

Can 24-Month-Old Toddlers Transfer Their Representational Insights

from Video to Pictures?

Eylon Liu

Vanderbilt University

Mentors: Georgene Troseth and Israel Flores

April 24, 2019

Abstract

Representational media are everywhere in children's daily lives: the photos on the wall, the videos shown on TV, and the picture books children read. In order to foster better learning and develop more age-appropriate interactive media for children, prior researchers have explored young children's ability to understand these representational media, including photos and videos. In the current study, 24-month-olds' ability to utilize and apply information from video and pictures to solve tasks in the real world was tested. Replicating two of the conditions in Troseth, Saylor, and Archer (2006)'s study, the current study investigated if a five-minute interactive, contingent video experience could facilitate children's use of information from video to solve an object retrieval task, compared to a controlled group. As a transfer task, children participated in an object retrieval task using photos on the following day. This task explored whether experience with contingent video (video chat) would promote children's understanding of representations more generally. A significant positive correlation was found between children's performance on the first day and the second day, and children's language skill also displayed a significant positive correlation with children's performance. However, children in the interactive video condition in the current study did not perform as well as in the Troseth et al. (2006) study. With the support from a follow-up CCTV condition (with a full-sized video screen), it was conjectured that the children's lower performance in the video condition resulted from a lack of parental support while co-viewing the video chat with their child.

Can 24-Month-Old Toddlers Transfer Their Representational Insights

from Video to Pictures?

Introduction

One of the most important ways that humans transmit knowledge is by using symbolic (or representational) media such as pictures and videos, scale models, diagrams, and maps. The connections and distinctions between representational media and their referents might seem obvious for adult, but to truly understand the representational values of media, one must recognize the duality of symbols (DeLoache, 2002). For example, a 4 x 6-inches two-dimensional picture of a couch in an IKEA catalogue has a low physical resemblance to the eight-feet-long couch that it represents. However, because of its representational value, the picture allows people to gain information about the real couch, including its color, style, and texture. This quality is the “dual nature” of representational media (DeLoache, 2002, p. 78): they are real-life objects in themselves with their own properties and simultaneously can abstractly stand for other objects or even ideas.

Children usually do not develop an understanding of the representational values of pictures or video until around 30 months, and previous research has consistently demonstrated that 24-month-olds cannot solve a real-world problem after seeing its solution on video or a picture (DeLoache, 1993; DeLoache & Burns, 1994; Troseth & DeLoache, 1998). In an object retrieval task, either by watching on a TV screen or looking through a window, 24-month-old children were shown by researchers where a toy was hidden in a familiar adjoining room. Children were then allowed into the adjoining room and asked to

find the toy. Children did not reliably find the toy using information from the video, but they were 100% successful if they saw the hiding event through the window (Troseth & DeLoache, 1998; Troseth, 2003). Since the only altered factor between two conditions was the source of the information, researchers could infer from these results that 24-month-old children did not use symbol-mediated information (from a 2-dimensional video image) in the same way that they used information gathered through direct experience.

This gap between children's ability to learn from video and from a live demonstration has been called the video deficit (Anderson & Pempek, 2005; Barr, 2010). For example, infants and toddlers are better at imitating behaviors from live demonstrations than from the ones seen on videos (Barr & Hayne, 1999; Hayne, Herbert, & Simcock, 2003). Specifically, in the study by Barr and Hayne (1999), infants were exposed to a series of behaviors either through live demonstration by a researcher or through pre-recorded videos of the same researcher. Infants as young as 12-month reproduced the target actions significantly more when they watched the demonstration live compared to through a video.

The difference in learning from video concerns parents and media policymakers due to the fact that very young children in the developed world have substantial experience with representational screen media. In fact, 38% of 0 to 1-year-old infants and 80% 2- to 4-year-olds were exposed to interactive mobile devices in 2013, and video-chatting also has become a more common experience for children (Common Sense Media, 2013, as cited in Choi, Kirkorian, Pempek, & Schroeder, 2016). With such a general increase in children's experience with media, learning about the nature of the video deficit and children's

understanding of video and other screen media has become a relevant topic among researchers. It has been a concern that children's screen time might replace their experiences with the real physical and social world, where more effective learning occurs. To develop plans for effective early education, parents and policymakers are especially interested in knowing at what age children should be exposed to screen media (such as video chatting with grandma) and whether children can learn from media experience (such as through educational videos) (Barr, 2010).

Perceptual or Conceptual?

Two possible hypotheses have been proposed by various developmental psychologists to understand the nature of the deficit in video learning when compared to learning from a live demonstration displayed by young children. The first one deals with the relatively poor perceptual cues of the two-dimensional videos compared to real three-dimensional stimuli. Research has shown that infants process information from 2D images slower compared to information from 3D objects (Carver, Meltzoff, & Dawson, 2006). Thus, the mismatched perceptual cues between 2D and 3D information could be potentially difficult for young toddlers to see the connection between information from a 2D video image and a 3D object that the video image represents (Barr, 2010). However, according to DeLoache, Strauss, and Maynard (1979), infants as young as 5 months old could transfer knowledge across dimensions and perceive similarities between a 2D picture and the 3D object that it represented. After being habituated to a doll, infants preferred to look at novel dolls than at either the original familiar doll or its picture. Despite the mismatched perceptual cues

