HENNING, KYLE JOSEPH, PhD., MAY 2022 PSYCHOLOGICAL SCIENCES

THE IMPACT OF METACOGNITIVE REPRESENTATIONS AND FEEDBACK ON CHILDREN'S DISAMBIGUATION PREDICTION (78 pp.)

Dissertation Advisor: William Merriman

Even one-year-olds show the so-called disambiguation effect, which is a tendency to select a novel object rather than a familiar object as the referent of a novel label. The strength of this effect increases over the preschool years. This age trend may be due in part to advances in metacognition. The accuracy of preschoolers' lexical knowledge judgment mediates the association between age and strength of the disambiguation effect. Also, their judgments about object label knowledge accounted for why most 4-year-olds, but only a few 3-year-olds could predict the solution to a new disambiguation problem before hearing the novel label (Henning & Merriman, 2019). Study 1 tested whether more preschoolers could make this kind of prediction if they were told that the labels were ones "you have never heard before." Results supported this hypothesis, but only for the younger children. Also, children's tendency to make these predictions was positively associated with their ability to give accurate reports of whether various words or pseudowords had known meanings. Study 2, which used an online rather than face-to-face testing procedure, demonstrated that 3-year-olds only learned how to solve the original prediction problem if they received direct rather than indirect feedback. When they receive helpful cues, most 3-year-olds can solve a disambiguation problem before hearing the novel label. Thus, most 3-year-olds can form metacognitive representations of the elements of the disambiguation problem and use these to draw inferences about the reference of a label.

THE IMPACT OF METACOGNITIVE REPRESENTATIONS AND FEEDBACK ON CHILDREN'S DISAMBIGUATION PREDICTION

A dissertation submitted

to Kent State University in partial fulfillment

of the requirements for the

degree of Doctor of Philosophy

by

Kyle Joseph Henning

May 2022

Dissertation written by

Kyle Joseph Henning

B.A., The University of Toledo, 2013

M.A., Kent State University, 2018

Ph.D., Kent State University, 2022

Approved by

William Merriman__________________, Chair, Doctoral Dissertation Committee

Jeffrey Ciesla______________________, Members, Doctoral Dissertation Committee

Bradley Morris______________________

Maria Zaragoza______________________

Accepted by

Maria Zaragoza_____________________, Chair, Department of Psychological Sciences

Mandy Munro-Stasiuk_______________, Dean, College of Arts and Sciences

LIST OF TABLES

Acknowledgements

I want to thank my wife, Jennifer, for being compassionate, patient, and encouraging throughout this challenging experience. Without your continued support, this body of work could not have been accomplished.

Additionally, I want to thank my boys, Gunner, Deacon, and Rorke, for their enduring love and positive impact that they have made on me.

The Impact of Metacognitive Representations and Feedback on Children's Disambiguation Prediction

Chapter 1. Introduction

In any typical setting individuals have hundreds if not thousands of objects surrounding them in their field of view. Young children experience these same complex environments and are frequently tasked with determining the reference of names for novel objects without explicit instruction. Children can attend to whether a speaker is looking or pointing at an object while naming to help solve this problem (Baldwin, 1993; Paulus & Fikkert, 2014). However, speakers do not always provide these cues, children may fail to detect them, and the cues themselves may not be sufficiently precise. Medina, Snedeker, Trueswell, & Gleitman (2011, Exp. 1) analyzed everyday parent-child interactions and found that only 14% of instances in which the parent uttered a common noun were highly informative referential acts. Even in specific scenarios such as object play between a parent and child, where there was greater referential clarity, only 58% of instances involved transparent referential cues (Yurovsky, Smith, & Yu, 2013, Exp. 1). Thus, children may need strategies for determining a novel word reference in many situations.

Children tend to map novel labels onto novel rather than familiar objects. This tendency, which is known as the *disambiguation effect*, can be observed in children as young as 16 months (Mervis & Bertrand, 1994: Halberda, 2003; Markman, Wasow, & Hansen, 2003). This strategy is quite successful for children, especially when no other referential cues are available (Markman & Wachtel, 1988; Merriman & Bowman, 1989). However, the strategy is not infallible because the speaker's intended referent could be an object for which the child already has an alternative label. Nevertheless, this strategy may play a crucial role in word learning due to the frequency of encounters with unclear references of novel labels.

In addition to typically-developing children, children from many special populations also show the disambiguation effect. These include bilingual and multilingual children (Merriman & Kutlesic, 1993; Davidson, Jergovic, Imami, & Theodos, 1997; Byers-Heinlein & Werker, 2009, 2013; Houston-Price, Caloghiris, & Raviglione, 2010), late talkers (Mervis & Bertrand, 1995a; Choi & Hwang, 2014), children with SLI (Specific Language Impairment) (Beverly & Estis, 2003, Estis & Beverly, 2015), children with ASD (autism spectrum disorder) or who are at risk for ASD (Preissler & Carey, 2005; de Marchena, Eigsti, Worek, Ono, & Snedeker, 2011; Bedford et al., 2013), children with various forms of mental disability (Mervis, 1995b; Romski et al., 1996; Wilkinson, & Green, 1998), and children who are deaf or hard-of-hearing (Lederberg & Spencer, 2008).

Findings from a recent meta-analysis indicated that the strength of the disambiguation effect (i.e., frequency of selecting the novel object as the referent of the novel label) increases from infancy through 5 years of age (Lewis, Cristiano, Lake, Kwan, & Frank, 2020). However, studies that include referential cues that conflict with the effect (e.g., the speaker pointing at the familiar object while providing the novel label) tend to find the opposite developmental trend (Grassmann & Tomasello, 2010; Gollek & Doherty, 2016; Scofield, Merriman & Wall, 2018) and studies that provide corrective feedback after every trial tend to find no developmental trend (Marazita & Merriman, 2004; Henning & Merriman, 2019).

The general goal of this dissertation was to advance our understanding of why the basic disambiguation effect – the one observed when no conflicting cues or corrective feedback are provided – tends to increase in strength over early childhood. Specifically, I examined whether metacognitive processes that develop during early childhood are at least a partial cause of this increase in strength.

In a seminal paper, Flavell (1979) coined the term *metacognition* to refer to a person's knowledge and cognition about cognitive phenomena. Metacognition is now typically defined more broadly as "thinking about thinking" or "knowledge about knowledge". It involves the interplay of two processes—monitoring and control (Nelson & Narens, 1990). Metacognitive monitoring involves evaluating the present state or ongoing progress of a cognitive activity. Metacognitive control involves regulating some aspect of an ongoing cognitive activity, such as deciding to continue it, stop it, or change it while in progress.

In this dissertation, I addressed the hypothesis (Slocum & Merriman, 2018; Henning & Merriman, 2019) that the disambiguation effect tends to be stronger when the child adopts metacognitive representations of various elements of the disambiguation problem. That is, a child tends to show a stronger tendency to map a novel label onto a novel rather than a familiar object when they represent the label and the novel object as unknown and represent the familiar object as known. For example, consider the child who must decide whether the novel word *whisk* refers to a whisk or a spoon. A child who did not represent the elements of this problem metacognitively might still respond correctly. They might note the mismatch between the label *whisk* and the label that they themselves would use for the spoon (*spoon*) and interpret this mismatch as indicating that the spoon is the wrong choice. My hypothesis is that if the child also represented both the label *whisk* and the whisk itself as unknown and represented the spoon as known, their tendency to select the whisk would be even stronger. Thus, I propose that children will tend to display a stronger disambiguation effect if they tend to represent the epistemic status (i.e., known vs. unknown) of each element of the disambiguation problem. These representations promote deciding to select the unknown kind of object rather than the known kind of object as the referent of the unknown label.

This first chapter will review the research that forms the foundation of the current study. I will review current theoretical accounts of the disambiguation effect, then summarize the developmental research on children's metacognitive processes that could potentially influence the effect. Finally, I will discuss recent studies which demonstrated that some preschool-age children can solve two purely metacognitive versions of the disambiguation problem (Slocum & Merriman, 2018; Henning & Merriman, 2019). These versions of the disambiguation problem were created with the goal of forcing children to rely on metacognitive processes to arrive at the same solution. Specifically, children needed to compare metacognitive representations to be successful and could not solve these problems by noting a mismatch between labels or by comparing the activation of objects.

Chapter 2 will provide an overview of two experiments conducted for this dissertation. Both experiments investigated the potential role of metacognitive processes in the disambiguation effect. I hypothesized that the tendency to adopt metacognitive representations of the elements of the disambiguation effect emerges with development. Specifically, as children's ability to reflect on their lexical knowledge develops, they should also develop a metacognitive representation of their own tendency to map novel labels onto novel rather than familiar objects.

Current Accounts of the Disambiguation Effect

There are three leading accounts of the disambiguation effect: Mutual Exclusivity; Pragmatic Inference; and Competitive Activation.

Mutual Exclusivity. According to the mutual exclusivity account, the disambiguation effect is the result of children's tendency to adopt the default assumption that two labels do not have exemplars in common (Markman & Wachtel, 1988; Merriman & Marazita, 1995). For example, if a novel label (e.g., *whisk*) is presented in a situation where it could refer to a familiar, nameable object (e.g., spoon) or an unfamiliar, as-yet-unnameable object (e.g., whisk), the child will reject the familiar object as the possible referent because it already has a known label and will ultimately choose the unfamiliar object as the intended referent.

One seeming problem for the mutual exclusivity account is that very young children can learn multiple labels for the same exemplar (Clark, 1997; Deák & Maratsos, 1998; Waxman & Hatch, 1992). However, according to the account, children relax the default assumption of mutual exclusivity in specific scenarios if given sufficient evidence to do so. For example, Grassmann & Tomasello (2010) found that when the adult requesting the referent of the novel label pointed at the familiar object while looking back and forth between the familiar object and the child, 2- and 4-year-olds selected the familiar object as the referent of the novel label instead of the unfamiliar object. Thus, the speaker's referential act needs to be ambiguous for the disambiguation effect to occur. Jaswal and Hansen (2006) found that in some circumstances, when a speaker used a more subtle cue such as gazing or pointing toward a familiar object while using a novel label, preschoolers tended to select the novel object next to it as the label's referent.

Another situation in which children will relax the mutual exclusivity assumption is when it would require them to reject some strongly-held belief. Merriman and Bowman (1989, Experiment 2) showed 2- and 3-year-olds an object that was either a typical exemplar (e.g., a typical-looking car) or an atypical exemplar (e.g., a hybrid of a car and truck) of a familiar label and provided a novel label for the object (e.g., *zav*). If the object was a typical exemplar, nearly all the children decided that both the novel label and the known label (e.g., *car*) were acceptable names for the object. However, when the object was an atypical exemplar, approximately half of the children rejected both of the two known labels for the object. For example, they judged that the hybrid of a car and a truck was not a car or a truck; it was a zav. Note that in a control condition in which no novel name was introduced for the hybrid object, no children rejected both known labels as a name for it. Waxman and Senghas (1992) also found evidence that children's decision to consider two novel labels to be mutually exclusive was based on the similarity or dissimilarity of the objects that they had heard called by these labels.

Pragmatic Inference. There are two leading pragmatic inference accounts – pragmatic contrast and mental files. According to the pragmatic contrast account, the disambiguation effect results because children expect speakers to be cooperative when referring to things (Gathercole, 1989; Clark, 1990; Bloom, 2000; Tomasello, 2000; Diesendruck & Markson, 2001). If some way of referring to an object is mutually known by the speaker and the child, then the child expects the speaker to use this expression to refer to the object. So, if the speaker uses some other expression, they infer that some other object must be the intended referent. For example, in a scenario where a child is presented with two objects, one familiar (e.g., cup) and one unfamiliar (e.g., corkscrew), the child selects the corkscrew as the possible referent of *blicket*. The child infers that if the speaker had intended to refer to the known object, then the speaker would have used the mutually-known label for it. Because the speaker did not do this, the speaker must intend to refer to the other object. This account essentially applies one of Grice's (1975) maxims of cooperative conversation, Quantity, to the disambiguation problem. According to this maxim, the participants in a conversation should give only as much information as is needed, and no more, to accomplish their communicative purposes.

Using this logic, Diesendruck and Markson (2001) argued that the disambiguation effect should occur for all novel referring expressions, not just labels. In support of this claim, they found that 3- and 4-year-olds showed the disambiguation effect for factual descriptions of objects. Children were presented with pairs of novel objects, and for each trial, the researchers provided a novel fact about one of the objects in the pair (e.g., “My uncle gave this to me.”) Children were later asked to select an object from the same pair based on a newly presented novel fact (e.g., “Can you give Percy the one my cat likes to play with?”). Children successfully selected the previously unmentioned object and disambiguated as frequently as in a test that involved labels. This result was taken as evidence that the disambiguation effect should be regarded as a domain-general phenomenon --- one evident with all referring expressions, not just labels.

However, subsequent research has found that children do not have the same expectations for factual descriptions as they do for labels. For example, whereas children readily extend a novel object label to all members of an object’s category, they tend to restrict a novel fact to the individual object that they heard the fact used to describe (Gelman & Coley, 1990; Waxman & Booth, 2000). Also, children’s tendency to disambiguate in the facts condition but not in the labels condition was disrupted by subtle socio-pragmatic cues such as gazing or pointing (Jaswal & Hansen, 2006).

Scofield and Behrend (2007) replicated Diesendruck and Markson’s (2001) original finding that the strength of the disambiguation effect in 4-year-olds was as strong for facts as for labels. However, Scofield and Behrend found that younger children only disambiguated in the label condition. They argued this evidence supported the mutual exclusivity account rather than the pragmatic contrast account of the disambiguation effect, at least for younger children. Kalashnikova, Mattock, and Monaghan (2014) tested two age groups of children (4- & 5-years-olds) and a group of adults on disambiguation of labels and facts. The two child groups disambiguated similarly in the labels and facts conditions, replicating previous findings. Adults, however, disambiguated significantly more frequently in the labels condition than in the facts condition, suggesting the domain-generality of the disambiguation effect changes over the course of development. This evidence further supports the mutual exclusivity account, which attributes the disambiguation effect to a default assumption about the semantic relations between pairs of words, over the pragmatic contrast account, which attributes the effect to an expectation that speakers will abide by Grice's maxim of Quantity when engaged in referential communication.

