance, $t(19) = 0.00$, $p = 1.00$. There were four 3-year-old children who performed above chance, however there were too few to analyze in a statistically meaningful manner. These tests remain significant even after applying a Bonferroni correction. In contrast, children's accuracy was not significantly above chance in the Another One Label Condition or the No Label Condition (all *p*s > .452).

To evaluate whether children learned about the concept of opposite within the study (across the six study trials), children's accuracy in the first set of three trials (block 1; $M = 0.52$, $SD = 0.31$) was compared to their accuracy in the second set of three trials (block 2; $M = 0.57$, $SD = 0.30$) with a 2 (block: 1, 2) x 3 (age: 3-, 4-, 5-years-old) x 3 (condition: Opposite Label Condition, Another One Label Condition, No Label Condition) mixed design ANOVA. Accuracy on the two blocks of trials was not significantly different, $F(1, 195) = 2.39$, $p = .124$. Furthermore, accuracy as a function of block did not interact with either condition or age (all *p*s > .903). Together this suggests that children's accuracy did not improve over the course of the Opposite Task, suggesting that children did not learn about the concept of opposites from the Opposite Task.

To evaluate whether some word pairs were easier or more difficult than other word pairs (i.e., item effects due to differences in ages at which different opposite word pairings are acquired), children's accuracy on each word pair was assigned a score (0-1). If the child selected the target the child received a score of 1 for that word pair. If the child selected the distractor the child received a score of 0 for that word pair. Mean accuracy for each word pair in each of the conditions is presented in Table 2.

Paired samples t tests were used to compare children's performance. In the Opposite Label Condition, children's accuracy was best with the word pair *big – small* and worst with the word pair *awake – asleep*, $t(67) = 2.55$, $p = .013$. Additional comparisons were not evaluated because the Bonferroni correction for making so many comparisons would be prohibitively conservative. Since *big – small* seemed to be the easiest word pair for children to learn, I also examined the accuracy of the 3-year-old group on this pair alone relative to chance performance to test for earlier appreciation of the opposite concept. Results showed that 3-year-old children's performance on the *big – small* word pair was not significantly above chance, $t(19) = 0.89$, $p = .385$.

Correlations & Demographics

The relationships between children's accuracy and the measures of working memory, receptive vocabulary (raw score on the TELD3), and communication skills were evaluated with partial correlations, correcting for children's age. Since children appeared to be employing distinctly different strategies in the three conditions, and performance was at chance in the Another One Label Condition and the No Label Condition, the subsequent analyses were only performed on the Opposite Label Condition data. Children's scores for each of these measures in the Opposite Label Condition are presented in Table 3, in the Another One Label Condition in Table 4, and the No Label Condition in Table 5.

Condition differences on the TELD3, the object memory task, the general composite score on the CCC-2, the syntax subscale of the CCC-2 and the context

Table 2

Children's Mean Proportions of Target Item Selections for Each Word Pair

		Age							
		3		4		5		Overall	
	Word Pair	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Opposite Label Condition	Awake – Asleep	0.25	0.44	0.57	0.50	0.75	0.44	0.53	0.50
	Clean – Dirty	0.50	0.51	0.59	0.50	0.90	0.31	0.66	0.48
	Happy – Sad	0.50	0.51	0.71	0.46	0.85	0.37	0.69	0.47
	Old – Young	0.55	0.51	0.59	0.50	0.80	0.41	0.64	0.48
	Small – Big	0.60	0.50	0.75	0.44	0.75	0.44	0.71	0.46
	Tall - Short	0.55	0.51	0.61	0.50	0.70	0.47	0.62	0.49
Another One Label Condition	Awake – Asleep	0.20	0.41	0.29	0.46	0.30	0.47	0.26	0.44
	Clean – Dirty	0.20	0.41	0.46	0.51	0.45	0.51	0.38	0.49
	Happy – Sad	0.60	0.50	0.71	0.46	0.55	0.51	0.63	0.49
	Old – Young	0.55	0.51	0.57	0.50	0.80	0.41	0.63	0.49
	Small – Big	0.60	0.50	0.81	0.40	0.65	0.49	0.70	0.46
	Tall – Short	0.5	0.51	0.39	0.50	0.53	0.51	0.40	0.50
No Label Condition	Awake – Asleep	0.45	0.51	0.29	0.46	0.15	0.37	0.29	0.46
	Clean – Dirty	0.55	0.51	0.57	0.50	0.58	0.51	0.57	0.50
	Happy – Sad	0.45	0.51	0.68	0.48	0.80	0.41	0.65	0.48
	Old – Young	0.45	0.51	0.36	0.49	0.60	0.50	0.46	0.50
	Small – Big	0.40	0.50	0.57	0.50	0.55	0.51	0.51	0.50
	Tall – Short	0.50	0.51	0.43	0.50	0.35	0.49	0.43	0.50

Table 3

Children's Mean Scores on Additional Measures and Demographic Characteristics in the Opposite Label Condition

	Age							
	3		4		5		Overall	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Object Memory	6.00	5.71	14.35	5.00	21.79	3.57	15.23	7.33
TELD3								
Percentile Rank	88.10	12.77	77.26	28.10	51.70	30.43	72.87	28.96
Raw Score	22.70	4.41	27.78	4.41	30.80	3.62	27.16	5.22
CCC-2 GCC								
Percentile Rank	–	–	49.52	29.52	50.95	25.41	50.11	27.61
Raw Score	–	–	78.70	15.85	80.26	11.70	79.35	14.16
CCC-2 Syntax								
Percentile Rank	–	–	58.22	28.55	59.37	24.22	58.70	26.58
Raw Score	–	–	3.41	3.84	2.21	2.84	2.91	3.48
CCC-2 Context								
Percentile Rank	–	–	51.31	28.54	46.32	22.40	49.20	25.97
Raw Score	–	–	6.59	3.38	5.79	2.37	6.26	3.00
Age Started Preschool/Daycare	1.25	0.62	2.65	0.81	2.65	0.86	2.21	1.00
Number of Children	2.05	0.52	2.15	0.78	2.45	0.83	2.22	0.74

Note. The CCC-2 is not intended for use with children under the age of 4.

Table 4

Children's Mean Scores on Additional Measures and Demographic Characteristics in the Another One Label Condition

	Age							
	3		4		5		Overall	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Object Memory	14.80	4.02	15.52	5.12	21.16	2.54	17.30	5.00
TELD3								
Percentile Rank	87.85	13.35	90.89	15.03	52.30	27.47	78.65	25.51
Raw Score	22.90	4.76	28.96	2.76	28.95	2.87	27.18	4.42
CCC-2 GCC								
Percentile Rank	–	–	51.26	29.22	48.50	26.97	50.09	28.02
Raw Score	–	–	80.22	15.71	79.00	15.44	79.70	15.44
CCC-2 Syntax								
Percentile Rank	–	–	58.89	21.68	51.50	29.23	55.74	25.14
Raw Score	–	–	2.85	2.33	3.40	3.78	3.09	3.00
CCC-2 Context								
Percentile Rank	–	–	51.00	23.57	52.25	25.15	51.53	24.00
Raw Score	–	–	6.89	2.90	5.20	2.95	6.17	3.01
Age Started Preschool/Daycare	1.64	0.74	2.68	0.72	2.47	0.87	2.34	0.88
Number of Children	1.50	0.63	2.04	0.74	2.45	0.76	2.03	0.80

Note. The CCC-2 is not intended for use with children under the age of 4.

Table 5

Children's Mean Scores on Additional Measures and Demographic Characteristics in the No Label Condition

	Age							
	3		4		5		Overall	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Object Memory	15.22	3.56	18.54	3.40	20.95	4.27	18.87	4.19
TELD3								
Percentile Rank	90.30	8.52	81.93	23.95	85.85	16.23	85.54	18.40
Raw Score	23.15	4.58	28.04	3.74	31.25	2.29	27.54	4.80
CCC-2 GCC								
Percentile Rank	–	–	34.74	25.18	63.00	23.60	46.41	28.06
Raw Score	–	–	70.96	13.77	85.53	13.39	76.98	15.29
CCC-2 Syntax								
Percentile Rank	–	–	43.96	22.05	64.26	23.23	52.35	24.47
Raw Score	–	–	4.85	3.44	1.84	2.54	3.61	3.42
CCC-2 Context								
Percentile Rank	–	–	37.28	28.53	53.21	20.31	43.86	26.43
Raw Score	–	–	8.78	4.37	5.16	2.29	7.28	4.04
Age Started Preschool/Daycare	1.27	0.70	2.34	0.90	2.76	1.09	2.24	1.06
Number of Children	2.11	0.81	1.81	0.74	2.35	0.88	2.06	0.82

Note. The CCC-2 is not intended for use with children under the age of 4.

subscale of the CCC-2 were each evaluated with a one-way (condition: Opposite Label Condition, Another One Label Condition, No Label Condition) between subjects ANOVA. There was a significant main effect of condition on the TELD3, $F(2, 200) = 4.47$, $p = .013$. Children in the No Label Condition scored significantly higher percentile ranks than children in the Opposite Label Condition, $t(197) = 2.99$, $p = .003$. Children in the Another One Label Condition did not differ from children in the Opposite Label Condition or the No Label Condition (all *p*s > .105). There was also a significant main effect of condition on the object memory task, $F(2, 164) = 5.76$, $p = .004$. Children in the No Label Condition scored significantly higher than children in the Opposite Label Condition, $t(161) = 3.38$, $p = .001$. Children in the Another One Label Condition scored marginally higher than children in the Opposite Label Condition ($t(161) = 1.93$, $p = .055$) but did not differ from children in the No Label Condition ($t(161) = 1.46$, $p = .147$). There was no significant main effect of condition on the general composite score of the CCC-2 or either of the subscales evaluated (all *p*s > .334).

Accuracy on the Opposite Task was negatively correlated with scores on the CCC-2 Context subscale, $r(41) = -.327$, $p = .029$. This relationship suggests that children who displayed better performance on the Context subscale (a lower raw score) also performed better on the Opposite Task. In contrast, children's performance on the Opposite Task was not correlated with their scores on the CCC-2 GCC or the CCC-2 Syntax subscale (all *p*s > .215). Likewise, children's performance on the Opposite Task in was not correlated with their performance on the object memory task or the TELD3 (all *p*s > .578).

