Young Children's Representational Understanding:

The Effect of Experience with Live Video

By

Colleen E. Russo

Dissertation

Submitted to the Faculty of the
Graduate School of Vanderbilt University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY
in
Psychology

May, 2017
Nashville, Tennessee

Approved:

Georgene Troseth, Ph.D.

John Rieser, Ph.D.

Megan Saylor, Ph.D.

Leona Schauble, Ph.D.

ACKNOWLEDGEMENTS

Thank you to the many family members, friends, and mentors who inspired me to pursue this career path and area of research. In particular, my grandmothers and parents for always placing the highest value on education, and my young nieces and nephew for providing me with countless hours of video chat experience with toddlers! And thank you to my supportive husband, puppy, and (upcoming) baby girl for being a constant source of love and motivation during this dissertation process.

Thank you to the National Science Foundation for providing funding through a graduate student fellowship, and to Vanderbilt for additional funding support. Thank you to the entire Vanderbilt faculty, support staff, and my dissertation committee for their guidance along the way. Thank you to my Vanderbilt cohort family- Sara Beck and Sarah Wiesen, for being my rocks and best friends the past 5 years. Thank you to Israel Flores for being the best research partner on this project and running lab the past year. And thank you to the dozens of brilliant research assistants in the Early Development Lab who have invested their time and energy into my many research projects over the years.

Finally, thank you to my incredible advisor, Georgene Troseth, for always believing in me and giving me the freedom to explore my passions. Georgene, you have celebrated my successes with the glee of a proud parent, and supported me through my failures without ever making me feel bad for trying something new. You never once doubted that I could finish my dissertation while getting married, moving to Canada, and having a baby all in the same year. You are a true feminist and inspiration. Thank you.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION

Infants are first exposed to many objects not directly, but through representations. For example, most infants will see a picture of a zebra in an "ABC" book before encountering one in real life. To effectively gather information, humans of all ages constantly rely on their understanding of representational systems, including: language, numerical expressions, maps, calendars, and pictures (both moving and still). Even before children learn to speak, they rely on gestures to communicate (Acredolo & Goldwyn, 1985). Similarly, because iconic images (those that closely resemble their referent) convey meaning without spoken language, they are unique in their universal ability to communicate for all ages (Gombrich, 1974).

Importantly, children are able to learn from certain representations from a young age. For example, infants can successfully transfer learning from picture book photos to their corresponding real life referents by 15 months of age (Ganea, Pickard, & DeLoache, 2008)—they can learn that the label "lion" is associated with pictures of lions, which are representative of actual lions. This beginning understanding of the representational nature of pictures likely begins from such experiences with parents labeling and talking about pictures while reading picture books. Fully understanding and using pictures and other representations, however, requires *representational insight*—children must detect, understand, and finally mentally represent the relationship between the representation and its referent (DeLoache, 1991). Achieving representational insight has proven to be a challenging task for young children, particularly in novel situations.

Research does clearly demonstrate that children's ability to use representations emerges rapidly in the first few years of life (Rakoczy, Tomasello, & Striano, 2005). Observing a representation being used by someone else (e.g. waving or watching a sibling draw) is the first step of understanding its meaning, and is followed by children's imitation of modeled representational actions (e.g. infants learning baby sign language and gestures). The onset of verbal language not only allows infants to better understand others' representational actions, but also helps young children ascribe meaning to their own representations and communicate them to others. For example, children can now explain how their seemingly random scribble is actually a picture of their dog. However, despite decades of research on children's representational development, questions pertaining to *how* children develop representational reasoning still persist (DeLoache, 2011).

The current series of experiments and accompanying literature review aim to address that question by exploring how naturalistic experiences with newer forms of representational media (digital photos and live video) may affect the development of children's representational understanding of these media. These experiments provide important information into children's ability to *learn* from such media, but more importantly, they aim to shed light on the cognitive processes behind the development of children's representational understanding and use-- specifically, the degree of flexibility in young children's representational insight, and the specific experiences that impact it.

One milestone in developing a mature understanding of pictures and video is what DeLoache (1987) called *dual representation.* As the renowned psychologist James Gibson stated, ‘‘A picture is both a surface in its own right and a display of information about something else.” (Gibson, 1979, as cited in Troseth, 2010, p. 156). Dual representation requires thinking about a picture or video as an object or medium with its own particular qualities (e.g. a small photo of a zebra in a picture book) as well as thinking about what it represents (an actual, real life zebra). When children show signs of dual representation for photos and videos, it suggests that they understand (in that situation) that their mental representation of the portrayed physical referent as an object is separate from the "picture/video of" the referent (Troseth, Pierroutsakos, & DeLoache, 2004). In fact, many of the portrayed object's physical attributes do not apply to the "picture/video" of the referent. For example, children eventually understand that an actual, physical glass of milk would spill if turned upside down, but a photo or video of a glass of milk would not (Flavell, Flavell, Green, & Korfmacher, 1990).

Judy DeLoache (1991) argues that the ability to use and represent symbols with ease is a major benchmark in children's cognitive development. Although the current experiments follow DeLoache's research paradigm, I depart from her use of the term "symbol", opting to use "representation" instead. DeLoache defines a symbol as "something that someone intends to stand for or represent something other than itself" (DeLoache, 2002, p. 75). She highlights four important deconstructed elements of her definition. (1) *Someone* refers to the fact that we, as a symbolic species, must learn to use symbol systems; (2) there is inherent *intention* built into symbol-referent relations; (3) *something* speaks to the fact that nearly anything can serve as a symbol; and (4) the symbol *represents or stands for* something *other than itself.* In the following section, I use the viewpoints from philosophers and psychologists, dictionary definitions of various terms, and the reality of our current technology, to argue against DeLoache's definition and support my reasoning for using the term "representation" in regards to digital photos and live videos.

Philosophy Background and Definitions

"Representation" and "symbol" have been used interchangeably; however, there are important differences both in their dictionary definitions and how they have been previously used by philosophers and theorists. In fact, according to Deacon (1998), "more philosophic ink has been spilt over attempts to explain the basis for symbolic reference than over any other problem" (p. 43). Merriam-Webster's dictionary defines representation as, "one that represents, such as an artistic likeness or image" (Representation, n.d.). This definition is admittedly vague, but it without question applies to photos and videos (including live video). Meanwhile, although the Merriam-Webster dictionary definition of symbol is more explicit, it includes examples that most certainly do not apply to photos and videos--"Something that stands for or suggests something else by reason or relationship, association, convention, or accidental resemblance, especially a visible sign of something invisible (e.g. the lion is a symbol of courage)" (Symbol, n.d.). Specifically, the mention of representing invisible qualities or traits does not apply to photos and video, and is a reminder of the various ways in which the term symbol is used and defined.

Philosophical support for DeLoache's definition stems from Potter's (1979) definition, "symbolization is the representing of an object or event by something other than itself" (p. 41), as well as Goodman's (1976) assertion that "almost anything may stand for almost anything else" (p.5). But when it comes to representational media such as photos and videos, and particularly newer technology such as live video chat, I would argue that there are clear flaws in using the term "symbol" based on DeLoache's (2002) definition. DeLoache's first and third assertions do work for photos and video, that we, as symbolic species, must learn to use symbols, and almost anything can serve as a symbol. Her second and fourth deconstructed points, however, are more difficult to apply.

DeLoache's second point is that there is inherent intention built into symbol-referent relations. Tomasello (1999) echoes the importance DeLoache places on intentionality in assigning symbolic meaning, due to the fact that symbol usage is inherently a social process. Children begin to understand pictorial intent between two and three years of age. In a study where children of this age were presented with ambiguous ink-blots, children were more likely to ascribe meaning to a blot and interpret it as a picture when told that someone worked hard to create it, as opposed to it simply being spilled paint (Gelman & Ebeling, 1998). Similarly, children rely on and trust their own symbolic intent. When three- and four-year-old children were asked to draw a picture of two perceptually similar objects (e.g. a balloon and a lollipop), they maintained their representational intent when later labeling the drawings, despite the two drawings appearing nearly identical (Bloom & Markson, 1998). Bloom (1996) further stresses that intention of the symbol creator is crucial in defining something (including a photo) as a symbol.

I would argue that most photos and videos undoubtedly have intention. Creators make clear choices both in what they shoot in the moment (e.g. lighting, framing, juxtaposition) and how they later edit their work (e.g. for photos and videos—after effects, cropping, printing, pacing, background music). Indeed, Goodman (1976) makes the point that photographers and videographers naturally edit their representations simply by choosing what to include and focus on in frame. Thus, Goodman and others argue that photos and videos are not mere "copies" of their referents but, like all representations, reflect their creator's intentions (see Gombrich 1974; Rakoczy, Tomasello & Striano, 2005; Werner & Kaplan, 1984).

Although intent is often present in picture and video representations, unedited photos and videos (live video chat in particular) leave undeniable ambiguity as to whether or not the photographer or videographer is always imposing intent for the medium be a representation of what is being depicted. Consider the "technology-impaired" grandparent who has little grasp on what their laptop or phone's built-in camera is even capturing. Indeed, many can relate to phrases such as, "Grandma, I just see your laptop keys! Tilt your screen up!" or "Grandpa, you're using your phone's back-facing camera. You need to switch to the front facing camera so I actually see you!" Surely Grandma isn't *intending* for you to see the laptop keys instead of her smiling face, and Grandpa isn't *intending* to give you a view of his wall instead of himself, yet nonetheless I would argue that the 2D display of light coming through on your end is still a representation of Grandma's actual laptop keys and Grandpa's actual wall.

I believe that the *resemblance* between the representation and referent is inherent in photos and videos, and more important to being a representation than intention is.

Goodman (1976) and Peirce (1955) argue the opposite when determining whether or not something is a symbol, stating that resemblance between symbol and referent is not necessary. For example, Goodman would argue that a smudge that happens to look exactly like a flower would not be considered symbolic of a flower, whereas a child's scribble that they proudly claim is a flower would (Rakoczy et al., 2005). Thus, the importance placed on intention coupled with the fact that I'd argue intention is not necessarily inherent in live video puts the intention requirements of DeLoache's symbol definition at odds with the media discussed in the current experiments.

DeLoache's fourth point is that a symbol stands for or represents something *other than itself*. Although I believe that any photo or video is indeed a representation of what it is depicting, I do not believe that it is always the intention of the representation creator *or* the user that they are creating/viewing a representation that stands for something else. Rather, the creator's intent might instead be to create a method to see the referent "directly" (i.e. ignoring or "seeing through" the medium-- Troseth & DeLoache, 1998). Indeed, Elisabeth McClure, who conducted her dissertation work on young children and skype, noted that watching toddlers and grandparents interact through video chat during a naturalistic study led her to believe that both parties were, for the most part "seeing through" the screen (personal communication, January 19, 2017). If this is how children and even adults are interpreting video chat (i.e. that they're "seeing through" it and not actively distinguishing the two at all times), then it would not always be considered a symbol according to DeLoache's requirements.

Whether or not something is *interpreted* as a representation depends on many factors, such as experience, context, cognitive development, and more. For example, an infant who lacks experience with representational media might act on a video image as if it was the graspable object being depicted (Pierroutsakos & Troseth, 2003). I would argue that even though the infant is not distinguishing the representation from the referent in the way that a child or adult would, it does not mean the moving images on the video screen is *not* a representation when used by the infant but *is* when used by the child or adult. Instead, it is always a representation, and what changes is simply whether or not the user interprets and uses it as one. Even within the same video chat call, an adult might oscillate between "seeing through" the medium, and then suddenly remembering they are communicating through a representational medium when the screen buffers or temporarily loses visual quality, as is common with video chat.

DeLoache further stresses her fourth point by highlighting the importance that the symbol and referent are clearly different from each other, and that if the two entities cannot be distinguished, the representational medium is not considered to be a symbol. Specifically, DeLoache (2002) notes, "a necessary element of symbolization is being able to distinguish between symbol and referent. Indeed, if for some reason one cannot tell two entities apart, one cannot be a symbol for the other." (p. 76). This was the case with Troseth and DeLoache's (1998) experiment when they placed a television screen behind a window, and told children they were looking through a window (in contrast to a condition where children were clearly watching a live video feed on a television or actually looking through a window). The researchers concluded that the majority of the children in the "fake window" condition believed they were watching through a window and not a television screen, because their task performance (in an object retrieval search task) was more similar to those who watched through the window than those who watched the live television feed. Those children likely did not distinguish the two entities, i.e. the representational medium (the moving images of the hiding event on the television screen) from the referent (the actual hiding event). Deacon (1998) stresses that the interpretation of a symbol is paramount in determining whether or not it is a symbol. In this case, according to DeLoache's definition, the moving images on the television screen would be considered a symbol only in the case when the television screen was *not* behind the window, when children clearly distinguished the medium from the referent. However, I would argue that the 2D display of light creating the moving images on the television screen is still a representation of the actual room when it is behind the window, even if the child did not distinguish the two.

Icon is another term that has been used to refer to photos and videos. In fact, Troseth (2003b) has referred to photos and videos as "iconic symbols", denoting a subcategory of symbols specific to these media (though the restrictions of DeLoache's overarching symbol definition would still apply). Charles S. Peirce (1955), actually argues for calling pictures *icons* rather than *symbols,* suggesting that it is language that is better categorized as symbolic. Bruner, Oliver, and Greenfield (1966) agree with Peirce, stating that pictures and images are iconic and language is symbolic. However, Peirce (1955) notes that icons simply resemble their referents and have no further connection to them. According to Merriam-Webster's first definition, an icon is indeed "a usually pictorial representation" (Icon, n.d.); therefore, pictures fit this definition. However, there are other common definitions of the word icon that easily come to mind, which make me less eager to adopt it, such as its strong tie to religious figures (the second definition), and even the common usage of icons within a computer screen (the fifth definition).

Peirce (1955) and Bruner et al. (1966) are not alone in classifying symbols as referring to language; indeed, DeLoache (2011) acknowledges that the term symbol has been most often used by scholars to refer to language. Similarly, Werner and Kaplan (1984) explain that the *linguistic word* “tree” is more of a symbol for the concept of tree than an image of a particular tree, which they would categorize as a representation. They believe that symbols are more abstract (i.e. language) and representations are more concrete (a particular image). Thus, Werner and Kaplan would likely argue that the image of Grandma on video chat is a representation rather than a symbol.