between 2D and 3D information, infants could, in fact, perceive the similarities between a 2D image and the 3D object it represents. Another study by Preissler and Carey (2004) showed that children as young as 18 months old could grasp the perceptual similarity between an object and the picture that depicted the object. After being familiarized with a picture of the target object, a whisk, 18-month-old toddlers reached for either just the target object or both the target object and its picture 100% of the time when the researcher asked for the “whisk”. Although never exposed to 3D whisk, 18-month-old toddlers can make association between the real target object and the picture of it. In another trial when the researcher asked for the “whisk”, these toddlers always picked the picture of “whisk” instead of the non-target object, a garbage disposal crusher. This suggested that 18-month-olds could not only match the image of an object with the actual object, but also differentiate other objects from the picture of the target object. As a result, past research has consistently demonstrated that the video deficit is unlikely to be a perceptual issue, in that children can recognize the perceptual similarities and differences between images of objects and the actual objects they represent from a very young age.

Another hypothesis suggests that the video deficit is based on a conceptual rather than a perceptual issue. A video image is both an object itself (i.e., a pattern of light behind a glass screen surrounded by a case) and a representation of something else (i.e., the content of the image) (Troseth & DeLoache, 1998). Because of the dual nature of video images and other representational media, it is possible that young toddlers only see the representational media as objects themselves and fail to comprehend their representational values, which can provide

information about the real world (DeLoache, 1987). There are many research studies supporting this hypothesis. In a study by Troseth and DeLoache (1998), 2-year-old toddlers were convinced by the researcher that a live TV screen was actually a window to the adjoining room. In contrast to toddlers' consistent poor performance on the object-retrieval task shown in the past studies (Troseth & DeLoache, 1998; Troseth, 2003), children successfully used information from the video to retrieve a hidden object in the adjoining room. In this case, children no longer had to contemplate the dual nature of the video images to succeed in the task, because they were convinced that they were watching a real event instead of a representation of the event. However, the images from TV were still 2D images with the absence of depth cues and motion parallax (Troseth, 2010). Thus, the fact that toddlers could succeed in the task suggested that the video deficit is more of a conceptual issue than a perceptual one.

Factors Leading to Representational Insight

“Representational insight” refers to the basic realization of the symbolic relationship and connection between a symbol and its referent (DeLoache, 1995a, 1995b). While young children might have difficulties to achieve such insight, developmental progress occurs rapidly with relevant experience. For example, young toddlers can better understand the representational value of media when they have more experience showing them that symbolic media, such as video, can provide accurate information about the real world (Troseth, Casey, Lawver, Walker, & Cole, 2007). In a research by Troseth (2003), a video-camera was connected to the family TV, and 24-month-old children watched themselves at home on TV

in real time for two weeks. Through this experience, children were able to see the relevance of the video images displayed on the TV to the real life. After this two-week training, these experienced toddlers could successfully use information from video (where they saw the researcher hide a toy in the room shown on TV) to solve the object retrieval task in the lab. Later, when the experimenter showed these children photos of the hiding places, they could also use this different kind of pictorial representation to find the hidden toy, even though they had received no special training with photos. The picture-version of this object retrieval task has been shown to be difficult for 24-month-olds (DeLoache & Burns, 1994), as children cannot spontaneously comprehend the representational value of the pictures. In the year 2017, 24-month-old children still consistently failed the object-retrieval task with printed pictures when no additional training was provided (only 11% correct), even though 95% the children had seen their own pictures on a screen (e.g. phone, tablet, computer, camera) during the previous month before the study (Russo Johnson, Troseth, & Flores, 2018). However, in Troseth's study (2003), after the two-week exposure to live video of themselves on TV, toddlers appeared to transfer their representational insights into the video as a source of information about the real world onto a new symbolic medium, pictures.

Increasing social contingency can also facilitate children's learning from video (Barr, 2011). For example, it was found that 2-year-old children are significantly more likely to imitate the actions of a socially responsive adult displayed in a live video achieved by a close-circuit TV system than an adult in a noncontingent, pre-recorded video (Nielsen, Simcock, & Jenkins, 2008). Moreover, a decade ago, researchers determined that a five-

minute socially interactive video experience can also help 2-year-olds succeed in the object retrieval task. Before the object-retrieval task, Troseth and her colleagues (2006) exposed 2-year-old toddlers to a five-minute contingent video experience. Through "video chatting" achieved by connecting two closed-circuit TVs with cameras, a researcher displayed on the TV screen had a conversation with the child and the parent, played Simon Says and sang a song with the child, and finally told the child to find a sticker under the chair that the child was sitting on. After the five-minute experience with this responsive, helpful video partner, these trained toddlers were able to solve the object retrieval task after contingent researcher on TV told them the hiding locations through video chat.

It might seem like the toddlers in this study gained representational insights about videos through the five-minute socially contingent video chat experience. However, because children in this study were not asked to "transfer" their use of information across different media (Troseth, 2003), it was not clear whether they understood the video-chatting person's behavior as symbolizing, or "standing for", a real event happening elsewhere. It was possible that such a success in the object retrieval task was a result of children "seeing through" the screen (Troseth, Flores, & Stuckelman, 2019, p. 5), or just perceiving the responsive person as if they were really there. In this case, 24-month-old toddlers might not recognize that they were responding to a mere representation of the reality due to the high degree of iconicity between the video of researcher and the actual researcher as well as the contingent responses they received through the live video. As the person on the screen responded to toddlers contingently with social cues, they might be tricked into thinking there is an actual person

behind the screen, or they lost sight of the fact that they were watching on video and just responded to the person. If this were true, the video would lose its representational value, and watching the video would be similar to when children thought they were looking through a window at a real event.