The relationships between children's accuracy in the Opposite Label Condition and demographic characteristics were evaluated with chi-square analyses and Pearson correlations. There was a significant association between children performing above chance (four or more trials out of six) on the Opposite task and whether parents reported having books or games that are focused on opposites in the home, $\chi^2(1) = 6.12$, $p = .013$. Children whose parents reported that they had books or games focused on opposites were more likely to perform above chance than those who did not have access to these materials in the home. There was a marginally significant association between children performing above chance on the Opposite task and whether parents reported playing opposite games with their child ($\chi^2(1) = 3.22$, $p = .073$). Children whose parents reported that they played opposite games with their children were somewhat more likely to perform above chance than those who reported that they did not play opposite games with their children. There was no association between children performing above chance on the Opposite Task and the presence of other children in childcare ($p > .597$). There was also no relationship between accuracy on the Opposite Task in the Opposite Label Condition and the number of children residing in the home ($r(63) = .19$, $p = .130$). Finally, there was no relationship between accuracy and the age at which children began preschool or daycare ($r(49) = .05$, $p = .706$), after correcting for children's age.

Processing Measures: Response Latency and Eye Gaze

Similar to the correlational and demographic analyses, eye gaze analyses were only performed on the Opposite Label Condition data. Recall that the Early phase refers to the time period between when the experimenter began voicing the request to when the child initiated movement, the Middle phase refers to the time period between the initiation of movement to when the child touched the image, and the Late phase refers to the time period between when the child touched the image to when the child placed the image in the box. Eye gaze and latency measures were included in the design of this study to evaluate whether children show evidence of an implicit understanding of antonymy before they are able to perform above chance. Differences as a function of age, depending on the particular pattern of the effect, might be indicative of implicit understanding. For example, if 3-year-old children spent proportionally more time looking at the target image on trials that ended with an incorrect selection than did 4- and 5-year-old children it would suggest that they had some awareness of the correct answer but were unable to overcome whatever other rule or preference was guiding their behaviour.

Response latencies. Children's response latencies (overall, Early phase, Middle phase, and Late phase evaluated individually) were evaluated with one-way ANOVAs to test for effects of age, with trials ending with a correct and incorrect selection evaluated separately. Mean latencies to each timing point on trials ending with a correct selection are presented in Table 6.

On trials ending with a correct selection there was no main effect of age on total response latency, $F(2, 62) = 1.09$, $p = .342$, or in the first two time phases (all ps > .217). There was a marginally significant effect in the Late phase, $F(2, 62) = 2.87$, $p = .064$. Five-year-old children were significantly faster than 3-year-old children, $t(62) = 2.31$, $p = .025$. Four-year-old children were marginally faster than 3-year-old children, $t(62) = 1.83$, $p = .072$, but did not differ from 5-year-old children, $t(62) = 0.65$, $p = .519$. Mean latencies to each timing point on trials ending with an incorrect selection are presented in

Table 6

Children's Mean Response Latencies (in ms) in the Opposite Label Condition on Trials Ending with a Correct Selection

	Early phase		Middle phase		Late phase		Overall	
Age	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
3	1218.00	880.46	1195.70	695.24	1869.36	1093.95	3810.40	960.61
4	1444.55	865.27	1187.69	661.33	1477.89	498.14	3949.63	1278.07
5	1227.92	1046.04	897.02	465.12	1343.60	445.00	3407.14	1470.30
Overall	1313.35	921.40	1100.47	623.69	1544.98	722.25	3744.15	1266.90

Note. The sum of the rows do not equal the overall value because some children were missing all data in the Early phase because they always initiated movement prior to the voicing of the response request.

Table 7. On trials ending with an incorrect selection there was no main effect of age on total response latency or any of the three individual time phases (all *p*s > .221).

Eye Gaze Fixations. Children's average numbers of fixations were evaluated with a one-way ANOVA to test for effects of age, with trials ending with a correct and incorrect selection evaluated separately. Mean number of fixations and mean proportions of first looks to the correct image (of the two images they could select) for trials ending with a correct or an incorrect selection are presented in Table 8.

Children's mean numbers of fixations and mean proportions of first looks to the correct image were each evaluated with a one-way ANOVA to test for effects of age for trials ending with a correct and incorrect selection separately. There was no main effect of age for the average number of fixations made on trials ending with a correct or an incorrect selection (all *p*s > .154). In contrast to the overall average number of fixation data, a marginally significant main effect of age was found for the mean proportions of first looks to the correct image (of the two images children could select) made on trials ending with a correct selection ($F(2, 62) = 2.82$, $p = .067$). Five-year-old children made a significantly greater mean proportion of correct first fixations than 3-year-old children $t(62) = 2.32$, $p = .024$. Four-year-old children did not differ from 5- or 3-year-old children in the mean proportion of correct first looks made on trials ending with a correct selection ($t(62) = 1.65$, $p = .103$ and $t(62) = 0.87$, $p = .386$ respectively). There was no main effect of age for the mean proportion of first looks to the correct image made on trials ending with an incorrect selection ($F(2, 43) = 1.69$, $p = .197$).

Table 7

Children's Mean Response Latencies (in ms) in the Opposite Label Condition on Trials Ending with an Incorrect Selection

	Early phase		Middle phase		Late phase		Overall	
Age	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
3	1310.43	745.52	1099.16	830.43	2078.81	883.42	4334.23	1143.61
4	1750.29	1280.93	1248.50	698.72	1782.83	869.27	4519.08	1828.58
5	2087.59	948.60	872.67	396.64	1625.56	1312.72	4205.09	1873.87
Overall	1652.81	1059.33	1119.78	706.42	1861.44	966.75	4393.42	1578.94

Note. The sum of the rows do not equal the overall value because some children were missing all data in the Early phase because they always initiated movement prior to the voicing of the response request.

Table 8

Mean Number of Fixations and Mean number of Correct First Fixations in the Opposite Label Condition

	Correct selection				Incorrect selection			
	Number of fixations		Number of correct first fixations		Number of fixations		Number of correct first fixations	
Age	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
3	4.00	1.08	0.62	0.45	4.37	1.27	0.28	0.39
4	4.46	1.61	0.71	0.32	5.29	1.85	0.43	0.40
5	4.36	1.55	0.87	0.16	5.46	1.71	0.56	0.33
Overall	4.30	1.45	0.74	0.34	4.98	1.66	0.40	0.39

Eyegaze Total looks. Eye gaze was also examined in terms of duration of looks to specific locations (target, distractor, experimenter, and image selected by the experimenter; as a proportion of the total looking time). A higher proportion reflects more time spent looking at a particular location and less time looking at other locations. Mean proportions of time spent looking at the target image on trials ending with a correct selection are presented in Table 9.

Table 9

Mean Proportion of Time Spent Looking at the Target Image in the Opposite Label Condition on Trials Ending with a Correct Selection

	Early phase		Middle phase		Late phase		Overall	
Age	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
3	0.43	0.34	0.69	0.24	0.46	0.18	0.51	0.17
4	0.35	0.23	0.67	0.19	0.28	0.18	0.44	0.13
5	0.32	0.28	0.72	0.20	0.27	0.14	0.39	0.12
Overall	0.36	0.28	0.69	0.21	0.33	0.19	0.44	0.15

Children’s proportions of time spent looking at each location, coded in each time phase of the trial (overall, Early phase, Middle phase, and Late phase evaluated individually), was evaluated with one-way ANOVAs to test for effects of age, with trials ending with a correct and incorrect selection evaluated separately. On trials ending with a correct selection there was a marginally significant main effect of age for the proportion of time spent looking at the target over the full trial, $F(2, 62) = 3.01$, $p = .056$. Five-year-old children spent proportionally less time looking at the target than 3-year-old children, $t(62) = 2.44$, $p = .018$. Four-year-old children did not differ from 5- or 3-year-old children in proportion of looking time spent on the target ($t(62) = 1.03$, $p = .307$ and $t(62) = 1.61$, $p = .114$ respectively). There was also a significant main effect of age for the proportion of time spent looking at the target in the Late phase of the trial, $F(2, 62) = 7.15$, $p < .001$. Five- and 4-year-old children spent proportionally less time looking at the target than 3-year-old children ($t(62) = 3.31$, $p = .002$ and $t(62) = 3.37$, $p = .001$ respectively). Five- and 4-year-old children did not differ from each other, $t(62) = 0.17$, $p = .864$. There were no main effects of age in the Early or Middle phases (all *p*s > .496). Mean proportions of time spent looking at the target image on trials ending with an incorrect selection are presented in Table 10.

Table 10

Mean Proportion of Time Spent Looking at the Target Image in the Opposite Label Condition on Trials Ending with an Incorrect Selection

	Early phase		Middle phase		Late phase		Overall	
Age	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
3	0.26	0.26	0.16	0.20	0.04	0.11	0.13	0.14
4	0.36	0.30	0.15	0.20	0.02	0.08	0.15	0.13
5	0.25	0.18	0.10	0.13	0.00	0.00	0.14	0.10
Overall	0.30	0.26	0.14	0.19	0.02	0.08	0.14	0.13

On trials ending with an incorrect selection there was no main effect of age on the proportion of time spent looking at the target over the full trial or in any of the individual time phases (all *p*s > 459).

Mean proportions of time spent looking at the distractor image on trials ending with a correct selection are presented in Table 11.

Table 11

Mean Proportion of Time Spent Looking at the Distractor Image in the Opposite Label Condition on Trials Ending with a Correct Selection

	Early phase		Middle phase		Late phase		Overall	
Age	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
3	0.16	0.22	0.13	0.20	0.00	0.01	0.11	0.13
4	0.21	0.25	0.11	0.18	0.02	0.05	0.10	0.12
5	0.10	0.11	0.03	0.05	0.02	0.07	0.06	0.06
Overall	0.16	0.21	0.09	0.16	0.02	0.05	0.09	0.11

On trials ending with a correct selection there was no main effect of age for the proportion of time spent looking at the distractor over the full trial, or in any of the individual time phases (all *p*s > .107). Mean proportions of time spent looking at the distractor image on trials ending with an incorrect selection are presented in Table 12.