Because I believe that term “symbol” is limiting when applying it to photos and video (in particular, live video), moving forward, I will refer to photos and videos simply as *representations*, rather than symbols. In the case of photos and videos, I define representation as the following: *Any medium that portrays an image (or series of moving images), depicting and resembling a referent, regardless of intention by either the depiction creator or user.* Although I believe that all photo and video representations communicate information, this communication does not need to be dually recognized by both creator and user (i.e. it can be a one-way, self-communication with just the creator or just the user). Finally, I believe that representations are always representations. Creator and user intention and interpretation may vary at times, but the particular photo or video’s classification as a representation does not.

DeLoache’s line of research (which the current experiments stem from) and her contributions to understanding young children’s symbolic (representational) development, including ideas/terms such as dual representation and representational insight, are still applicable with the term “representation”. The term representation, however, is admittedly extremely broad, consisting of multiple uses even within describing children's understanding of representations. For instance, in defining dual representation (discussed in detail below), DeLoache (1987) describes how children must create *mental representations* of the *representational mediums/objects.* To further complicate matters, they are simultaneously holding a mental representation of the information/event presented through the medium (just as they would have to do if viewing the event directly), hence "dual representation" (DeLoache, 1987). Therefore, to avoid confusion, I will clarify the type of representation whenever necessary.

Scale Models as Representations

DeLoache (1987) and her colleagues developed a program of research to investigate children's understanding of representations using a unique object retrieval/search task paradigm. They began by creating a scale model to serve as a representation of a larger laboratory room, which was set up as a living room complete with a sofa, chair, rug, and end table. The scale model room (approximately 2 feet wide) had the exact same furniture as the lab room in proportionately scaled down dimensions, and was housed in an adjoining control room.

All of DeLoache's search task experiments (e.g., 1987; 1991) follow a similar procedure. First the children are introduced to the contents of the model and the larger room, with the experimenter explicitly pointing out the connections between the two—for example, highlighting that the chair in the little room was "just like" the chair in the big room. The child is also introduced to "big Snoopy" in the larger room, and "little Snoopy" in the scale model. Next, the child completes a placement trial to further highlight the relationship between the scale model and larger room; the experimenter places little Snoopy in a location in the scale model, and simply asks the child to place big Snoopy in the same location in the larger room. Finally, the test trials begin. The experimenter hides the little Snoopy in a location in the scale model (e.g. behind the pillow) and tells the child that her assistant was going to hide big Snoopy in "the same place" in the bigger room. The child is then asked to retrieve big Snoopy from the larger room next door. Success on the search task (errorless retrievals) suggests that the child understands the relation between: the two spaces (scale model and large room), toys (big and little Snoopy) and events (hiding of the two toys in their respective places), and that the child can effectively *use* the representational medium (the scale model).

Results from several studies repeatedly demonstrate that 2½-year-old children fail at using the representational information from the scale model to find Snoopy in the larger room, whereas 3-year-olds are successful in doing so (DeLoache, 2011). With success rates of 15% and 77% respectively, representational development appears rapid over this six-month age period (DeLoache, 1987). Notably, this difference is not due to memory problems in the 2½-year-old children, after failing to search correctly in the larger room, they *were* able to accurately recall where in the model room the representational, scaled-down object was hidden, yet could not make the connection to where the analogous larger toy was hidden in the room—a task that requires representational insight (DeLoache, 1991). An inability to simply draw correspondences between the scale model and larger room referent was also ruled out as a potential crux. Although the 2½-year-olds were able to understand that the smaller chair relates to the larger chair; this lower level processing was not enough for children to succeed in extracting and applying information about the location of a hidden object from the correspondence, which also relied on representational insight (Troseth, Bloom Pickard, & DeLoache, 2007).

Dual Representation

Achieving representational insight, however, requires *dual representation* (DeLoache, 2000). Representational objects possess an inherent "dual reality" due to being both concrete (the actual object) and abstract (representing something other than itself--DeLoache, 1991; Troseth, 2003a). Dual representation, therefore, refers to the ability to understand that a symbol (e.g. the scale model) is both an object itself as well as a representation of another entity (the larger room), and to mentally represent the object in both ways. When children make the representational connection, they form a higher-level cognitive representation of the relationship between the representation and its referent (Troseth et al., 2007).

One reason children have difficulty achieving dual representation when using scale models as representations is due to the fact that a scale model is a three-dimensional object itself with multiple uses. For example, a child might more readily associate the model room with a play dollhouse before interpreting it as a potential symbol. To demonstrate this, DeLoache and her colleagues manipulated the salience of the scale model as an object, (i.e. the degree to which the scale model was treated as an object), thus making it easier or harder to achieve dual representation. First, the researchers decreased the scale model's salience as an object by placing it behind a glass window and not allowing the child to interact with it, reasoning that young children would be less likely to think of the model as an object to be played with. Indeed, 2½-year-old children were more successful at the search task in this condition (DeLoache, 1993). Conversely, when 3-year-old children (who typically succeed in the search task) played with the scale model *first* (thus further increasing the salience of the model as an object), their symbol-based search performance decreased (DeLoache, 1993). Thus, to successfully extract relevant information from a representational object, children must achieve a certain amount of psychological distance from it. In the case of the scale model, focusing on the representation as an object (a dollhouse) appears to get in the way of their achieving dual representation (constructing the "stands for" relation between the model and the larger room--DeLoache, 1993).

In a clever experimental design, DeLoache and her colleagues were able to strip away the need for any dual representation. They told the children that they had a shrinking machine that could make a room big or little. Through a creative set up complete with sound effects, the researcher "demonstrated" the machine to the child. The smaller toy was hidden in the scale model, and the child and experimenter left the room after turning on the special machine. When they returned, the scale model was gone and the full-size room was in its place. In this condition, when the child believed that the scale model and the larger room were the *same room*, the 2½-year-olds had no problem finding the hidden object in the larger room based on where it was hidden "when the room was small" (DeLoache, Miller, & Rosengren, 1997). Rather than create a representational relation between the two rooms, children were able to make a simpler identity relation—they represented it as one room that simply got bigger or smaller. In contrast, when making a representational relation, children had to think about the model as both an object *and* a representation at the same time—they had to both represent the model as a "dollhouse" and represent that it stood for the big room. Thus, this study supports the notion that achieving dual representation is a likely source of failure for 2½-year-olds use of scale models in object retrieval tasks.

Photographs as Representations

Using photos as the representational medium, on the other hand, makes it easier to achieve dual representation. For young children, photos often serve as representations in the first place (e.g., in picture books, on clothes), rather than as objects themselves (DeLoache, 1991). Being two-dimensional, a photograph of the room is less salient (i.e. noticeable) as an object compared to the scale model. Indeed, when an experimenter simply pointed to a photo of where the object was hidden instead of using the scale model, 2½-year-olds *were* successful (DeLoache, 1987; 1991). This pattern of results held true not only for photos of individual hiding spots, but also for wide-angle views of the entire room, as well as less iconic line drawings of the room and hiding locations (DeLoache, 1993). On the other hand, when photographs were treated as objects to increase their salience as an object *rather* than as a representation, (i.e. the experimenter hid the little Snoopy *behind* the actual printed photograph to show where big Snoopy was hiding in the room), 2½-year-old children (who typically succeed) were unsuccessful in the search task (DeLoache, 1991).

Generalizing representational meaning beyond the 2D photo is crucial to successful use of the pictorial representation. For example, when a very young child sees a picture of a dog and hears the label "dog", are they simply making an association between the word and the *picture* of a dog? Or are they making a deeper referential mapping, understanding that the picture is a representation of an actual, real dog? Preissler and Carey (2004) shed light on this question by teaching very young children (18- and 24-month-olds) the word "whisk" using a line drawing of a whisk. Afterwards, they presented children with the line drawing and an actual whisk and asked the child to identify the whisk. Both age groups were more likely to choose the actual object whisk than the drawing from which they had learned the word. Importantly, a control condition assured that children were not biased to always choose any real object over a 2D photo. Thus, these results suggest a deeper learning that generalized beyond the 2D picture.

Although 2-year-old (24 months) children display some representational understanding in Preissler and Carey's (2004) object label identification task, they struggle using information from pictures to retrieve a hidden object (a task that 2½-year-olds are able to do—DeLoache & Burns, 1994). DeLoache and Burns ran a series of follow-up experiments to try to help children grasp the representational relation between the photo and hiding locations, but no experimental manipulations succeeded in doing so. For example, children's search performance did not improve when they were shown a photo of just one hiding place at a time compared to a photograph depicting all the hiding places (DeLoache & Burns, 1994). However, children were successful when told verbally where to search, suggesting that children's failure is not solely due to perseveration errors (returning to the previous hiding location) or lack of motivation to search.

Children may have an easier time using the pictorial representations in Preissler and Carey's (2004) object labeling task because it is a familiar use of a representation. For example, it is very similar to a parent labeling pictures in books and talking about the referent items in the real world (Troseth, 2010). The search task, on the other hand, requires toddlers to use photos in a novel way--to provide information about where to search for an object in real time. When children were simply asked to *place* the toy in a particular location (i.e. a future action) rather than search for a pre-hidden toy, they were successful, suggesting a struggle in using the photo to guide search behavior for something that has already occurred (DeLoache and Burns, 1994). Additionally, children's possible preconceptions of photos might work against them when trying to use photos during the search task—this will be discussed in detail later.

Videos as Representations

Moving pictures (i.e., video) are another representational medium that contains information for toddlers, with countless television programs and videos targeting this young audience. To investigate children's understanding and use of information on video, the same object retrieval search task has been conducted numerous times using a live, closed circuit television set-up (Troseth, 2010). In these tasks, children must take information they view on the representational medium (the televised room, person, and toy) and apply it to its referent (the actual room, person, and toy). After the same general orientation to the searching room and its contents discussed before, children in the video condition receive a brief orientation to live video, where a researcher provides explicit connections between what the camera is recording and what is simultaneously displayed on the screen. For example, the child waves to the camera and sees their wave show up on the television screen. The experimenter also explicitly points out the correspondences between the hiding locations on the screen and in the hiding room. Next, with the child sitting in the adjoining control room, the experimenter tells the child that she is going to hide Snoopy in his room, and to watch her on TV to see where she hides him. Unlike photographs, which merely show where the toy was hidden, video allows the child to actually *watch* the toy be hidden on-screen.

Similar to the results with photos, 2½-year-olds were successful in using the information on the video screen to subsequently retrieve the hidden Snoopy, but 2-year-old children (24 months) were not (Troseth & DeLoache, 1998). This failure is not due to an inability to complete the object retrieval task, as 2-year-olds performed perfectly when they watched the hiding event occur directly through a TV-screen-sized window instead of a TV monitor, even though the two conditions showed roughly the same view (Troseth & DeLoache, 1998). The same procedure and pattern of results was replicated by Deocampo and Hudson (2005).

Dual representation provides an explanation for this difference in performance between watching live through a window and watching on video. When watching on video, children must mentally represent both what they see in the 2-dimensional image on the screen, and the actual hiding event occurring in the adjoining room. When simply watching through the window, however, the need for dual representation vanishes since children are watching the hiding occur directly.

Schmitt and Anderson (2002) also replicated this video search task procedure and found that although 2½-year-olds were twice as successful as 2-year-olds, they were still only achieving a chance success rate. However, both 2-and 2½-year-olds were successful when watching through a window. Three-year-old children were successful when watching both on video and through the window. The authors hypothesized that the drastic differences between those who viewed the hiding event on video compared to through the window stemmed from a perceptual issue, meaning it was too difficult to go from a 2D televised representation to a 3D referent. Specifically, they state, “television provides a limited and poorly resolved spectrum of the cues that ordinarily allow perception in depth, especially motion parallax, texture and shadow gradients, and stereopsis” (p. 70). The go on to posit that the cognitive burden of these perceptual differences hinders task performance.

This hypothesized account has been refuted in multiple experiments. Troseth and DeLoache (1998), for example, placed the television behind a window to make 2-year-old children think they were watching the hiding event live through a window. Thus, if children believed the manipulation, they would have no need for dual representation, even though they were still viewing the event on a 2D screen. On the other hand, if the problem did stem from a perceptual 2D to 3D transfer issue, children would perform just as badly in the “fake window” condition compared to the actual video condition, since they were still watching a 2D image. However, children in the “fake window” condition performed significantly better than those who knew they were watching a television set, showing that, at least for this task, a perceptual issue was not the crux of children’s difficulty with video.

In another experiment, the referent of the video image was changed from a room with 3D hiding places to a 2D felt board with four different flat felt pieces (e.g. a gift box, a birthday cake) as hiding places for a sticker. The idea was that the 2D symbolic representation (a video of the sticker being hidden somewhere on the felt board) would be a better perceptual match to its referent (the real, flat felt board with flat hiding places and flat hidden object); however, 2-year-olds were still unsuccessful (Schmidt, Crawley-Davis, & Anderson, 2007), indicating that the main problem was not perceptual dissimilarity.

Finally, two different groups of experimenters removed the need to spatially map between a 2D representation (on screen) and a 3D room. Instead of watching the experimenter hide the object, the child watched the experimenter (either in person or standing against a blank wall on video) say where she hid the object; thus, because the information to be transferred from video to reality was presented verbally, there was no need for visual 2D to 3D mapping between the video image and room. Two-year-olds in both experiments were successful when receiving this information verbally in person, but were still unsuccessful when receiving the information via video (Schmidt, Crawley-Davis, & Anderson, 2007; Troseth, Saylor, & Archer, 2006), once again suggesting that children's failure to use information presented on video more likely stems from a problem with dual representation rather than perceptual issues as proposed by Schmitt and Anderson (2002).

Multiple follow-up studies have been conducted to explore why 2-year-old children struggle using information from video to find a hidden object. For example, even when two toys were hidden in each spot and children first watched an assistant successfully retrieve one of them, children did not succeed in retrieving the second toy from the same location (Troseth, 2003a). Providing a goal incentive did not increase retrieval rates either (Deocampo & Hudson, 2005). Finally, when 2-year-olds were given the opportunity to watch the hiding not only on live video, but also simultaneously through an open door, they were successful. However, immediately following that success, when they viewed the event solely on the live video, 2-year-old children returned to a below chance success rate (Troseth, 2003a). All of this suggests that there is something specific about *video* as a representational medium that hinders 2-year-old children from applying information presented on screen to the real-life search task situation.