Current Study

The goal of the current study is to determine if the five-minute interactive video experience truly facilitated children's understanding of video as a representation (Troseth et al., 2006). A previous study by Troseth (2003) demonstrated that if children grasped the representational value of one representational medium, they could transfer their knowledge to another one. Thus, in the current research, children were asked to generalize or transfer their knowledge of using information from a contingent video of a person talking to them to static photos of the hiding locations. If children can achieve representational insight with video, then they may be able to transfer that understanding to a second, similar object-retrieval task with photos. This study seeks to clarify whether or not 24-month-olds can achieve representation insight with a limited-length training on symbolic experience. The potential implications of this study can help parents know how much support is needed for their toddler to learn from video, including video chat. If parents think that the connection between video chat and the real world is obvious and that toddlers understand it, parents might not help these young children benefit from it.

In this study, the main research question to be answered is whether the five-minute contingent interaction through video chat could really promote 24-month-olds' understanding

of video. More specifically, the current study is a two-day experiment, in which two slightly different versions of the object retrieval task are used during both days of the study (Troseth & DeLoache, 1998; Troseth, 2003). The first day includes a replication of the two conditions in the Troseth et al. (2006) research, and the sources of information for the object retrieval task are different for the two conditions. In the video condition, children are given information to retrieve a hidden object through video-chatting with a research assistant. In the current study, phones and tablets, devices that are more commonly employed for video-chatting, are used in place of the close-circuit TVs set used in the original Troseth et al. (2006) study. In the control condition, children are directly told in-person about the retrieval location. On the second day for both conditions, children experience the same object retrieval task, but the researcher shows them various photos of the hiding places instead of using video chat. This difference in procedure between the first and second day could reveal if children can achieve representational insights with the first medium (video) and transfer that understanding to the second kind of representation (pictures).

In the original Troseth et al. (2006) study, children who were directly told where to find the hidden toy in the “direct control” were more successful than children in the interactive video condition who experienced the five-minute interactive training. This was another instance of the “video deficit” in children’s learning from video compared to from direct experience (Anderson & Pempek, 2005; Barr, 2010). The same difference in children’s performance in the video and control conditions during the first day is also predicted in this current study. If children’s experience with contingent video-chatting can improve their

understanding of representational media, however, it was predicted that children in the video condition would perform significantly better than children in the control condition while using pictures on the second day. This is because children in the control condition would have received practice on the object retrieval task but no experience using a representation for information on the first day. Moreover, since the object retrieval task involves verbal instructions, it is predicted that children with a higher level of language skill will generally perform better in the task for both the first and the second day. Thus, a measure of children's vocabulary level is also included in the current study.

An alternative hypothesis is that the five-minute socially contingent video experience cannot help children better develop representational insights on video; instead, children in the Troseth et al. (2016) study happened to "look through" the screen (Troseth et al., 2019, p. 5). If this were true, children in the control condition (who hear the person face-to-face) will still perform better on the first day than those in the video condition, following the pattern of video deficit. However, children in both conditions will have similar performance with the picture task on the second day. Previous studies reported a practice effect, where children generally performed slightly better on the symbolic object retrieval on the second day after having experience with a similar object retrieval task on the first day (DeLoache, 1993). Children's better performance after practicing the task might result from their increased symbolization experience, their better general understanding of the rules and the social pragmatics of the task, or both factors (Troseth et al., 2019). If the five-minute contingent video experience cannot facilitate children's symbolic understanding, children in both

conditions will merely receive practice in the search task on the first day. As a result, although the object retrieval tasks on the first day for both conditions provide children with information from different sources (directly in-person or through video chat), children in both conditions are predicted to perform equivalently on the second day with their general better understanding on the rules of the task after practicing on the first day.

A factor that makes it difficult to predict children's performance and their responses to the live, contingent video today is that children's exposure to video chat in daily life has changed drastically since the original research. When Troseth et al. (2006) conducted their study, talking to a person on TV who responded contingently was a novel experience for all participants. It is not clear how children in the current study conducted in 2017-2019, with much more prior experience video chatting with relatives, would respond in the experiment. It is possible that toddlers would learn just as much from the video as from direct interactions with adults due to their abundant experience of video chatting. On the other hand, the experience that toddlers have had with video chat prior to participating in the research might hinder their performance. This is because for the toddlers who are exposed to video chat frequently, the interactive video experience might lose its novelty. This could lead to children's reduced attention and interests during the interactive video portion of the study. Moreover, according to the parent questionnaires collected in this study, 34% of the time their children have "never" or "rarely" seen the people they video chatted with, such as their grandparents, in real life. As a result, children might not spontaneously connect a person whom they have met on a video screen to real life. This research is important because the

American Academy of Pediatrics media guidelines (American Academy of Pediatrics: Council on Communications and Media, 2016) singled out video chatting as the one kind of screen media acceptable for toddlers, in part based on Troseth et al.'s (2006) study. Thus, it is crucial for researchers to understand if children could actually learn information from the video chatting experience and apply such knowledge in the real world.

Method

Participants

Participants for this study were 48 typically developing, native English-speaking toddlers (M age = 24.33 months, range = 23.2-25.5 months). Families were recruited in the Nashville general area through Tennessee state birth records and through flyers distributed at the Nashville Zoo and on the Vanderbilt University campus. Parents were contacted by telephone or emails and were invited to visit the Early Development Lab at Vanderbilt University. All the parent-child dyads completed two consecutive days of testing. Children were given a storybook and a small toy for participation.