On trials ending with an incorrect selection there was a main effect of age on the proportion of time looking at the distractor over the full trial, $F(2, 43) = 4.86$, $p = .013$. Five-year-old children spent proportionally less time looking at the distractor than 3-year-

Table 12

Mean Proportion of Time Spent Looking at the Distractor Image in the Opposite Label Condition on Trials Ending with an Incorrect Selection

	Early phase		Middle phase		Late phase		Overall	
Age	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
3	0.33	0.23	0.71	0.23	0.50	0.27	0.49	0.16
4	0.27	0.23	0.64	0.26	0.34	0.18	0.40	0.15
5	0.23	0.17	0.53	0.28	0.28	0.17	0.31	0.10
Overall	0.28	0.22	0.64	0.25	0.39	0.23	0.42	0.16

children, $t(43) = 3.06$, $p = .004$. Four-year-old children did not differ from 5- or 3-year-old children in proportion of looking time spent on the distractor ($t(43) = 1.62$, $p = .113$ and $t(62) = 1.85$, $p = .071$ respectively). There was also a significant main effect of age for the proportion of time spent looking at the distractor in the Late phase of the trial, $F(2, 43) = 3.66$, $p = .034$. Five- and 4-year-old children spent proportionally less time looking at the distractor than 3-year-old children ($t(43) = 2.38$, $p = .022$ and $t(43) = 2.18$, $p = .034$ respectively). Five- and 4-year-old children did not differ from each other, $t(43) = 0.65$, $p = .516$. There was no main effect of age in the Early or Middle phases (all *p*s > .211).

Mean proportions of time spent looking at the experimenter on trials ending with a correct selection are presented in Table 13. On trials ending with a correct selection there was no main effect of age for the proportion of time spent looking at the experimenter over the full trial, or any of the individual time phases (all *p*s > .141).

Table 13

Mean Proportion of Time Spent Looking at the Experimenter in the Opposite Label Condition on Trials Ending with a Correct Selection

	Early phase		Middle phase		Late phase		Overall	
Age	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
3	0.32	0.33	0.08	0.15	0.07	0.13	0.14	0.15
4	0.24	0.26	0.14	0.19	0.15	0.18	0.16	0.16
5	0.29	0.29	0.14	0.15	0.18	0.19	0.21	0.16
Overall	0.28	0.29	0.12	0.17	0.14	0.18	0.17	0.16

Mean proportions of time spent looking at the experimenter on trials ending with an incorrect selection are presented in Table 14. On trials ending with an incorrect selection there was a main effect of age on the proportion of time looking at the experimenter in the Middle phase, $F(2, 43) = 4.97$, $p = .011$. Five-year-old children spent proportionally more time looking at the experimenter than 4- and 3-year-old children ($t(43) = 2.78$, $p = .008$ and $t(43) = 2.97$, $p = .005$ respectively). Four- and 3-year-old children did not differ from each other in proportion of time spent looking at the experimenter, $t(43) = 0.32$, $p = .750$. There was no main effect of age over the full trial, or in the Early or Late phases (all *p*s > .105).

Mean proportions of time spent looking at the image selected by the experimenter on trials ending with a correct selection are presented in Table 15. On trials ending with a correct selection there was no main effect of age for the proportion of time spent looking at the image selected by the experimenter over the full trial, or in any of the individual

Table 14

Mean Proportion of Time Spent Looking at the Experimenter in the Opposite Label Condition on Trials Ending with an Incorrect Selection

	Early phase		Middle phase		Late phase		Overall	
Age	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
3	0.22	0.29	0.05	0.09	0.12	0.20	0.15	0.15
4	0.15	0.19	0.07	0.15	0.18	0.20	0.16	0.15
5	0.32	0.36	0.26	0.29	0.27	0.24	0.28	0.17
Overall	0.21	0.27	0.10	0.19	0.17	0.21	0.18	0.16

Table 15

Mean Proportion of Time Spent Looking at the Image Selected by the Experimenter in the Opposite Label Condition on Trials Ending with a Correct Selection

	Early phase		Middle phase		Late phase		Overall	
Age	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
3	0.10	0.25	0.10	0.13	0.47	0.16	0.24	0.09
4	0.20	0.19	0.08	0.10	0.54	0.23	0.29	0.12
5	0.27	0.28	0.10	0.08	0.50	0.17	0.31	0.10
Overall	0.20	0.24	0.09	0.11	0.51	0.19	0.28	0.11

time phases (all *p*s > .122). Mean proportions of time spent looking at the image selected by the experimenter on trials ending with an incorrect selection are presented in Table 16. Similarly, on trials ending with a incorrect selection there was no main effect of age for

Table 16

Mean Proportion of Time Spent Looking at the Image Selected by the Experimenter in the Opposite Label Condition on Trials Ending with an Incorrect Selection

	Early phase		Middle phase		Late phase		Overall	
Age	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
3	0.19	0.25	0.08	0.15	0.34	0.18	0.23	0.12
4	0.20	0.23	0.13	0.16	0.46	0.17	0.26	0.08
5	0.20	0.17	0.10	0.11	0.45	0.23	0.28	0.10
Overall	0.19	0.22	0.10	0.15	0.41	0.19	0.25	0.10

the proportion of time spent looking at the image selected by the experimenter over the full trial, or any of the individual time phases (all ps > .152).

Discussion

The purposes of the present experiment were a) to determine the development of understanding for the concept of antonymy with a novel, less verbal task than used in previous research, b) to determine whether there are other cognitive developments related to the development of the concept of antonymy, and c) to evaluate whether there was evidence of a sensitivity to, or an implicit understanding of, the concept of antonymy prior to behavioural evidence of understanding.

Development of the Concept

The main predictions, that 4- and 5-year-old children would demonstrate an understanding of antonymy in the Opposite Label Condition of the Opposite Task, and that accuracy on the task would improve with age, were supported. Four- and 5-year-old children, but not 3-year-old children, performed above chance. Furthermore, 5-year-old children outperformed 4-year-old children, who in turn outperformed 3-year-old children. Although the task difficulty may have been minimized by eliminating the need for a verbal response and by encouraging children's engagement with the task through the use of pictures of animals, the youngest children still did not demonstrate an understanding of antonymy. This is consistent with the conclusions made in previous metalinguistic studies of antonymy that 4-year-old children, but not 3-year-old children, understand the concept of antonymy (E. V. Clark, 1972; Morris, 2003). Like these studies, I found that the age of acquisition was considerably earlier than suggested by the results of the initial metalinguistic study conducted by Kreezer and Dallenbach (1929). I also found that children's accuracy improved between the ages of 4 and 5 which is consistent with Clark (1972) and with the general pattern of acquisition found by Kreezer and Dallenbach (1929) but inconsistent with the findings of Morris (2003).

It was further hypothesized that older children would be faster than younger children to make responses in the Opposite task. This was partially supported. The only effect of age was found in the Late response phase on trials in which the child selected the target image, and the effect was only marginally significant. Five-year-old children responded more quickly in this phase than 3-year-old children. Four-year-old children were marginally faster in this phase than 3-year-old children but did not differ from 5-year-old children.

Predictions about children's performance in the Another One Label Condition and the No Label Condition were also tested. In the Another One Label Condition, it was expected that children would perform at chance if they needed the concept label "opposite" in order to understand the task, but that they would perform above chance if a relationship between the adjectives was made salient by the pictures themselves or some other aspect of the task. The former was found to be true; children of all ages performed at chance levels when adjective labels were provided but they were not explicitly asked for the "opposite one". Furthermore, performance did not vary as a function of age. Anecdotally, a few children in the Another One Label Condition did behave as if the antonymy relation was too salient to be ignored. These children tended to explicitly state that they were selecting the opposite and performed above chance. Overall though, these children were a very small group and therefore could not be analyzed in more depth.

In No Label Condition, it was expected that children would perform at chance if the images used in the task did not bias children's responses such that they performed above chance. The goal was to design stimulus images that did not convey too much information about the adjectives; given that children of all ages performed at chance in No Label Condition, and that children's performance did not vary as a function of age, I believe that this goal was achieved. Thus, there was nothing in the stimulus materials themselves that made the opposite relationship salient. Children were only able to demonstrate appreciation for the opposite relation when explicitly asked to do so (in the Opposite Label Condition).

The design of my task allowed me to evaluate antonym appreciation in children younger than 4 years, which addressed one of the major limitations of previous metalinguistic studies of antonymy. Previous studies had concluded that children as young as 4 years had at least a nascent understanding of antonymy but had not evaluated younger children and thus were unable to conclude whether younger children might also understand the concept. In a pilot study, Clark (1972) found that no children under the age of 4 could understand her task, therefore they were not further evaluated in the two experiments she reported. The design of my task enabled the participation of even the youngest children I tested, and therefore allowed me to establish with more confidence a baseline age at which most children did not understand the concept of antonymy.

The adjective list used in the present study was limited by the design of my task; adjectives needed to be depicted using a single animal per image without additional objects in the image. For example, the adjective pair *full – empty* could not be depicted within these constraints. The adjective list was also limited by the age at which words are typically acquired. For example, while the adjective pair *wide – narrow* could be depicted within the aforementioned constraints, the words themselves are acquired at a much older age and therefore would be unnecessarily difficult for the younger children. What these limits mean in terms of the order of acquisition hypothesis is that, with few exceptions (*big – small* and *tall – short*; *big – small* and *old - young*), the adjective pairs came from different conceptual domains. This, in turn, means my conclusions are restricted to the age of acquisition of the words themselves and not inferences like those of Clark (1972) and others, which were based on the acquisition of pairs within a conceptual domain.

It was hypothesized that there would be an order of acquisition, with children performing most accurately on the word pairs *big – small*, *clean – dirty*, and *happy – sad*. Support for this hypothesis was not found. In contrast with some previous research (Bartlett, 1976; Brewer & Stone, 1975; E. V. Clark, 1972; Eilers et al., 1974; L. B. Smith et al., 1986), there was not strong evidence of an order of acquisition in my set of antonym pairs. Performance was indeed best on the word pair *big – small* and was significantly worse on the word pair *awake – asleep* but no other comparisons between the word pairs could be made because of the statistical corrections that would be necessary to account for the number of comparisons made. It was, however, possible that children's overall below chance performance was masking above chance performance on a few word pairs. In fact, prior research suggests that a few of the word pairs included in the present research, *big – small* and *happy – sad* in particular, should be among the first antonym pairs learned by children (E. V. Clark, 1972; Dale & Fenson, 1996). As such, one final comparison was made to evaluate 3-year-old children's performance on the word pair on which they displayed the best performance (*big – small*). Even on the word pair on which they performed best, they did not perform above chance. This suggests that the overall chance performance by the youngest children was not the result of the majority of word pairs being too difficult. However, it is also important to note that item differences were confounded with differences with the pictures because different pictures and different animals were used for each word pair and this was not counterbalanced between children.