Toddlers' Learning from *Passive* Video

When looking at infant and young children's interactions with and learning from television (or *passive video*), interesting age effects emerge. Although 2-year-old children are capable of learning from television, they almost always learn better from a real-life demonstration (Hayne, Herbert, & Simcock, 2003). Indeed, decades of research on young children and television support the notion that live instruction leads to superior learning outcomes for young children than video (Kirkorian, Wartella & Anderson, 2008). However, infants tend to learn equally well from video and live presentations in both imitation (Barr, Muentener, & Garcia, 2007) and word learning tasks (Krcmar, Grela, & Lin, 2007). Between the ages of two and three, children tend to struggle most in obtaining information from pre-taped video compared to real life, a common phenomenon termed the "video deficit" (Anderson & Pempek, 2005). The type and difficulty level of the measured task determine the age at which children typically start learning equally well from video, but generally by their 3rd birthday children "overcome" the video deficit and are competent at learning from video (Anderson & Pempek, 2005).

A child's preconception of the medium proves to be crucial. In an in-lab experiment comparing imitation of a person on video to a person present in the room, 24-

month-olds in both conditions were able to successfully imitate behaviors (Strouse & Troseth, 2008). However, when 24-month-olds viewed the video on their *home* television sets (compared to the in-lab television), they imitated significantly fewer behaviors—importantly, in-person imitation was equal across both home and lab settings (Strouse & Troseth, 2008).

Children's failure to learn from their home television set might be caused by their conception of television's *purpose* and *use*. For example, the children in this study may have concluded that what they view on their home television is a source of entertainment, *not* a representation of reality or something to be learned from. Even though toddlers often have extensive experience watching television, they are not used to receiving information about real situations from video. In fact, preschool television shows are full of content that actually highlights the distance from reality: talking animals and puppets, various animated worlds, and cartoons that defy laws of gravity (e.g. Bugs Bunny slowing down on a free fall just before landing); even young infants know unsupported objects fall (Baillargeon, Needham, & DeVos, 1992). Moreover, when characters "talk" to the at home viewers (e.g. *Mr. Rogers, Blue's Clues, Dora*), they do not provide contingent responses to the children viewing the program. In other words, these characters are "quasi-interactive", and children eventually realizes that even if they do not respond, the character will continue as if they had (Hollenbeck and Slaby, 1979).

Toddlers' Learning from *Live* Video

Because children's television typically depicts clearly *fictional* events occurring in various temporal contexts, the ability to apply video information to present reality (e.g. in the object retrieval search tasks) is particularly difficult for children as it is a relatively novel use of video. In other words, the video search task "requires children to respond *flexibly,* to a *familiar* medium" (Troseth, Pierroutsakos, & DeLoache, 2004, p. 23). To succeed in the search task, children are suddenly required to respond differently than they are accustomed to (i.e. flexibly) to this content appearing on a television screen (i.e. the familiar medium). If 24-month-old children lack sufficient general experience interacting with representational media in such a way, they might benefit from experiences that connect the symbolic medium to their current reality, either specific training or naturalistic exposure (Troseth, 2010).

According to a regression analysis of data from questionnaires completed by parents whose children participated in object retrieval tasks, exposure to live video (defined as children seeing themselves live on camcorder screens flipped to face the child and/or seeing themselves on store security monitors) best predicted successful performance in the search task when the information was presented via video (Troseth, Casey, Lawver, Walker & Cole, 2007). These results persisted even when several control variables were taken into account (e.g. child's vocabulary, birth order, primary caregiver education, and parent occupation), suggesting that exposure to *live video* may enhance children's representational insight into the link between video and reality.

To further investigate the effect of experience viewing oneself on live video, Troseth (2003b) provided 24-month-old children with 2 weeks of experience seeing themselves on television in their home. This was done by having parents hook up a camcorder to their television so children were seeing themselves on the television screen in real time as they played with an assortment of toys. This group of 2-year-olds was subsequently successful in using a live video presentation to find the hidden object in the search task (77% success rate) compared to the control group that spent the same amount of time playing with the toys but without seeing themselves on the television screen (23% success rate). In-lab experience with live video has also been shown to aid 2- and 3-year-old children's self-recognition on live video (Suddendorf, 1999; Demir & Skouteris, 2010).

The *naturalistic* experience with live video that children received back in the 1990's and early 2000's when the research by Troseth and her colleagues was completed was very different from children's current experience. For example, Troseth et al. (2007) defined experience with live video as children seeing themselves on store security monitors and on camcorder screens flipped to face the child. At the time, those were the primary opportunities for children to see themselves on live video. Similarly, experience seeing themselves on recorded video was restricted to the rare occurrences when (if, ever) parents would haul out home videos. Digital cameras that allowed for instant photo viewing and video playback on the device itself became popular during the 2000's. Today, young children are inundated with video images of themselves, both live and recorded. This is highlighted by the fact that almost all present-day cell phones have instant picture/video viewing capabilities. In fact, the notion of having to "wait" for film to be developed to see photos is a foreign concept to most young children. Anyone who spends time with young children can attest to their desire to both independently take photos as well as instantly see the photo or video just taken. Indeed, in my personal experience, videos of young children often end with the child saying, “can I see it?” Furthermore, built-in webcams and front facing cameras (on devices such as computers and smartphones) provide the unique opportunity for children to actually see themselves on “video” *in real time* as they are taking a “selfie” photo or video.

Contingency in Live Video

We live in the age of live video chatting, during which children see not only themselves in real time on live video (often in a smaller picture-in-picture frame on screen), but another person as well (in the main screen). This person, unlike the characters that typically appear on television, actually responds contingently to the child; the information is accurate and relevant to the child and the partner responds immediately. Research demonstrates that infants as young as 4 months prefer responsive interactions both in real life (Bigelow & Birch, 1999) and on video (Hains & Muir, 1996). Thus, video chatting provides a unique and potentially beneficial combination of screen media and social interaction.

There are multiple reasons why socially contingent interactions are beneficial to children’s engagement and learning. Strouse, O’Doherty, and Troseth (2015) outline four: 1) the expected contingency from a teacher (e.g. eye contact, taking turns in an exchange); 2) the appropriate feedback based on the child’s actions; 3) the established relevance to the child’s life; and 4) the formed relationship between the two parties. I would add a fifth reason to this list: temporal relevance. The immediate, real-time responses the child receives signal that the given information also has relevance to their *current* reality. This is particularly important for the object retrieval tasks, which require children to apply information presented on video to their current reality, as opposed to other tasks featuring more enduring aspects of the world such as learning object labels.

Before the prevalence of video chat, Troseth, Saylor, and Archer (2006) created a similar set-up using closed-circuit television. In baseline conditions without any prior contingent video chat, children who were verbally told where a toy was hidden on video performed significantly worse in the search task than those who heard the verbal information in person. However, children who experienced a mere five minutes of socially contingent live video interaction with the experimenter, and then were verbally told the hiding location on video, performed equally well in the search task as those who received the verbal information in person. The experimenter in the contingent video established relevance by addressing the child by name and referencing facts about the child, commenting on things in the room, and carrying on a conversation with the child and parent. The accuracy of this contingent interaction proved critical, as children who watched a yoked video of another participant's contingent interaction (i.e. the same length and type of information, but lacking the social contingency and accurate information) did not succeed in the object retrieval task. Thus, providing experience with socially contingent live video chat appears to create a sense of trust and potential relevance of subsequent information presented.

The benefit of social contingency on video is not specific to object retrieval tasks. Roseberry, Hirsh-Pasek and Golinkoff (2014) found a similar benefit of contingent social interaction that preceded a word-learning task with 2– to 2.5-year-olds. Children who received a contingent interaction with the experimenter via Skype learned just as many verbs as those who learned the verbs in a live interaction, whereas those who received a yoked video learned significantly less. Notably, the socially contingent video interaction was only a minute long, and yet still produced this effect. Further supporting the importance of social contingency, Nielsen, Simcock, and Jenkins (2008) found that 2-year-olds were more likely to imitate on-screen actions from an experimenter they could communicate with contingently (one who provided interactive feedback) than one who did not. In fact, children who could interact with the experimenter via video chat performed no differently than those who interacted with the experimenter in person. Even younger infants recognize and benefit from social contingency via video chat; 12- to 25-month-olds who experienced interactive video chat conversations throughout a week (compared to a pre-recorded video) preferred and recognized their partner more, learned more novel patterns, and the older children even learned more novel labels (Myers, LeWitt, Gallo, & Maselli, 2015).

On-Screen Social Relationships

Embedding socially contingent interactions into a video experience signals to young children that what they view on the screen is relevant to their real-world situation. In fact, researchers suggest that understanding the social significance and relevance of screen content could help overcome the video deficit (Richert, Robb, & Smith, 2011). Social relevance includes not only contingency but *social meaningfulness* (Calvert, Richards, & Kent, 2014). Children are indeed able to perceive socially meaningful individuals on screen. For example, young children are able to be comforted by their caregivers when connected to them over a live video feed, and have similar interactions with their caregivers as they would in person (Tarasuik, Galligan, & Kaufman, 2011).

Simply having an audio feed, however, is not sufficient for the infant to feel secure (Tarasuik, Galligan, & Kaufman, 2013). Additionally, infants (ages 21-24 months) are more likely to imitate actions demonstrated on video by their mother than by a stranger (Krcmar, 2010).

When children create social relationships with an on-screen TV character, known as parasocial relationships, they are more likely to learn from that character. For example, young American children (21-months) learned significantly more math concepts from a known on-screen character (Elmo) compared to an equally appealing novel character from Taiwanese TV (Lauricella, Gola, & Calvert, 2011). When children were given the opportunity to become familiar and play with the unknown character (in stuffed animal form), however, their learning increased; this was particularly prominent for the children who showed signs of developing a parasocial relationship in their play with the character (e.g. tucking the stuffed animal in for bed).

Adding contingency to a meaningful social relationship has the potential to lead to even greater learning. This is demonstrated by a 3-month long study by Calvert, Richards and Kent (2014). Infants at age 18-months were given an interactive plush toy they had never before been exposed to, that was either programmed to be personalized to the child (e.g. say the child's name, know their favorite things, same gender) or non-personalized (e.g. called the child "pal", opposite gender). The infants were then tested at 21 months with seriation tasks presented on video by the character they had been playing with the past three months. Infants who had been interacting with the personalized toy showed signs of an increased parasocial relationship with the character (e.g. giving hugs, tucking it into bed) and subsequently performed better on the outcome seriation tasks than infants in a control group who only received the outcome tasks. An increase in performance was not seen, however, for the infants who interacted with the non-personalized characters. Notably, this study investigated a contingent relationship established with a physical, 3-D stimulus (a stuffed animal). Future research should explore the effect of embedded on-screen contingent relationships with media characters into formally “passive television”, which will soon be possible using artificial intelligence technology.

Although technology now allows for on-screen characters to interact contingently with the child, the inclusion of characters that attempt to simulate this contingency is nothing new. As previously mentioned, shows like *Dora the Explorer, Blue’s Clues,* and *Mr. Rogers Neighborhood* feature “quasi-interactive” characters who pause for responses from their at-home audience. Although these embedded one-way interactions lack all five of the prior mentioned benefits of social contingency, they have still been shown to successfully engage children with the program content (Anderson et al., 2000), as well as increase comprehension (Calvert, Strong, Jacobs, & Conger, 2007) and word learning (Strouse, O’Doherty, & Troseth, 2013).

Having a truly interactive on screen partner, however, will likely lead to even greater learning. Research on the benefit of parental co-viewing and scaffolding on young children’s learning from screens further supports the potential of on-screen contingent partners, as similar techniques can be embedded into interactions (Roseberry, Hirsh-Pasek, Parish-Morris, Golinkoff, 2009; Strouse & Troseth, 2014; Strouse, O’Doherty, & Troseth, 2013; 2015). As technology that responds contingently to young children becomes more advanced and prevalent, the possibilities for increased attention and learning through on screen social partners become endless. This recent research has actually led the American Academy of Pediatrics (AAP) to modify their screen guidelines, stripping away any age limit for video chat (compared to a recommendation of 18-months for passive screen time--AAP Council on Communications and Media, 2016).

Toddlers' Preconceptions of Representational Media

Experience is crucial in developing and understanding the affordances of various media--particularly those that looks a great deal like what they stand for. Research shows that 9-month-old infants manually investigate and grasp at objects in both pictures (DeLoache, Pierroutsakos, Uttal, Rosengren, & Gottlieb, 1998) and video images (Pierroutsakos & Troseth, 2003). By 19 months, infants begin to demonstrate an increased understanding of what pictorial media are used for, as they begin pointing at images instead of grasping (DeLoache et al., 1998). They are gradually developing the realization that the image they see is not an actual "X", but a "picture of X" (Troseth et al., 2004).

However, even children sometimes express confusion when determining how images on the television screen relate to their current reality. For example, over half of 3- and 4-year olds thought that when Ronald McDonald came on the TV screen he could see them in their homes (Atkin, Hocking, & Gantz, 1979). Similarly, 3-year-old children were more likely to believe objects on screen could be acted upon, whereas 4-year-olds typically understood that what they saw on screen was merely a picture of the object (Flavell, Flavell, Green, & Korfmacher, 1990). Over time, through experiences with particular representational media, children develop preconceptions of what the various media do and do not afford.

Indeed, the “picture concept” and “video concept” refer to conceptual knowledge about pictures and videos learned through experience (DeLoache, Pierroutsakos, & Troseth, 1996). DeLoache and Burns (1994) hypothesized that 2-year-old children hold a rigid conception of pictures, which inhibits them from seeing the relation between the representation (the picture) and its real-time referent (what was depicted in the picture) during object retrieval search tasks. Supporting this notion is the fact that pictures are often presented in decontextualized ways—a child opens his or her picture book before bedtime and sees a lion, but the child’s bedroom context is *not* the same as the lion’s (e.g. a jungle), removing any contextual relevance between the photo and the child’s life. Additionally, any pictures children might have seen depicting real, familiar referents—themselves or others in family photos, for example—still lack the important relation to *current* reality.

Yet one of the most crucial aspects of the object retrieval task is the requirement for children to recognize the relation between the *representation* and their *current reality.* In other words, the child must apply the mental image of what they have learned from the representational medium (i.e. the photo or video), to what it represents *in real time* (Troseth & Deloache, 1998). This adds to the difficulty of dual representation because the child has to not only mentally depict the referent (hiding location in the room), but also understand that the hiding event is presently occurring (Troseth, 2010). Adding to this difficulty is the fact that, as previously mentioned, pictures typically “provide no immediate real-world information” (Ittelson, 1996, p. 173). This is illustrated in a study where children watched an experimenter squirt ketchup on a stuffed pig, then were asked which of three photos the pig resembled. One photo showed the clean pig, another showed the pig with ketchup, and a third showed the pig with a white spot. Children under the age of two were not successful in identifying that the picture of the pig with ketchup best resembled the pig in real life (Harris, Kavanaugh, & Dowson, 1997). Understanding the temporal relevance between a representation (the photo) and its referent (the pig with ketchup) is clearly difficult for young children.