The data of five additional participants had to be excluded from analysis: one child had a developmental delay and four were very distracted during the study and did not pay attention during the orientation to the game. Of the children included in the study, some data from individual test trials needed to be dropped due to parental interference or experimenter error. The final sample consisted of 24 participants in each condition (video and control) with equal numbers of both sexes and the same average age of 24.33 months.

Generally, these toddlers came from families with relatively high socioeconomic status, with the average annual family income falling into the range between \$50,000-\$100,000, and the average education level of their parents was between some graduate work and a master's degree. Most of the participants were Caucasian (82.54%), a few of them were Asian (3.17%) and African American (1.59%), and the rest of them had a mixed background (9.52%).

Materials

The study took place in two adjoining room in the lab space. A hiding room (6.8 x 3.1m) contained a table covered by a green tablecloth, a couch with a red pillow on it, a blue chair, a basket on a table, and a white rug on the floor. These five pieces of furniture were used as the five potential hiding places. For the first 15 participants, the five pieces of furniture (hiding locations) for the second day's testing were a white bag on a small chair, a table covered by black tablecloth, a pink chair with red blanket on it, a container on a stool, and a brown blanket. The intent for this change was to make the transfer task more challenging. Later during testing, the same hiding places as on the first day were used instead of the different set of furniture for the second day. No significant difference was found in toddlers' performance between the two different room setups for the transfer task on the second day. Thus, to simplify the procedure, the room setting stayed the same for both days for the remaining 37 participants.

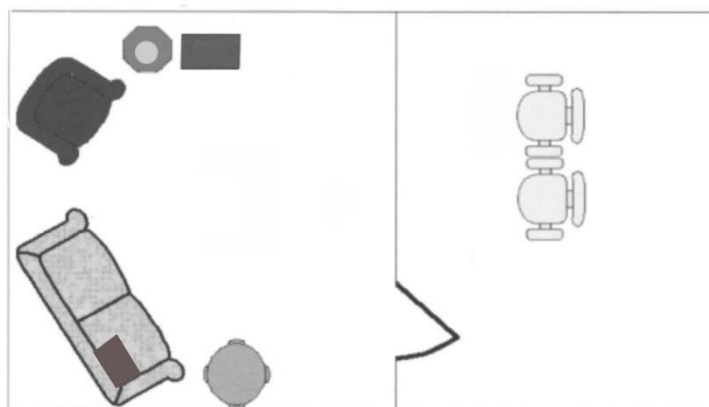


Figure 1: Layout of the laboratory rooms. This figure displays the hiding room on the left, and the control room on the right. Furniture in the hiding room from top to bottom: white rug, basket on a table, blue chair, red pillow on a couch, and green table.

The control room (on the right in Figure 1) was adjacent to the hiding room. In the video condition, children and their parents sat on two chairs in the control room with either an iPhone 5C or an iPad (based on which device was more commonly used by each family at home). The research assistant would use the other phone or tablet to video chat with the parent and child in the adjoining hiding room. In the control condition, the research assistant stood right in front of the parent and the child to have the same chat face-to-face.

During both days, a small Piglet stuffed animal (20 cm high) was used as the object to be hidden. For the first 15 participants, a small Big Bird stuffed animal (20cm high) was used for the transfer task on the second day. Since no significant difference in toddlers' performance using different toys was found for the transfer task, the hidden object (Piglet) stayed the same for both days for the remaining 37 participants. The hiding room contained a video camera to videotape the search task.

Procedure

There were two conditions in this study: a video condition and a control condition. Toddlers were randomly assigned to one of the conditions, with age and sex equated across condition.

Day one. During the first day, children were introduced to the experimental room and were familiarized with the furniture names and rules of the object retrieval task in both conditions through an orientation in the hiding room. After the orientation, there were four placement trials, where the researcher would verbally label each of the four hiding locations and ask the child to place Piglet. For example, the researcher would ask the child, "Can you put Piglet on his red pillow?". If the child did not place Piglet at the correct location, researcher would help the child put Piglet at the right place and say, "See, this is Piglet's red pillow." The purpose of the placement trials was to examine if the child understand the labels for the hiding locations.

Before the actual test trials, children participated in a five-minute contingent interaction with the researcher either through Facetime or directly face-to-face. The interaction in both conditions contained the same elements as in the earlier study conducted by Troseth and her colleagues (2006). Specifically, the researcher talked to the parent and the child about events in the child's life (e.g., their birthday or siblings), then played Simon Says with the child, and finally asked the child to find a sticker in a box under their chair. The intent was to help children see the connection between what a person on screen talked about to current reality.

After the contingent interaction, the four test trials started. On each trial, the child and the parent sat in the control room. The researcher would enter the hiding room and hide Piglet, a stuffed animal toy, in or under one of the four hiding locations. Depending on the condition, the child was told three times either through FaceTime or a face-to-face direct speech, where Piglet was hidden in the hiding room. The child would then be asked to search for Piglet in the hiding room. If Piglet was not found on the first search, the child would be encouraged to continue searching at different locations; however, only a correct first search would be considered as a successful trial. Different prompts with more and more explicit hints would be given to the child until the child successfully retrieve the hidden toy in order to sustain the child's motivation to continue participating (DeLoache, 1993).