A related goal of the current study was the evaluation of whether there were other cognitive developments related to the development of the concept of antonymy. It was hypothesized that certain skills would be related to the development of the concept. It was predicted that accuracy on the Opposite Task would be positively correlated with working memory and receptive vocabulary. Neither of these hypotheses was supported, perhaps due to the restricted range of scores observed in the sample, particularly with the receptive vocabulary task. There was a difference between the conditions on these two measures, however in both cases children in the No Label Condition outperformed children in the Opposite Label Condition. This makes it all the more striking that children in the No Label Condition did not perform better on the Opposite Task. Children in the Another One Label Condition also performed marginally better than children in the Opposite Label Condition on the working memory task, a result that again makes it more striking that they did not perform better. It was further predicted that accuracy would be negatively correlated with the general composite score on the CCC-2, as well as the CCC-2 syntax and context subscales. Only the scores on the CCC-2 use of conversational context subscale were significantly correlated with accuracy in the Opposite Task. Children whose parents judged them better able to make use of situational context to derive meaning from utterances also had better accuracy on the Opposite Task, supporting the idea that one way that children learn about antonymy is through conversations and communicative context (e.g., Jones & Murphy, 2005; Murphy & Jones, 2008). None of the measures associated with the CCC-2 differed between the three conditions.

Demographic data provided by parents revealed one significant association. Having access to books or games about opposites in the home was significantly associated with above chance performance on the Opposite Task. In addition, children who had experience playing games about opposites were marginally more likely to show above chance performance in the Opposite Task. One interpretation of this finding is that books and games about the concept of antonymy support children's acquisition of the concept. Alternatively, it may be that books and games stimulate conversations about opposites, which, in turn, helps children learn about antonymy. Of course, there could also be some additional unmeasured variable that is responsible for the relationship between exposure to books and games on opposites and children's appreciation of opposites. Other information collected, such as childcare and birth order, were not associated with performance on the Opposite task. This relationship could be further explored in future research via a training study involving shared book reading.

Implicit Understanding

Several of the processing and behavioural measures were included in order to evaluate whether the acquisition of antonymy was gradual or sudden. If children's understanding of antonymy developed gradually I might have been able to detect evidence of implicit understanding in the Opposite Label Condition even when children were unable to demonstrate their explicit understanding by choosing the target (i.e., on trials ending with the selection of the distractor, otherwise referred to as incorrect trials). Of particular interest was the behaviour of the 3-year-old children on trials ending with the selection of the distractor in the Opposite Label Condition. These trials could show sensitivity towards the selection of the target that the child then ignores or discounts in favour of the selection of the distractor. I found no evidence of implicit understanding in the eye gaze or behavioural measures; therefore it appears that children's understanding of antonymy develops quite suddenly. Of course, it is also possible that my measures were simply not sensitive enough to detect implicit awareness of the opposite concept.

It is also notable that children of all ages did not show improvement within the study itself. Performance on the first three trials did not differ from performance on the final three trials, suggesting that children did not learn the opposite concept within the study. As such, children did not behave as if they had an understanding of antonymy that was lying dormant that could be awakened through exposure to the Opposite Task.

The final piece of behavioural evidence against children showing gradual understanding of antonymy was that 3-year-old children did not perform above chance on *big – small*, the word pair that had the highest overall accuracy and was originally hypothesized to be among the word pairs that children should learn first, based on previous research and on the age at which the individual words are acquired. If these children had performed above chance on this word pair it would have suggested that perhaps the other word pairs in the task were too difficult for the youngest children, or that perhaps children's understanding of antonymy begins with a few word pairs and then grows gradually by expanding this understanding to additional word pairs.

Additional evidence against the gradual acquisition of the concept of antonymy comes from the eye gaze results (recall that only the eye gaze data from the Opposite Label Condition were analyzed). Five-year-old children had a significantly higher mean proportion of first looks to the target image than 3-year-old children on trials ending with the selection of the target. This suggests that the older children had a more developed concept of antonymy than the younger children because it shows a faster decision process, a result that is echoed in the pattern of results found in the proportion of time spent looking at the target and the distractor. The older children may have been able to generate the answer to the task while hearing the adjectival labels. This would not be particularly surprising given that previous research has shown that 4- and 5-year-old children can successfully perform in an opposite task in which they are required to generate the opposite on their own (E. V. Clark, 1972; Morris, 2003). The older children in the present study may have been able to come up with the correct answer earlier in the process, perhaps when the label of the middle image was confirmed.

There was no effect of age on the mean proportion of first looks to the target image on trials ending with the selection of the distractor, providing no evidence of implicit understanding of antonymy. If 3-year-old children had a higher mean proportion of first looks to the target image on trials ending with the selection of the distractor it would have been suggestive of a sensitivity to the correct answer in the absence of explicit behavioural evidence of antonymic understanding. It might have suggested that the children were tending towards selecting the correct answer but were unable to overcome whatever other rule might have been guiding their behaviour (e.g., select the image on the left). The lack of a greater mean proportion of first looks in the older children also suggests that the older children did not have a nascent understanding of antonymy for the word pairs that they got incorrect. If they had shown this pattern it would have suggested that their concept of antonymy, which was demonstrated by above chance performance on the Opposite task, was still developing and that additional word pairs were being added to their concept of antonymy in a gradual fashion.

In terms of duration of looks to a specific location (target, distractor, experimenter and image selected by the experimenter), younger children tended to look proportionally longer at the image that they would ultimately select. On trials ending with the selection of the target, 5-year-old children spent proportionally less time looking at the target image than 3-year-old children over the full response. In the Late phase of the response both 4- and 5-year-old children spent proportionally less time looking at the target image than 3-year-old children. The pattern was the same for the proportion of time spent looking at the distractor on trials ending with the selection of the distractor. There was an effect of age on the proportion of time spent looking at the distractor over the full response. Five-year-old children spent proportionally less time looking at the distractor than 3-year-old children. In the Late phase of the response, again, both 4- and 5-year-old children spent proportionally less time looking at the distractor image than 3-year-old children. This pattern suggests that the younger children were more uncertain of their decisions and that older children were especially more decisive at the end of their response process. Once the older children had touched the image that they would select they did not spend much additional time looking at their selection.

There was no effect of age for the proportion of time spent looking at the target on trials ending with the selection of the distractor or the proportion of time spent looking at the distractor on trials ending with the selection of the target. The hypothesis of an implicit understanding of antonymy in 3-year-old children would have been supported if they had spent proportionally more time looking at the target on trials ending with the selection of the distractor. Similar to the pattern of results with the first look behaviour, this would have suggested that the younger children were sensitive to what the correct answer was but were unable to overcome whatever preference or bias was guiding their behaviour. Instead, similar to the first look results, this pattern of results suggests that the older children are more confident in their responses than the 3-year-old children and does not support the hypothesis that 3-year-old children have an implicit understanding of antonymy in the absence of explicit behavioural responses.

Unexpectedly, there was an effect of age on the proportion of time spent looking at the experimenter in the Middle phase on trials ending with the selection of the distractor. Five-year-old children spent proportionally more time looking at the experimenter than 4- and 3-year old children. Perhaps this occurred because 5-year-old children have a more developed theory of mind and therefore may have hoped that the experimenter might be able to provide clues as to which image to select.

On balance, the results tend to favour the interpretation that the acquisition of antonymy is through sudden insight rather than being a gradual process; there was little evidence of an order of acquisition of the word pairs and there was no evidence of implicit awareness in the eye gaze and response latency measures in the absence of explicit behavioural evidence. This inconsistency with previous literature on the topic is at least in part because of how sudden insight was defined and in part because of how the acquisition of antonymy was evaluated. Previous research evaluated the acquisition of antonymy with several word pairs from the same conceptual domain, and argued that acquisition was gradual based on results showing the progression of antonymy being first applied to more general word pairs and then to more specific word pairs within the conceptual domain. In contrast, the majority of my word pairs came from different conceptual domains, and I base my conclusions on eye gaze behaviours instead of the progression of acquisition with multiple word pairs. In addition I defined sudden insight differently than Kreezer and Dallenbach (1929) in that I do not consider sudden insight to be equivalent with the acquisition antonymy as it relates to all possible antonym pairs. Rather, for me, sudden insight is considered to be the rapid acquisition of antonymy for the majority of antonym pairs for which the child knows both members of the pair. Also, in contrast to Kreezer and Dallenbach (1929), I do not consider the use of negation as evidence of an understanding of antonomy. Therefore, although Kreezer and Dallenbach (1929) and I use similar phrases to describe the acquisition of antonymy we are not describing the same pattern of acquisition. However, my results also diverge somewhat with the more gradual acquisition of antonymy found by other research groups because I did not find evidence of an order of acquisition within my set of word pairs. However, this difference may be due to methodological differences; the word pairs utilized in the present study, with few exceptions, were from different conceptual domains. It is possible that I would have found item differences supporting a more gradual acquisition process if I had more word pairs from fewer conceptual domains. Because of the differences between my research and previous research I do not believe that direct comparisons are advisable.

Limitations

One limitation that I have alluded to is the fact that the word pairs evaluated in the present study belonged to a variety of conceptual domains. As mentioned, this limits the comparisons that can be made in regards to whether the concept of antonymy develops gradually or with sudden insight. The word pairs included in the present study were selected based on previous research as well as the age at which children could be expected to have acquired the individual terms and the ease with which the words could be depicted using animals without the use of a second animal in the image or the use of inanimate objects. Future research should investigate children's understanding of antonymy with a larger set of word pairs. Manipulating the complexity of the images by including an inanimate object in the image of each animal, or by including images that do not contain animals at all, would allow for the inclusion of a greater number of word pairs (e.g., *full – empty*). This manipulation might reduce task difficulty because children are able to understand the influence of context on gradable adjectives (Syrett et al., 2010). Children understand that an object (or animal) that is considered tall in one context can also be considered short in a different context. The inclusion of additional cues could make the dimension of interest and the context of comparison more salient. The inclusion of additional word pairs would increase the comparisons that could be made with previous research in terms of whether the acquisition of antonymy should be considered a gradual process or one that occurs more suddenly. However, before expanding the set of word pairs evaluated, it may be advisable to re-evaluate children's performance on the current set of adjective pairs with the more complex images.