Aside from polaroid "instant" pictures, printed pictures rarely represented events occurring in real time. Even when DeLoache and Burns (1994) used an instant Polaroid camera during the photo search task, 2-year-old children were not successful. A lack of experience with and understanding of Polaroid cameras, however, could account for this finding. In contrast, young children today who have experience taking instantaneous photos on smartphones, where they can see the photo immediately after taking it, might have a different preconception of pictures in relation to their current reality than children a decade ago. In other words, it might be *easier* for a 2-year-old today to extract information from a pictorial or video and apply it to their current reality.

Changes in Technology

Because previous research has demonstrated that experience with the appropriate affordances of the symbolic medium (e.g. live video) leads to an increased learning from the medium, it is possible that children today might show earlier learning from screen media. The reviewed previous experiments were conducted before the time of live video chat software and the common use of portable screens on tablets and smartphones. As previously mentioned, the front "selfie" camera of a smartphone or tablet now allows users to *see themselves "live" on a 2D screen* as they take a picture or video, and the ease of video recording on phones affords instant (and frequent) playback of home videos. Video chat adds contingent interaction with a person on screen who usually is not physically present (e.g. an out of town relative, or a parent on a business trip). Even if young children are not actively taking photos or video chatting themselves (though they are often capable of doing so independently), they are likely being exposed to others using these devices in such ways.

Therefore, all of this may give toddlers a different concept about the relation between events on a screen and reality, compared to toddlers in research conducted a decade and more ago. It is possible that young children's experience with video chat in their daily lives is pervasive enough that they would not need special training with live video to use information from video in the lab. In other words, children today might learn equally well from video as they do from a person who addresses them "face to face".

The notion of a type of "secular" or "historical" trend in children's success with video object retrieval tasks was posited by Georgene Troseth and her colleagues (2007) a decade ago. Troseth noted an overall increase in children's success rate in the 2000's (38% of children achieving success on 3 of 4 search task trials) compared to mid 1990's (25% achieving success on 3 of 4 trials), and hypothesized that this could be due to a change in technology (i.e. camcorders with reversible screens). Because the shift in technology from the 2000's to today is even more drastic, it is quite possible that current children's performance will reflect this change.

Testing Representational Understanding

Previous researchers have focused on the presence or absence of *representational insight.* However, I prefer a spectrum term of *representational understanding,* which I believe children develop over time (eventually leading to “representational insight”), and that their level of representational understanding depends on the complexity and context of the representation and its relation to its referent. Indeed, typically developing children at this young age of 24-months undoubtedly come into the lab experiment with *some* representational understanding, depending on their prior naturalistic experiences and cognitive development. Toddlers who succeed at the search task with no extra help likely already have the representational insight to use photos or videos in such a context. Others might require extra in-lab experience with the medium, experimenter scaffolding, and/or training to aid in their representational understanding. However, I believe that finding success after in-lab manipulations is not suggestive of an immediate insight achieved that did not exist prior. Instead, I believe it is a sign of a child’s growing, representational understanding that can easily be strengthened (or weakened) under the right conditions. The current research aims to to further explore such conditions. Understanding how and when young children’s representational understanding is strengthened ultimately provides researchers crucial information regarding the cognitive mechanisms underlying children’s representational development.

The object retrieval search task provides an ideal vehicle for testing representational understanding and insight. Typically, children are given four test trials, and the number of errorless retrievals are added across those four trials for each participant. However, whether or not success on the first trial suggests a child having representational insight is a topic of debate. Multiple researchers have reported higher performance on the first trial (e.g. Schmitt & Anderson, 2002; Melzer & Daehler, 2003; Sharon & DeLoache, 2003; Suddendorf, 2003). Some researchers hypothesized this demonstrates that children have fragile insight that is easily "swamped" by perseverative errors (returning to the location they previously found the toy) (Kuhlmeier, 2005; Schmitt & Anderson, 2002).

Although perseverative errors are extremely common in the object retrieval search task with video, and pictures, research on these errors actually suggests that the urge to perseverate is not the *source* of failure in search tasks (Kuhlmeier, 2005). When researchers removed the opportunity to perseverate by either removing previous hiding location (O'Sullivan, Mitchell, & Daehler, 2001) or altering it in a way that made it clear the toy was not hidden there (Sharon & DeLoache, 2003), children's performance did not increase. Thus, perseveration is simply young children's best "educated guess" rather than an overwhelming behavior that diminishes their representational insight. Indeed, when children *do* have the insight, they know where to search, and thus succeed on the task. Fatigue was also ruled out as a possible of failure on subsequent trials; when Suddendorf (2003) changed the search room (and thus hiding locations) for each trial (using photos as the representational medium) with 24-month-olds, thus removing the possibility of perseveration, children were successful across all four trials. Therefore, Suddendorf (2003) calls for researchers to conduct a separate analysis of trial 1 data.

What, then, explains young children's common success on trial 1, and all 4 trials in Suddendorf's multi-room manipulation? Even if they do not truly have representational insight, they could still succeed on trial 1—when faced with the need to pick a hiding location, if they have any memory of seeing or hearing a location, they would likely to search there (Troseth, 2003b). On the first trial of the search task, the children have no competing information for where they last searched for Piglet (and no chance to perseverate). On the subsequent trials, however, the information presented representationally must compete with the very strong memory the child now holds of retrieving the toy (even if they are not successful, the experimenter eventually helps them retrieve it).

Children's encoding of contextual information provides further support for this reasoning. The context in which children encode their memories is crucial to remembering, particularly when children are very young (Barr, 2010). Even though the television source is different than the context in the actual room, on Trial 1, when children have no other memory to act from, they can likely rely on the perceptual similarities between what they saw on the television and what is in the room, even if they do not realize that the image on the screen represents the real situation. However, on subsequent trials, they now have a source of information that matches their test context--the location of the toy in the previous trial. Thus, without representational insight, the source-matched room information overrides any information presented on the mismatched source (the television). Even young infants show evidence of encoding information about the context of an event, as seen in Rovee-Collier's (1999) seminal mobile kicking study--the context that children viewed the actions in (such as the crib bumper) influenced their memory and behavior. When children's understanding goes beyond perceptual matching to achieving representational insight, they understand that what they see on the television, although a mismatched source, stands for the adjoining room, and therefore know to use that information rather than rely on their previous search behavior.

Current Research

The two experiments presented here build upon the work of DeLoache, Troseth, and colleagues, by exploring how new technology might affect children's representational understanding. Both experiments used the object retrieval search task paradigm with 24-month-old children. In Experiment 1, I investigated children's concepts of pictures shown on a digital device (an iPhone) compared to traditional printed photos, to see whether a child today may be more likely to use information from a photo on a smartphone over printed photos because of their experience taking and instantly viewing photos on such devices. Additional conditions varied the amount and kind of in-lab experience children received observing photos taken in real time. In the second experiment, I tested whether children today are able to utilize verbal information presented to them in a live video (with no prior in-lab training) as successfully as they could use verbal information presented to them in person.

CHAPTER II

EXPERIMENT ONE

Introduction

If children's experiences with new picture technology have indeed impacted their conception of photos compared to children in past decades—specifically, the ability of photographs to provide temporally relevant information about real situations—then 24-month-old children today might use information from photographs more readily to search for hidden objects than their counterparts did over 10 years ago. Further, if children's "picture concept" (Troseth, Pierroutsakos, & DeLoache, 2004) is tied to the devices they most frequently use to view and take photos, young children might use a photo shown on an iPhone, for example, with greater ease than a printed photo. On the other hand, it could be that young children require that an adult highlight the temporal relevance of the photo during the experiment. This notion is supported by a previous study using the search task with scale models that demonstrated how children benefited from a researcher highlighting the intentional origins and intended function of the model (Sharon, 2005). Therefore, in Experiment 1, children participated in one of five conditions varying in the way information about the hiding location was presented, resulting in four planned condition comparisons listed in Table 1.

Table 1. The 5 conditions and 4 planned comparisons in Experiment 1.

Condition	Brief Description	Planned Comparison
Printed photo	Identical to previous photo search tasks	
iPhone photo	Same but photos displayed on iPhone	Printed photo
Live iPhone photo	Experimenter takes iPhone photos in "real time" during orientation and testing	iPhone photo
Live iPhone photo + Training	Live iPhone with extra training procedure	Live iPhone photo
Printed photo + Training	Printed photos with same extra training procedure	Printed photo

Method

Participants

Seventy-one children from a city in the southeastern United States participated, but 11 children did not complete the task due to fussiness or boredom. Thus, analyses included data from 60 children (80% Caucasian; 33 males) spread equally across the five conditions, with an average age of 24.5 months (range: 23.0 to 26.2 months), similar to that in previous research. Participant families were recruited via phone using state birth records, from the same middle-class population as in the earlier research. Participants were randomly assigned to condition, with sex matched across condition and no statistical condition differences in children's age or parental socioeconomic status (measured by self-reported income and education level).

Materials

The experiment took place in two adjoining rooms: a waiting room with chairs for the child and parent, and the search space (a.k.a. “Piglet’s room”— 6.8m x 3.1m), which featured a round side table with a green tablecloth, a small flowered sofa with a red blanket draped on it, a blue upholstered chair, a wooden end table with a straw basket on top, and a red floor pillow. As in previous research, printed photos of the furniture (see Figure 1) were displayed on glossy paper in 4in x 6in plastic frames. Digital photos of the furniture were shown on an iPhone 5S with a 2in x 3in screen (either accessed from the camera’s photo library or taken in real time). Finally, the object to be hidden was a 20 cm-tall Piglet stuffed animal (see Figure 1). The entire session was recorded with two video cameras (one in the search room and one in the waiting room) in a picture-in-picture setup.

Parents completed a questionnaire during and/or after the experiment, which gathered information about their family’s socioeconomic status and race/ethnicity, as well as their child’s media exposure, including interactive media (e.g. taking pictures with smartphones, video chatting). Because previous research demonstrates that exposure to live video predicted how well children used photos in the lab, it was imperative to include this experience.

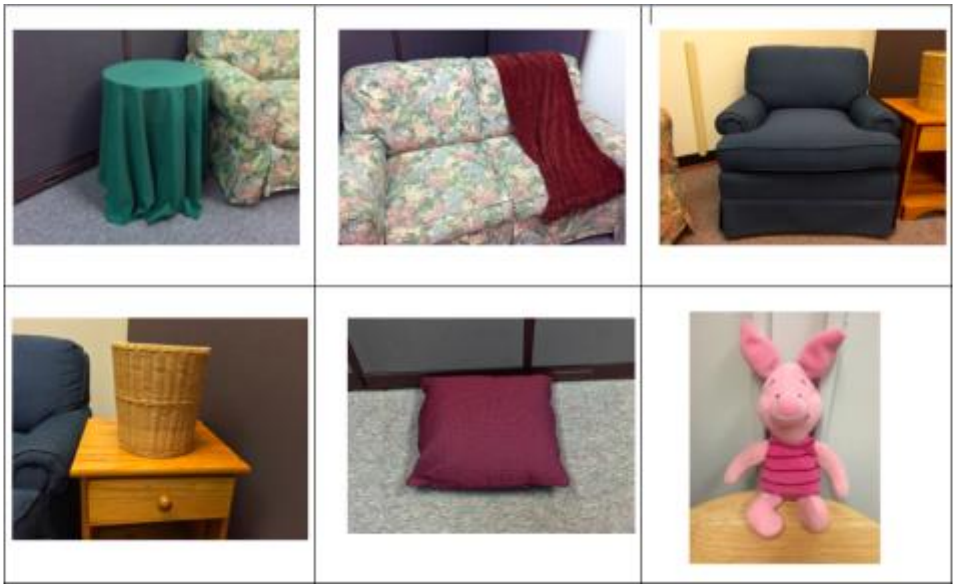

Figure 1. The hiding locations and toy (Piglet) from Experiment 1 and Experiment 2.

Study design and hypotheses

The first of the 5 conditions was nearly identical to that used by DeLoache and Burns (1994) and Troseth (2003b): the experimenter hid a Piglet stuffed animal and then showed children a printed photo of the hiding location to tell them where to search. In the second condition, the experimenter showed the child the same pre-taken photograph on an iPhone, accessed from the phone's photo library. Notably, because the size of the *printed* photo was consistent with that used in past research, this meant that the photo on the iPhone was half the size of the printed photo due to the screen's smaller size. Thus, children in the iPhone conditions were at a slight perceptual disadvantage. I hypothesized, however, that due to children's probable experience with people taking and instantly viewing photos on digital devices, they might perform better (have more errorless retrievals) in conditions using iPhone photos, even with the smaller photo size.

Due to the young age of the participants (just at their second birthday), I was cognizant that the cumulative experience they had with iPhones might be too limited to see the hypothesized increased performance over that in the printed photo condition. Thus, in the third condition, I scaffolded children's experience in the task by highlighting the temporal contingency between the photo and the physical hiding location. To do this, I capitalized on an important feature of iPhones that sets them apart from printed photos: their ability to take and view photos in real time. To highlight the relevance of the photo, the experimenter led the child to believe that (s)he took the photograph of the hiding location right after hiding Piglet (but the child was actually shown the same iPhone photo used in the previous condition, for consistency). I hypothesized that children in the *Live iPhone Photo* condition who had the temporal contingency emphasized during the orientation would achieve more errorless retrievals than children who were exposed to the pre-taken iPhone photo.

However, simply highlighting the temporal contingency in the orientation and alluding to it during the test trials still might not help these young children to grasp the relation between the photo and hiding location. Therefore, I designed a fourth, training condition that further emphasized the temporal contingency between the iPhone photo and the hiding location, as well as a photo's representational role in communicating information: on four training trials, the child helped the experimenter hide Piglet and take photos of the hiding locations for an assistant who then successfully used the photo to retrieve Piglet. I hypothesized that this training would significantly boost children's subsequent use of photos as information to retrieve Piglet, compared to children who only received the *Live iPhone Photo* orientation but no extra training.

To tease apart the effect of the experience actually taking the photos in real time with the experimenter from the effect of seeing the photos intentionally and successfully used for information by an adult, children in a fifth condition experienced the same novel training procedure and successful adult searching with the printed photographs. The results of this comparison help in interpreting any effects of the temporal contingency training in the previous condition.