Day two. On the second day, participants in both conditions participated in a photo transfer task, in which researcher first gave the child a simple orientation by showing the child the correspondence between each hiding place and a photo of each hiding place. After the orientation, there was a placement trial. Researcher would ask the child to place Piglet on the white rug by only showing the child the photo of the white rug without labeling the location. More and more explicit hints would be given to the child until the child placed Piglet on the white rug, even though only the child's first try would be counted. Lastly, the test trials started. During each of the four test trials, researchers showed the child a photo of the hiding location and told children "I am going to hide Piglet in/under this place in his room". While the primary researcher entered the hiding room to put Piglet in one of the four hiding locations, a secondary researcher showed the child the same picture of the hiding

location again. After Piglet was hidden, the primary researcher would come back to the control room to remind the child again of the hiding location of Piglet by showing the child the same photo for the third time. Finally, children would enter the hiding room and search for Piglet.

For each participant, the primary researcher was the same person for both days. The order of hiding places was counterbalanced between participants and across days.

Measures

Performance on the object retrieval task. To assess how well children performed in the retrieval tasks on both days, the number of each participant's errorless retrievals was counted. A test trial would be counted as an errorless retrieval trial only when children found Piglet on their first search without any assistance. In rare cases, when parents or the researcher accidentally gave hints to the participants (verbally or by looking at the hiding place), a trial was excluded from the data analysis. Since there were four test trials in total, children could receive a score ranging from zero to four, depending on how many errorless retrievals they had. Two researchers independently coded the videotapes with a 99.99% inter-rater reliability. Disagreements on coding were resolved by discussions. In the end, the percentages of each participant's errorless retrievals out of the total number of valid trials were calculated for their performance on each day.

Expressive vocabulary. The linguistic ability of each participant was measured using the MacArthur-Bates Child Development Inventory (CDI) (Fenson et al., 2007), one of the most widely used parental report forms for language assessments in younger children with

relatively high validity (Dale, 1991). Parents would mark on a checklist of 100 vocabulary words about which words their child had explicitly said. Based on norms tables and the child's sex and age in months, the percentile of each participant's vocabulary score was then coded ranging from 0% to 100%.

Parent questionnaires. A demographic questionnaire was filled out by parents to collect information relating to their family's socioeconomic status and their children's race and ethnicity. Parents also completed another media questionnaire about their children's prior media exposure and video-chatting experience. For example, this questionnaire included questions on the age when the child was first exposed to video chat and the amount of time the child spent on video-chatting each week. The demographic information and children's prior experience with representational media were used to ensure statistically equivalent levels of family's SES and children's media background between the two condition groups.

Results

Various statistical analyses were conducted to test the two proposed hypotheses. Both hypotheses predicted that children with higher CDI percentile would generally have better performance on both days regardless of the conditions they were in. Also, a video deficit on the first day was predicted by both alternative hypotheses, in which children in the control condition would perform significantly better on the first day than children in the video condition after controlling other variables. The two alternative hypotheses differed on the predictions for children's performance on the second day. The first hypothesis was based on previous findings that children can transfer or generalize their representational insight from

one kind of representation to another. If the five-minute interactive video experience on the first day could help children with their understanding of images as symbols, it was hypothesized that on the second day, children in the video condition would perform significantly better on the picture task than those in the control condition. In contrast, if the five-minute experience had no effect on children's symbolic understanding, the alternative hypothesis predicted that children in both groups would perform equivalently on the second day.

Demographics and Experiment Implementation

To eliminate the potential effects of other confounding variables, statistical analysis was first conducted to see if there were any systematic differences across conditions in other potential predictive factors, including the age of participants, the household income level, parents' joint education level, children's prior video chat and other media experience, and the age children first used video chatting. With several two-sample t-tests, no significant difference was found on any of these variables across conditions.

Table 1

2-Sample t-test on predictors for children's performance between conditions

Predictors	<i>df</i>	<i>t</i>	<i>p</i> (two-tailed)
age	46	-.187	ns
Parents' joint education	44	-.636	ns
Weekly Video Exposure	46	-1.06	ns
Video Exposure Age	46	1.117	ns
Prior media experience	40	-.897	ns

Note. Some variables had lower degree of freedom due to the nonresponses by parents.

Looking at children's interactive experience in the experiment, no statistically significant difference was found on the average length of the interactive experience between

the researcher and the child across conditions, $U = 176$, $p = ns$, (video condition: $M = 6:23$ minutes, $SD = 1:23$ minutes; control condition: $M = 5:36$ minutes, $SD = 01:05$ minutes). This suggested that children across conditions were given statistically equivalent amount of time to have contingent interaction with the researcher either directly face-to-face or indirectly through a phone or a tablet.

Dropped Participants and Data Exclusion

As mentioned before, there were the additional five participants excluded from the general data analysis, one because of development delay, four due to their fussiness. Because of the different sample sizes, bootstrapped One-way ANOVA tests were conducted to determine if a systematic difference existed between the participants who were excluded from the main data analysis and those who were included. A significant difference in children's CDI percentile was found between dropped and kept participants, $F(1) = 5.08$, $p < .05$ (dropped participants: $M = 19$, $SD = 23.56$; kept participants: $M = 43.37$, $SD = 29.24$). This indicates that the participants who were excluded from the later data analysis had been observed to have less expressive vocabularies by their parents. Such a result suggested that this object retrieval task (which required children to get a verbal hint from the researcher) might have been too difficult for children with low expressive vocabulary. Additionally, parents' joint education level was significantly lower for the dropped participants (between a college degree and some graduate work) compared to those who were included in the data analysis (between some graduate work and a master's degree), $F(1) = 6.11$, $p < .05$.