Another limitation of the current design of the Opposite Task is that, in principle, a child could perform above chance simply by understanding that there is some kind of association between the words that form the antonym pair without necessarily understanding that the association is one of antonymy. This is possible because the distractor is unrelated to the word pair. This was a strategic choice because I wanted to make the Opposite Task as easy as possible for this initial evaluation, with the goal that future studies would make systematic changes to the Opposite Task in order to evaluate this and other limitations of the present design. The simpler version of the Opposite Task was evaluated first to ensure that the results of previous studies, in particular above chance performance by 4-year-old children, could be replicated with my novel task. Although children could be simply noticing an association between the word pairs, evidence from the Another One Label Condition in the present study suggests that this is not likely the case. When children were presented with the adjectival labels but not the concept name they did not perform above chance; the adjectival labels themselves were not sufficient for above chance performance. If the children in the Opposite Label Condition were simply noticing the relationship between the words in an antonymous word pair, the children in the Another One Label Condition should have also performed above chance. That they did not suggests that children in the Opposite Label Condition were paying attention to the specific relationship mentioned in the instructions. A more stringent evaluation of this limitation could involve the use of a related distractor word, such as a synonym of the stimulus word (e.g., *big dog – little dog – small dog*), within the general format of the Opposite Task. In this situation the stimulus word (*little dog*) would have a lexical association with both the target word (*big dog*) and the distractor word (*small dog*). Such a design would make it less likely that children could perform above chance by simply observing there was an association between the stimulus word and the target word.

The final methodological limitations I wish to address with regards to the design of the present experiment are the timing of the eye gaze analysis and the visual resolution limitations, especially with regards to the coding of eye gaze when children are looking at the image selected by the experimenter. The Early phase of the processing measured in the present study does not capture the earliest time at which the child could begin to make a decision about which image to select. Processing was measured beginning with the request made by the experimenter ("Can you put the opposite one in the box?"). Given that the experimenter always put the middle image into the box, the child was given all the information they needed to make an image choice once the images were labelled. Therefore, the Early response phase of the Opposite task is only early relative to the Middle and Late response phases of the task and does not represent the earliest point in time during which the child could make decisions about their image selection. In fact, it was not unusual for children to begin making a reach before the experimenter began voicing the request. The delayed timing of the eye gaze analysis might have hidden evidence of a latent understanding of antonymy. If children have largely decided which image they are going to select by the time the experimenter makes the request, then it would not be surprising that there were few differences as a function of age and that there was little evidence of latent understanding of antonymy. If I had been able to measure children's eye gaze analysis at an earlier point in the task I might have been able to capture more of the child's decision making process as it unfolded.

One way to address this limitation would be to include more practice trials so that children would become more familiar with the Opposite task. With an increased number of practice trials children could be trained to respond after the experimenter placed an image into the box, and to do so without waiting for the explicit prompt. In theory, the eye gaze analysis could then begin before the request was vocalized. This would allow measurement of the Early phase of processing to begin earlier. To increase the amount of training the children receive the list of word pairs used would need to be expanded, thus this modification would be best implemented in conjunction with the expansion of the set of word pairs evaluated that I mentioned previously.

Children were considered to be looking at the image that the experimenter selected if they were looking at the image itself or at the response box once the image had been placed in the box. Given the resolution of the video, it was not possible to differentiate whether children were looking at the image or whether they were looking at the box once the image had been placed in the box. As a result, it is not surprising that there was no effect of age on the proportion of time spent looking at the image selected by the experimenter, over the full trial or in any of the time phases, on trials ending with the selection of the target or those ending with the selection of the distractor. This could be addressed in a couple of ways. One method would be to upgrade the video camera used so that the resolution of the video is improved and therefore the person coding the video could use a digital zoom while evaluating the recording. This modification would not change the child's experience and would not be expected to alter the Opposite task itself. An alternative method of addressing this limitation could be altering the presentation format of the task so that it was presented on a tablet computer through eye tracking software such as Eye Link II or Tobii. This would also improve the resolution of the eye gaze analysis but this change to the methodology would need to be evaluated with a pilot study to ensure that the Opposite Task can successfully be migrated to a different response methodology.

Further Future Directions

In addition to the ideas for follow-up studies that I mentioned in relation to the limitations of the present study, I believe at least two other directions are worthy of note. Future research on children's understanding of antonymy could include an evaluation of a) whether understanding the concept of antonymy helps children understand verbal irony and b) whether the development of the concept of antonymy differs across cultures.

In the simplest form of verbal irony the literal meaning of a remark is the opposite of the intended meaning (Climie & Pexman, 2008; Pexman, 2008). It seems reasonable that understanding the concept of antonymy could aid in the appreciation of verbal irony, though it is unlikely that it is the sole contributing factor to said appreciation. Since the concept of antonymy is unlikely to be the sole conceptual development that supports the development of an appreciation of verbal irony, other cognitive measures that could be related to the appreciation of verbal irony, such as theory of mind, executive function tasks and language skills, would need to be evaluated in conjunction with the opposite task in this follow-up study (Carlson, 2005; Carlson & Moses, 2001; Filippova & Astington, 2008; Pexman, 2008). To test irony comprehension, I suggest using a procedure similar to that developed by Climie and Pexman (2008) in which children indicate their judgments of speaker intent non-verbally by placing objects in an answer box while being video recorded. By this procedure, children's eye gaze can be monitored as they listen to the ironic utterance and make their interpretation of speaker intent. Thus, it would be possible to determine how quickly children look to the object that corresponds to ironic intent and thereby assess what they are "considering" and when they begin "considering" the alternative responses. It is possible that an appreciation for the concept of antonymy affects the earliest moments of processing an ironic utterance. This evaluation of irony comprehension, in conjunction with an evaluation of children's understanding of antonymy, would allow me to determine if this is the case.

Based on the results of the present study, I believe that the current design of the Opposite Task is best suited for 4-year-old children. Therefore, the major challenge associated with this direction of research would be to alter the difficulty of the Opposite Task and the irony task used by Climie and Pexman (2008) such that children would show sufficient variability and success in both tasks. This would involve an increase in the difficulty of the Opposite Task and a decrease in the difficulty of the irony task. To achieve the former the Opposite Task could be altered to include a synonym distractor. Alternatively the images included in the Opposite Task could be altered so that each image would depict a relationship between two animals. To succeed, children would need to abstract the relationship in the target image (e.g., *sad rabbit – happy rabbit*) in order to correctly select between the correct image (e.g., *clean rabbit – dirty rabbit*) and the distractor image (e.g., *small rabbit – wet rabbit* or *small rabbit – little rabbit)*. In this version of the Opposite Task the experimenter would label both adjectives within an image. It is predicted that this would increase the difficulty of the Opposite Task because it would involve abstracting the opposite relationship from one word pair to another word pair. The irony task could be made less difficult by increasing the number of cues to speaker intent that were provided to children. I suggest that most efficient and effective way of increasing the cues provided to children would be to present videotaped scenarios of adults interacting. This would allow for the inclusion of cues, like facial expression, that are not possible when the actors are puppets (as has been the case in many of the developmental studies of irony appreciation).

Another avenue for future research is to explore whether the acquisition of antonymy varies cross culturally. To date, very few studies have been conducted that look at antonymy in non-English speaking children, and there are reasons to believe that there may be cross-cultural differences. Based on cultural and language differences, as well as some non-developmental research that has been conducted, it is my opinion that cross-cultural research on antonymy should begin with native Japanese-speaking children. Culturally, there are differences between English- and Japanese-speakers in terms of the pragmatics of polite conversation (Kim et al., 1990). There is a tendency for Japanese-speakers, relative to English-speakers, to prefer to use their conversation partner's words in their own responses, which could result in the preference of negation over antonymy. Similarly, there is a stylistic difference between speakers of the two languages whereby English-speakers avoid the use of negative sentences while Japanese-speakers do not. Japanese-speakers may even favour the use of negative sentences in order to express more subtle graduations of meaning (Akiyama, 1985). In addition, the greater popularity of dialectical thinking in Eastern cultures (the concept that what was good will become bad and what was bad will cycle to become good) may influence antonym use by Japanese speakers if they view antonyms to be on opposite ends of a continuum (Peng & Nisbett, 1999; Spencer-Rodgers, Williams, & Kaiping Peng, 2010). The alternative would be that Japanese speakers might use antonyms more if they use them to express the full scale (termed Coordinated Antonymy by Jones, 2002). A preliminary study on the discourse functions of antonymy in written Japanese suggests that Coordinated Antonymy is used at a much lower rate than in English or Swedish texts, though the authors warn against direct comparisons due to the limited set of antonyms evaluated in the Japanese texts as well as differences in the source of the Japanese corpora (Muehleisen & Isono, 2009; Murphy, Paradis, Willners, & Jones, 2009). Together these cultural and language differences may influence the frequency with which children are exposed to antonymy, which may, in turn, influence the developmental trajectory of the concept.

Previous cross-cultural research related to antonymy has evaluated differences in statement denial. Akiyama (1985) found that English-speaking children produced antonyms when generating statement denials much more frequently than did Japanese-speaking children. However, he concluded that the results implied language-specific differences in the process of creating a denial rather than differences in antonym knowledge because the two groups of children had similar proportions of incorrect antonym use. Future research should more directly evaluate the antonym knowledge in Japanese-speaking children and compare their performance to that of English-speaking children. Additional cultures of interest would include Chinese-speaking children because of the prevalence of dialectical thinking in that culture (Peng & Nisbett, 1999) and Korean-speaking children because they have been shown to be more similar to English-speaking children than Japanese-speaking children in the generation of statement denials (Kim et al., 1990). If cross-cultural differences in the acquisition of antonymy are found, they could have implications for the broader understanding of concept development in children. At the very least, they would suggest that researchers would need to be more aware of the potential influence of cross-cultural differences on their research.

Conclusions

Despite some limitations, the development of my novel Opposite Task represents a major contribution to the study of the development of antonymy. The results of the present study provide several unique contributions. To evaluate children's understanding of antonymy I modified a task used to evaluate children's understanding of irony, developed by Climie and Pexman {#Climie:2008ik}. This study involves a novel application of a task and its associated measures. Previous metalinguistic research on children's understanding of antonymy had been largely limited to verbal response tasks; therefore this study also demonstrates that the conclusions about children's understanding of antonymy can be generalized to other tasks. This modification of the task allowed me to expand the type of data collected beyond behavioural responses to include processing measures. In addition, to my knowledge this study is the first metalinguistic study that to includes 3-year-old children. Therefore, the conclusion made by previous researchers about the lack of understanding in 3-year-old children is now supported by empirical results. By expanding the type of data collected and decreasing the age of children evaluated, the results allow me to conclude that the majority of children under the age of 4 do not understand antonymy, nor do they demonstrate a latent appreciation of the concept. The results further suggest that at least some aspects of children's understanding of the concept of antonymy develop as if with sudden insight. The final unique contribution of the present study is that it is the first to collect demographic data on the participants. As such, it is the first to report a relationship between children's access to books and games in the home and children's understanding of antonymy. This finding could be a fruitful avenue of research, which could influence how children are taught the concept of antonymy in the future. Additional research using variants of my Opposite task will allow for a more thorough understanding of antonymy and permit further theory development and evaluation.