Procedure

The procedure lasted approximately 30 minutes and included: a warm-up/consenting process, an orientation phase, a placement trial, and 4 search trials. First the child and experimenter spent approximately 5 minutes playing with foam blocks in the waiting room while an assistant described the experiment to the parent/guardian and obtained informed consent. Next, the experimenter introduced the child to Piglet and invited the child and parent into "Piglet's room" for the orientation. In the first three conditions, the child continued on to the test trials directly following the orientation. In the latter two conditions, the child took part in the extra training trials after the orientation, and then proceeded to the test trials.

Orientation phase. The experimenter made an explicit connection between each piece of furniture and its photograph by individually approaching it with the child, labeling it, and highlighting the relation between the 2D photograph and 3D object. In the conditions using printed photos and the first condition with digital iPhone photos (where the experimenter accessed each photo in the phone's photo library), the experimenter remarked: "This is a picture of Piglet's [furniture item]. And look, here is Piglet's [same actual furniture item]! See, they're the same!" In the conditions that involved taking iPhone photos in real time, the experimenter said: "I'm going to take a picture. [Takes photo]. Let's look at the picture I just took," then made the identical comparison between picture and furniture item. To complete the orientation and set up the placement trial, the experimenter demonstrated the connection between a fifth item (the floor pillow) and its photograph. Afterwards, the experimenter accompanied the child and parent back to the waiting room.

Placement trial. In all conditions, this phase was designed to emphasize how a picture can be used to give information. The experimenter showed the child the photo of the pillow and told the child, "Piglet wants to go to his room and sit right here. You take him and help him sit right here." If the child did not follow the direction, then the experimenter gave hints, re-emphasizing the relation between the photo of the pillow and the actual pillow. After the child placed Piglet on the pillow (with or without experimenter help), the child and experimenter returned to the waiting room.

The parent was asked to stay in the waiting room and work on filling out the demographic questionnaire and a media questionnaire about his/her child. If the child required his/her parent to be in the search room for comfort, the parent was instructed to not provide any hints (e.g., to look only at the child and not at any hiding locations, let the child lead, etc.).

Test phase. The experimenter showed the child a printed photo of a hiding place or the same photo on the iPhone, saying, "I'm going to hide Piglet right here in his room, and when I come back you can go and find him!" The experimenter then handed the photo/iPhone to the assistant and went into Piglet's room, closing the adjoining door.

While the experimenter was gone, the assistant pointed to the hiding location on the photo and reminded the child, "[experimenter] is hiding Piglet right here in his room!" Furniture items were not labeled during the test phase, so children had to rely on the pictures rather than on verbal cues for information.

In the "Live" iPhone conditions, the procedure was adapted slightly. The experimenter told the child, "I'm going to hide Piglet in his room and take a picture of where I hide him. Then, I'll come back and show you the picture of where I hid Piglet, and you can go and find him there!" The experimenter took the iPhone and entered Piglet's room, closed the adjoining door, and hid Piglet. Meanwhile the assistant reminded the child, "[experimenter] is hiding Piglet right now in his room. (S)he's taking a picture to show us where (s)he hid him." However, the experimenter actually selected the identical image of the hiding location from the iPhone's photo library used in the other conditions.

Next, the experimenter re-entered the waiting room, pointed to the photo, and said, "Piglet is hiding right [in/under] here in his room. Go find him!" If the child went to the incorrect location, the experimenter provided a series of increasingly explicit prompts ("Not there, where else do you want to look?"; "Do you remember where I pointed to in the photo? Piglet is hiding here in the same place"; and finally, "I think Piglet is hiding somewhere over here.") If necessary, the experimenter helped the child retrieve Piglet. The hiding positions were: under the green tablecloth, under the red blanket, under the bottom flap of the blue chair, and inside the straw basket. The order of hiding locations was counterbalanced across participants.

Extra training phase. In the final two conditions, after the placement trial and before the search trials, children participated in an extra training phase in which they helped the experimenter hide Piglet for an adult assistant to find. This phase was originally designed to further highlight the temporal contingency between photo and hiding event for children using photos taken in real time, but also highlights the *intentional creation* or *selection* of a particular photo by the experimenter, and provides children a successful modeling of the task solution (*use* of the photo for information) by the adult assistant.

The experimenter's comments to the assistant matched the language directed to the children in the test phase for the respective condition. The only modification was the inclusion of the child's name when directing the talk to the assistant, for example, "[Child's name] and I are going to go hide Piglet…" The child accompanied the experimenter when hiding Piglet, at which point the experimenter either highlighted the temporal contingency between photo and search task by hiding Piglet and taking a photo of the hiding location with the iPhone camera in real time, or explicitly pointed out the connection between the printed photo and Piglet hiding in that predetermined location. After returning to the waiting room, the experimenter addressed the assistant following the same, condition-specific script described in the test phase. On each of the four training trials, the assistant demonstrated understanding by looking carefully at the photo, nodding, and successfully retrieving Piglet immediately. After the fourth training trial, the experimenter told the child that now it was his/her turn to search for Piglet, and the test phase began.

Measures

Children's search behavior was recorded on a data collection sheet during the session (by the assistant), and later by a different trained research assistant using the video recording of the session. In the rare case of a disagreement, the video data was re-analyzed and an agreement was reached. Consistent with previous research, only errorless retrievals on the first search were considered correct. Children received a score from 0 to 4 across the four search trials, with 4 representing 100% errorless retrievals. First trial performance was also analyzed separately, resulting in an additional score based on performance on trials 2-4, thus ranging from 0 to 3 across the last three trials.

Results

Prior media exposure

Media questionnaire data was missing for one participant. According to their parents, the majority of the toddlers had been previously exposed to the kind of live video afforded by new technology. At least once a month, 90% of the toddlers engage in video chat, with 17% of them video chatting daily or multiple times a week, and another 27% video chatting weekly. The average age of exposure to video chat was around seven months of age (M = 7.22 months, SD = 5.77), and parents believed their infant was actually engaging with the video chat experience around their first birthday (M = 12.8 months, SD = 6.10).

Parents reported that in the past month, 95% of the toddlers had seen themselves on a screen (e.g. phone, tablet, computer, camera) while someone filmed a video or took a photo of them, with 42% experiencing this exposure daily or multiple times a week. Nearly half (44%) of the toddlers had actually taken a photo or video selfie *themselves* in the past month. Finally, in the past month, 97% of the toddlers had seen themselves on a screen while watching the "replay" of a video of photo that was taken at an earlier time, with 44% viewing such playback daily or multiple times a week.

No correlations were discovered between naturalistic live video experience and search task performance, likely due to three factors: 1) high live video exposure across all participants; 2) generally low search task performance overall; and 3) small sample size. There were no significant differences in naturalistic exposure across conditions; however, children in the *iPhone* and *Live iPhone with Training* conditions--two of the conditions that I expected to provide more support for children's use of photos--had somewhat lower average amounts of exposure from their daily lives compared to children in the other conditions. Thus, if anything, less prior exposure (through random assignment) worked against my hypothesis of better performance in these conditions.

Search performance

In each comparison across conditions, I looked at overall performance as well as children's success on the first trial and on the remaining 3 trials: after searching for and finding Piglet on Trial 1, children have competing information on Trials 2-4 between their memory of where they found the toy (with or without help) on Trial 1 and a memory of information they were just shown in a photo. Thus, the latter comparison provides a stringent test of children's understanding that the picture provides the relevant information about the hidden toy's location. See Table 2 for a comparison of all search score means and standard deviations.

Table 2. Mean number of errorless retrievals in Experiment 1 (with the percentage in parentheses to the right, and the standard deviation in parentheses underneath).

	Number of errorless retrievals out of 4 trials	Number of errorless retrievals out trials 2, 3, & 4
Printed Photo:	0.42 (11%) (1.00)	0.25 (8.3%) (0.62)
iPhone Photo	0.92 (23%) (1.56)	0.67 (22%) (1.15)
Live iPhone Photo	0.92 (23%) (0.79)	0.33 (11%) (0.49)
Live iPhone Photo + Training	1.83 (46%) (1.19)	1.33 (44%) (0.98)
Printed Photo + Training	1.33 (33%) (1.55)	1.08 (36%) (1.31)

Printed photo vs. iPhone photo. There was no significant condition difference in toddlers' errorless retrieval rates when looking at all four trials for the *Printed photo* (11% correct) versus the *iPhone photo* (23% correct) conditions. The same non-significant pattern emerged on trials 2-4 (*Printed photo*—8.3%; *iPhone photo*—22%). Similarly, in the first trial with no competing information, 17% of toddlers successfully used printed photos in the search task, compared to 25% with digital iPhone photos.

Thus, merely seeing photos on an iPhone (compared to seeing printed photos) did not help children to more successfully use the photos in the search task.

iPhone photo vs. live iPhone photo. The next comparison compared iPhone photos that either were in the iPhone photo library (i.e., previously taken) or (supposedly) taken by the experimenter right after (s)he hid Piglet. There was no significant condition difference in toddlers' errorless retrieval rates across all four trials (*iPhone photo*—23%; *Live iPhone photo*—23%) nor on trials 2-4 (*iPhone photo*—22%; *Live iPhone photo*—11%). However, on Trial 1, 58% of toddlers who were told the experimenter had just taken the iPhone photo were successful, compared to a 25% success rate in the other condition with the pre-taken iPhone photo.

Live iPhone photo vs. live iPhone photo + training. Before the search trials, toddlers in the *Live iPhone Photo + Training* condition had helped the experimenter hide Piglet and take an iPhone picture of the hiding place for the adult confederate. Toddlers who received this experience had significantly more errorless retrievals across the four trials (46% success rate) compared to toddlers in the Live *iPhone Photo* condition (23% correct; $t(22) = 2.22$, $p = .037$, $d = 0.95$). The toddlers in these conditions had similar success rates on Trial 1 (50% and 58% correct, respectively). However, their performance was vastly different when looking at the subsequent trials: toddlers with the training experience had a 44% errorless retrieval rate, significantly higher than the 11% correct achieved by toddlers without this helpful experience, $t(22) = 3.15$, $p = .005$, $d = 1.34$.

Printed photo vs. printed photo + training. It was important to determine whether the positive effects of the training depended on the type of photo used. Toddlers who used printed photos and received the extra training did perform better (33% errorless retrieval rate) than toddlers who used printed photos without the training (11%), but this difference did not reach the conventional level of significance—$t(22) = 1.716$, $p = .10$. Similarly, the errorless retrieval rate for trials 2-4 was 36% with training compared to 8.3% without, a marginally significant difference with a large effect size, $t(15.7) = 1.99$, $p = .064$, $d = 0.81$ (degrees of freedom correction from 22 to 15.7 according to Levene's test for equality of variances).

Discussion

Previous research documents the struggle that 24-month-olds have using photographs to correctly search for hidden objects, regardless of the manipulations implemented to help children notice the representational nature of the photos (DeLoache & Burns, 1994; Troseth, 2003). The current experiment adds to the literature by investigating: 1) whether naturalistic experience with photos that highlight temporal contingency ("selfies" and other "live" smartphone photographs) helps children to use photos as representations in the search task; 2) whether emphasizing the temporal contingency of digital photos *during* the search task helps children achieve the representational understanding needed to succeed; and 3) whether also emphasizing the *intentionality* behind creating/showing the photos leads to even greater representational understanding.

Toddlers of today who participated in the traditional (DeLoache & Burns, 1994) version of the search task (with printed photos) did not show any gain in representational understanding compared to the 1994 children's level of success, despite exposure to digital photos afforded by modern technology. Even when using the medium by which they often experience digital photos (a smartphone), toddlers today performed no differently than those in previous decades.

When the temporal contingency between photo and current event was highlighted by taking photos of the furniture during the orientation and having the experimenter refer to taking the photo of the hiding place in real time on each test trial, toddlers found the toy on the first trial at rates much higher than usual--more than twice as many toddlers were successful as when pre-taken photos (printed or on the iPhone) were shown to children. However, on the next three trials, toddlers' use of the live photos plummeted. Looking at it from another angle, toddlers who used pre-taken iPhone photos performed relatively consistently across trials—25% errorless retrieval rate on Trial 1, compared to 22% on the following three trials. The success rate of toddlers shown live iPhone photos, however, dropped from 58% on Trial 1 to 11% on the next 3 trials. The sudden dip in performance on subsequent trials suggests that the scaffolding may have helped them make the initial surface-level *identity* relation between the photo and real life hiding location on Trial 1, but did not lead to a strengthened *representational understanding* of the relation between the photo and referent--toddlers were more likely to perseverate than use the information provided in the photos.

The training procedure gave toddlers additional experience with the temporal contingency between the iPhone photos and the real-life locations. Instead of merely having iPhone photos of the furniture taken in real time during the orientation and then alluding to the taking of live photos during the test trials, this novel training procedure gave children the chance to "shadow" and help the experimenter hide Piglet in each location and take a photo of that hiding place to help an adult confederate. Notably, this training may also have highlighted for the children the experimenter's *intention* for the photo being taken to serve as a representation of the hiding location to help the searcher find Piglet. Indeed, toddlers in this condition performed significantly better than those without the extra training. Although toddlers in both *Live iPhone* conditions achieved success on Trial 1, it was only those with the experience helping make representations for the adult who maintained that success across all four test trials. On trials two through four, toddlers with training were four times more likely to find Piglet than toddlers who only saw live photos being taken during the orientation. This suggests that toddlers with the special training achieved a higher level of representational understanding, as opposed to a those in the *Live iPhone* condition, who likely grasped an initial identity relation (based on similarity) on Trial 1, but did not have a strong enough level of representational understanding to reject their memory of their previous search, and use the information from the photos on subsequent trials. Instead, the memory of where they found the toy on the previous trial was a stronger source of information than the representation. Because the toddlers with training better understood the representational nature of the photo, they were able to use the photo for information *even though* there was a source mismatch between photo and search room (Barr, 2010).

In another condition, children participated in the novel training condition but using the printed photographs of the furniture. This comparison teases apart the effect of seeing the (iPhone or printed) photographs intentionally and successfully used for information by the confederate, from the experience of actually creating the depiction in real time (taking the iPhone photo) which highlighted the temporal contingency between the photo and referent. Unlike the significant boost from the Live iPhone training condition, there was no significant difference between toddlers' success using printed photos with and without the additional training. This result suggests that both aspects of the Live iPhone photo training together gave young children greater representational understanding of the relation between the photo depiction and its referent, including: 1) seeing digital photographs intentionally taken and then directly utilized as a representation to help another individual, which 2) highlighted the relevant temporal contingency of the photos.