Prior Media Exposure

Including the dropped participants whose parents completed the media questionnaires, 97% of the children had been exposed to video chat before coming into the lab. The mean age when children were first exposed video chat was 8.2 months. In fact, only two of the participants in this study were described by their parents as having never experienced any sort of video chatting. Both of these participants were randomly assigned into the video condition and had 0% errorless retrieval on both days' performance. Remarkably, one of the two participants was noted to put Piglet on the picture of a white rug twice after the researcher pointed to the picture and asked the child to place Piglet at the same place in the hiding room.

Outcome Measures

Before statistical tests were conducted on children's search performance, a significant condition difference was found in the participants' CDI percentile between the control and the video conditions $t(46) = -1.724, p < .10$). Through a two-tailed t -test, a higher average of CDI percentiles for children in the control condition ($M = 50.50, SD = 31.60$) was found than those in the video condition ($M = 36.25, SD = 25.34$). This difference could lead to systematic better performance for children in the control condition, considering the fact that later linear regressions showed that there were significant correlations between children's CDI percentiles and their performance on both days.

Table 2

Children's performance on both days by conditions and sex

Mean errorless retrieval	Conditions		Sex	
	Video	Control	Female	Male
Day 1 performance	.33	.59	.49	.41
Day 2 performance	.34	.49	.48	.35

Note. Mean errorless retrieval is the mean of the number of times of errorless retrievals out of the total number of effective trials completed.

The descriptive data on children's performance between the two conditions and sexes are shown in Table 2. The ratio of the number of children's errorless retrievals out of the total number of effective trials they completed was calculated and the means of such ratios in all conditions and between sexes were computed and compared. Children in the control condition performed on average better on both days compared to those in the video condition. Generally, female participants had higher average performance on both days compared to male participants. Children's performance in both days in both conditions had non-normal distributions. T-tests and ANOVA could not be conducted due to their reliance on the assumption of a normal distribution. Therefore, a nonparametric statistical test was used in place of ANOVA and *t*-tests.

Mann-Whitney U Test was first conducted to evaluate if differences existed in children's performance on both the first and second day between the two conditions. A significant difference was found in children's performance on the first day ($U = 188.5, p < .05$), while no significant difference was found in children's performance on the second day ($U = 212, p = .107$). Linear regression analysis was also used to control the effect of the statistically unequal distribution of children's CDI percentiles between the two conditions

and to better understand how the predictors contributed to children's performance during both days.

Table 3
Linear Regression on children's performance on both days

Predictors	Day 1		Day 2	
	β	p	β	p
Conditions	.281	.022*	.081	ns
Age	.152	ns	.291	.010**
CDI percentile	.511	.000***	.563	.000***
Sex	-.077	ns	-.081	ns
Parents' joint education	-.107	ns	-.219	ns

Note. Asterisks indicate significant coefficients in the Linear Regression Model in a two-tail test.

* $p < .05$. ** $p < .01$. *** $p < .001$.

Table 3 displays the results of the linear regression tests on children's performance on both days. The five predictors, including the different conditions that the children were in, age, CDI percentile, sex, and parent's joint education, displayed a good linear relationship with children's performance on the first day as the dependent variable through the Linear Regression model ($R^2_{adj} = .399$, $p < .001$). With all the predictions in the model, only two of the predictors had significant predictive effects on children's first-day performance: the conditions children are in and their CDI percentile. This implied that the differences in children's mean performance on the first day between the two conditions were statistically significant after controlling the effect of other potential variables. Such a result confirmed the hypothesis that after controlling other factors, children in the control condition performed significantly better than those in the video condition.

The linear regression model also fit well for children's performance on the second day as the dependent variable ($R^2_{adj} = .462$, $p < .001$). In this model, the conditions children were

in had no significant predictive effect on children's performance on the second day. Among all the other potential predictors, children's CDI percentile and their age had significant predictive effects on children's performance on the second day. The result confirmed the hypothesis that a higher CDI percentile was significantly positively correlated to better performance on both days and supported the alternative hypothesis proposed in the introduction.

A two-tailed Pearson Correlation test was conducted to see if children's performance on the first day has a predictive effect on their second-day performance. A significant positive correlation between the two was found at the 0.001 level ($\rho = 0.558, p < 0.001$).

Discussion and Limitations

The results of the current study confirmed the second hypothesis. While children in the control condition performed significantly better than those in the video condition on the first day, no significant difference was found between the two conditions on children's performance on the second day. In fact, although not statistically significant, children in the control condition performed slight better than those in the video condition in the transfer task with pictures on the second day. The picture transfer task on the second day has also been shown to be a difficult task for 24-month-olds, with less than 15% errorless retrievals when given no external assistance or training in Russo Johnson et al.'s (2018) research. However, in the current study, children in both conditions had on average better performance than 24-month-olds typically achieved in prior studies (DeLoache & Burns, 1994). Since children's performance on the first day and the second day were highly correlated, the hypothesis that

just practicing the searching task successfully could boost children's later performance was supported.