References

Akiyama, M. M. (1985). Denials in Young Children from a Cross-linguistic Perspective. *Child Development*, *56*, 95–102. doi:10.1111/j.1467-8624.1985.tb00089.x

Baldwin, D. A. (1992). Clarifying the role of shape in children's taxonomic assumption. *Journal of Experimental Child Psychology*, *54*, 392–416. doi:10.1016/0022-0965(92)90027-4

Barner, D., & Snedeker, J. (2008). Compositionality and Statistics in Adjective Acquisition: 4-Year-Olds Interpret Tall and Short Based on the Size Distributions of Novel Noun Referents. *Child Development*, *79*(3), 594–608. doi:10.1111/j.1467-8624.2008.01145.x

Bartlett, E. J. (1976). Sizing things up: The acquisition of the menaing of dimensional adjectives. *Journal of Child Language*, *3*, 205–219. doi:10.1017/S0305000900001458

Bonthoux, F., & Kalenine, S. (2007). Preschoolers' Superordinate Taxonomic Categorization as a Function of Individual Processing of Visual Vs. Contextual/Functional Information and Object Domain. *Cognition, Brain & Behavior*, *11*(4), 1–20. doi:10.3758/PBR.15.3.66

Booth, A. E., & Waxman, S. R. (2003). Mapping words to the world in infancy: Infants' expectations for count nouns and adjectives. *Journal of Cognition and Development*, *4*(3), 357–381. doi:10.1207/S15327647JCD0403_06

Bracken, B. A. (1988). Rate and Sequence of Positive and Negative Poles in Basic Concept Acquisition. *Language, Speech and Hearing Services in Schools*, *19*, 410.

Brewer, W. F., & Stone, J. B. (1975). Acquisition of spatial antonym pairs. *Jounal of*

Experimental Child Psychology, *19*, 299–307. doi:10.1016/0022-0965(75)90092-2

Bryant, P. E. (1992). Arithmetic in the Cradle. *Nature*, *358*, 712–713. doi:10.1038/358712a0

Carlson, S. M. (2005). Developmentally Sensitive Measures of Exectuive Function in Preschool Children. *Developmental Neuropsychology*, *28*(2), 595–616. doi:10.1207/s15326942dn2802_3

Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development*, *72*(4), 1032–1053. doi:10.1111/1467-8624.00333

Caselli, M. C., Bates, E., Casadio, P., Fenson, J., Fenson, L., Sanderl, L., & Weir, J. (1995). A cross-linguistic study of early lexcial development. *Cognitive Development*, *10*, 159–199. doi:10.1016/0885-2014(95)90008-X

Charles, W. G., & Miller, G. A. (1989). Contexts of antonymous adjectives. *Applied Psycholinguistics*, *10*, 357–375. doi:10.1017/S0142716400008675

Cicchetti, D. V. (1994). Guidelines, Criteria, and Rules of Thumb for Evaluating Normed and Standardized Assessment Instruments in Psychology. *Psychological Assessment*, *6*(4), 284–290. doi:10.1037/1040-3590.6.4.284

Clark, E. V. (1971). On the acquisition of the meaning of before and after. *Journal of Verbal Learning and Verbal Behavior*, *10*, 266–275. doi:10.1016/S0022-5371(71)80054-3

Clark, E. V. (1972). On the Child's Acquisition of Antonyms in two Semantic Fields. *Journal of Verbal Learning and Verbal Behavior*, *11*, 750–758. doi:10.1016/S0022-5371(72)80009-4

Clark, E. V. (1973). Non-linguistic strategies and the acquisition of word meanings. *Cognition*, *2*, 161–182. doi:10.1016/0010-0277(72)90010-8

Clark, E. V. (2004). How language acquisition builds on cognitive development. *Trends in Cognitive Sciences*, *8*(10), 472–478. doi:10.1016/j.tics.2004.08.012

Clark, E. V., Carpenter, K. L., & Deutsch, W. (1995). Reference states and reversals: Undoing actions with verbs. *Journal of Child Language*, *22*, 633–662. doi:10.1017/S0305000900009983

Clark, H. H. (1970). The primitive nature of children's relational concepts. In J. R. Hayes (Ed.), *Cognition and the Development of Language* (pp. 269–278). New York: Wiley.

Climie, E. A., & Pexman, P. M. (2008). Eye Gaze Provides a Window on Children's Understanding of Verbal Irony. *Journal of Cognition and Development*, *9*(3), 257–285. doi:10.1080/15248370802247939

Close, J., Hahn, U., Hodgetts, C. J., & Pothos, E. M. (2010). Rules and similarity in adult concept learning. In D. Mareschal, P. C. Quinn, & S. E. G. Lea (Eds.), *The Making of Human Concepts* (pp. 29–52). New York: Oxford University Press.

Dale, P. S., & Fenson, L. (1996). Lexical Development Norms for Young Children. *Behavior Research Methods, Instruments & Computers*, *28*(1), 125–127. doi:10.3758/BF03203646

Diesendruck, G., Hall, D. G., & Graham, S. A. (2006). Children's Use of Syntactic and Pragmatic Knowledge in the Interpretation of Novel Adjectives. *Child Development*, *77*(1), 16–30. doi:10.1111/j.1467-8624.2006.00853.x

Donaldson, M., & Balfour, G. (1968). Less is more: A study of language comprehension in children. *British Journal of Psychology*, *59*(4), 461–471. doi:10.1111/j.2044-8295.1968.tb01163.x

Ebeling, K. S., & Gelman, S. A. (1994). Children's use of context in interpreting "big" and ‘little’. *Child Development*, *65*, 1178–1192. doi:10.1111/j.1467-8624.1994.tb00811.x

Eilers, R. E., Oller, D. K., & Ellington, J. (1974). The acquisition of word-meaning for dimensional adjectives: The long and short of it. *Journal of Child Language*, *1*, 195–204. doi:10.1017/S0305000900000647

Fellbaum, C. (1995). Co-Occurrence and antonymy. *International Journal of Lexicography*, *8*(4), 281–303. doi:10.1093/ijl/8.4.281

Filippova, E., & Astington, J. W. (2008). Further Development in Social Reasoning Revealed in Discourse Irony Understanding. *Child Development*, *79*, 126–138. doi:10.1111/j.1467-8624.2007.01115.x

Gelman, S. A., & Taylor, M. (1984). How two-year-old children interpret proper and common names for unfamiliar objects. *Child Development*, *55*, 1535–1540. doi:10.2307/1130023

Graham, S. A., Cameron, C. L., & Welder, A. N. (2005). Preschoolers’ extension of familiar adjectives. *Journal of Experimental Child Psychology*, *91*(3), 205–226. doi:10.1016/j.jecp.2005.03.001

Graham, S. A., Welder, A. N., & McCrimmon, A. W. (2003). Hot dogs and zavy cats: Preschoolers“ and adults” expectations about familiar and novel adjectives. *Brain and Language*, *84*, 1–22. doi:10.1016/S0093-934X(02)00518-7

Hall, D. G. (1994). Semantic constraints on word learning: Proper names and adjectives.

Child Development, *65*, 1299–1317. doi:10.1111/j.1467-8624.1994.tb00818.x

Hall, D. G., & Graham, S. A. (1999). Lexical form class information guides word-to-object mapping in preschoolers. *Child Development*, *70*(1), 78–91. doi:10.1111/1467-8624.00007

Hall, D. G., & Moore, C. E. (1997). Red Bluebirds and Black Greenflies: Preschoolers' Understanding of the Semantics of Adjectives and Count Nouns. *Journal of Experimental Child Psychology*, *67*, 236–267. doi:10.1006/jecp.1997.2404

Hallgren, K. A. (2012). Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial. *Tutor Quantitative Methods Psychology*, *8*(1), 23–34.

Heidenheimer, P. (1975). The Strategy of Negation and the Learning of Antonymic Relations. *Developmental Psychology*, *11*(6), 757–762. doi:10.1037/0012-1649.11.6.757

Heidenheimer, P. (1978). Logical Relations in the Semantic Processing of Children between Six and Ten: Emergence of Antonym and Synonym Categorization. *Child Development*, *49*, 1243–1246. doi:10.1037/0012-1649.11.6.757

Herold, D. S., Nygaard, L. C., Chicos, K. A., & Namy, L. L. (2011). The developing role of prosody in novel word interpretation. *Journal of Experimental Child Psychology*, *108*(2), 229–241. doi:10.1016/j.jecp.2010.09.005

Holland, V. M., & Palermo, D. S. (1975). On learning “less”: Language and cognitive development. *Child Development*, *46*, 437–443. doi:10.1111/j.1467-8624.1975.tb03330.x

Huettig, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *ACTPSY*, *137*(2), 151–

171. doi:10.1016/j.actpsy.2010.11.003

Jones, S. (2002). *Antonymy: A corups-based approach*. London: Routledge.

Jones, S. (2006). A lexico-syntactic analysis of antonym co-occurrence in spoken English. *Text & Talk*, *26*(2), 191–216. doi:10.1515/TEXT.2006.009

Jones, S. (2007). "Opposites" in discourse: A comparison of antonym use across four domains. *Journal of Pragmatics*, *39*(6), 1105–1119. doi:10.1016/j.pragma.2006.11.019

Jones, S., & Murphy, M. L. (2005). Using corpora to investigate antonym acquisition. *International Journal of Corpus Linguistics*, *10*(3), 401–422. doi:10.1075/ijcl.10.3.06jon

Justeson, J. S., & Katz, S. M. (1991). Co-occurrences of Antonymous Adjectives and Their Contexts. *Computational Linguistics*, *17*(1), 1–19.