Previous research using video suggests that merely watching an adult confederate succeed is not enough--when two toys were hidden in each spot and children first watched an assistant retrieve the toy successfully, there was no boost in performance (Troseth, 2003a). Another video training procedure had children first watch the object hidden both through the doorway and on the live television feed. Although children were successful in finding the toy during the training situation, as soon as the door was closed and children had to rely just on the video information, they were no longer successful (Troseth, 2003a). In the current training manipulation, however, children *helped create* the representation with the experimenter for the assistant to use, all while having this representational intention continually emphasized (both while taking the picture and when the experimenter communicated the intention to the adult confederate).

Training with the printed photos did boost performance; compared to children who used printed photos without the special experience, toddlers showed a marginally significant increase in performance on trials 2–4 with a large effect size. With more participants, a significant difference might emerge for those trials. There was also no significant difference between toddlers' overall success rate in the two training conditions when directly compared.

Moreover, a closer look at the trial-by-trial performance of children who received various levels of support reveals an interesting story. As previously discussed, toddlers in the *Live iPhone* condition received a small amount of information about temporal contingency when the experimenter took photos of the furniture during the orientation. These children appeared to catch on immediately on Trial 1 and make the connection between the photo and referent (58% errorless retrieval rate), but their performance plummeted on the subsequent trials (11%). Conversely, toddlers in the *Printed Photo + Training* condition did not show high performance on the first trial (25% errorless retrieval rate), but instead slowly increased in performance across the next three trials (36%). Toddlers in the *Live iPhone Photo + Training* condition who had the temporal contingency highlighted both during the orientation and while creating the photos with the experimenter for the adult confederate started out successful and did consistently well across trials (50% on Trial 1, 44% on the following three). This suggests that the printed photo training (which highlighted he experimenter's intentional use of the pre-taken photos, along with the assistant's successful use of them for information), might have helped children gradually catch on to the representational nature of the printed photos.

Although toddlers demonstrated the highest overall performance with the training using the live iPhone photos, this finding is counterintuitive in terms of cognitive load or working memory. Children using the live iPhone photos (with or without training) only viewed the hiding location once during the test trials when the experimenter returned to the waiting room to display the photo (s)he "just took" after hiding Piglet. In contrast, in the other three conditions (and consistent with previous research), the child had the hiding location pointed out three times on the previously-taken digital or printed photo (before, during, and after the experimenter's hiding of the toy). Thus, success in the *Live iPhone Photo + Training* condition is even more impressive. Additionally, by chance, toddlers in this condition had the lowest prior exposure to live video (as reported in the parent media questionnaires) compared to those in the other conditions. Based on research reporting correlations between naturalistic live video exposure and photo search task success (Troseth et al., 2007), children's success may have been even higher had their prior live video experience been greater.

CHAPTER III

EXPERIMENT TWO

Introduction

The second experiment explored children's responses to another representational medium—live video. As discussed earlier, previous research demonstrates that young children (24 months) easily learn and apply information provided "in person" by an individual to solve a problem (find a hidden toy), but not information from the same individual on video (Troseth et al., 2006). However, when toddlers were given just 5 minutes of experience with a live, contingent video interaction with the person (similar to video chat), they successfully used information offered by the person on video to succeed at the search task (Troseth, et al., 2006; Troseth, 2003b). Similarly, after children saw themselves on live video at home (for an hour total, across 2 weeks), they were successful at using video-presented information in the lab on a search task (Troseth, 2003b). Finally, a correlational study also supports the positive relationship between naturalistic experience with live video and use of information presented on video and pictures (Troseth, et al., 2007). Notably, all of these studies were conducted before the popularization of live video exposure from video chat (e.g. Skype, FaceTime) and instant playbacks afforded by smartphones.

The primary aim of Experiment 2 was to explore whether toddlers today still require specific video chat training to support their learning from video, or if, due to their higher naturalistic exposure to live video, they will come to the lab able to learn equally well from an individual on video and in person. Children (24-months) were told verbally where a toy was hidden by an experimenter who addressed them either directly in-person, or from a video screen (an iPhone screen) via video chat from an adjoining room. The child was then asked to retrieve the toy. The procedure is identical to one used in prior research that occurred in the same lab space (Troseth, et al., 2006). Without special in-lab video chat experience, children in the 2006 research rarely used information from video. However, given the everyday exposure to new technology (such as video chat on computers and smartphones) that children now experience, I hypothesized that they might learn from an individual on video just as well as if the individual gives them the information in person. Parents completed the same questionnaire from Experiment 1 about their children's exposure to video chat and other media, providing information as to just how much naturalistic exposure toddlers have today.

An alternative hypothesis is that the experience toddlers have with live video in their everyday lives is not specific and supportive enough to help them in the lab task. Notably, in the past studies, a co-viewing parent would often scaffold the child's learning by also engaging with the experimenter via video chat, or pointing out the connections between real life and events on screen (Troseth, 2003b; Troseth et al., 2006). It may be that today's toddlers (particularly at the young age of 24 months) still need specific parental support and/or extra in-lab training to recognize the connection between events on a video screen and a real-world situation.

Method

Participants

Thirty-two children participated, but data from eight were dropped from analysis for not completing the task (e.g. fussiness or boredom). Analyses include data from 24 children (13 male), with an average age of 24.7 months (range of 23.8 to 26 months). Participants were recruited the same way as in Experiment 1 from the same Southeastern U.S. city, and randomly assigned to one of the two conditions. Sex was matched as closely as possible across conditions, and there were no statistical condition differences in children's age or self-reported parental income. There was, however, a difference in parental reported education level, with children in the *Direct* (in-person) condition coming from families with a higher education level than children in the *Video Chat* condition-- $t(22) = 2.20$, $p = .039$. However, all parents in both conditions had at least some college education: 61% parents of children in the Direct condition held a Masters or Doctorate degree, compared to 46% of parents of children in the Video Chat condition. Most participants were Caucasian (79%).

Materials

The rooms, hiding locations, iPhone, and demographic/media parent questionnaire were identical to the ones used in Experiment 1. In the video chat condition, the experimenter used an Apple 13" MacBook Air for the video chat call, which toddlers saw both during the brief orientation and when they entered the search room on each trial.

Procedure

The same warm up and consent procedure from Experiment 1 was used.

Orientation and practice trials. The toddlers first received an orientation to the four hiding locations in Piglet's room to ensure they understood the labels being used—green table, blue chair, basket, and pillow. Then, to introduce the hiding game, the experimenter hid Piglet in/under each of the hiding places in front of the child, and asked the child to retrieve Piglet after a momentary delay caused by walking away from the furniture (breaking the child's visual access to the hiding place) and "counting to five".

Placement trials. Next, the experimenter asked the toddler to place Piglet on/in each of the four hiding locations, with the experimenter reinforcing the verbal labels of each location as they went (following procedures from Schmidt, Crawley-Davis, & Anderson, 2007; Troseth, et al., 2006). Because the child did this for each of the four hiding locations (rather than just once, as in Experiment 1), the placement trial score (from 0 to 4 errorless placements) served as a pretest of children's understanding of the basic task structure, since it occurs before any condition manipulation. Children in general were able to complete this task with ease, so a score of 2 or less was considered to be "not passing" the placement task--an indication that children did not understand the verbal labels, were not listening to the verbal instructions, and/or were not cooperating with the experimenter's game.

Video chat orientation. For participants in the *Video Chat* condition, the same brief (approximately one minute) orientation to live video used by Troseth et al. (2006) was used. This took place in the search space where the experimenter's laptop was set up for the video chat call. The iPhone to be used by the child during the test phase was temporarily brought from the adjoining waiting room into the search room for the sake of this orientation. The child was asked to identify themselves and their parent on both the laptop and iPhone screens. The experimenter then demonstrated how she could talk to the child through the video and in person at the same time, by waving and alternating between looking directly at the child and at the laptop camera.

Test trials. Next the child and parent went to the adjoining waiting room. In the *Direct* condition, the experimenter told the child that she was going to go hide Piglet and would come back to tell the child where she had put him. While the experimenter hid Piglet, the assistant said to the child, "[Experimenter's name] is hiding Piglet right now!" When the experimenter returned, she said, "I hid Piglet under/in the _____. He's under/in the _____. Can you go find him? Remember, he's under/in the ____." Then the child was allowed to search in the room. The procedure and wording for the *Video Chat* condition was identical, except all of the communication between the experimenter and child took place over video chat. The child sat on his or her parent's lap in front of a cart which held the docked iPhone showing the experimenter's face against the side wall of the lab. Before going off screen to hide Piglet, the experimenter first reminded the child to keep watching because she would "come back on the iPhone" to tell them where she hid Piglet. When needed in the Video Chat condition, the assistant (in the waiting room with the child and parent) re-directed the child's attention to the screen. To clarify, children did not *watch* Piglet hidden, they simply received the verbal information afterward from the experimenter (in person or on the iPhone screen). To test whether performance improved across trials, 8 test trials took place (instead of the typical 4), using the four hiding locations each repeated twice (with two orders counterbalanced across participants).

Measures

The same measures from Experiment 1 were used.

Results

Media Questionnaire

According to parent reports, all but one of the 24 toddlers (96%) had been previously exposed to video chat. (That toddler had been randomly assigned to the video chat condition.) Parents reported that their infants were first exposed to video chat around 6 months of age (M = 6.17 months, SD = 6.69), and started interacting more with it just after their first birthday (M = 13.2 months, SD = 6.59). Compared to the toddlers in Experiment 1, the majority of the children in Experiment 2 were lighter video chat users, with just 12.5% of them video chatting daily or multiple times a week (54% reported video chatting on a weekly basis and 29% on a monthly basis).

All but one toddler (96%) had previously seen themselves on a screen (e.g. phone, tablet, computer, camera) while someone filmed a video or took a photo of them, with this occurring frequently (daily or multiple times a week for 46% of the toddlers). Additionally, 38% of the toddlers had previously taken a selfie photo or video of themselves. Finally, all but one toddler had previously seen him or herself on a screen while watching the “replay” of a video of photo that was taken at an earlier time. This was a very frequent activity, with 50% of the toddlers doing so daily or multiple times a week, and an additional 41% doing so weekly.

No correlations were discovered between naturalistic live video experience and search task performance. This is likely due to the same three factors as in Experiment 1: 1) high exposure to live video across all participants, 2) generally low search task performance overall, and 3) small sample size. There were no significant differences in exposure between the two conditions.

Search Task Performance

The dependent variable was the number of errorless retrievals. All 24 toddlers completed the first four trials. Only 18 of the participants (9 in each condition) completed the subsequent four trials due to a procedural change later in data collection to begin data collection on a related transfer task (not reported here).

The first hypothesis was that there would be a significant difference in the number of errorless retrievals between the two conditions. For the same reasons as in Experiment 1, Trial 1 success was analyzed both separately and with the other trials, as the subsequent trials provide more meaningful information about representational insight. Independent sample t-tests were run, comparing performance in the Direct and the Video Chat conditions. Regardless of how the trials were analyzed, however, there were no significant (nor marginal) differences in the number of toddler's errorless retrievals between the two conditions. Participants who received the hiding information in the direct condition had a 58% success rate across the first four trials, and 60% across eight trials, and those who received the information via video chat had a 43% success rate across the first four trials, and 51% across eight trials.

However, toddlers in the video chat condition underperformed those in the direct condition on the *placement* trial, *before* any condition manipulations were introduced. Four participants in the video chat condition failed on 2 or more of the 4 placement trials, compared to just one participant in the direct condition. All five were males. When data from those five participants were removed from analyses, success rates in the Direct condition were 64% on all four trials and on all eight trials, whereas the success rate in the Video Chat condition was 50% on the first four trials, and 67% across all eight trials. Because five participants were removed from this analysis, more participants (who successfully pass the placement trial) are needed to further explore similarities and differences in search task performance across conditions.

A secondary hypothesis was that performance might increase across the 8 trials. Toddlers had a clear tendency to reach fatigue/distraction by the seventh trial. Therefore, I decided before conducting the analyses to test this hypothesis by comparing performance on two early trials (2 and 3) to two later trials (5 and 6). Trial 1 was excluded as being less informative about representational insight (since children might succeed by noticing similarity, as described earlier), and Trials 7 and 8 were not considered due to reported child fatigue on these trials (later confirmed by a dip in performance on those trials). I excluded Trial 4 to match the same number of earlier (Trials 2-3) and later (Trials 5-6) responses.

Paired sample t-tests were run, comparing the average success rate on trials 2 and 3 to trials 5 and 6 in the two conditions. In the Direct condition, participants' average success rate rose from 56% (SD = 0.46) to 83% (SD = 0.35) which was a marginally significant increase, $t(8) = 2.29$, $p = .051$. In the video condition participants' average success rate increased from 33% (SD = 0.35) to 61% (SD = 0.33), a statistically significant increase, $t(8) = 3.16$, $p = .013$.

Discussion

Toddlers were just as successful using verbal information provided on Video Chat (displayed on an iPhone) as they were using information given to them in person. This result could point to a historical trend, where naturalistic experience with live video as afforded by new technology affects how young children use information from video. The high exposure to and use of live video, as noted by the parent self-report media questionnaire of their children's experiences, supports this notion. Furthermore, performance was higher in the current video chat condition than the comparable condition from Troseth et al. (2006) over a decade ago, which had a 27% success rate across the four trials (compared to approximately 50% today).

One possibility for this change is that toddlers today are not *interpreting* the video chat as a representation for the person on screen, and are instead "seeing through" the screen based on their prior experience with live video. This could lessen the demand for children to achieve dual representation. Children might instead be making a simple identity relation between the video representation and actual person, similar to the reasoning provided by the shrinking room study (where children thought the big and little room were the same room rather than the small room being a representation for the big room--DeLoache et al., 1997), and when the television screen was placed behind a window (causing children to think they were looking through a window, thus removing the need also represent and process that what they were viewing was occurring on television--Troseth & DeLoache, 1998).

Thinking in functional terms of children *using* information from screens, is the naturalistic experience children receive in their daily lives enough to overcome the video deficit when interacting with live video? Indeed, this was a possibility proposed when the term video deficit was first coined by Anderson and Pempek (2005). Citing the results of Troseth's (2003) experiment that found when toddlers viewed themselves on live video at home for two weeks prior to completing the task, they were more successful using video in the search task, Anderson and Pempek hypothesized that the video deficit in young children was flexible, being easily affected by their experiences with video. The results from Troseth et al. (2006), as well as the current experiment, further support this notion. Recent research, however, demonstrates toddlers normally affected by the video deficit learn equally well from someone over video chat compared to a live demonstration (Roseberry et al., 2014; Myers et al., 2015). More research is needed to determine whether this finding is specific to socially contingent video chat situations, or whether the naturalistic experience with live video that young children receive today is modifying how they interpret and use information on video broadly, including passive video (such as TV).