The alternative hypothesis was that symbolic experience—not just practice in the task—would help the video group alone do better on day 2. In prior studies, after experiencing success in the object-retrieval task using one type of medium, 24-month-olds (Troseth, 2003), and 30-month-olds (DeLoache, 1993) displayed better-than-typical performance in the object-retrieval task using information from a completely different type of representational medium. In these studies, an easier symbolic object retrieval task would be on the first day and a more difficult one on the second day; for the older children, success in the relatively easy (for them) picture task improved performance on the model task beyond what was typical for their age. This improvement in children's performance on the more difficult task might be a result of both their symbolization experience in the first task and improved understanding of the instructions and format of the game (Troseth et al., 2019). In other words, the easier symbolic task on the first day allowed children to develop a better understanding of the social pragmatics and the rules of the object-retrieval task, enabling them to devote more cognitive energy to the symbolic elements of the task on the second day. However, since the direct-experience control group was even more successful on day 2, there was no evidence that children in the video chat condition had received additional benefit of symbolic experience, beyond the benefit of practice with the search game. Rather than video chatting giving children insight into the video-reality symbolic relation, they may have merely responded to the contingent person on the screen.

The significant positive correlations between children's CDI percentile and their performance on both days confirmed our hypothesis. The current study involved a highly verbal task for toddlers and the key to succeed in the object retrieval task was to understand not only the labels of the hiding locations but also the instructions given by the researchers. Thus, a higher linguistic ability could be advantageous for children to perform better in the task, whether the information came from video, pictures, or direct experience. Moreover, better language ability could also be a general indicator of children's general intelligence.

The current study did not fully replicate the results of the original study by Troseth et al. (2006). The mean performance for children in the interactive video condition was 69% errorless retrievals in Troseth et al.'s (2006) study, while those in the current study only had an average of 33% errorless retrievals on the first day. There are three possible reasons proposed for such a difference in children's performance. The first one deals with a generation change in media use. Compared to 2006, when any form of video chat was rare, video chat in all formats is very prevalent in the current digitized society. Specifically, 97% of the children participating had video chat experience prior to the study. Since video chat is very common today, the interactive video experience may have lost its novelty for children. This could have resulted in children's lower interests and attentiveness during the interactive video chat with the researcher in this current study compared to the Troseth et al. (2006) study, where children had a high responsiveness to the on-screen person during the five-minute contingent video experience. Unfortunately, we were unable to code attentiveness to

video chat by recording the video output of the phone or tablet in the current study (see below), so this is an area for future research.

The second possibility concerns parental behaviors during the video chat. Due to experience with video chat, parents in this study may have viewed a contingent person on screen differently than parents did a decade ago. As an initial test of this idea, in an independent follow-up project, I systematically re-coded several videos from the original Troseth et al. (2006) study and found that parents in that study generally displayed good modeling behaviors for their children during the interactive video chat. For instance, most parents responded socially to the person displayed on the TV. Since video-chatting through a closed-circuit TV was also a novel experience for the parents, parents in the original study paid general good attention to the interactive video chat and consistently encouraged their children to engage with the researcher displayed on the TV screen. Due to a breakdown of the recording-system in the control room, no video of parents' and children's behaviors during the interactive video chat was captured for the current study. Although no systematic coding could be done, the researchers reported anecdotally that parents in the current study might have responded differently to the interactive video chat. Because FaceTime and Skype calls are common experiences for parents and children today, parents in our current study might have reached a false conclusion that their children could understand and perform video chatting by themselves. This could lead to less parental scaffolding when the researcher was video chatting with children through an iPhone or iPad in the lab. Such a reduction in parent-child interaction could contribute to children's lower performance in the search task, as

research suggested that parents' active participation when co-viewing videos could facilitate more effective learning for children (Strouse, O' Doherty, Troseth, 2013; Strouse, Troseth, O'Doherty, & Saylor, 2017). Therefore, parental behaviors could be systematically different between the original Troseth et al. (2006) study and the current study.

Last but not least, the devices used to deliver the video chat were different between the two studies. In the study by Troseth et al. (2006), the interactive video chat between the researcher and the child along with the parents was achieved through two connected closed-circuit TV monitors. However, an iPhone or an iPad was used instead in the current study, with the original purpose to conduct the interactive video chat in a more familiar way to children's and parents' former video-chatting experience. This could lead to potential perceptual problems, as the screen of a TV monitor was larger than the screen of an iPhone 5C or an iPad. Compared to a phone or a tablet, the person shown on a TV screen resembles a real person more in terms of the size, and the facial details of the person could be presented more clearly to the children by a bigger TV screen. As a result, the differences in screen sizes between the devices used by current study and the one used by Troseth et al. (2006) could also contribute to the different results of the two studies.

To explore the three possible explanations, a follow-up study is needed. A TV monitor is used instead of an iPhone or an iPad to replicate procedures in the Troseth et al. (2006) study more closely. If using a TV monitor in place of an iPhone or iPad could improve children's performance on the first day from 33% errorless retrieval to a higher rate closer to 69% as in the original Troseth et al. (2006) study, then the three proposed hypotheses could

be tested. Does the attentiveness of the parent and the child or the sizes of the screen lead to different performance patterns among 2-year-olds on the first and second day? The follow-up study with a CCTV condition explored these possibilities.