Recently, Doherty and Perner (2020) have proposed a different kind of pragmatic account of the disambiguation effect. According to their account, the disambiguation effect is a result of children's tendency to create only one representation of an object, called a mental file, when comprehending discourse that refers to the object. According to these authors, "a mental file can be characterized by the following five features (based on Recanati, 2012):

(1) *Representation*: A file is a conceptual representation of a particular object, its referent.

(2) *Reference*: Information about a referent, via perception or verbal testimony, causes a file to be created. This relation fixes the file's referent.

(3) *Function*: A file's function is to (a) track its referent over time and (b) acquire and update knowledge about its referent. This knowledge is registered on the file, thus (c) binding the

object's properties to the object, and (d) making the information available for thinking and action.

(4) *Tracking*: To fulfil a file's tracking function three processes are needed:

(*i*) *Recognition*: the information on the file is used to (re-)identify an encountered object as the file's referent.

(*ii*) *Attention*: Once (re-)identified the object will be tracked by an attentional mechanism and need not be continuously re-identified.

(*iii*) *Tracking Constraint*: Some mechanism is required to prevent potential confusion that could result from mixing different conceptualizations of objects within the same conversation. We propose that only one file is allowed to track an object at a particular time.

(5) *Conceptuality*: Mental files are constituents of thought, hence conceptual representations. They represent their referent as something."

Doherty and Perner (2020) propose that the tracking constraint, viewed as a pragmatic principle governing referring expressions (both labels and factual descriptions), prevents the use of more than one file to track an object. Multiple files, or multiple perspectives, on the same object in a single discourse could potentially cause confusion regarding whether the same object is being tracked during the discourse. Thus, children are initially unable to consider more than one label or factual description for an object, producing the disambiguation effect for both labels and facts. For example, upon attending to the familiar object (e.g., spoon) the child reidentifies or updates the existing mental file containing the familiar label (e.g., *spoon*) and uses this to track the spoon. Due to the tracking constraint, this file will prevent tracking by alternative files for the duration that the spoon is tracked. Once the unfamiliar object (e.g., whisk) is attended to and no file is able to be identified by the child, a new mental file is created, which contains the unfamiliar label (e.g., *whisk*) as well conceptual and perceptual information about the whisk, resulting in the child's selection of the whisk as the intended referent.

Doherty and Perner (2020) argue that the tracking constraint does not prevent the creation of a new mental file, which contains the novel label, from attaching to the familiar object unless the object is spontaneously tracked by a file with the familiar label. As a result, it differs from the mutual exclusivity account in that it is situation-specific and does not act on the lexicon to maintain one-to-one word-object associations. Alternatively, Doherty and Perner (2020), the mental files account is related to the pragmatic contrast account in that both concern expectations about communication. However, the accounts differ in that the pragmatic contrast account involves a higher level of mental state reasoning in comparison to the mental files account's tracking constraint, which is assumed to operate without any reasoning.

According to the mental files account, it is not until around the fourth birthday that children are able to assign multiple labels to the same object in a single episode of discourse. To accomplish this feat, they must merge two files into one file or *link* the files. Linking occurs because the child realizes that they have two files for a particular object. It is possible that the child could recognize this as a mistake and drop one of the files altogether. However, this could be problematic because one loses the information gained about the referent under the dropped file's perspective. Perry (2002) suggests that linking allows the information of each file to be assessed from the other file while preserving the different perspectives. Doherty & Perner (2020) propose that the emergence of the ability to link files around 4 years explains why this is the age at which most children first succeed on tests of their understanding of false belief and their ability to generate alternative names for things.

In the classic test of children's understanding of false belief, (Wimmer & Perner, 1983) a child is told a story about Maxi, a boy who put his chocolate away in a cupboard before leaving to play. While away playing, his mother transfers Maxi's chocolate to a different location. Upon Maxi's return, the child is asked, "Where will Maxi look for the chocolate." Most children below 4 years mistakenly indicate that Maxi will look in the new location. Presumably, they do not yet understand that Maxi has a different perspective than their own regarding where the chocolate is. The mental files account assumes that to pass this test, a child needs to create co-referential files for the chocolate (i.e., the child's own file, which represents everything that happened to the chocolate, and the child's 'vicarious' file, which only represents what happened to the chocolate that Maxi is aware of). They must also link these files to be able to represent how Maxi will reason about the location of the chocolate. Therefore, the ability to create and link co-referential files (i.e., attach more than one mental file to the same referent) around the age of 4 allows the child to answer correctly when asked where Maxi will look for the chocolate as well as where Maxi's mother put the chocolate.

Likewise, development of the ability to attach more than one mental file to a referent can explain why the ability to produce alternative names for things also develops around the age 4 (Doherty & Perner, 1998). The test of the latter ability uses pairs of alternative labels for an item that are familiar to the children (e.g., "bunny" and "rabbit"). Participants are instructed using a story character and a puppet as follows: "This is Tony, and he is going to tell us what these things are. But Puppet also wants to play but he does not want to say the same thing as Tony, so we are going to help him think of a different way of saying what these things are. Now, Tony says this is a rabbit. What could Puppet say it is? I know…he could say it is a bunny, because it is a rabbit and it is a bunny. They are different ways of saying the same thing. Test: Tony says:

"this is a rabbit." What could Puppet say it is?" Three-year-olds typically fail this task, most often by simply repeating the word that Tony said. The mental files account suggests these children cannot yet link files. Because Tony identified the object of interest as a "bunny", this object cannot also be a "rabbit" within this same perspective. Thus, they provide the incorrect answer of "bunny" despite the instruction to name the animal in a different way than Tony just did. However, 4-year-olds typically respond correctly with "rabbit". They can successfully follow the instructions of providing an alternative name because of their ability to link files. Linking makes the co-referential files for the object mutually available to the child and so, enables the child to select the mental file for the object that contains the different name.

Because the disambiguation effect initially results from the inability to attach more than one file to a referent, the strength of this effect should be inversely related to one's performance on tasks such as false belief and alternate naming that require the child to attach more than one file to a referent. Gollek and Doherty (2016) tested this hypothesis with 2- to 5-year-old children and found that standard disambiguation performance approached ceiling and was not associated with the other tasks. However, children's tendency to select the unfamiliar object in a pragmatic-cue version of the disambiguation task was found to be significantly negatively correlated with their performance on false belief and alternative naming tests, even after controlling for age, verbal mental age, and inhibitory control. Within this pragmatic-cue version, which was originally developed by Haryu (1991), the child is still asked to decide whether a familiar or an unfamiliar object (e.g., an apple and a lipstick holder) is the referent of a novel label. However, the child is told that a puppet "is hungry [or in some other need state]. I would like to give her the *heku* [a nonsense word]." Of the two object choices, only the familiar one could satisfy the puppet's need [e.g., the apple, but not the lipstick holder could satisfy the puppet's hunger.]

Thus, the child's tracking constraint conflicted with a rational cue (i.e., that a person will seek an object that satisfies their expressed need). Gollek and Doherty (2016) found that the tendency to select the unfamiliar object (i.e., maintain the tracking constraint rather than choose based on the rational cue) was negatively correlated with both false belief understanding and alternative naming ability.

All of these accounts of the disambiguation effect characterize children as rejecting the familiar object because they detect a mismatch between its known label and the novel label. The accounts differ regarding the principles that compel the children to infer from this mismatch that the familiar object is the wrong choice (Halberda, 2003; 2006). However, none of the accounts specifies exactly how children make this decision regarding the mismatch between the novel label and they label they know for the familiar object. Evidence suggests it is likely based on a comparison of the two labels' phonological representations (Merriman & Marazita, 1995; Jarvis, Merriman, Barnett, Hanba, & Van Haitsma, 2004; Marazita & Merriman, 2004; Wall, Merriman, & Scofield, 2015). For instance, imagine a child is tasked with finding the referent for an unfamiliar noun (e.g., "Find me the *pilson*."). The child will reject familiar objects such as spoon and fork while searching for the referent because the names the child knows for these object (i.e., *spoon* and *fork*) mismatch *pilson* (see Jarvis et al., 2004, for a review of research on how cases of phonological similarity between a known and novel name are resolved by children).

Competitive Activation. A third set of accounts differs from the others in that the disambiguation effect does not depend on either label retrieval or comparison of label representations. In various competitive activation models (Merriman, 1999; Regier, 2005; McMurray, Horst, & Samuelson, 2012) the disambiguation effect results from the activation of excitatory and inhibitory connections that have developed among representations. Upon hearing a novel label, children select a novel object as its intended referent simply because the representation of this object receives more activation than the representation of the familiar object. This activation differential results because past experiences have strengthened the connection between the familiar object's representation and its associated name.

Role of Metacognition in Current Accounts. No account of the disambiguation effect presumes that children have a metacognitive representation of the processes involved in successful disambiguation. The pragmatic contrast account comes the closest because it proposes that children reason about the speaker's intentions. However, even if children map factual descriptions in the way that is predicted by the pragmatic contrast account, they might not necessarily be reasoning about the speaker's intentions. Indeed, the pragmatic mental files account proposes that they are not; they are simply abiding by an unconscious tracking constraint on mental file creation. They may have simply developed a habit of rejecting an object as the referent of a referring expression (e.g., "My uncle gave me this one") whenever they hear or retrieve a differing referential expression for the object (e.g., "the one my cat likes"). Similarly, children might abide by the mutual exclusivity account if they have developed a habit of rejecting an object as the referent for a label whenever they retrieve a different label for the object. They may not represent this decision as being based on choosing a novel kind of object over a familiar kind of object. Likewise, from the perspective of the competitive activation models, children may select the novel object only because it receives more activation from the novel word than the familiar object does. They may not reflect on the epistemic status (i.e., known vs. unknown) of the novel word, novel object, or the familiar object.

Development of Metacognition Relevant to the Disambiguation Effect

The ability to represent the elements of the disambiguation problem on a metacognitive level develops during the preschool years. Children tend to give accurate reports of their lexical knowledge by age four. For example, when shown various familiar and unfamiliar kinds of objects and asked whether they know the names for them, 4-year-olds tend to respond correctly, whereas younger children often report knowing names for many of the unfamiliar kinds (Marazita & Merriman, 2004; Merriman & Lipko, 2008; Merriman, Lipko, & Evey, 2008; Wall, Merriman, & Scofield, 2015). Similarly, when asked whether they know highly familiar words as well as made-up words, 4-year-olds tend to respond correctly; in contrast, younger children frequently report knowing many of the made-up words (Smith & Tager-Flusberg, 1982; Merriman & Schuster, 1991; Chaney, 1992; Merriman et al., 2008). Since identifying novel kinds of objects and novel words as ones they do not know is a task that younger preschoolers often fail, it is unlikely that they spontaneously represent the disambiguation problem as a choosing whether an object they know or an object they do not know is the referent of a label they do not know.

A more robust disambiguation effect has been found to be associated with awareness of lexical knowledge. Merriman and colleagues administered a disambiguation paradigm in which every trial involved first asking the child whether they knew a particular novel label and then asking for its referent (e.g., "Do you know what a blicket is? Which one is a blicket?) In four studies (Merriman & Bowman, 1989, Studies 1 and 2; Merriman & Schuster, 1991; Merriman, Marazita, & Jarvis, 1993, Experiment 2), most of the children who were younger than 3 ½ years old reacted to the first question by either not responding or saying, "Yes." Slocum and Merriman (2018) reported that in these four studies the average correlation between children's tendency to say, “No,” and their tendency to map the novel labels onto an unfamiliar rather than a familiar object was .47, p = .0002.

Wall et al. (2015) administered an object “nameability” judgment task (Marazita & Merriman, 2004) to 3- and 4-year-olds. They asked the children whether they knew names for various objects that were either highly familiar (e.g., a toothbrush) or unfamiliar (e.g., a tube squeezer). During the post-test, children were asked to name any of the latter objects that they had identified as having a known name (i.e., responded “Yes” when asked whether they knew its name). Children tended to either not respond to this post-test question or described some characteristic of the object. In those few instances in which a child named one of the unfamiliar objects correctly or produced a plausible overextension, the object was not counted in the calculation of the accuracy of the child’s initial object nameability judgments. Specifically, the measure of accuracy equaled the average of the proportion of familiar objects they had judged correctly (i.e., answered “Yes” [I know its name]) and the proportion of non-excluded unfamiliar objects they had judged correctly (i.e., answered “No” [I do not know its name]).

In each of two studies, Wall et al. (2015) found that the accuracy of these judgments increased with age, replicating Marazita and Merriman (2004). Furthermore, judgment accuracy predicted the strength of children’s disambiguation effect in a cross-modal paradigm. This paradigm involved teaching children the name for a novel object and then asking them to examine two hidden objects with their hands. One of the hidden objects was identical to the previously novel object that they had just learned the name for, and the other hidden object was unfamiliar. For some of the trials, children were asked which hidden object was the referent of the label that they had just learned. The children nearly always responded correctly. For the other trials, children were asked to decide which hidden object was the referent a particular novel label (i.e., that did not match the label they had just learned). Thus, these trials were tests of the disambiguation effect. Only the children who had made highly accurate object nameability judgments showed the disambiguation effect on these trials. Children's object nameability judgment accuracy remained significantly correlated with successful disambiguation even after statistically controlling for age.

In a recent study (Wall & Merriman, 2020), an adult taught 3- and 4-year-olds a label for a visible novel object, then asked them to look inside a box that contained an unexpected copy of this object and an additional novel object. The adult then asked them to pick the one that was the referent of a novel label. The children showed a weaker disambiguation effect in this paradigm than in previous studies. The authors attributed this weakness to the effect's conflict with children's expectation to communicate about a discovery of mutual interest to themselves and the adult (i.e., the unexpected copy of the object that the adult had just taught them to name). In fact, in a separate condition, in which children were first asked to identify the object in the box that was the same as the training object, they showed a stronger tendency to select the novel object as the referent of the novel label.