There are important limitations of this study, all potentially lowering children's search performance in the Video Chat condition. First, by chance, more children in the Video Chat condition struggled in the placement trial, and children in this condition also came from families who had on average lower parental education (albeit still high). A larger sample size is therefore needed. Additionally, because the child sat on a chair (or on his/her parent's lap) with the iPhone screen on a cart in front of them, the small screen of the iPhone was likely at a greater distance from the child than is typical for everyday experiences with live video on smartphones or tablets (where the screen is often held by the child or parent). Finally, when going into the search room on their way to retrieve Piglet, several children were distracted by the keyboard on the laptop used by the experimenter to video chat between the search room (experimenter) and the waiting room (child). It will be important to rectify this problem (e.g., cover the keyboard) going forward.

CHAPTER IV

GENERAL DISCUSSION

Both picture and video representations are pervasive in most American children's lives, from picture books to educational television shows. Adults, including content creators, care providers, teachers, and parents, often assume that young children can learn equally well from representational media as they can from their real-life referents. Although previous research has debunked this assumption (i.e. the video deficit--Anderson & Pempek, 2005), the current research suggests that experience with digital photos and live video afforded by new technology may positively influence children's ability to understand, use, and thus learn from such representations.

The results from Experiment 1 suggest that, for 24-month-olds, being surrounded by instances of "live" photos on digital cameras and smartphones (where the photo is seen right after being taken) in their daily lives alone is not sufficient to understand the relationship between the photo and referent in a search task. That is, mere naturalistic exposure to the temporal relevance between a photo and its real-world referent, by itself, is not enough for very young children to see the immediate relevance of a photo to give information about a current situation. Even a reminder that the photos in the experiment were taken in real time (during the *Live iPhone* photo orientation and search trials) was not sufficient. However, when this exposure (along with other factors) was further highlighted through extra training, toddlers did demonstrate representational understanding of the connection between the photo and referent in the context of the search task. This is important information for those who might otherwise assume that the picture-reality connection is transparent; even in a world inundated with reminders of this connection, very young children greatly benefit from examples and scaffolding of how to use photos for information, just as they must learn to use other cultural tools (Tomasello, Striano, & Rochat, 1999).

The results from Experiment 2 suggest that the high amount of *naturalistic* experience with live video that young children today receive has indeed increased their representational understanding of how videos images relate to their referents in real time. Young children (24-months) today performed better in an identical search task (using verbal information presented on video) than children over a decade ago did. Additionally, the toddlers in the current Video Chat condition performed just as well as toddlers in the Direct, in-person condition, whereas toddlers a decade ago performed significantly better in the Direct in-person condition than the Video Chat condition (Troseth et al., 2006). This hypothesis and pattern of results is consistent with research findings discussed earlier, demonstrating that in-lab exposure to and naturalistic experience with live video aid in search task performance (e.g. Troseth, 2003b; Troseth et al., 2006; Troseth et al., 2007).

Compared to young children around their second birthday (Anderson & Pempek, 2005), infants showed less disparity in their learning from video and in-person presentations in research completed a decade ago (e.g. Barr et al., 2007; Krcmar et al., 2007). However, *recent* research (including the current Experiment 2) demonstrates that young children today learn equally well from contingent, live video as they do from in-person presentations (e.g. Roseberry et al., 2014; Myers et al., 2015). Additionally, results from Experiment 1 suggest that relevant scaffolding aids young children's understanding of photo representations. Therefore, I believe that the *experience* (both scaffolded and naturalistic) children receive with relevant representational media greatly influences their concepts of such media, and thus how they understand, use, and learn from them. The *type* of experience, however, particularly naturalistic, has the potential to either help or hinder young children's concepts and understanding. Moreover, many factors can impact this experience, including historical shifts (e.g. the technology that is available and affordable, trends in media content, and policy changes in media usage), as well as individual differences (e.g. children's specific experiences, family rules, socioeconomic status, geographical location).

For instance, watching television remains a fairly passive viewing experience compared to Skyping with Grandma (although this will likely change over time as technology such as artificial intelligence is incorporated into television programs). In addition to the (current) lack of interactivity, there are various formal features of children's television (e.g. animated characters) that may lead children to interpret the content as fictional and not directly related to their lives. Therefore, children might hold a concept for television that cues them to process the content as entertainment and fictional rather than trustworthy and relevant information. On the other hand, children may also view passive video content on a smartphone or tablet. However, their experience with these devices likely also includes activities that clearly delineate the tablet/smartphone as a communicative device, including participating in video chat (e.g. FaceTime, Skype). These experiences (e.g. viewing a trustworthy and known individual on the screen while video chatting) might lead to a generalized concept that these devices *do* provide information that is relevant, meaningful, and accurate to the child. Thus, my prediction regarding differences in individual experience is that a young child who spends the majority of her time watching animated television programs that have little relation to their current reality will likely have a different concept of video images than another child who spends the majority of his screen time video chatting with relatives, an experience directly related to his current reality.

This prediction is based on the concept of *functional fixedness* (Duncker & Lees, 1945), wherein experience with and/or observation of an object (or medium) in a particular way influence one's perceived use of such object to a degree that it "blocks" one from seeing other possible affordance and uses of the object (Luchins & Luchins, 1970, as cited in National Research Council, 2000). Thus, the hypothetical child mentioned above who primarily watches passive television would need to overcome the cognitive hurdle of a familiar medium being used and acted upon in such a different way. The hypothetical child who is primarily exposed to video chat, on the other hand, might "see through" all video, taking it at "face-value".

Although this could help overcome typically observed video deficit effects, there is certainly value in knowing when to interpret video as real or as fictional. The more likely scenario, however, is that children today receive exposure to both live video chat experiences and passive television, promoting a more flexible view of video that it is sometimes real and sometimes fake. This would require that children begin to distinguish the formal features that separate documentary from fictional television (a skill that children typically develop around age 4 or 5 -- Wright, Huston, Reitz, & Piemyat, 1994). It may be that exposure to contrasting kinds of video will promote earlier development of this conceptual information about video. However (as discussed below) children's ability to recognize multiple functions of video will be constrained by factors such as the development of cognitive flexibility and other executive functions (Zelazo et al., 2003). For example, to succeed in mentally representing both a representational object and its referent, a child must have the working memory capacity to hold two mental representations at one time (Halford, Cowan, & Andrews, 2007) and the flexibility to switch between different functions of a representational medium.

In summary, I believe that how young children interpret and use representational media is a flexible cognitive process, heavily influenced by: individual media exposure, the given context, cultural and historical trends, and scaffolding that they receive highlighting the relations between representational media and real life, combined with their current level of cognitive flexibility. Results from both experiments support this. In Experiment 1, relevant scaffolding (the in-lab training that primarily highlighted the photo's temporal contingency) increased children's representational understanding of photos. In Experiment 2, changing technology coupled with individual experience (the high amount of naturalistic experience children today receive with live video) likely accounts for the increase in representational understanding of live video compared to children a decade ago.

Future Directions: Transfer of Learning

If these types of experiences are indeed affecting children's representational understanding, or sensitivity (i.e. their readiness to interpret something as a representation; DeLoache, 1995), then we should observe an increase in children's ability to use representational media more generally. This can be measured in studies that test transfer of learning. In addition to demonstrating a deeper and more generalized learning, transfer is a crucial skill for applying meaning to otherwise arbitrary tasks (Barnett & Ceci, 2002). Transfer is often spoken of in the binary terms *near*, meaning very similar to the original task, and *far*, meaning a more abstract and generalized application. However, transfer is more accurately reflected on a spectrum between the two sides. Barnett and Ceci (2002) identified six different domains by which transfer can vary. The *physical context* refers to the location; does a task that is taught in the lab transfer to the child's home? *Temporal context* refers to lapse of time; can a skill taught be maintained the next, or several days later? *Modality* can refer to the actual medium being used; can children transfer training on video to a different representational medium, such as photos? *Social context* involves the person the child is learning from; can children transfer skills learned from a known person to someone unknown, or to an animated character? *Knowledge domain* refers to what they are being taught; can a child transfer success on an imitation task to success on a search task? Finally, *functional context* refers to the purpose of the information; can a child transfer information learned in an academic setting to their play?

In the previously mentioned study in which 2-year-old children saw themselves on live video at home for two weeks before coming to the lab, the prior experience helped them use both video and picture in the search tasks (Troseth, 2003b). On the first day, children successfully used video to find the hidden toy compared to a control group who had not gotten the live video exposure. This, alone, demonstrates four types of transfer: physical context (training at home versus testing in lab), temporal context (training and testing on different days), social context (seeing themselves and family members on video during training versus a researcher on video during testing), and functional context (naturalistic playing at home versus a tasked game in lab). On the following day, children returned to the lab and were shown a photo of the hiding place on each trial. Despite not getting experience using photos in this way, the children who received the live video exposure at home succeeded using the photos in the search task compared to the control group. Thus, the experience with live video transferred across representational media, signaling a modality transfer (video to photos), in addition to a further temporal transfer. Similarly, naturalistic experience with live video, as measured through questionnaires, predicted success on search tasks using both video and photos (Troseth et al., 2007). Broadly, this suggests that specialized training in one domain increases not only children's successful representational use within the specific domain (e.g. video) but their overall representational understanding/sensitivity.

Research also demonstrates that experience with an easily understood representational medium aids in understanding a more advanced representational medium. For example, 2½-year-olds, who typically struggle using a scale model in object retrieval tasks, were able to succeed using the model after first using photographs in the search task a day prior, a representational medium children of that age are more readily able to understand (DeLoache, 1991). The same effect occurred with video; 2½-year-olds who experienced the search task with video on the first day did significantly better on the search task using a scale model two days later, compared to children who used the model on both days (Troseth & DeLoache, 1998).

When tests of transfer are not conducted, it is difficult to determine whether a particular experience taught children more information about representations generally, thus increasing their overall representational understanding, or whether it merely taught them to be successful in that one domain. For example, when a child is video chatting with an experimenter, it is possible that they are experiencing increased success because they are “seeing through” the representation (video screen) and going right to the referent (the person), completely missing the representational relation between the two (Troseth, Pierroutsakos, & DeLoache, 2004).

An interesting transfer question using photos is whether in-lab training on the temporal relevance of digital photos leads to success only when children are tested using digital photos (as in the fourth condition of Experiment 1), or whether such training leads to a more generalized understanding of pictorial symbols, with children extracting and applying information from *printed* photos with greater ease.

Another question to explore is whether these potentially varying concepts about learning from different screen media are specific to the particular device a child has experience with, or whether such experiences result in more general increased representational insight across devices. This question can be tested by providing in-lab training on the affordances of a particular device, and measuring whether this training transfers to learning from a different device. When Troseth and colleagues (2006) provided 5 minutes of video chat with the experimenter before an object retrieval task, both the training and testing were done using the same television screen. What if young children instead receive the training via a different device (e.g. an iPhone) before being tested using the television screen? Will their experience of an interactive video chat give them the needed representational insight into the relation between video and reality that will allow them to transfer this insight to a different medium? Or is the training they receive (and any skill they receive at using information from a person on video) specific to the device? However, in our ever-changing technological landscape, it is crucial to consider children's individual experiences with various media. For example, a child who regularly video chats using his or her family television, which is customary in some households who connect their computer to their TV monitor, would likely hold a different conception of a television screen than a child who solely watches passive, fictional content.

Troseth et al. (2007) emphasized the importance of investigating whether a potential shift toward earlier representational understanding of screen media could generalize to increased learning from educational TV. To test this question, transfer tasks that vary the knowledge domain, social context, and functional context will be particularly helpful. For example, to investigate transfer between knowledge domains, one could test whether success on a search task translates to success on a word-learning task. In the social context, one could test whether children develop a broad enough representational understanding from receiving a contingent video chat interaction with one experimenter that they subsequently use information provided by different experimenter (or even an animated character) on video during the testing phase.

Finally, in our modern world of increasing technology, it is crucial to further investigate how various types of contingent interactions affect learning and children's representational understanding. New technology, for instance, allows for *physical* contingency (e.g. a child can tap a touchscreen or move their body with a motion capture device to create a desired outcome). When 2-year-old children were instructed to touch the tablet during an object labeling video they learned more novel words than children who did not receive this contingent interaction. Further, those who were instructed to touch the specific location of the labeled object performed best (Kirkorian, Choi, & Pempek, 2015). Another study using an object retrieval task found that playing an interactive computer game that demonstrated the hiding location led to greater success than merely watching the video of the hiding event (Lauricella, Pempek, Barr, & Calvert, 2010). These studies suggest that physical contingency (a cause and effect interaction) is beneficial even outside of the more typical social and verbal contingency. Virtual reality further blurs the lines between medium and referent. In a recent review paper considering the effects of virtual reality in children's lives, the authors explain, "From a psychological standpoint, VR is the feeling of non-mediation; the sensation that there is no technology between the user and his or her sensory experience" (Bailey & Bailenson, 2017, p. 2). Thus, the degree to which children treat virtual reality as a representation should be explored.

As technology continues to blend physical and social contingency through increasingly sophisticated motion capture, virtual reality, and artificial intelligence technology, timely research is urgently needed to both investigate how children understand, use, and are affected by these representational experiences, as well as capitalize on the shifting media landscape to answer unique questions about the cognitive development of young children's representational understanding.

REFERENCES

AAP Council on Communications and Media (2016). Media and Young Minds. *Pediatrics. 138*(5). doi:10.1542/peds.2016-2591

Acredolo, L. P., & Goodwyn, S. W. (1985). Symbolic gesturing in language development. *Human development*, *28*(1), 40-49.

Anderson, D. R., Bryant, J., Wilder, A., Santomero, A., Williams, M., & Crawley, A. M. (2000). Researching Blue's Clues: Viewing behavior and impact. *Media Psychology*, *2*(2), 179-194.

Anderson, D. R., & Pempek, T. A. (2005). Television and very young children. *American Behavioral Scientist*, *48*(5), 505-522.

Atkin, C. K., Hocking, J., & Gantz, W. (1979). How young children perceive television commercials. In *70th annual Convention of the Eastern Communication Association, Philadelphia*.