Kennedy, C., & McNally, L. (2005). Scale structure, degree modification, and the semantics of gradable predicates. *Language*, *81*(2), 345–381. doi:10.1353/lan.2005.0071

Kim, J. L., Shatz, M., & Akiyama, M. M. (1990). Effects of language and task on children's patterns of sentence verification and denial. *Developmental Psychology*, *26*(5), 821–829. doi:10.1037/0012-1649.26.5.821

Klatzky, R. L., Clark, E. V., & Macken, M. (1973). Asymmetries in the Acquisition of Polar Adjectives: Linguistic or Conceptual? *Journal of Experimental Child Psychology*, *16*, 1–15. doi:10.1016/0022-0965(73)90060-X

Klibanoff, R. S., & Waxman, S. R. (2000). Basic level object categories support the acquisition of novel adjectives: Evidence from preschool-aged children. *Child*

Development, *71*(3), 649–659. doi:10.1111/1467-8624.00173

Kreezer, G., & Dallenbach, K. M. (1929). Learning the Relation of Opposition. *The American Journal of Psychology*, *41*(3), 432–441. doi:10.2307/1414683

Kuczaj, S. A., II, & Maratsos, M. P. (1975). On the acquisition of front, back, and side. *Child Development*, *46*, 202–210. doi:10.1111/j.1467-8624.1975.tb03291.x

McGraw, K. O., & Wong, S. P. (1996). Forming Inferences About Some Intraclass Correlation Coefficients. *Psychological Methods*, *1*(1), 30–46. doi:10.1037/1082-989X.1.1.30

Morgan, J. T., & Greene, T. R. (1994). An Analysis of Categorization Style in Preschoolers. *Psychological Reports*, *74*, 59–66. doi:10.2466/pr0.1994.74.1.59

Morris, B. J. (2003). Opposites attract: the role of predicate dimensionality in preschool childrens processing of negations. *Journal of Child Language*, *30*(2), 419–440. doi:10.1017/S0305000903005610

Muehleisen, V., & Isono, M. (2009). Antonymous adjectives in Japanese discourse. *Journal of Pragmatics*, *41*(11), 2185–2203. doi:10.1016/j.pragma.2008.09.037

Murphy, M. L. (2003). *Semantic relations and the lexicon: Antonymy, synonymy, and other paradigms*. Cambridge: Cambridge University Press.

Murphy, M. L., & Jones, S. (2008). Antonyms in children's and child-directed speech. *First Language*, *28*(4), 403–430. doi:10.1177/0142723708091047

Murphy, M. L., Paradis, C., Willners, C., & Jones, S. (2009). Discourse functions of antonymy: A cross-linguistic investigation of Swedish and English. *Journal of Pragmatics*, *41*(11), 2159–2184. doi:10.1016/j.pragma.2008.09.040

Nilsen, E. S., & Graham, S. A. (2009). The relations between children’s communicative perspective-taking and executive functioning. *Cognitive Psychology*, *58*(2), 220–249. doi:10.1016/j.cogpsych.2008.07.002

Norbury, C. F., Nash, M., Baird, G., & Bishop, D. V. M. (2004). Using a parental checklist to identify diagnostic groups in children with communication impairment: a validation of the Children's Communication Checklist—2. *International Journal of Language & Communication Disorders*, *39*(3), 345–364. doi:10.1080/13682820410001654883

Nygaard, L. C., Herold, D. S., & Namy, L. L. (2009). The Semantics of Prosody: Acoustic and Perceptual Evidence of Prosodic Correlates to Word Meaning. *Cognitive Science*, *33*(1), 127–146. doi:10.1111/j.1551-6709.2008.01007.x

O'Dowd, S. C. (1980). Semantic Differential Studies of Children's Language Development: Some Methodological Problems. *Journal of Experimental Child Psychology*, *29*, 36–49. doi:10.1016/0022-0965(80)90089-2

Owens, R. E. (2008). *Language development: An introduction* (7 ed.). Boston.

Palermo, D. S. (1973). More about less: A study of language comprehension. *Journal of Verbal Learning and Verbal Behavior*, *12*(2), 211–221. doi:10.1016/S0022-5371(73)80011-8

Palermo, D. S. (1974). Still more about the comprehension of "less." *Developmental Psychology*, *10*(6), 827–829. doi:10.1037/h0037248

Peng, K., & Nisbett, R. E. (1999). Culture, Dialectics, and Reasoning About Contradiction. *American Psychologist*, *54*(9), 741–754. doi:10.1037/0003-066X.54.9.741

Pexman, P. M. (2008). It's Fascinating Research: The Cognition of Verbal Irony. *Current*

Directions in Psychological Science, *17*(4), 286–290. doi:10.1111/j.1467-8721.2008.00591.x

Ravn, K. E., & Gelman, S. A. (1984). Rule Usage in Children's Understanding of "Big" and 'Little'. *Child Development*, *55*, 2141–2150. doi:10.1111/j.1467-8624.1984.tb03910.x

Simcock, G., & Hayne, H. (2003). Age-related changes in verbal and nonverbal memory during early childhood. *Developmental Psychology*, *39*(5), 805–814. doi:10.1037/0012-1649.39.5.805

Smith, L. B., Cooney, N. J., & McCord, C. (1986). What is "high"? the development of reference points for "high" and "low." *Child Development*, *57*, 583–602. doi:10.1111/j.1467-8624.1986.tb00229.x

Smith, M., Johnston, D. R., & Coop, R. H. (1979). The acquisition of dimensional and expressive terms in young children. *Child Study Journal*, *9*(4), 239–249.

Spencer-Rodgers, J., Williams, M. J., & Kaiping Peng. (2010). Cultural Differences in Expectations of Change and Tolerance for Contradiction: A Decade of Empirical Research. *Personality and Social Psychology Review*, *14*(3), 296–312. doi:10.1177/1088868310362982

Syrett, K., Bradley, E., Kennedy, C., & Lidz, J. (2006). Shifting standards: Children's understanding of gradable adjectives. In K. Ud Deen, J. Nomura, B. Schulz, & B. D. Schwartz (Eds.), *Proceedings of the Inaugural Conference on Generative Approaches to Language Acquisition - North America* (pp. 353–364). Cambridge, Mass: UConn Occasional Papers in Linguistics 4.

Syrett, K., Kennedy, C., & Lidz, J. (2010). Meaning and Context in Children's

Understanding of Gradable Adjectives. *Journal of Semantics*, *27*(1), 1–35. doi:10.1093/jos/ffp011

Townsend, D. J. (1976). Do children interpret “marked” comparative adjectives as their opposites? *Journal of Child Language*, *3*, 385–396. doi:10.1017/S030500090000725X

Trueswell, J., & Gleitman, L. (2004). Children's Eye Movements during Listening: Developmental Evidence for a Constraint-Based Thory of Sentence Processing. In F. Henderson & J. M. Ferreira (Eds.), *The Interface of language, vision, and action: Eye movements and the visual world* (pp. 319–346). New York : Psychology Press.

Vilette, B. (2002). Do young children grasp the inverse relationship between addition and subtraction? Evidence against early arithmetic. *Cognitive Development*, (17), 1365–1383. doi:10.1016/S0885-2014(02)00125-9

Waxman, S. R., & Markow, D. B. (1998). Object properties and object kind: Twenty-one-month-old infants' extension of novel adjectives. *Child Development*, *69*(5), 1313–1329. doi:10.1111/j.1467-8624.1998.tb06214.x

Weiner, S. L. (1974). On the development of more and less. *Journal of Experimental Child Psychology*, *17*, 271–287. doi:10.1016/0022-0965(74)90072-1

Appendix A

Word Pairs

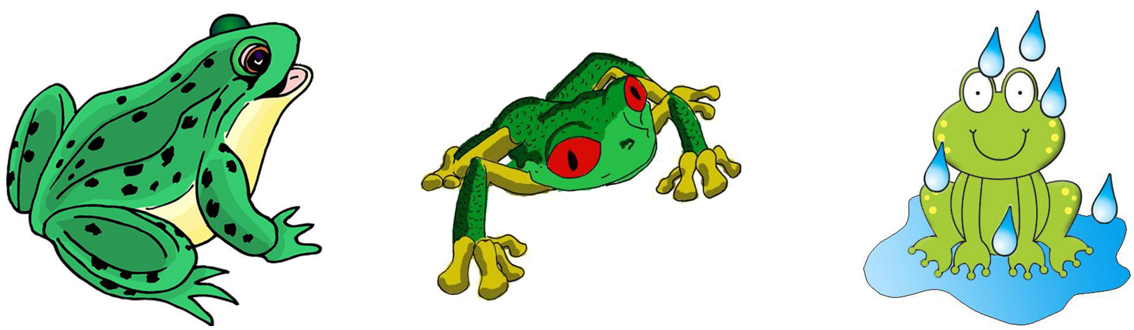
Training Trials	dry - wet
	fat – thin
Test Trials	awake – asleep
	clean – dirty
	happy – sad
	old - young
	small – big
	tall – short

Note: the first item of each pair is the image selected by the experimenter.

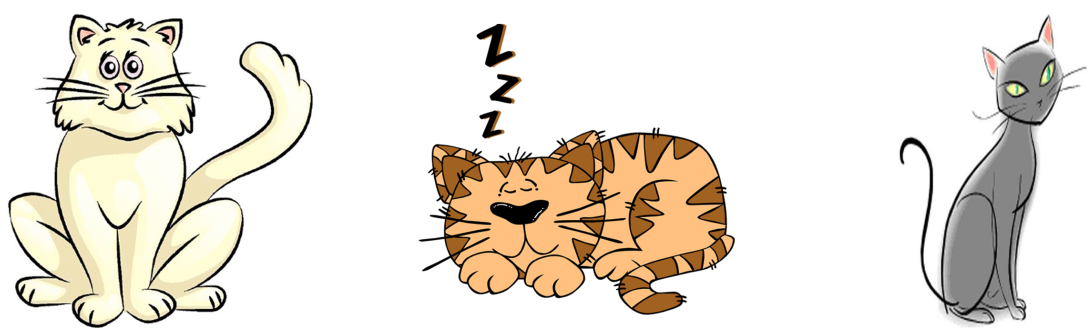
Appendix B

dry dog – wet dog – thin dog

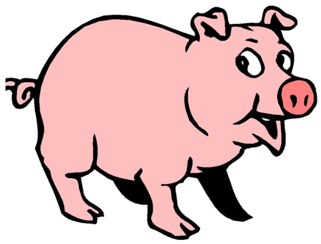
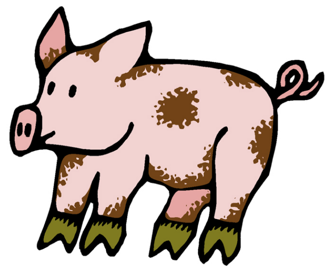
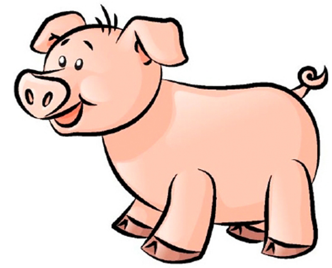
fat frog – thin frog – wet frog