Wall and Merriman (2020) also found that regardless of whether children were first asked to identify the object that matched the training object, their tendency to map the novel label onto the novel object was positively correlated with their performance in the object nameability judgment task. That is, those who showed a stronger disambiguation effect also showed a stronger tendency to say, "No" when asked whether they knew the names of unfamiliar objects and say "Yes," when asked whether they knew the names of familiar objects. This correlation was significant even after statistically controlling for children's performance on measures of inhibitory control such as the Kansas Reflection-Impulsivity Scale for Preschoolers (KRISP; Wright, 1971).

This last finding is important because it challenges an alternative explanation for the correlation between awareness of lexical knowledge and the disambiguation effect. According to this explanation, the correlation does not reflect the tendency for awareness of lexical knowledge to promote mapping novel labels onto novel objects. Rather, the correlation merely reflects individual differences in impulsivity. According to this explanation, children who make lexical knowledge judgments impulsively (e.g., saying, “Yes” when asked whether they know the label for some object or the meaning of some label) also tend to answer disambiguation problems impulsively. Conversely, those who inhibit a tendency to make impulsive lexical knowledge judgments also tend to inhibit a tendency to give impulsive answers to disambiguation problems. If this explanation were valid, however, then measures of inhibitory control should have completely mediated the relationship between object nameability judgment and the disambiguation effect. As already noted, they did not.

Younger preschoolers are also less likely than older ones to provide an appropriate justification for the disambiguation effect. Merriman and Schuster (1991) analyzed the data of Merriman and Bowman’s (1989) final experiment, in which children were asked to justify their solutions to the standard disambiguation effect. For example, they asked a child who had selected an unfamiliar object rather than a spoon as the referent of *pilson*, “How did you know this was the *pilson*?” Nearly all of the younger children chose not to respond or referenced some property of the unfamiliar object (e.g., “because it has holes”), whereas many 4-year-olds pointed to the familiar object and said, “Because that one is a [familiar label, e.g., spoon].”) Those 4-year-olds who offered these justifications were more likely to have acknowledged their ignorance of the novel label at the beginning of the test trial, compared to the 4-year-olds who offered either no justification or an inappropriate one.

In addition, those 4-year-olds who offered appropriate justifications for the disambiguation effect also tended to make more accurate lexical knowledge judgments than the other 4-year-olds did. That is, the former group responded more accurately than the latter group when judging either word knowledge ("Do you know what a dack/car is?) or object nameability (e.g., "Do you know what this [familiar/unfamiliar object] is called?"). In 2 ½-year-olds, the tendency to offer an appropriate justification for disambiguation was also correlated with the accuracy of word knowledge judgments, but not with the accuracy of object nameability judgments. Taken together, these findings suggest that as children become aware of what they know and what they do not know, they also develop the ability to think metacognitively about the types of words and objects that are involved in the disambiguation effect.

Two Purely Metacognitive Versions of the Disambiguation Problem

<u>The Hidden Objects Problem.</u> Slocum and Merriman (2018) developed a purely metacognitive version of the disambiguation problem and found that 3- and 4-year-olds' tendency to solve it was associated with their awareness of lexical knowledge. The problem involved first helping the children to sort objects according to familiarity, putting familiar kinds of objects into a bucket for things "I know" and unfamiliar kinds of objects into a bucket for things "I don't know." After sorting the objects, the children were asked to recall the objects that they had put in the "I know" bucket. Typically, children could only remember one or two objects. The researcher then removed the recalled objects from the "I know" bucket and removed an equal number of unfamiliar objects from the "I don't know" bucket to keep the number of objects in each bucket equal. At this point, children knew that the two buckets contained an equal

number of objects and could not recall names for any of the objects. Finally, children received trials in which the researcher told them he had seen a "[label]" in one of the buckets and then asked the children to decide which bucket it was in. On all trials except one, the label was novel (i.e., a made-up word such as *terch* or *cobe*).

Whether children solved this problem most likely depended on their ability to represent the novel label as being a name they did not know and then reasoning that because it was a name that they did not know, the "I don't know" bucket must contain its referent. Since children could not recall the labels of the objects in the "I know" bucket, it is unlikely that they would reject this bucket based on noticing that the novel label mismatched the specific labels for the things in the bucket. It is also unlikely that the "I know" bucket received less activation than the "I don't know" bucket because all of the objects were hidden, and the children could not recall the names of any object in the "I know" bucket. On the other hand, if they represented the label as being one they did not know, it is likely that they would decide that its referent must be in the "I don't know" bucket.

The majority of 4-year-olds tended to select the "I don't know" bucket, whereas 3-year-olds did not favor one bucket over the other (Slocum & Merriman, 2018). Furthermore, children's performance on this task was positively associated with their awareness of their own lexical knowledge, as assessed by word knowledge judgments (i.e., children's tendency to report that they knew the meanings of various familiar words, but not the meanings of various made-up words), even after controlling for age and receptive vocabulary size.

These findings support the claim that as children develop a reflective awareness of their lexical knowledge, they tend to represent the elements of the disambiguation problem in metacognitive terms. Representing the elements metacognitively causes a strengthening of their tendency to select the unfamiliar object over the familiar object as the referent of the label.

Despite the evidence that 4-year-olds can solve a purely metacognitive version of the disambiguation problem, there is no evidence that they ever consult metacognitive representations when solving ordinary disambiguation problems. That is, in circumstances in which they can recall the name for the familiar object, they may avoid selecting it as the referent of a novel label solely because of the mismatch they perceive between the object's known name and the novel label (Halberda, 2006) and/or because the novel label activates the unfamiliar object more than the familiar object (McMurray et al., 2012). In doing so, they might not also consider that they know the familiar object and do not know the novel label or the novel object.

The Disambiguation Prediction Problem. Henning and Merriman (2019) investigated 3- and 4-year-olds' tendency to solve a different version of a purely metacognitive disambiguation problem. In the initial phase of this so-called *disambiguation prediction* task, children were asked to solve four standard disambiguation problems (i.e., decide whether a visible unfamiliar object or a visible familiar object was the referent of a novel label). Each problem was presented as an example of how a game was played. While playing the game, children received corrective feedback regarding their decisions (i.e., if they chose the unfamiliar object, they were told they were correct and if they chose the familiar object, they were told the other object was correct). After four rounds of the game, the prediction problems were presented. The children were told that they were going to continue playing the same game and were shown new pairs of familiar and unfamiliar objects. The only difference was that after a pair was presented, the children were asked to predict which object they thought would be "the right one". Following their prediction, the researcher said, "Now, let's see," and then asked them to select the one that was the referent of some novel label (e.g., a *hust*). That is, the researcher made the same kind of request as in the four first rounds of the game. After the child made a selection, the researcher provided corrective feedback, again just as in the first four rounds of the game.

Henning and Merriman (2019) reasoned that in order for children to solve this disambiguation prediction problem (i.e., consistently predict that the correct choice would be the unfamiliar object), they would have to represent their solutions to the first four standard disambiguation problems in metacognitive terms. They would need to realize that in the first four rounds, the correct object had always been one that evoked a particular kind of cognitive experience (e.g., a feeling of novelty and/or inability to retrieve a label) and the incorrect object had always been one that had evoked a contrasting kind of cognitive experience (e.g., a feeling of familiarity and/or successful retrieval of a label). Additionally, on a prediction trial, children would need to reason that because the pair of objects contrasted in the same manner as the previously presented pairs of objects, the correct choice was the object that evoked the same kind of cognitive experience that the previous correct choices had evoked.

Henning and Merriman (2019) found that both 3- and 4-year-olds performed well in the initial phase of the task (i.e., on the four standard disambiguation problems), but only the 4-year-olds made accurate predictions in the second phase of the task. Slocum and Merriman (2018) had observed this very same age difference in their hidden objects task. Both findings support the hypothesis that as children develop reflective awareness of their lexical knowledge, they also develop an ability to represent the elements of the disambiguation problem in metacognitive terms.

The prediction problem developed by Henning and Merriman (2019) differs from Slocum and Merriman's (2018) hidden objects problem in several important ways. In the prediction problem, the objects are visible, but the novel label is not presented. The opposite is true for the hidden objects problem. That is, the objects are placed inside opaque buckets, but the novel label is presented. Furthermore, solving the prediction problem does not necessarily require children to consciously represent the objects as ones “I know” or “I don’t know”. It may be possible for children to solve it using unconscious metacognitive representations (Karmiloff-Smith, 1986). For instance, children might make correct predictions based on the observation that the correct answer had always been the object that evoked a feeling of novelty rather than a feeling of familiarity. They might not necessarily represent these contrasting feelings as indicating that one object is one they do not know, and the other object is one they know.

Children’s tendency to solve Slocum and Merriman’s (2018) hidden objects problem was found to be associated with the accuracy of word knowledge judgment (e.g., “Do you know what a dack/car is?”). In contrast, children’s tendency to solve Henning and Merriman’s (2019) prediction problem was not related to word knowledge judgment. However, it was significantly related to the accuracy of children’s object nameability judgments (e.g., Do you know what this [familiar/unfamiliar object] is called?). Children who had judged the nameability of every test object correctly made more correct predictions than children who misjudged the nameability of at least one test object (typically, claiming to know the name of an unfamiliar object). Thus, solving the prediction problems may depend on representing the metacognitive contrast between familiar and unfamiliar objects, whereas solving the hidden objects problems may depend on representing the metacognitive contrast between familiar and unfamiliar labels.

Although Henning and Merriman (2019) found that most 4-year-olds could make correct disambiguation predictions, it remains to be seen whether children this age consult general metacognitive representations when solving a single instance of the disambiguation problem. It is possible that children only consult metacognitive representations of objects or labels when they are compelled to consider what multiple instances of the disambiguation problem have in common (i.e., how each instance in the disambiguation prediction task is another round of the same "game"). However, Henning and Merriman (2019)'s results do add further empirical support for some of the presuppositions of the claim that development of metacognitive representations explains why the basic disambiguation effect tends to become stronger with age. The results demonstrated that 4-year-olds can recognize a metacognitive pattern that multiple instances of the disambiguation problem follow, and then use this pattern during a future instance to infer what its solution will be. Neither of these conclusions could be drawn from 4-year-olds's ability to solve Slocum and Merriman's (2018) hidden objects problem.

Chapter 2. Overview of the Investigation

My goal was to examine whether certain experiences might increase preschoolers' tendency to adopt a metacognitive representation of the disambiguation problem, and consequently, increase their ability to solve disambiguation prediction problems (Henning & Merriman, 2019). In Study 1, I examined 3- and 4-year-olds' solutions to these prediction problems when the tester provided them with a general metacognitive representation of the novel labels presented in the initial disambiguation phase. In Study 2, I tested the effect of providing 3-year-olds with clearer corrective feedback regarding their disambiguation predictions.

In Study 1, 3- and 4-year-olds performed a modified version of the disambiguation prediction task. The procedure was identical to the one used by Henning and Merriman (2019, Study 1) except that prior to each trial, the children were told that the label they were about to hear was "a word that you have never heard before." This description presents a metacognitive representation of the label because it describes the label in terms of a general cognitive experience of the child (i.e., never having heard it before). I hypothesized that providing this way of representing the label would promote children's tendency to represent their solution to the initial disambiguation problems in metacognitive terms. Most likely, it would promote their tendency to represent the unfamiliar objects they selected as being ones they had never seen before. If so, they would be more likely to predict during the second phase of the game that the correct choice would also be ones they had never seen before. Therefore, I hypothesized their accuracy during the prediction phase of the task would exceed that of children in Study 1 of Henning and Merriman (2019).

In addition to the disambiguation prediction task, children completed two tests of their awareness of their lexical knowledge -- word familiarity judgment and object nameability judgment. I predicted that children who demonstrated a greater awareness of their lexical knowledge would make more correct disambiguation predictions than children who demonstrated less such awareness.

Study 2 examined whether 3-year-olds might be able to learn to solve disambiguation predictions problems if they received clearer feedback about the accuracy of their predictions. One reason why the 3-year-olds in Henning and Merriman (2019)'s studies did not show improvement over these trials may have been that the feedback they received about their responses was not delivered immediately after they had made a prediction. Rather, after they made a prediction, the experimenter said, "Let's see," then presented a novel label, asked the children to select its referent, then told them whether they had selected the correct referent. Some children might have only used this feedback to decide whether the object they had just selected for the novel label was correct, but not have used it to decide whether the prediction they had made before this selection was correct.

In Study 2, half of the 3-year-olds were tested with the feedback that Henning and Merriman (2019) used in their experiments. I will refer to this as the Indirect Feedback condition because although the feedback directly concerned the accuracy of their mapping of the novel label, it only indirectly concerned the correctness of their initial predictions. The remaining children were tested in the Direct Feedback condition in which they were told immediately after making a prediction whether or not the prediction was correct.

Chapter 3.

Study 1: Predicting the Mapping of a "Word That You Have Never Heard Before"

The study was designed to test two hypotheses. The first was that children would solve more disambiguation prediction problems (Henning & Merriman, 2019) if they heard each novel label described as "a word that you have never heard before." The performance of these children will be compared to that of the children in Study 1 of Henning and Merriman (2019). In that study, no descriptions of the novel labels were provided. The second hypothesis was that children who demonstrate a greater awareness of their lexical knowledge will perform better in this disambiguation prediction task than those who demonstrate less such awareness.

I expected children's prediction performance to exceed that of the children in Henning and Merriman (2019) study because of the new instructions. Successful disambiguation prediction involves detecting a metacognitive invariant in the first four rounds of the game. The invariant is that the correct choice is always an object that is an unfamiliar kind and/or that has no known label and/or that feels unfamiliar, and the incorrect choice is always an object that has the opposite epistemic quality or qualities. Telling children that each novel label is a word that they had never heard before might increase their likelihood of representing the correct choices as ones that match the epistemic quality of the label. It might help them to realize that on each trial they mapped a word that they had never heard before onto a kind of object they had never seen before and/or a kind of object whose name they had never heard before.