Bailey, J. O., & Bailenson, J. N. (2017). Considering virtual reality in children's lives. *Journal of Children and Media*, 1-7.

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn?: A taxonomy for far transfer. *Psychological bulletin*, *128*(4), 612.

Barr, R., Muentener, P., & Garcia, A. (2007). Age-related changes in deferred imitation from television by 6- to 18-month-olds. *Developmental Science*, *10*(6), 910–921.

Bigelow, A. E., & Birch, S. A. J. (1999). The effects of contingency in previous interactions on infants' preference for social partners. *Infant Behavior and Development*, *22*(3), 367–382.

Bloom, P. (1996). Intention, history, and artifact concepts. *Cognition, 60,* 1-29.

Bloom, P., & Markson, L. (1998). Intention and analogy in children's naming of pictorial representations. *Psychological Science*, *9*(3), 200-204.

Calvert, S. L., Strong, B. L., Jacobs, E. L., & Conger, E. E. (2007). Interaction and participation for young Hispanic and Caucasian girls' and boys' learning of media content. *Media Psychology, 9*, 431–445. doi: 10.1080/15213260701291379

Calvert, S. L., Richards, M. N., & Kent, C. C. (2014). Personalized interactive characters for toddlers' learning of seriation from a video presentation. *Journal of Applied Developmental Psychology*, *35*(3), 148-155. doi: http://dx.doi.org/10.1016/j.appdev.2014.03.004

Deacon, T. W. (1998). *The symbolic species: The co-evolution of language and the brain.* New York: Norton.

DeLoache, J. S. (1987). Rapid change in the symbolic functioning of very young children. *Science*, *238*(4833), 1556-1557.

DeLoache, J. S. (1991). Symbolic functioning in very young children: Understanding of pictures and models. *Child development*, *62*(4), 736-752.

DeLoache, J. S. (1993). Distancing and dual representation. In R. R. Cocking & K. A. Renninger (Eds.), *The development and meaning of psychological distance* (pp. 91-107). Hillsdale, NJ: Psychology Press.

DeLoache, J. S. (2000). Dual representation and young children's use of scale models. *Child Development, 71*(2), 329–338.

DeLoache, J. S. (2002). The symbol-mindedness of young children. *Child psychology in retrospect and prospect: In celebration of the 75th anniversary of the Institute of Child Development, 32,* 73-101.

DeLoache, J. S. (2011). Early development of the understanding and use of symbolic artifacts. In U. Goshwami (Ed.), The Wiley-Blackwell Handbook of Childhood Cognitive Development (2nd edition) (pp. 312–336). Oxford, UK: Wiley-Blackwell Publishers.

DeLoache, J. S., & Burns, N. M. (1994). Early understanding of the representational function of pictures. *Cognition*, *52*(2), 83-110.

DeLoache, J. S., Miller, K. F., & Rosengren, K. S. (1997). The credible shrinking room: Very young children's performance with symbolic and nonsymbolic relations. *Psychological Science, 8,* 308–313. doi:10.1111/j.1467-9280.1997.tb00443.x

DeLoache, J. S., Pierroutsakos, S. L., & Troseth, G. L. (1996). The three R's of pictorial competence. *Annals of child development*, *12*, 1-48.

DeLoache, J. S., Pierroutsakos, S. L., Uttal, D. H., Rosengren, K. S., & Gottlieb, A. (1998). Grasping the nature of pictures. *Psychological Science, 9*(3), 205-210.

Demir, D., & Skouteris, H. (2010). Self- recognition in live videos by young children: does video training help?. *Early child development and care, 180*(3), 327-333.

Deocampo, J. A., & Hudson, J. A. (2005). When seeing is not believing: Two-year-olds' use of video representations to find a hidden toy. *Journal of Cognition and Development*, *6*(2), 229–260. doi: 10.1207/s15327647jcd0602_4

Duncker, K., & Lees, L. S. (1945). On problem-solving. *Psychological monographs*, *58*(5), i.

Flavell, J. H., Flavell, E. R., Green, F. L., & Korfmacher, J. E. (1990). Do young children think of television images as pictures or real objects? Journal of Broadcasting and Electronic Media, 34(4), 399-419.

Ganea, P. A., Pickard, M. B., & DeLoache, J. S. (2008). Transfer between picture books and the real world by very young children. *Journal of Cognition and Development*, *9*(1), 46-66.

Gelman, S. A., & Ebeling, K. S. (1998). Shape and representational status in children's early naming. *Cognition*, *66*(2), B35-B47.

Gibson, J. J. (1979). The ecological approach to the visual perception. New York: Psychology Press.

Gombrich, E. H. (1974). The visual image. In D. R. Olson (Ed.), *Media and Symbols: The forms of expression, communication, and education* (pp. 241-270). Chicago: University of Chicago Press.

Goodman, N. (1976). Reality remade. In *Languages of Art.* (2nd Ed.) pp 3-43. Indianapolis, IN: Hackett.

Hains, S. M. J., & Muir, D. W. (1996). Effects of stimulus contingency in infant–adult interactions. *Infant Behavior and Development, 19*(1), 49-61.

Halford, G. S., Cowan, N., & Andrews, G. (2007). Separating cognitive capacity from knowledge: A new hypothesis. *Trends in cognitive sciences*, *11*(6), 236-242.

Harris, P. L., Kavanaugh, R. D., & Dowson, L. (1997). The depiction of imaginary transformations: Early comprehension of a symbolic function. *Cognitive Development, 12,* 1-19.

Hayne, H., Herbert, J., & Simcock, G. (2003). Imitation from television by 24- and 30 month- olds. *Developmental Science, 6,* 254–261. doi:10.1111/1467-7687.00281

Hollenbeck, A. R., & Slaby, R. G. (1979). Infant visual and vocal responses to television. *Child Development*, 41-45.

Icon. (n.d.). In *Merriam Webster Online,* Retrieved from https://www.merriam-webster.com/dictionary/icon

Ittelson, W. H. (1996). Visual perception of markings. *Psychonomic Bulletin & Review, 3,* 171-187.

Kirkorian, H. L., Wartella, E. A., & Anderson, D. R. (2008). Media and young children's learning. *The Future of Children, 18*(1), 39-61

Kirkorian, H. L., Choi, K., & Pempek, T. A. (2015). Toddlers' word learning from contingent and non-contingent video on touchscreens. Manuscript submitted for publication.

Krcmar, M., Grela, B. G., & Lin, Y.-J. (2007). Can toddlers learn vocabulary from television? An experimental approach. Media Psychology, 10, 41–63.

Krmar, M. (2010). Can social meaningfulness and repeat exposure help infants and toddlers overcome the video deficit?. *Media Psychology, 13*(1), 31-53.

Kuhlmeier, V. (2005). Symbolic insight and inhibitory control: Two problems facing young children on symbolic retrieval tasks. *Journal of Cognition and Development*, *6*(3), 365-380.

Lauricella, A. R., Pempek, T. A., Barr, R., & Calvert, S. L. (2010). Contingent computer interactions for young children's object retrieval success. *Journal of Applied Developmental Psychology, 31*(5), 362-369. doi:10.1016/j.appdev.2010.06.002

Lauricella, A. R., Gola, A. A. H., & Calvert, S. L. (2011). Toddlers' learning from socially meaningful video characters. *Media Psychology*, *14*(2), 216-232.

Melzer, D. K., & Daehler, M. W. (2003, April). First trial performance and error corrections on symbolic insight tasks. Poster presented at the 2003 Meeting of the Society for Research in Child Development, Tampa, FL.

Myers, L. J., LeWitt, R. B., Gallo, R. E., Maselli, N. M. (2015). Baby facetime: Can toddlers learn from online video chat? Manuscript submitted for publication.

National Research Council. (2000). *How people learn: Brain, mind, experience, and school: Expanded edition*. National Academies Press.

Peirce, C.S. (1955). Logic as semiotic: The theory of signs. In J. Buchler (Ed.), *The philosophical writings of Peirce* (p. 98-119). New York: Dover.

Pierroutsakos, S. L., & Troseth, G. L. (2003). Video verite: Infants' manual investigation of objects on video. *Infant Behavior and Development*, *26*(2), 183-199.

Potter, M. C. (1979). Mundane symbolism: The relations among objects, names, and ideas. In N. R. Smith & M. B. Franklin (Eds.), *Symbolic functioning in childhood* (p. 41-65). Hillsdale, NJ: Lawrence Erlbaum Associates.

O'Sullivan, L. P., Mitchell, L. L., & Daehler, M. W. (2001). Representation and perseveration: Influences on children's representational insight. Journal of Cognition and Development, 2, 339–365.

Preissler and Carey (2004). Do both pictures and words function as symbols for 18 and 24-month-old children? *Journal of Cognition and Development, 5*(2), 185-212.

Rakoczy, H., Tomasello, M., & Striano, T. (2005). How children turn objects into symbols: A cultural learning account. In L. Namy (Ed.), Symbol use and symbol representation (pp. 67-97). New York: Erlbaum.

Representation. (n.d.). In *Merriam Webster Online,* Retrieved from https://www.merriam-webster.com/dictionary/representation

Richert, R. A., Robb, M. B., & Smith, E. I. (2011). Media as social partners: The social nature of young children's learning from screen media. *Child Development, 82*(1), 82-95.

Roseberry, S., Hirsh-Pasek, K., Parish-Morris, J., & Golinkoff, R. M. (2009). Live action: Can young children learn verbs from video? *Child Development, 80,* 1360–1375. doi:10.1111/ j.1467-8624.2009.01338.x

Roseberry, S., Hirsh- Pasek, K., & Golinkoff, R. M. (2014). Skype me! Socially contingent interactions help toddlers learn language. *Child development*, *85*(3), 956-970.

Rovee-Collier, C. (1999). The development of infant memory. *Current Directions in Psychological Science*, *8*(3), 80-85.

Schmidt, M. E., Crawley-Davis, A. M., & Anderson, D. R. (2007). Two-year-olds' object retrieval based on television: Testing a perceptual account. *Media Psychology, 9*(2), 389-409.

Schmitt, K. L., & Anderson, D. R. (2002). Television and reality: Toddlers' use of visual information from video to guide behavior. *Media Psychology, 4,* 51-76.

Sharon, T. (2005). Made to symbolize: Intentionality and children's early understanding of symbols. *Journal of Cognition and Development*, *6*(2), 163-178.

Sharon, T., & DeLoache, J. S. (2003). The role of perseveration in children's symbolic understanding and skill. Developmental Science, 6, 289–296.

Strouse, G. A., O'Doherty, K., & Troseth, G. L. (2013). Effective coviewing: Preschoolers' learning from video after a dialogic questioning intervention. *Developmental psychology*, *49*(12), 2368.

Strouse, G. A., & Troseth, G. L. (2008). “Don’t try this at home”: toddlers’ imitation of new skills from people on video. *Journal of Experimental Child Psychology, 101,* 262–80. doi:10.1016/j.jecp.2008.05.010

Strouse, G. A., & Troseth, G. L. (2014). Supporting toddlers’ transfer of word learning from video. *Cognitive Development, 30,* 47-64.

Strouse, G. A., Troseth, G. L., O’Doherty, K. D., & Saylor, M. M. (2015). Co-viewing and contingency support toddlers’ word learning from video. Manuscript submitted for publication.

Suddendorf, T. (1999). Children’s understanding of the relation between delayed video representation and current reality: A test for self-awareness? *Journal of Experimental Child Psychology, 72,* 157–176.

Suddendorf, T. (2003). Early representational insight: Twenty-four-month-old use a photo to

find an object in the world. *Child Development*, *74*(3), 896-904.

Symbol. (n.d.). In *Merriam Webster Online,* Retrieved from https://www.merriam-webster.com/dictionary/symbol

Tarasuik, J. C., Galligan, R., & Kaufman, J. (2011). Almost being there: Video communication with young children. *PLOS ONE, 6*(2). doi:10.1371/journal.pone.0017129

Tarasuik, J. C., Galligan, R. F., & Kaufman, J. (2013). Seeing is believing, but is hearing? Comparing audio and video communication for young children. *Frontiers in Psychology, 4,* 1-6. doi: 10.3389/fpsyg.2013.00064

Tomasello, M. (1999). The cultural ecology of young children’s interactions with objects and artifacts. *Ecological approaches to cognition: Essays in honor of Ulric Neisser*, 153-170.

Tomasello, M., Striano, T., & Rochat, P. (1999). Do young children use objects as symbols?. *British Journal of Developmental Psychology*, *17*(4), 563-584.

Troseth, G. L. (2003a). Getting a clear picture: Young children's understanding of a televised image. *Developmental Science*, *6*(3), 247-253.

Troseth, G. L. (2003b). TV guide: Two-year-old children learn to use video as a source of information. *Developmental Psychology, 39*(1), 140–150.

Troseth, G. L. (2010). Is it life or is it memorex? Video as a representation of reality. *Developmental Review, 30*(2), 155-175.

Troseth, G. L., & DeLoache, J. S. (1998). The medium can obscure the message: Young children's understanding of video. *Child Development, 69,* 950-965.

Troseth, G. L., Pierroutsakos, S. L., & DeLoache, J. S. (2004). From the innocent to the intelligent eye: The early development of pictorial competence. In R. Kail (Ed.), *Advances in child development and behavior, Vol. 32*. New York: Academic Press.

Troseth, G. L., Saylor, M. M., & Archer, A. H. (2006). Young children's use of video as a source of socially relevant information. *Child Development, 77*(3), 786-799. doi:10.1111/j.14678624.2006.00903.x

Troseth, G. L., Bloom Pickard, M. E., & DeLoache, J. S. (2007). Young children's use of models: Testing an alternative to representational insight. *Developmental Science, 10*(6), 763-769.

Troseth, G. L., Casey, A. M., Lawver, K. A., Walker, J. M., & Cole, D. A. (2007).

Naturalistic experience and the early use of symbolic artifacts. *Journal of Cognition and Development*, *8*(3), 309-331.

Werner & Kaplan. (1984). The organismic basis of symbol formation (pp. 2-39). In *Symbol formation*. Hillsdale, N.J.: Erlbaum.

Wright, J. C., Huston, A. C., Reitz, A. L., & Piemyat, S. (1994). Young children's perceptions of television reality: determinants and developmental differences. *Developmental Psychology, 30*(2), 229.

Zelazo, P. D., Müller, U., Frye, D., Marcovitch, S., Argitis, G., Boseovski, J., ... & Carlson, S. M. (2003). The development of executive function in early childhood. *Monographs of the society for research in child development*, i-151.