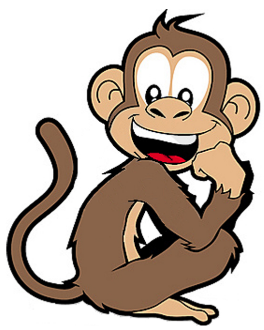
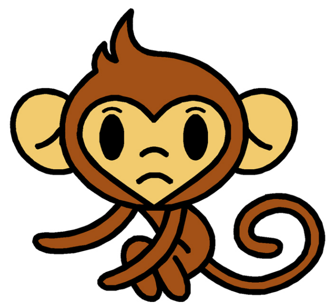
awake cat – asleep cat – tall cat



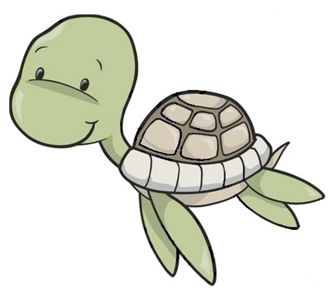
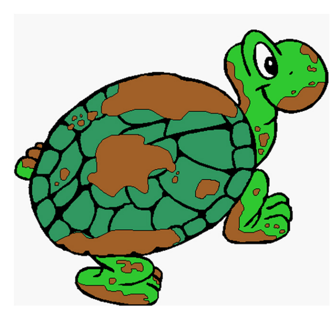
clean pig – dirty pig – small pig

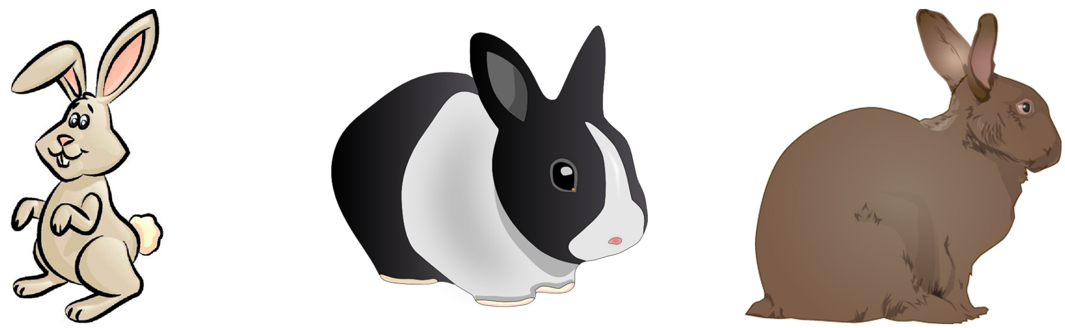
happy monkey – sad monkey – asleep monkey

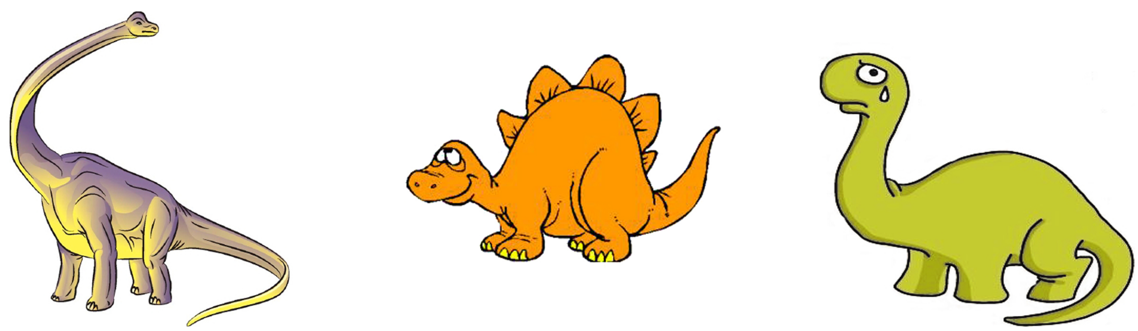

old turtle – young turtle – dirty turtle

small rabbit – big rabbit – old rabbit



tall dinosaur – short dinosaur – sad dinosaur



Note: Stimuli are not presented to scale.

Appendix C

Script for the Opposite Task – the Opposite Label Condition

"The first game we're going to play today is a game with animals. Part of the game is that you keep your hands in your lap until it is your turn. Can you put your hands in your lap? Now I'm going to show you how to play the rest of the game" *Demonstrate with the first set of practice pictures during the course of the explanation of the game.*

> "In this game I'm going to put 3 pictures on the table, like this. Then I'll tell you about each picture "This is a *thin frog*, this is a *fat frog*, and this is a *wet frog*". Then I'll put one of the pictures in the box, like this. Your job in this game is to put the picture that is the opposite of my picture into the box. Can you do that for me?"
>
> > If the child's choice was correct, the experimenter praised the child and moved on to the next set of practice images.
> >
> > If the child's choice was incorrect, the experimenter first praised the child for putting a picture into the box, then the experimenter put the correct image in the box and corrected the child saying "I think that the opposite of *fat* is *thin*. Let's practice some more."

The second practice trial was completed as follows

> "This is a *thin dog*, this is a *dry dog*, and this is a *wet dog*. If I put the *dry* dog in the box, can you put the opposite picture in the box?"

The practice trials were followed by 6 test trials that were identical to the final practice trial except that no feedback was provided. If the child requested feedback during the test trials he was told that the experimenter would tell him how he performed at the end of the game. If necessary, the child was reminded to put his hands on his lap between trials.

Script for the Opposite Task – the Another One Label Condition

"The first game we're going to play today is a game with animals. Part of the game is that you keep your hands in your lap until it is your turn. Can you put your hands in your lap? Now I'm going to show you how to play the rest of the game" *Demonstrate with the first set of practice pictures during the course of the explanation of the game.*

> "In this game I'm going to put 3 pictures on the table, like this. Then I'll tell you about each picture "This is a *thin frog*, this is a *fat frog*, and this is a *wet frog*". Then I'll put one of the pictures in the box, like this. Your job in this game is to put another one into the box. Can you do that for me?"
>
> > If the child put one picture into the box, the experimenter praised the child and moved on to the next set of practice images.
> >
> > If the child put more than one picture in the box, or did not put a picture in the box, the experimenter corrected the number of images in the box and corrected the child saying "Remember the game is to put another one picture in the box. Let's practice some more."

The second practice trial was completed as follows

> "This is a *thin dog*, this is a *dry dog*, and this is a *wet dog*. If I put the *dry* dog in the box, can you put another one in the box?"

The practice trials were followed by 6 test trials that were identical to the final practice trial. If necessary, the child was reminded to put his hands on his lap between trials.

Script for the Opposite Task – the No Label Condition

"The first game we're going to play today is a game with animals. Part of the game is that you keep your hands in your lap until it is your turn. Can you put your hands in your lap? Now I'm going to show you how to play the rest of the game" *Demonstrate with the first set of practice pictures during the course of the explanation of the game.*

"In this game I'm going to put 3 pictures on the table, like this. Then I'll tell you about each picture "This is a *frog*, this is a *frog*, and this is a *frog*". Then I'll put one of the pictures in the box, like this. Your job in this game is to put another one into the box. Can you do that for me?"

If the child put one picture into the box, the experimenter praised the child and moved on to the next set of practice images.

If the child put more than one picture in the box, or did not put a picture in the box, the experimenter corrected the number of images in the box and corrected the child saying "Remember the game is to put another one picture in the box. Let's practice some more."

The second practice trial was completed as follows

"This is a *dog*, this is a *dog*, and this is a *dog*. If I put the *dry* dog in the box, can you put another one in the box?"

The practice trials were followed by 6 test trials that were identical to the final practice trial. If necessary, the child was reminded to put his hands on his lap between trials.

Appendix D

Instructions:
Okay _______ For this next game I am going to show you some pictures, I want you to say aloud what the pictures are and try to remember the last picture you see. So after you say the pictures I will ask you to tell me what the last picture in each row is.

For example – what are these pictures here (leaf, frog, shoe) – okay, now ***what was the last picture in that row?*** *(shoe) – very good.*

And in this example there are two rows – so what are the pictures here (ball, sock, cake; and leaf, sun, pig) – ***so what were the last pictures of each row?*** *(cake, pig)*

> *(if correct) – very good – you told me the last picture of each row*
> *(if not correct) – that's not quite right, in this example you saw sock, ball, cake and leaf, sun, pig (show full view of card) – so the last picture in the rows would be cake and pig.*

Okay, now we are going to try the real game.
You are to say the pictures aloud and then tell me the last picture on each row.
What are the pictures here?
What was/were the last picture(s) of the/each row?

Note: stop after 4 consecutive incorrect answers (i.e., pt. could not remember even one last item)

Working Memory Task Response Sheet

1. (ball) _______	**10. (dog, flower) ________**
2. (sock) _______	**11. (frog, shoe) ________**
3. (leaf) _______	**12. (pig, sock) _______**
4. (fork) _______	**13. (ball, tree) ________**
5. (sun, dog) _______	**14. (fork, leaf) _______**
6. (flower, shoe) _______	**15. (dog, sun, tree) ________**
7. (cake, frog) ________	**16. (pig, flower, cake) ________**
8. (pig, leaf) ________	**17. (shoe, fork, sock) ________**
9. (tree, sun) _______	**18. (ball, leaf, frog) ________**

Items correct (i.e., in correct order; /18): _____
Total numbers correct (/36): _____

Appendix E

Participant #: ___________________ Date (d/m/y): ___________________

Parent Questionnaire

It would be helpful if you could provide some additional information about your child who is participating in today's study. Providing this information is voluntary and will be used only for the purposes of this study.

1. What is the primary language spoken in your household?

   ____

2. Please indicate any other languages spoken in your household.

   ____

3. Please tell us about your child's childcare experience. Have they spent their time away from the family *primarily* at:

   _____ home, with a primary caregiver, and no other children present
   _____ home, with a primary caregiver, and other children present
   _____ a daycare centre or preschool, part time
   _____ a daycare centre or preschool, full time
   _____ a dayhome, with no other children present
   _____ a dayhome, with other children present

4. If applicable, at what age did your child begin attending preschool or daycare?

   ____

5. Does your child have access (in the household) to books or games focused on opposites?

   ____

6. Do you, or someone in your household, play word games involving opposites with your child?

   ____

7. How many children currently reside in your household?

____

8. What are the ages of each child?

____

9. What is the participating child's birth order? (e.g., only child, first born, last born, etc.)?

____