Regarding the second hypothesis, awareness of lexical knowledge will be assessed with two tasks – word familiarity judgment and object nameability judgment. These measures were expected to correlate significantly with disambiguation prediction performance even after statistically controlling for child age and receptive vocabulary size. It was not clear whether these measures would show the same relation to disambiguation prediction performance. On the one hand, object nameability judgment (i.e., accuracy of reporting whether various familiar and unfamiliar objects have known names) might show the stronger relation because this was what Henning and Merriman (2019) found. Both object nameability judgement and disambiguation prediction involve visible objects and a "hidden" (not-yet-presented) label. So, the more relevant kind of metacognitive awareness might be one that concerns visible objects rather than spoken labels.

Alternatively, because of the change in the instructions, word knowledge judgment (i.e., accuracy of reporting whether various familiar and unfamiliar words have known meanings) might show the stronger relation to disambiguation prediction performance. Because the current study presented children with a metacognitive representation of the label, the more relevant metacognitive ability might be one concerning labels.

Method

Participants. Forty-eight children were recruited from preschools in the middle- to upper-class regions of Northeast Ohio. There were 24 3-year-olds (M_{age} = 3;8, range = 3;3–3;11, 8 girls) and 24 4-year-olds (M_{age} = 4;6, range = 4;2–4;11, 11 girls). Nearly all were Caucasian, and all were monolingual speakers of English. Each child received a few stickers for participating.

Materials. A different pair of objects was used on each of eight trials in the disambiguation prediction task. The objects used in the current study were the same objects used by Henning and Merriman (2019). One member of each pair was familiar (comb, scissors, flower, spoon, pencil, cup, glove, or ball) and one was unfamiliar (single serve coffee filter, cable hider, tea infuser, PVC plug, conduit cover, T-clip, lamp cover, or pourer spout). The objects could be easily handled by the children due to their small size. Eight pseudowords (*jat*, *terch*, *lerb*, *borp*, *mig*, *geet*, *dack*, and *cheed*) were used as novel labels (Henning & Merriman, 2019).

The materials used in object nameability judgment task included six familiar objects (flashlight, key, sock, toy car, fork, and toothbrush) and six objects that were likely to have unknown labels (tube squeezer, egg slicer, plate hanger, spouncer, latch hook, and heel cushion). These objects were used in previous administrations of the object nameability judgment task (see Hartin, Stevenon, & Merriman, 2016; Henning & Merriman, 2019). The word knowledge judgment task included five familiar words (*shoe*, *dog*, *truck*, *chair*, and *house*) and five pseudowords (*hust*, *mave*, *gock*, *prad*, and *blim*).

Procedures. Children participated in a 15-30 minute one-on-one session at their preschool. The tester and the child sat opposite of each other at a small table in a quiet room. After the child was comfortable and assented to participate, the disambiguation prediction task, the object nameability judgment task, the word knowledge judgment task, and the Peabody Picture Vocabulary Test – Fourth Edition (PPVT-4; Dunn & Dunn, 2007) were presented. The disambiguation prediction task was always presented first, with the order of the other tasks counterbalanced.

Disambiguation Prediction. The child was told that he or she was going to play a game. The tester said, “In this game, I am going to show you two things and ask you to pick the one that I name. I am going to say a word that you have never heard before. Let’s give it a try.” The tester placed one of the pairs of objects on the table and asked the child to select the referent of a novel label (e.g., “Which one is the *jat*?”). If the selection was correct (i.e., the unfamiliar object), the tester confirmed it (e.g., “Good job! That is right.”). If the selection was incorrect, the tester corrected it (e.g., “No, that is not the *jat*. This one (pointing to the unfamiliar object) is the *jat*. This one’s right.”). The objects were removed, and the tester stated, “Let’s play that game again.” The same procedure was repeated for three additional trials, each involving a different pair of objects and different novel word. Following the fourth trial, children were asked to make their first prediction. The tester placed another pair of objects on the table and announced, “Let’s play that game again. I am going to say a word you have never heard before. Which one do you think is going to be right?” After the child made a prediction, the tester said, “Now let’s see,” and then asked the child to select the referent of a novel label (e.g., “Which one is the *terch*?”). After the child’s selection, the tester either confirmed or corrected the selection, depending on whether it was correct. The remaining three trials followed the same procedure as the fifth trial, except that each involved a different pair of objects, a different novel label and children were asked to justify their selection (i.e., “Why do you think it’s that one?”). To summarize, the children were asked to predict which of the two objects was “going to be right”, told “now let’s see,” and asked which one was the referent of a novel label, and asked to justify the object they chose as the referent.

Object Nameability Judgment. The child was told, “Are you ready to play a game? I’m going to show you an object and you’re going to tell me if you know the name for it. If you know the name, say “Yes”. If you don’t know the name, say “No”. I’ll show you how it’s done.” The tester then demonstrated to the child a familiar object trial (e.g., flashlight) and said, “Let’s see, do I know the name for this? Yes, I do. It’s a flashlight. So I would say, “Yes, I know its name. I know what it’s called. Let’s try another one.” The tester then demonstrated to the child an unfamiliar object trial (e.g., plater hanger) and said, “Hmm, do I know the name for this? No, I don’t. I don’t know what this thing is called. So, I would say, ‘No, I don’t know its name.’ Now, you try it.” The tester then presented six familiar objects and six unfamiliar objects one at a time in a random order to the child and asked the child, “Yes or no. Do you know the name for this?” The tester responded (e.g., “Very good”) regardless of the child’s response. Once the responses were recorded, the tester conducted a post-test in which children were asked to name any of the unfamiliar objects that they had identified as having a known name (e.g., “You said you knew the name for this. What is it called?”). Children tended to either not respond to this post-test question or described some characteristic of the object. In those instances, in which a child named one of the unfamiliar objects correctly or produced a plausible overextension (e.g., calling the spouncer a “sponge”), the object was not counted in the calculation of the accuracy of the child’s initial object nameability judgments. (See Hartin et al., 2016, for further details about the procedures in this task.)

Word Knowledge Judgment. The child was told, “I’m going to play a Yes-No game with you. I’m going to say some words. Listen carefully because some of the words are ones that you know and some are ones that you don’t know. I’m going to say a word, and then I’m going to ask you whether you know it. Just say “Yes” or “No.” Let me show you how to play.” The tester then demonstrated with one familiar word and one pseudoword. “The first word is book. Do you know what a book is? Say, “Yes.” Book is a word that you know. You’ve heard that word before. You know what a book is. Are you ready for another one? The next word is *zimbiddy*. Do you know what a *zimbiddy* is? Say, "No." *Zimbiddy* is not a word that you know. It's a made-up word. You've never heard that word before. There is no such thing as a *zimbiddy*." The word knowledge judgment task followed. The child was asked, "Do you know what a _____ is?" regarding the five familiar words and five pseudowords. The child's accuracy score was the proportion of the ten judgments made correctly. Order of presentation was random, except that words of the same type never occurred more than twice in a row. The tester conducted a post-test for any pseudoword the child claimed to have known, saying "You said you know what a _____ is. What is a _____? Can you tell me anything about it?" If it was evident that the child had misidentified a pseudoword as a similar-sounding familiar word (e.g., "wave" for *mave*), this trial was excluded from computation of the accuracy score. (See Hartin et al., 2016, for further details about the procedures in this task.)

Results

Disambiguation Prediction. Children's performance in the disambiguation prediction task was analyzed in three parts: selections on the initial four trials, predictions made on the final four trials, and selections made on the final four trials. Note that each of the final four selections were made in response to a specific novel label. The unfamiliar object was considered the correct selection or prediction on every trial (see Table 1). The critical measure was how frequently the children made correct predictions. If they were able to make these at a rate significantly above chance, this would be evidence that they had noted a metacognitive commonality among their correct selections in four initial disambiguation trials and applied it to the disambiguation prediction problems.

On the initial disambiguation trials, both age groups performed quite well (*M* correct = 3.69 out of 4), which is consistent with previous results in which corrective feedback has been delivered after every selection in a disambiguation task (see Marazita & Merriman, 2004; Henning & Merriman, 2019). Performance on these trials did not differ for the two age groups, t (46) = .78, p = .44, d = .23. Both 3- and 4-year-old's mean number correct exceeded chance (2.00) (4-year-olds: t (23) = 16.13, p < .001, d = 3.29; 3-year-olds: t (23) = 13.83, p < .001, d = 2.82). These results replicated those of Henning and Merriman (2019). Mean number of correct selections on the initial trials did not differ from that reported by Henning and Merriman (2019, Study 1) for either 3-year-olds, t (46) = .42, p = .68, d = 0.12, or 4-year-olds, t (46) = .53, p = .60, d = 0.15. This result is important for ruling out a potential explanation for incorrect responses on the prediction trials. If children had not made consistently correct selections on the initial trials, then one could argue that they went on to make incorrect responses on the prediction trials because there was not a consistent metacognitive property to their own selections in the task. That is, because they had not consistently selected the unfamiliar object, they would not have had an opportunity to notice that every correct selection had a metacognitive property that contrasted with a metacognitive property of the incorrect selections (e.g., every correct object evoked a feeling of novelty, and every incorrect object evoked a feeling of familiarity). If they did not notice this pattern, then they would not be able to infer the solutions to the disambiguation prediction problems.

On the critical disambiguation prediction trials, mean number correct was significantly greater in the 4-year-olds (3.00 out of 4) than in the 3-year-olds (2.21), t (46) = 2.29, p = .03, d = 0.67. Four-year-old's performance exceeded chance (2.00), t (23) = 4.15, p < .001, d = .84, whereas 3-year-old's performance did not differ from it, t (23) = .84, p = .41, d = .17. Most 4-

year-olds (17 of 24) made at least three correct predictions, whereas only about half of 3-year-olds (11 of 24) did so. Henning and Merriman (2019, Study 1) found a similar age difference. Mean number of correct predictions in the current study differed from those in Henning and Merriman (2019, Study 1) for 3-year-olds, one-tailed t (46) = 1.75, p = .05, d = 0.51, but not for 4-year-olds, t (46) = -.13, p = .90, d = 0.04. Thus, there was evidence that the revised instructions of the current study led to an overall increase in 3-year-old children's tendency to make accurate disambiguation predictions.

The very first prediction trial has special significance because it is the only prediction trial that is free from possible carryover effects from previous prediction trials. For example, when asked to predict whether a familiar or an unfamiliar object would be correct in some task, children might have some disposition to select the familiar object. This disposition might increase from one trial to the next if it was not fulfilled on the first trial. That is, it might be difficult to maintain a tendency to correctly predict that the unfamiliar object would be correct after the first trial if it became increasingly difficult to continue to inhibit a disposition to select the familiar object. In addition, some children might have an expectation that the familiar object would become the correct choice at some point in the game. On the other hand, because children received feedback on their performance on prediction trials, they might learn from this feedback and show an increase over trials in their tendency to make correct predictions. Therefore, performance on the very first prediction trial may be a better measure than performance on the other trials of whether a child had noticed the metacognitive quality that the correct answers to the initial disambiguation problems had in common. (See Evey and Merriman (1998) for a similar argument for why the first trial of a standard disambiguation problem is better than the subsequent trials for gauging the prevalence of the disambiguation effect.)

On the very first prediction trial, most 3-year-olds (19 of 24) made a correct prediction. This proportion exceeded chance, $\chi^2 (1) = 8.17$, p = .008, and did not differ from the proportion of 4-year-olds (22 of 24) who made a correct prediction on this trial, two-tailed Fisher's exact test $p = .42$. In Henning and Merriman (2019, Study 1), the proportion of 3-year-olds who made a correct prediction on the first trial (12 of 24) did not differ from chance, and was less than the proportion of 4-year-olds (23 of 24) who did so, two-tailed Fisher's exact test $p = .001$. Also, the proportion of 3-year-olds who made correct predictions on this trial in Henning and Merriman (2019, Study 1) was lower than in the current study, one-tailed Fisher's exact test $p = .034$.

Immediately following a child's prediction, the tester said, "Now let's see," then presented a novel label and asked the child to select its referent. Both age groups tended to select the unfamiliar object ($M = 3.31$ out of 4) and did not differ in how often they made this selection, $t(46) = .72$, $p = .47$, $d = .21$. Mean number correct exceeded chance in both 4-year-olds, $t(23) = 6.09$, $p < .001$, $d = 1.24$, and 3-year-olds, $t(23) = 7.11$, $p < .001$, $d = 1.45$. The number of correct selections following predictions did not differ from the number made in the initial (non-prediction) phase of the task in either age group. Once again, mean number of correct selections following predictions matched those of Henning and Merriman (2019, Study 1) for both 3- and 4-year-olds, $t(46) = -.32$, $p = .75$, $d = 0.09$; $t(46) = -.56$, $p = .58$, $d = 0.16$, respectively.

Henning and Merriman (2019, Study 1) found that 3-year-olds made significantly more correct predictions on the first two than on the final two prediction trials. (The 4-year-olds showed a similar trend, but it was not significant.) Thus, the 3-year-olds' prediction performance declined even though they tended to make correct selections when the novel label was presented (after the prediction) and were told whether these selections were correct. Three-year-olds in the current study also tended to make more correct predictions on the first two than on the final two prediction trials (*M* change = -0.19, *SD* = .46), one-tailed $t\ (23) = 1.99$, $p = .029$, $d = .41$. Among 4-year-olds, there was no significant difference (*M* change = -0.08, *SD* = .41), one-tailed $t\ (23) = 1.00$, $p = .16$, $d = .20$. As already noted, a decline in performance over prediction trials may be due to an increase over trials in the difficulty of inhibiting a disposition to select the familiar object and/or to the expectation that the familiar object would become the correct choice at some point in the game.

Other measures and task intercorrelations. Children's performance on the two lexical knowledge judgement tasks and the PPVT-4 is summarized in Table 2. Regarding object nameability judgment, both age groups performed quite well and did not differ, $t\ (46) = .84$, $p = .41$, $d = .24$. However, 4-year-olds outperformed 3-year-olds on both word knowledge judgment, $t\ (39) = 3.87$, $p < .001$, $d = 1.13$, and PPVT-4 (scores not standardized by age), $t\ (44) = 4.93$, $p < .001$, $d = 1.45$. In both lexical knowledge judgment tasks, children's primary error was to claim to know unfamiliar kinds of stimuli rather than to deny knowing familiar kinds of stimuli. This error pattern is typical for these kinds of judgments (Marazita & Merriman, 2004; Henning & Merriman, 2019; Hartin et al., 2016) as well as for other kinds of knowledge judgments (Aguiar, Stoess, & Taylor, 2012; Sodian & Wimmer, 1987).

Table 3 summarizes intercorrelations among disambiguation prediction (both overall and first trial performance), age (in months), object nameability judgment accuracy, word knowledge judgment accuracy, and vocabulary size. Disambiguation prediction first trial performance was only associated with word knowledge judgment accuracy and overall disambiguation prediction performance was only associated with word knowledge judgment accuracy and age. As predicted, children who responded more accurately when asked whether they knew various words and pseudowords also made more correct disambiguation predictions. Age did not fully account for the association between overall disambiguation prediction performance and word knowledge judgment accuracy; the correlation between the latter two remained significant even after associations with children's age were partialled out, partial r (44) = .31, p = .04. Also, there was no evidence that age predicted performance on the disambiguation prediction test beyond what could be predicted by word knowledge judgment accuracy, partial r (44) = .20, p = .18.

Performance on the two lexical knowledge judgment tasks was significantly intercorrelated, even after statistically controlling for age and vocabulary, partial r (41) = .38, p = .01, which also replicates previous findings (Merriman & Lipko, 2008; Lipowski & Merriman, 2011; Hartin et al., 2016). Because two-thirds of the children made no errors in the object nameability judgment task, restriction of range was a concern. Therefore, I examined correlations for a dichotomized version of this measure, where errorless performance was scored as 1 and less-than-perfect performance scored as 0. This measure was not strongly correlated with overall disambiguation prediction performance, r (46) = .13, p = .38, first trial disambiguation prediction performance, r (46) = -.04, p = .78, or word knowledge judgment accuracy, r (45) = .24, p = .10.

Discussion

The main prediction was that more children would solve a disambiguation prediction problem if they were told that the label they were about to hear was "a word that you have never heard before." This prediction was supported in the 3-year-olds. This age group's mean number of correct disambiguation predictions was greater than that of the 3-year-olds in Study 1 of Henning and Merriman (2019). The procedures in the latter study were the same except that children were told nothing about the label they were about to hear. Additionally, the number of

3-year-olds who solved the very first prediction problem was greater than chance in the current study but was right at chance in Study 1 of Henning and Merriman (2019).

Although a substantial majority of the 3-year-olds in the current study solved the very first prediction problem, they did not maintain this tendency over the remaining prediction trials (*M* proportion correct on these trials = .47, *SD* = .35; chance = .50). A smaller decline was evident in the 3-year-olds in Study 1 of Henning and Merriman (2019), where 50% solved the very first prediction problem and *M* proportion correct on the remaining trials was .35 (SD = .35). So, the new instructions primarily affected 3-year-olds' performance on the very first prediction problem. Even the 3-year-olds who made the correct prediction on this problem showed little tendency on subsequent trials to predict that a novel object rather than familiar object would be the referent of a "word you have never heard before" (*M* proportion = .51).

There was no evidence that 4-year-olds benefitted from the new instructions. One reason for this is that nearly every 4-year-old solved the very first prediction problem in Study 1 of Henning and Merriman (2019). So, there was little room for improvement (i.e., a ceiling effect) in this age group on the very problem that was most affected in the 3-year-olds. Regarding the remaining prediction problems, although there was room for the 4-year-olds to improve on these, this age group performed no better on them in the current study than in Study 1 of Henning and Merriman (2019). (*M* correct was .69 in both studies.)

The information conveyed by the new instructions (i.e., that the word was one that they had never heard before) may not have been news to most of the 4-year-olds. That is, they may have already noticed by the end of the initial phase of the procedure that every trial had involved a word that they had never heard before. So, when asked to make their first disambiguation prediction, even the 4-year-olds in Study 1 of Henning and Merriman (2019) anticipated that the solution would be the referent of a novel label.

Some of the 4-year-olds and most of the 3-year-olds who made a correct prediction on the very first prediction trial went on to make an incorrect prediction on at least one of the remaining prediction trials. Even in the current study, where they were explicitly told that the solution to each problem would be the referent of a novel word, they often predicted that a familiar object would be the right choice. After the first prediction trial, children may experience an increase in their preference to select a familiar rather than a novel object and/or an increase in their expectation that the familiar object will become the correct choice.

The decline in prediction performance after the first prediction trial in both the current study and Study 1 of Henning and Merriman (2019) are somewhat surprising because children received corrective feedback after every trial. However, this feedback may have been too indirect. The feedback directly concerned whether the child had just selected the correct object as the referent of the novel label and only indirectly concerned whether their initial predictions had been correct. Children who mistakenly predicted that a familiar object "was going to be right", but then when asked which object was a "pilson" [or some other novel label] tended to select the unfamiliar object. When told that they had made the correct selection, they might not have inferred that this meant that their prediction had been incorrect.

One measure of awareness of lexical knowledge was positively associated with solving disambiguation prediction problems. Word knowledge judgment accuracy, but not object nameability judgment accuracy, predicted overall and first trial disambiguation prediction performance. The opposite pattern of results was found in Henning and Merriman (2019, Study 2). The correlation between word knowledge judgment accuracy and overall disambiguation prediction performance in their study (r = .11, z_r = .11) differed from the correlation in the current study (r = .43, z_r = .46), one-tailed $z_{difference}$ = 1.65, p = .049. This pattern may reflect how success in both the word knowledge judgment task and the current version of the disambiguation prediction task may depend on representing novel labels as unknown (see Slocum and Merriman, 2018, for similar results). The cue in the word knowledge judgment task is a single label (e.g., "Do you know what a truck/hust is?") and the additional cue in the current version of the disambiguation prediction task concerns a single label (i.e., "word that you have never heard before"). In Henning and Merriman's (2019) version of the disambiguation prediction task, the child does not receive a description of the novel label in either the initial disambiguation selection trials or in the disambiguation prediction trials. The only cue in each prediction trial is a pair of objects, one familiar and one unfamiliar. Thus, the relevant metacognition in the original version of the disambiguation prediction task concerns the epistemic contrast between such objects (that is, the realization that one is known and the other is not). Children must focus on this same contrast to make accurate responses in the object nameability judgment task.

Chapter 4. Study 2: The Effect of Direct vs. Indirect Feedback

In Study 1 and in Henning and Merriman's (2019) studies, 3-year-olds' tendency to solve disambiguation prediction problems declined over trials. The main goal of Study 2 was to examine whether giving 3-year-olds clearer feedback might help them to improve, or at least maintain, their prediction performance.

As already noted, the previous studies have used an indirect form of feedback. After children made a prediction, they were asked to select the referent of a novel label and then were told whether their selection was correct. Thus, some might have only learned whether their selection was correct, but not have inferred whether the prediction that they made prior to selecting the referent of the novel label was correct.

In the current study, 3-year-olds in the direct feedback condition were told immediately after making a prediction whether it was correct. The remaining children received the indirect form of feedback that had been used in the previous studies. The main prediction was that children in the direct feedback condition would outperform those in the indirect feedback condition once they began to receive feedback (i.e., after they received feedback about their first prediction).

One major change from Study 1 and the studies of Henning and Merriman (2019) was that the children were tested virtually rather than face-to-face, due to the COVID-19 pandemic. The switch to virtual testing entailed making several procedural changes. Another change was to present only 20 items from the PPVT-4. Preschoolers find the full PPVT very challenging to complete even when tested face-to-face.

Method

Participants. Forty-eight 3-year-olds were recruited via online sources (e.g., lab website, children's science website, and flyers on social media). All of the children were monolingual speakers of English. Twenty-four were assigned to the direct feedback condition (M_{age} = 3;6, range = 3;0-3;11, 13 girls) and 24 were assigned to the indirect feedback condition (M_{age} = 3;7, range = 3;0-3;11, 11 girls). Each child received a gift card for participating.

Materials. All tasks were administered remotely using an Alienware Aurora R7 Desktop. To participate, individuals were required to use a computer, tablet, or smartphone with access to a reliable internet connection. Zoom, a videoconferencing software program, was used to communicate and interact with the participants for the duration of the testing period. The disambiguation prediction task, object nameability judgment task, and the abridged version of the PPVT-4 (Dunn & Dunn, 2007) (which I shall refer to as the Kent PPVT20) required the tester to share their computer screen with the participants in order for the participants to complete these tasks. These three tasks were programmed and run using Microsoft PowerPoint. The objects and labels used in the tasks were the same as in Study 1. High-resolution photos of the objects were presented instead of the physical objects themselves due to concerns with video quality based on participants' internet connection speeds. The materials used for the word knowledge judgement task were the same as in Study 1.

The Kent PPVT20 was developed by examining how the 3- and 4-year-olds in Study 2 of Henning & Merriman (2019) performed on each item of the PPVT-4. Because the items in the

PPVT-4 become increasingly challenging as the test progresses, I started with an item that nearly every child got right and then included every third item until I had a set of 20 items. Items that had low point-biserial correlations with total score on Kent PPVT20 and would increase Cronbach's alpha if removed from the scale were replaced with other items from the PPVT-4. For the final 20 item set, *M* number correct = 13.54, *SD* = 4.73, *range* = 0 to 19, and Cronbach's alpha = .86. Performance on this set was significantly correlated with overall performance on the PPVT-4, $r = .88$, $p < .001$, as well as performance on the PPVT-4 items that remained after the 20 items were removed, $r = .80$, $p < .001$.

The scale was cross-validated by examining how children in Study 1 of the current investigation performed on the items from the scale. Results were similar. *M* number correct = 13.15, *SD* = 4.63, *range* = 1 to 20, and Cronbach's alpha = .85. Performance was significantly correlated with overall PPVT-4, $r = .94$, $p < .001$, as well as performance on the PPVT-4 items that remained after the 20 items were removed, $r = .90$, $p < .001$. Children's performance on each item in order of task completion is summarized in Table 4.

Procedures. The tester and the child participated in a videoconference session for approximately 20 to 30 minutes. Once the child was comfortable and assented to participate, the disambiguation prediction task, the object nameability judgment task, the word knowledge judgment task, and the Kent PPVT20 were completed. The disambiguation prediction task was always completed first, with the order of the other tasks being counterbalanced.

<u>Disambiguation Prediction.</u> Children were tested in one of two conditions. The tester told the child that they were going to play a game and said, "In this game, I am going to show you two things and ask you to pick the one that I name. Let me show you how to play." The tester presented the child with a training trial which showed a vertically split screen, with the left side

of the screen being blue and the right side yellow. On the blue side of the screen was a centered black circle and on the yellow side of the screen was a centered black triangle. The tester asked the child to point to, "Which one is the circle?", followed by, "Which one is the triangle?". The tester made sure the child could correctly identify and select each shape before moving on to the testing phase.

Direct Feedback Condition. To begin the testing phase, the tester stated, "Now that you know how to play the game, let's play. Remember that I am going to show you two things and ask you to pick the one that I name. Let's give it a try." The tester presented one of the pairs of objects and asked the child to select the referent of a novel label (e.g., "Which one is the *jat*?"). If the selection was correct (i.e., the unfamiliar object), the tester confirmed it (e.g., "Good job! That is right. That is the *jat*."). If the selection was incorrect, the tester corrected it (e.g., "No, that is not the *jat*. The *jat* is on the yellow side. Can you point to the *jat*?"). Objects were then removed, and the tester announced, "Let's play that game again." After the first trial, the tester just asked the child, "Which one is the *lerb* (or whatever the next novel word is)?" The same procedure was repeated for three additional trials, each involving a different pair of objects and different novel word. Following the completion of the fourth trial, children were asked to make their first prediction. The tester presented another pair of objects and announced, "Let's play that game again. Which one do you think is going to be right?" After the child made his or her prediction, the tester immediately provided feedback. If the child's prediction was correct, the tester confirmed it (e.g., "Yes, that's right because I was going to ask you which one is the *terch* and that one is the *terch*."). If the prediction was incorrect, the tester corrected it (e.g., "No, that's not right because I was going to ask you which one is the *terch* and the *terch* is on the blue side.

Can you point to the *terch*?”). The remaining three trials followed the same procedure as the fifth trial, except that each trial involved a different pair of objects and different novel label.

Indirect Feedback Condition. The tester followed a procedure that was the same as the one used in Henning and Merriman’s (2019, Study 1) except for modifications needed to administer it online. The procedure for the first four trials was the same as in the direct feedback condition. The final four trials, which were the prediction trials, each involved the child making their prediction first, followed by the child being asked to select the referent of the novel label (e.g., “Let’s see, which one is the *terch*?”). In summary, children were asked to predict which of the two objects was “going to be right”, told “now let’s see,” and asked which one was the referent of a novel label and informed whether their selection was right or wrong.

The procedures for the two judgment tasks were identical to those of Study 1 with the exception that photos of objects were used as stimuli in the object nameability judgment task. The Kent PPVT20 was administered using Microsoft PowerPoint. Its items were presented one slide at a time. Each slide showed color drawings of four objects or four events that were arrayed in a 2 x 2 matrix. The child was asked to select the picture that was the referent of a test word. A different colored star was positioned next to each picture. Because it was not always clear which picture the child had selected, they were asked to say the color of the star next to the picture they selected.

Results

Disambiguation Prediction. As in Study 1, children’s performance in the indirect feedback condition was analyzed in three parts: selections on the initial four trials, predictions made on the final four trials, and selections made (after each prediction) on the final four trials (see Table 5). The indirect feedback condition served as a direct replication of Henning and Merriman (2019, Study 1) because the procedures and stimuli were nearly identical. Thus, I expected that the accuracy of predictions in the indirect feedback condition would not exceed chance, replicating Henning and Merriman (2019)'s finding for 3-year-olds. I also expected to replicate their finding that the number of children who responded correctly on the very first prediction trial did not exceed chance.

The direct feedback condition served as a comparison group to determine whether the type of feedback affected children's prediction performance. If 3-year-olds were able to learn to make disambiguation predictions from direct feedback, then I expected that the accuracy of the predictions in the direct feedback condition after the first prediction trial (i.e., after a chance to receive feedback) would exceed chance and would exceed that of the children in the indirect feedback condition. The accuracy of the first prediction trial was not expected to differ in the two feedback conditions.

On the initial disambiguation trials, children performed quite well in both the direct feedback (M correct = 3.58 out of 4) and indirect feedback (M = 3.63) conditions. Performance on these trials did not differ by condition, t (46) = -.21, p = .83, d = 0.06. Mean number correct exceeded 2.00 (chance) in both conditions (direct: t (23) = 11.86, p < .001, d = 2.42; indirect: t (23) = 11.20, p < .001, d = 2.29). As expected, mean number of correct selections on these trials in the indirect condition did not differ from what Henning and Merriman (2019, Study 1) reported for 3-year-olds, t (46) = .39, p = .70, d = 0.11.

Performance on the first prediction trial was not expected to differ by condition. In both conditions, children only received feedback about a prediction after they had made their first prediction. Results supported this expectation. In both conditions, 13 of 24 made the correct prediction. This proportion was similar to the proportion of 3-year-olds in Henning and Merriman (2019, Study 1) who made the correct prediction on the first trial (12 of 24).

On the final three critical prediction trials, mean number correct was significantly greater in the direct feedback condition (1.71) than in the indirect feedback condition (1.21), $t(46) = 2.05$, $p = .046$, $d = 0.59$. These results support the prediction that the 3-year-olds would benefit from receiving direct rather than indirect feedback about their predictions. The performance of children in the indirect feedback condition on the three critical prediction trials essentially replicated that of the 3-year-olds in Henning and Merriman (2019, Study 1) (M correct = 1.04). The difference between the mean scores of these two groups was not significant, $t(46) = .63$, $p = .53$, $d = 0.18$. The mean scores were significantly less than 2.00 (chance) in both groups (indirect feedback condition: $t(23) = -4.96$, $p < .001$; Henning & Merriman: $t(23) = -4.52$, $p < .001$). In contrast, mean scores on the critical prediction trials in the direct feedback condition were not different from chance, $t(23) = -1.56$, $p = .15$.

The effect of feedback on children's predictions was evident on the first prediction they made after receiving feedback (i.e., on the second prediction trial). Seventeen of 24 in the direct feedback condition made the correct prediction, compared to only 8 of 24 in the indirect feedback condition, two-tailed Fisher's exact test $p = .02$. The number in the direct feedback condition who made the correct prediction exceeded chance, the exact binomial calculation of the chance probability of this number or higher is .032. The effect was short-lived, however, as only 11 in the direct feedback condition made the correct prediction on the next trial, compared to 13 in the indirect feedback condition, two-tailed Fisher's exact test $p = .77$. On the final prediction trial, 13 in the direct feedback condition and 8 in the indirect feedback condition made the correct prediction. Although this difference was in the expected direction, it was not significant, two-tailed Fisher's exact test $p = .24$.

Change in performance from the first prediction to the second prediction was further evaluated by separating those who had made the correct prediction on trial 1 from those who made the incorrect prediction on this trial. Among the first group (correct initial predictors), 10 of the 13 in the direct feedback condition and 7 of 13 in the indirect feedback condition also predicted correctly on the second trial. This difference was not significant, two-tailed Fisher's exact test $p = .41$. Among the children whose initial predictions were incorrect, however, 7 of 11 in the direct feedback condition and 1 of 11 in the indirect feedback condition predicted correctly on the second trial. This difference was significant, two-tailed Fisher's exact test $p = .024$. Thus, the immediate effect of feedback was primarily due to the behavior of children who made an incorrect first prediction. They were highly likely to repeat this mistake on the next trial if they had received an indirect form of correction. The probability that those receiving indirect feedback would repeat the mistake exceeded chance. The exact binomial calculation of the chance probability of at least 11 of 12 repeating the mistake is .003. In contrast, among those who received direct feedback, about as many repeated the mistake (5) as corrected it (7), which is close to the number (6) expected by chance. These differential effects of feedback were limited to just the first prediction that children made after receiving feedback (i.e., responses on the second prediction trial).

Only in the indirect feedback condition were children asked to select the referent of a novel label during the prediction phase of the procedure. As already noted, they were asked to do this immediately after making a prediction. They tended to select the unfamiliar object ($M = 3.17$ out of 4) more often than would be expected by chance, $t\ (23) = 6.58$, $p < .001$, $d = 1.34$. The mean number of these selections did not differ from that of the 3-year-olds (3.29) in Henning and Merriman (2019, Study 1), t (46) = -.48, p = .64, d = -.14. It did differ from the mean number of correct selections that the indirect feedback condition made during the initial selection phase of the procedure (*M change* = - .46*,* *SD* = .98, t (23) = 2.30, p = .03, d = .47.)

Other measures and task intercorrelations. Performance on the two lexical knowledge judgment tasks and the Kent PPVT20 is summarized in Table 6. Scores on these measures did not differ by condition: object nameability judgment, t (45) = .67, p = .51, d = .19; word knowledge judgment, t (45) = -.44, p = .66, d = .13; Kent PPVT20, t (45) = -.55, p = .59, d = .16. Children's primary error in both lexical knowledge judgment tasks was the same as in Study 1, which was to claim to know unfamiliar kinds of stimuli rather than to deny knowing familiar kinds of stimuli. The mean level of performance on word knowledge judgment (.86) was better than that by the 3-year-olds in Study 1 (.72), t (68) = 2.89, p = .005, d = .74; but was no different than the 3-year-olds in Henning & Merriman (2019, Study 2) (.78), t (69) = 1.57, p = .12, d = .40. The mean level of performance on object nameability judgment (.89) did not differ from the 3-year-olds in Study 1 (.89), t (70) = .06, p = .96, d = .01; nor did it differ from the 3-year-olds in Henning & Merriman (2019, Study 2) (.89), t (70) = .06, p = .95, d = .02. Object nameability judgment performance for the 3-year-olds approached ceiling in all three studies. The mean level of performance on Kent PPVT20 (13.31) was better than that by the 3-year-olds in Study 1 (10.46), t (70) = 3.07, p = .003, d = .77; and was better than that of the 3-year-olds in Henning & Merriman (2019, Study 2) (11.21), t (70) = 2.17, p = .033, d = .54. Taken together, these findings indicate that the current sample recruited for the virtual testing procedure may have been more advanced than the samples in previous studies. Even so, the disambiguation prediction results in the indirect feedback condition were no different from what Henning and Merriman (2019) found using indirect feedback.

Table 7 summarizes intercorrelations among disambiguation prediction (overall & first trial performance), age (in months), object nameability judgment accuracy, word knowledge judgment accuracy, and vocabulary size for all children in Study 2. Neither overall nor first trial disambiguation prediction performance was associated with the other measures. Because feedback affected children's performance on the final three prediction trials, correlates of performance on these trials were examined separately for each feedback condition. In neither condition was performance on these trials correlated with any other measure (all $p > .10$).

Henning and Merriman (2019, Study 2) found that for both 3- and 4-year-olds, those who made no errors when judging object nameability were more likely to respond correctly on the very first disambiguation prediction trial compared to their agemates ($\Phi = .49$, $p < .0001$). This finding did not replicate ($\Phi = .04$, $p > .10$).

Discussion

The main hypothesis was supported. The 3-year-olds who received direct feedback made more correct predictions than those who received indirect feedback. This difference was evident in the first prediction that children made after receiving feedback (i.e., on the second prediction trial). In the direct feedback condition, the number who responded correctly on this trial exceeded chance. In the indirect feedback condition, the number who responded correctly on this trial fell below chance.

Children who received indirect feedback might not have interpreted it as confirming or rejecting the prediction they had made before the novel word was presented. They might have only interpreted it as confirming or rejecting the object they had just chosen as the referent of the novel word. That is, they did not think about what the feedback implied about the accuracy of their initial prediction.

A second possibility is that because the children in the indirect feedback condition received no feedback immediately after making their prediction, some may have interpreted the absence of feedback as a cue that they may have made the wrong prediction. In an analysis of transcripts of interactions between young children and adults, Bonawitz et al. (2020) found that when adults followed up a young child's statement by repeating all or part of the question, rather than by directly confirming or rejecting the child's statement, the probability that the child's statement was false was very high. Krahenbuhl, Blades, and Eiser (2009) found that when 4-year-olds were asked about their memory for a witnessed event, they often changed their answers in response to repeated questioning, resulting in a decline in accuracy (see also Howie, Sheehan, Mojarrad, & Wrzeinska, 2004; Poole & White, 1991). In several experiments by Bonawitz et al. (2020), preschool-age children often switched their answers in a guessing game when a speaker who knew the correct answer responded to their answers by asking, "Is that your final guess?"

In the indirect feedback condition of Study 2 (as well as Study 1 and the two studies reported by Henning and Merriman, 2019), the tester did not repeat the request to predict the correct answer. However, when the tester responded to a child's answer by saying, "Let's see," then presenting the novel word and asking the child to pick its referent, some may have interpreted this reaction as the tester offering them a chance to change their prediction. If so, it may have caused some to question whether their prediction had been correct.

One rather unusual phenomenon is consistent with this suggestion. In studies in which toddlers or preschoolers have received feedback about their disambiguation responses, performance has tended to be very good (Evey & Merriman, 1998; Marazita & Merriman, 2004). In the initial phase of Studies 1 and 2 and the studies by Henning and Merriman (2019), children also received disambiguation trials with feedback. As expected, their performance on these trials was very good. Even in the younger age group, average rate of incorrect response tended to be around .10. In these same four studies (excluding the direct feedback condition of Study 2), a subgroup of 3-year-olds showed a higher error tendency on the disambiguation trial that followed up their first prediction. Among the 3-year-olds whose first prediction was correct, 26% (15 of 57) responded incorrectly to the disambiguation request that followed up their prediction. In contrast, among the 3-year-olds whose first prediction was incorrect, 7% (3 of 39) responded incorrectly to the follow-up disambiguation request (two-tailed Fisher exact probability = .032). It was strange to observe children who predicted, for example, that a conduit cover rather than a cup was "going to be right", but when the tester said, "Let's see. Which one is the *terch*?", responded by selecting the cup! This inconsistent behavior makes sense if these children interpreted the tester's response to their prediction as indicating that their response was wrong and as inviting them to change their answer.

The effect of feedback in the current study was not sustained over the final two prediction trials. As already suggested, the desire to select the familiar object and/or the expectation that the familiar object will become the correct choice may increase over trials. Even when 3-year-olds received direct feedback that clearly indicated that the novel object was the correct choice on both the first and the second prediction trial, this feedback may not have been sufficient to counteract an increasing desire/expectation that favored picking the familiar object as the "right" one.

Two results from Henning and Merriman (2019) replicated. First, the number of 3-year-olds who selected the novel object on the first prediction trial did not differ from chance. Second, when they received only indirect feedback about their predictions, the children began to show a preference to select the familiar object. These two results replicated even though the current study involved a switch to testing children virtually (and recruiting them via social media) and to showing them photographs of objects rather than actual objects. The results replicated even though there was evidence that the 3-year-olds in the current study may have been more cognitively-advanced than the 3-year-olds in Study 1 or Henning and Merriman (2019). Thus, the two results for 3-year-olds appear to be robust.

One result from Henning and Merriman (2019) did not replicate. In their second study, the 3-year-olds who made no errors when judging object nameability were more likely than the other 3-year-olds to respond correctly on the very first disambiguation prediction trial. These groups did not differ in the current study. Although sampling error may account for some of this discrepancy, another contributor may be a methodological difference. The current study used the procedure from Study 1 of Henning and Merriman (2019) in which children performed four standard disambiguation trials before receiving the disambiguation prediction trials. In contrast, in Study 2 of Henning and Merriman (2019), children also observed a doll perform four other standard disambiguation trials before receiving the disambiguation prediction trials. Three-year-olds who have a strong awareness of object nameability may only outperform their agemates in disambiguation prediction if they are shown these additional instances of the disambiguation effect before the prediction phase. (Note that Henning and Merriman (2019) only tested children's object nameability judgment in their second study.)

Chapter 5. General Discussion

The current investigation examined whether certain experiences might cause preschool-age children's disambiguation prediction performance to improve. If such improvements were found, we could conclude that preschool-age children have more capacity to solve purely metacognitive disambiguation problems than had been demonstrated previously. Study 1 examined whether providing 3- and 4-year-olds with a metacognitive representation of the test word would increase the accuracy of their disambiguation predictions. Study 2 examined whether providing 3-year-olds with direct rather than indirect feedback about their predictions would also have this effect.

Both studies found evidence for their hypothesized effects. In Study 1, the testing procedures of Henning and Merriman's (2019, Study 1) were modified so that children received the additional information that the test word was going to be one "you have never heard before." The provision of this metacognitive representation was expected to promote the children's tendency to adopt a similar kind of representation of the disambiguation problem (e.g., as the choice between an unknown kind and a known kind). If so, their tendency to predict that the unknown kind would be correct would increase.

In support of this hypothesis, the overall accuracy of the predictions by the 3-year-olds in Study 1 exceeded that of the 3-year-olds in Henning and Merriman (2019). Moreover, a substantial majority of the former group chose correctly on the very first disambiguation prediction trial, whereas only 50% of the latter group did so. The first prediction trial has special importance because it is the only trial that cannot be affected by the predictions the child has made on previous trials and/or by the feedback they received about those predictions.

The effect documented in Study 1 was limited, however. There was no evidence that providing the 3-year-olds with a metacognitive representation of the test word caused them to favor correct over incorrect predictions after the first trial. Also, there was no evidence that their tendency to make correct predictions after the first trial was different from that of the 3-year-olds in Henning and Merriman (2019, study 1). Finally, there was no evidence that the 4-year-olds in Study 1 performed any differently than the 4-year-olds in Henning and Merriman's study. In both studies nearly every 4-year-old chose correctly on the first prediction trial. So, the only predictions by this age group in Study 1 that could have possibly shown increased accuracy were the predictions made after this trial. Even these predictions had been made with above-chance accuracy by the 4-year-olds in Henning and Merriman's study.

Study 2 provided evidence for the hypothesized benefit of giving 3-year-olds direct rather than indirect feedback about the accuracy of their predictions. Those children who were told immediately after making a prediction whether it was correct were more likely to predict the unfamiliar object as the 'right' choice on the next trial than those children who received indirect feedback. After children in the latter group made a prediction, they were told the novel word, asked to pick its referent, and then told whether this choice had been correct. Their poorer performance may reflect failure to infer from this feedback whether their prediction had been correct and/or a tendency to interpret this feedback as indicating that their prediction had been wrong and inviting them to change their answer.

Development of Disambiguation Prediction

Three-year-olds have greater capacity to solve a purely metacognitive language problem – one that requires representing objects as either known or unknown kinds -- than has been demonstrated in previous studies. Moreover, providing them with a metacognitive representation of the word that will be involved in this language problem (as "one you have never heard before") increases their chances of solving it.

These results lend credibility to the claim that when children as young as 3 years old face everyday disambiguation problems (i.e., opportunities to resolve the ambiguous reference of a novel word), there may be some circumstances in which they consult epistemic representations of these problems. This tendency may increase their chances of solving these problems (i.e., correctly identifying the intended referents). It may even increase their confidence in these solutions as well as their likelihood of retaining them. Retention is especially important. Even if a child decides correctly that the novel word *whisk* refers to a whisk rather than a fork, for example, this decision will have no impact on their vocabulary knowledge unless the child retains some memory that *whisk* is a name for a whisk.

The results of the current studies bode well for "metacognitive" strategies that parents or teachers might use to help young children solve everyday disambiguation problems. When children as young as 3 years attempt to solve such problems, they might benefit from interventions that highlight the epistemic status of words and/or objects. For example, they may benefit if before asking them to find the referent of some novel word (e.g., *whisk*), the requester said something like, "I bet you don't know what a whisk is. I think you can figure it out though. Which one is the whisk?)

The results of Study 2 indicate that it would be best for future studies of disambiguation prediction to use the direct feedback method rather than the indirect feedback method, especially with children who are 3 years or younger. Given the results of Study 1, it may be worthwhile to test the effect of combining the direct feedback method with the practice of providing a metacognitive representation of the test word (i.e., “one you have never heard before.”) Three- and 4-year-olds may show better prediction performance with these procedures than they have in past studies. On the other hand, if these age groups still show a marked decline in prediction accuracy after the first trial, it would be valuable to address possible explanations for this decline. For example, if the decline is due to an expectation that the familiar object will become the correct choice at some point, one might revise the test procedure by interjecting a trial or two in which children are asked to make irrelevant decisions about the pairs of object (e.g., Which one do you like better?), then tell them, “Now we are going to go back to that game we were just playing. That’s the game where I show you two things and ask you to pick the one that I name. The name will be a word that you have never heard before.” If the hypothesis is correct, they should tend to shift to selecting familiar objects during the interjected trials and shift back to selecting unfamiliar objects when the game is resumed.

Many important questions about young children’s disambiguation prediction remain unanswered. One unresolved issue is how awareness of lexical knowledge relates to disambiguation prediction. Although correlations have been found between measures of awareness of lexical knowledge and disambiguation prediction, these correlations have varied from study to study. Further research is needed to determine whether any of these correlations are robust and if so, why. Currently, we have only post hoc explanations. For example, in Study 1, children’s word knowledge judgment, but not their object nameability judgment was associated with the accuracy of their predictions for words they were told “you have never heard before.” If this correlation were to replicate, it would be valuable to test various explanations for it. One interesting possibility is that children who are better at judging whether various words are ones they know are also more likely to realize that a word that they are told they have never heard before has the same epistemic representation as an unfamiliar kind of object – both are unknown kinds of stimuli.

A major developmental question is whether children who are younger than 3 years can make disambiguation predictions with above-chance accuracy in any circumstances. A promising approach may be to first train 2-year-olds to make accurate judgments of word knowledge. As just noted, a measure of the ability to make these kinds of judgments was positively correlated with disambiguation prediction accuracy in Study 1 (albeit in 3-year-olds). Once 2-year-olds have been taught to make word knowledge judgments accurately, their disambiguation prediction should be tested using the combined methods of Studies 1 and 2.

Another promising approach may be to have 2-year-olds observe someone model how to make correct disambiguation predictions, and then ask the children themselves to make disambiguation predictions for a different set of materials. Lipko-Speed, Buchert, and Merriman (2018) found that this kind of modeling was effective in improving preschoolers’ accuracy in a more challenging kind knowledge judgment task

Limitations of the Research

There are many limitations to the research on young children’s disambiguation prediction. Only four studies have been conducted. Each has involved a small convenience sample of monolingual English-speaking predominantly Caucasian children from middle-to-upper class homes in the Great Lakes region of the United States. We know nothing about how variables such as race, ethnicity, native language, parental education, culture, multilingualism, or language disorder affect performance. Only cross-sectional designs have been used to measure developmental changes in performance. Longitudinal designs would not only shed further light on these changes, but also be useful for examining whether measures of awareness of lexical knowledge in very young children predict change over time in their disambiguation prediction performance.

One variable that is likely to impact disambiguation prediction is multilingualism. For example, there is evidence that bilingual infants begin to show the disambiguation effect a few months later than monolinguals (Byers-Heinlein & Werker, 2009; Houston-Price et al., 2010). Regarding preschool and school-age bilinguals, there is mixed evidence regarding the disambiguation effect. Some studies have found that bilinguals show a less robust effect than monolinguals (Bialystok et al., 2010; Davidson et al., 1997, Davidson & Tell, 2005; Diesendruck, 2005), whereas other studies have shown similar performance for the two groups (Byers-Heinlein, Chen, & Xu, 2014; Frank & Poulin-Dubois, 2002; Merriman & Kutlesic, 1993). Therefore, it may be likely that monolinguals' disambiguation prediction performance exceeds that of their bilingual counterparts since bilinguals are more likely to accept a novel word as the referent for a familiar object.

Alternatively, bilinguals tend to have greater metalinguistic awareness than monolinguals (Barac & Bialystok, 2011; Ben-Zeev, 1977; Campbell & Sais, 1995; Cummins, 1978; Galambos & Goldin-Meadow, 1990; Ricciardelli, 1992), which may mean that they develop awareness of lexical knowledge earlier. Having to keep track of what various family members know how to say in the languages they speak could promote a tendency to adopt epistemic representations of the word and/or objects in a disambiguation problem. Thus, bilinguals' disambiguation prediction performance may well exceed that of their monolingual counterparts due to their metalinguistic awareness advantages.

In sum, the current investigation demonstrated that 3-year-olds' disambiguation prediction performance could be improved by providing them with two different experiences. First, their performance improved when they received a metacognitive representation of the test word. Second, their performance improved when they received direct feedback rather than indirect feedback. Thus, when given helpful cues, most 3-year-olds can form metacognitive representations of the elements of the disambiguation problem and use these to draw inferences about the reference of a label.

I have outlined several possible explanations for the two effects that were documented in this dissertation. It will be important for future research to evaluate these explanations.

References

Baldwin, D. A. (1993). Early referential understanding: Infants' ability to recognize referential acts for what they are. *Developmental psychology*, *29*(5), 832.

Barac, R., & Bialystok, E. (2011). Cognitive development of bilingual children. *Language Teaching*, *44*(1), 36-54.

Bedford, R., Gliga, T., Frame, K., Hudry, K., Chandler, S., Johnson, M. H., ... & Basis Team. (2013). Failure to learn from feedback underlies word learning difficulties in toddlers at risk for autism. *Journal of Child Language*, *40*(1), 29-46.

Ben-Zeev, S. (1977). The influence of bilingualism on cognitive strategy and cognitive development. *Child development*, 1009-1018.

Beverly, B., & Estis, J. (2003). Fast mapping deficits during disambiguation in children with specific language impairment. *Journal of Speech Language Pathology and Audiology, 27*, 163–171.

Bialystok, E., Barac, R., Blaye, A., & Poulin-Dubois, D. (2010). Word mapping and executive functioning in young monolingual and bilingual children. *Journal of cognition and development*, *11*(4), 485-508.

Bloom, P. (2000). *How Children Learn the Meanings of Words*. Cambridge, MA: MIT Press.

Bonawitz, E., Shafto, P., Yu, Y., Gonzalez, A., & Bridgers, S. (2020). Children change their answers in response to neutral follow-up questions by a knowledgeable asker. *Cognitive science*, *44*(1), e12811.

Byers-Heinlein, K., Chen, K. H., & Xu, F. (2014). Surmounting the Tower of Babel: Monolingual and bilingual 2-year-olds' understanding of the nature of foreign language words. *Journal of Experimental Child Psychology*, *119*, 87-100.

Byers-Heinlein, K., & Werker, J. F. (2009). Monolingual, bilingual, trilingual: infants' language experience influences the development of a word-learning heuristic. *Developmental science*, *12*(5), 815-823.

Byers-Heinlein, K., & Werker, J. F. (2013). Lexicon structure and the disambiguation of novel words: Evidence from bilingual infants. *Cognition*, *128*(3), 407-416.

Campbell, R., & Sais, E. (1995). Accelerated metalinguistic (phonological) awareness in bilingual children. *British Journal of Developmental Psychology*, *13*(1), 61-68.

Choi, I. R., & Hwang, M. (2014). Korean Late-Talkers' Use of the Mutual Exclusivity Assumption on First versus Second Label Learning. *Communication Sciences & Disorders*, *19*(3), 285-293.

Clark, E. V. (1990). On the pragmatics of contrast. *Journal of Child Language, 17*, 417–431. doi:10.1017/S0305000900013842

Clark, E. V. (1997). Conceptual perspective and lexical choice in acquisition. *Cognition*, *64*, 1-37.

Cummins, J. (1978). Bilingualism and the development of metalinguistic awareness. *Journal of cross-cultural psychology*, *9*(2), 131-149.

Davidson, D., Jergovic, D., Imami, Z., & Theodos, V. (1997). Monolingual and bilingual children's use of the mutual exclusivity constraint. *Journal of child language*, *24*(1), 3-24.

Davidson, D., & Tell, D. (2005). Monolingual and bilingual children's use of mutual exclusivity in the naming of whole objects. *Journal of experimental child psychology*, *92*(1), 25-45.

Deák, G. O., & Maratsos, M. (1998). On having complex representations of things: Preschoolers use multiple words for objects and people. *Developmental Psychology*, *34*, 224‑240.

Diesendruck, G., & Markson, L. (2001). Children's avoidance of lexical overlap: A pragmatic account. *Developmental Psychology*, *37*, 630–641. http://dx.doi.org/10.1037/0012-1649.37.5.630

de Marchena, A., Eigsti, I., Worek, A., Ono, K., & Snedeker, J. (2011). Mutual exclusivity in autism spectrum disorders: Testing the pragmatic hypothesis. *Cognition*, *119*(1), 96–113. https://doi.org/10.1016/j.cognition.2010.12.011

Doherty, M. J., & Perner, J. (2020). Mental files: Developmental integration of dual naming and theory of mind. *Developmental Review*, *56*, 100909.

Dunn, D. M., & Dunn, L. M. (2007). *Peabody picture vocabulary test*: *PPVT-4A*. Minneapolis, MN: NCS Pearson.

Estis, J. M., & Beverly, B. L. (2015). Children with SLI exhibit delays resolving ambiguous reference. *Journal of Child Language*, *42*, 180–195. https://doi.org/10.1017/S0305000914000038

Evey, J. A., & Merriman, W. E. (1998). The prevalence and the weakness of an early name mapping preference. *Journal of Child Language*, *25*(1), 121-147.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist*, *34*(10), 906. https://psycnet.apa.org/doi/10.1037/0003-066X.34.10.906

Frank, I., & Poulin-Dubois, D. (2002). Young monolingual and bilingual children's responses to violation of the mutual exclusivity principle. *International Journal of Bilingualism*, *6*(2), 125-146.

Galambos, S. J., & Goldin-Meadow, S. (1990). The effects of learning two languages on levels of metalinguistic awareness. *Cognition*, *34*(1), 1-56.

Gathercole, V. C. (1989). Contrast: A semantic constraint? *Journal of Child Language, 16*, 685-702. https://doi.org/10.1017/S0305000900010795

Gelman, S. A., & Coley, J. D. (1990). The importance of knowing a dodo is a bird: Categories and inferences in 2-year-old children. *Developmental psychology*, *26*(5), 796.

Golinkoff, R. M., De Villiers, J. G., Hirsh-Pasek, K., Iglesias, A., Wilson, M. S., Morini, G., & Brezack, N. (2017). *User's Manual for the Quick Interactive Language Screener (QUILS): A Measure of Vocabulary, Syntax, and Language Acquisition Skills in Young Children*. Paul H. Brookes Publishing Company.

Gollek, C., & Doherty, M. J. (2016). Metacognitive developments in word learning: Mutual exclusivity and theory of mind. *Journal of experimental child psychology*, *148*, 51-69. https://doi.org/10.1016/j.jecp.2016.03.007

Grassmann, S., & Tomasello, M. (2010). Young children follow pointing over words in interpreting acts of reference. *Developmental Science*, *13*(1), 252-263. https://doi.org/10.1111/j.1467-7687.2009.00871.x

Grice, P. (1975). Logic and conversation. In Cole, P.; Morgan, J. (eds.). *Syntax and semantics*. 3: *Speech acts*. New York: Academic Press. pp. 41–58.

Halberda, J. (2003). The development of a word-learning strategy. *Cognition*, *87*, B23–B34. http://dx.doi.org/10.1016/S0010-0277(02)00186-5

Halberda, J. (2006). Is this a dax which I see before me? Use of the logical arguments disjunctive syllogism supports word-learning in children and adults. *Cognitive Psychology*, *53*, 310–344. http://dx.doi.org/10.1016/j.cogpsych.2006.04.003

Hartin, T. L., Stevenson, C. M., & Merriman, W. E. (2016). Preexposure to objects that contrast in familiarity improves young children's lexical knowledge judgment. *Language Learning & Development, 12,* 311-327. https://doi.org/10.1080/15475441.2015.1077123

Haryu, E. (1991). Continuity and the persistence of objects: When the whole is greater than the sum of the parts. *Japanese Journal of Educational Psychology*, *39*, 11‐20.

Henning, K. J., & Merriman, W. E. (2019). The Disambiguation Prediction Effect. *Journal of Cognition and Development*, *20*(3), 334-353.

Houston-Price, C., Caloghiris, Z., & Raviglione, E. (2010). Language experience shapes the development of the mutual exclusivity bias. *Infancy*, *15*(2), 125-150.

Howie, P., Sheehan, M., Mojarrad, T., & Wrzesinska, M. (2004). 'Undesirable'and 'desirable'shifts in children's responses to repeated questions: Age differences in the effect of providing a rationale for repetition. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *18*(9), 1161-1180.

Jarvis, L. H., Merriman, W. E., Barnett, M., Hanba, J., & Van Haitsma, K. S. (2004). Input that contradicts young children's word-mapping strategy affects their phonological and semantic interpretation of other words. *Journal of Speech, Language, and Hearing Research*, *47*, 392-406. http://dx.doi.org/10.1044/1092-4388(2004/032)

Jaswal, V. K., & Hansen, M. B. (2006). Learning words: Children disregard some pragmatic information that conflicts with mutual exclusivity. *Developmental Science, 9*, 158-165.

Johnston, A. M., Sheskin, M., & Keil, F. C. (2019). Learning the Relevance of Relevance and the Trouble with Truth: Evaluating Explanatory Relevance across Childhood. *Journal of Cognition and Development*, *20*(4), 555-572.

Kalashnikova, M., Mattock, K., & Monaghan, P. (2014). Disambiguation of novel labels and referential facts: A developmental perspective. *First Language*, *34*(2), 125-135.

Karmiloff-Smith, A. (1986). From meta-processes to conscious access: Evidence from children's metalinguistic and repair data. *Cognition*, *23*(2), 95-147.

Krahenbuhl, S., Blades, M., & Eiser, C. (2009). The effect of repeated questioning on children's accuracy and consistency in eyewitness testimony. *Legal and Criminological Psychology*, *14*(2), 263-278.

Lederberg, A. R., & Spencer, P. E. (2009). Word-learning abilities in deaf and hard-of-hearing preschoolers: Effect of lexicon size and language modality. *Journal of Deaf Studies and Deaf Education*, *14*(1), 44-62.

Lewis, M., Cristiano, V., Lake, B. M., Kwan, T., & Frank, M. C. (2020). The role of developmental change and linguistic experience in the mutual exclusivity effect. *Cognition*, *198*, 104191.

Lipko-Speed, A. R., Buchert, S., & Merriman, W. E. (2018). Observing an adult model can cause immediate improvement in preschoolers' knowledge judgments. *Cognitive Development*, *48*, 225-234.

Malone, S. A., Kalashnikova, M., & Davis, E. M. (2016). Is it a name or a fact? Disambiguation of reference via exclusivity and pragmatic reasoning. *Cognitive Science*, *40*, 2095‑2107.

Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meaning of words. *Cognitive Psychology*, *20*, 121–157. http://dx.doi.org/10.1016/0010-0285(88)90017-5

Markman, E. M., Wasow, J. L., & Hansen, M. B. (2003). Use of mutual exclusivity assumption by young word learners. *Cognitive Psychology*, *47*, 241–275. http://dx.doi.org/10.1016/S0010-0285(03)00034-3

Marazita, J. M., & Merriman, W. E. (2004). Young children's judgment of whether they know names for objects: The metalinguistic ability it reflects and the processes it involves. *Journal of Memory and Language, 51*, 458–472. doi: 10.1016/j.jml.2004.06.008

McMurray, B., Horst, J. S., & Samuelson, L. K. (2012). Word learning emerges from the interaction of online referent selection and slow associative learning. *Psychological Review*, *119*, 831–877. http://dx.doi.org/10.1037/a0029872

Medina, T. N., Snedeker, J., Trueswell, J. C., & Gleitman, L. R. (2011). How words can and cannot be learned by observation. *Proceedings of the National Academy of Sciences*, *108*(22), 9014-9019.

Merriman, W. E. (1999). Competition, attention, and young children's lexical processing. In B. MacWhinney (Ed.), *The emergence of language* (pp.331–358). Hillsdale, NJ: Erlbaum.

Merriman, W. E., & Bowman, L. L. (1989). The mutual exclusivity bias in children's word learning. *Monographs of the Society for Research in Child Development, 54*(3–4, Serial No. 220) doi: 10.2307/1166130.

Merriman, W. E., & Kutlesic, V. (1993). Bilingual and monolingual children's use of two lexical acquisition heuristics. *Applied Psycholinguistics*, *14*(2), 229-249.

Merriman, W. E., & Lipko, A. R. (2008). A dual criterion account of the development of linguistic judgment in early childhood. *Journal of Memory and Language, 58,* 1012–1031. doi: 10.1016/j.jml.2007.10.003

Merriman, W. E., Lipko, A. R., & Evey, J. A. (2008). How young children judge whether a word is one they know: A dual criterion account. *Journal of Experimental Child Psychology*, *101*(*2*), 83-98. http://dx.doi.org/10.1016/j.jecp.2008.06.001

Merriman, W. E., & Marazita, J. M. (1995). The effect of hearing similar-sounding words on young 2-year-olds' disambiguation of novel noun reference. *Developmental Psychology*, *31*, 973–984. http://dx.doi.org/10.1037/0012-1649.31.6.973

Merriman, W. E., Marazita, J., & Jarvis, L. H. (1993). Four-Year-Olds′ Disambiguation of Action and Object Word Reference. *Journal of Experimental Child Psychology*, *56*(3), 412-430. https://doi.org/10.1006/jecp.1993.1042

Merriman, W. E., & Schuster, J. M. (1991). Young children's disambiguation of object name reference. *Child Development, 62*, 1288–1301. doi: 10.1111/j.1467-8624.1991.tb01606.x

Mervis, C. B., & Bertrand, J. (1994). Acquisition of the novel name–nameless category (N3C) principle. *Child Development*, *65*, 1646–1662. http://dx.doi.org/10.1111/j.1467-8624.1994.tb00840.x

Mervis, C., & Bertrand, J. (1995a). Early lexical acquisition and the vocabulary spurt: A response to Goldfield & Reznick. *Journal of Child Language, 22*, 461-468. doi:10.1017/S0305000900009880

Mervis, C. B., & Bertrand, J. (1995b). Acquisition of the novel name-nameless category (N3C) principle by young children who have down syndrome. *American Journal on Mental Retardation, 100,* 231-43.

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. *The psychology of learning and motivation*, *26*, 125-141.

Paulus, M., & Fikkert, P. (2014). Conflicting social cues: Fourteen-and 24-month-old infants' reliance on gaze and pointing cues in word learning. *Journal of Cognition and Development*, *15*(1), 43-59.

Perry, J. (2002). *Identity, Personal identity and the self*. Cambridge: Hackett Publishing.

Poole, D. A., & White, L. T. (1991). Effects of question repetition on the eyewitness testimony of children and adults. *Developmental Psychology*, *27*(6), 975.

Preissler, M., & Carey, S. (2005). The role of inferences about referential intent in word learning: Evidence from autism. *Cognition, 97*, B13–B23. https://doi.org/10.1016/j.cognition.2005.01.008

Recanati, F. (2012). *Mental files*. Oxford University Press.

Regier, T. (2005). The emergence of words: Attentional learning in form and meaning. *Cognitive Science*, *29*, 115–135. https://doi.org/10.1207/s15516709cog0000_31

Romski, M. A., Sevcik, R. A., Robinson, B. F., Mervis, C. B., & Bertrand, J. (1996). Mapping the meanings of novel visual symbols by youth with moderate or severe mental retardation. *American Journal on Mental Retardation, 100*, 391–402.

Rosenthal, R. (1991). *Meta-analytic procedures for social research*. Newbury Park: CA: Sage.

Scofield, J., Merriman, W. E., & Wall, J. L. (2018). The effect of a tactile-to-visual shift on young children's tendency to map novel labels onto novel objects. *Journal of experimental child psychology*, *172*, 1-12. https://doi.org/10.1016/j.jecp.2018.03.003

Smith, C. L., & Tager-Flusberg, H. (1982). Metalinguistic awareness and language development. *Journal of Experimental Child Psychology*, *34*, 449–468. http://dx.doi.org/10.1016/0022-0965(82)90071-6

Tomasello, M. (2000). The social–pragmatic theory of word learning. *Pragmatics*, *10*, 401–413. doi:10.1075/prag.10.4.01tom

Vélez, N. & Gweon, H. (2020). Preschoolers use minimal statistical information about social groups to infer the preferences and group membership of individuals. *Cogsci Proceedings 2020*.

Wall, J. L., & Merriman, W. E. (2020). The pragmatics of discovery constrains children's tendency to map novel labels onto novel objects. *First Language*, *40*(3), 251-274.

Wall, J. L., Merriman, W. E., & Scofield, J. (2015). Young children's disambiguation across the senses. *Cognitive Development*, *35*, 163-177.

Waxman, S. R., & Booth, A. E. (2000). Principles that are invoked in the acquisition of words, but not facts. *Cognition*, *77*(2), B33-B43.

Waxman, S., & Hatch, T. (1992). Beyond the basics: Preschool children label objects flexibly at multiple hierarchical levels. *Journal of Child Language*, *19*, 153‑166.

Wilkinson, K., & Green, G. (1998). Implications of fast mapping for vocabulary expansion in individuals with mental retardation. *Augmentative and Alternative Communication, 14:3*, 162-170. doi: 10.1080/07434619812331278336

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*, 103‑128.

Wright, J. C. (1971). The Kansas Reflection-impulsivity Scale for Preschoolers (KRISP). St, Louis: Cemrel.

Yurovsky, D., Smith, L. B., & Yu, C. (2013). Statistical word learning at scale: The baby's view is better. *Developmental science*, *16*(6), 959-966.

Table 1. Study 1 disambiguation prediction task mean number of correct responses (*SD*).

AGE	Initial Selections	Predictions	Final Selections
		Study 1	
3	3.63 (.58)	2.21 (1.22)	3.21 (.83)
4	3.75 (.53)	3.00 (1.18)	3.42 (1.14)
Henning & Merriman (2019, Study 1)			
3	3.54 (.78)	1.54 (1.41)	3.29 (.95)
4	3.67 (.56)	3.04 (1.08)	3.58 (.93)

Note – Maximum score = 4

Table 2. Mean scores (*SD*) on other tests in Study 1.

AGE	Obj Nameability Judgment[a]	Word Knowledge Judgment[a]	PPVT-4[b]
3	.89 (.18)	.72 (.20)	69 (17)
4	.93 (.15)	.91 (.14)	91 (14)

Notes – [a] Proportion of judgments correct   [b] Raw score

Table 3. Intercorrelations in Study 1.

MEASURE	2	3	4	5
1. Disambiguation Prediction:				
Overall	.37**	.27	.43**	.26
1st Trial	.26	.10	.40**	.27
2. Age (months)	------	.15	.49**	.62**
3. Obj Nameability Judgment		------	.44**	.34*
4. Word Knowledge Judgment			------	.47**
5. PPVT-4				------

df = 46 for all correlations except WKJ (df = 45) & PPVT-4 (df=44)

* *two-tailed p < .05*

** *two-tailed p < .01*

Table 4. Percentage correct for each item in the Kent PPVT20 for Study 1 (n = 48).

Item	%
Squirrel	91.7
Cobweb	72.9
Fountain	77.1
Net	89.6
Roof	91.7
Envelope	64.6
Buckle	72.9
Picking	58.3
Dentist	60.4
Claw	81.3
Furry	64.6
Violin	47.9
Group	41.7
Vehicle	66.7
Chef	70.8
Chimney	72.9
Sorting	47.9
Vegetable	47.9
Trunk	60.4
Swamp	33.3

Table 5. Study 2 disambiguation prediction task mean number of correct responses (*SD*).

Condition	Initial Selections	First Prediction	Final 3 Predictions	Final Selections
Direct	3.58 (.65)	.54 (.51)	1.71 (.91)	-----
Indirect	3.63 (.71)	.54 (.51)	1.21 (.78)	3.17 (.87)
H&M	3.54 (.78)	.50 (.51)	1.04 (1.04)	3.29 (.95)

Note – Maximum score: for selections = 4, first prediction = 1, and final predictions = 3.

H&M – data from the 3-year-olds in Study 1 of Henning and Merriman (2019). which used indirect feedback are included for comparison.

Table 6. Mean scores (*SD*) on other tests in Study 2.

Condition	Obj Nameability Judgment[a]	Word Knowledge Judgment[a]	PPVT20[b]
Direct	.91 (.15)	.85 (.19)	13.12 (3.22)
Indirect	.88 (.15)	.87 (.20)	13.50 (3.41)

Notes – [a] Proportion of judgments correct [b] Raw score

Table 7. Intercorrelations in Study 2.

MEASURE	2	3	4	5
1. Disambiguation Prediction:				
Overall	.09	.22	.07	.11
1st Trial	.11	.14	.27	.16
2. Age (months)	------	.13	.10	.34*
3. Obj Nameability Judgment		------	.42**	.14
4. Word Knowledge Judgment			------	-.01
5. Kent PPVT20				------

df = 46 for all correlations except WKJ (df = 45)

$r > .29$, two-tailed $p < .05$ *$r \geq .37$, two-tailed $p < .01$*