Follow-up Condition

A new CCTV condition was added to test if the device used during the video chat was the key contributing factor of children's low performance in the video condition in the main study. In this condition, the only change in the procedure compared to the video condition in the main study was that a TV monitor was used for the video chat session in place of the iPhone or iPad. The rest of the procedure was completely the same as the video condition in the main study. Video chatting through a TV is still a novel experience for children even now. According to the parent questionnaires collected in this study, 75% of the participants have never used a TV to video chat. If children's performance in this condition were more similar to the 69% errorless retrievals as in the Troseth et al. (2006) study, it would suggest that children's lower performance in the video condition of the first study could result from the different device used for video chat. Statistical analysis will be conducted to see whether the screen sizes of the device used for video chat is a significant predictor of children's performance on both days. If the differences in screen sizes were not a predicting factor, the possible reasons for children's lower performance in the main study than in the Troseth et al. (2006) study would be narrowed down to the different behavior patterns of the parents and the child during video chat, resulted from the generational changes

on people's experience video chat. Moreover, parental behaviors during video chat are recorded and coded in this condition to compare with the Troseth et al. (2006) study.

Method

Participants. Twelve typically developing children whose native language is English will participate in this condition (M age = 24.66 months, range = 23.5-25.7 months). Four additional participants have been excluded from the data analysis thus far due to their low attention during the orientation and their inability to pass the placement trials. Failing the placement trials was an indicator that the children did not have the language skill to understand the labels of the hiding locations – one of the most important keys to succeed in the later test trials for the object-retrieval tasks.

Materials. The room setting, hiding places, and toy to be hidden are completely the same as the first study. The only change from the initial video condition is that a 28-inch TV monitor was set up on a table in the control room, right in front of the two chairs where the parent and the child would be sitting. The TV monitor was connected with a camera in the hiding room and displayed the live video of a researcher recorded by the camera to achieve video chat. A webcam is placed under the TV monitor, so that the researcher in the hiding room can see the child and the parent on a laptop screen connected to the webcam.

Procedures. The procedure in this CCTV condition are identical to the video condition in the initial study, including the use of the picture task on the second day.

Measures. The same measures were used again in this condition and a new measurement was added to systematically analyze parents' and children's behaviors during

the five-minute contingent video experience. The total number of times of parents' scaffolding behaviors were counted. Parents' scaffolding behaviors included when they actively answered the on-screen researcher's questions, encouraged their child to respond to the video, and translated their child's unclear words to the on-screen researcher. Children's responsiveness was also coded. Only when children showed a clear response to the on-screen researcher, that would be counted as one incidence. For example, answering on-screen researcher's questions, following on-screen researcher's instructions, or pointing to the screen were coded as incidences when children responded to the live video.

Results

Statistical tests were conducted to see if replacing the device used to display video chat images from iPhone or iPad in the main study into a 28-inch CCVT would result in a change in children's performance.

Table 4

Children's CDI percentile and performance by conditions

Condition	<i>N</i>	CDI percentile	Day 1 performance	Day 2 performance
Control (main study)	24	50.50	.59	.49
Video (main study)	24	36.25	.33	.34
CCTV (follow-up study)	12	43.67	.52	.52

Note. Data from the control and video conditions on the first study were also included as comparisons.

Children's performance on both days in the CCTV condition will be informally compared with the control and video conditions from the prior study for the purposes of exploratory analysis. Table 4 displays the children's CDI percentiles and performance on both days in all conditions. No statistical tests of mean performance have been conducted due to the different sample sizes. As is clear in the table, use of the CCTV is resulting in search

performance that is very similar to that of the direct experience control condition on day 1, and slightly higher but not significant success than the control group in the picture task on day 2.

Table 5

Children's performance on both days by the screen sizes

Types of device	<i>N</i>	CDI percentile	Day 1 performance	Day 1 performance
iPhone 5C (main study)	13	31.92	.33	.31
iPad (main study)	11	41.36	.32	.34
CCTV (follow-up study)	12	43.67	.52	.52

Note. Data were grouped and organized based on the screen sizes of the device children used for the interactive video chat. Screen sizes of the device increased from the top row to bottom.

Moreover, children's performance relating to the screen size of the device they used was also compared to test the hypothesis if the screen sizes had any predictive effect on children's performance on the first day. Since the screen of an iPad is larger than that of an iPhone, children in the video conditions from the first study were further divided into two groups, based on the device (iPhone or iPad) they used for video chat. Table 5 displays the mean performance of children based on the device children have used. Linear Regression tests were conducted, and the type of device used during video chat was not a significant predictor for children's performance on both the first and second days.

Table 6

Parent's and child's behaviors during interactive video chat

Conditions	<i>N</i>	Parent scaffolding	Child Responsiveness
CCTV (current)	3	19	14.25
Interactive video (Troseth et al., 2006)	3	13.25	14.75

Note. The numbers of time when a parent displayed positive modeling behaviors as well as when the child clearly responded to the researcher on the screen were coded for both the CCTV condition in the current study and the interactive video condition from the Troseth et al. (2006) study.

Finally, parent's and child's behaviors in this condition were compared with the ones in the interactive video condition in the Troseth et al. (2006) study. Although only three videos have currently been coded in each condition, equivalent amounts of parents' positive modeling behaviors and child responsiveness were found between the two conditions.

Discussion

Children in the CCTV condition had much better average performance on both days than children in the video condition in the main study. This indicated that the types of the device used during video chat had an effect on children's performance on the object retrieval task while all the other factors stayed the same. However, no significant correlation could be found between the screen sizes and children's performance. Children using an iPad to video chat had almost the same performance as those using an iPhone, even though the screen size of an iPad is much bigger than an iPhone one. This suggested that children's increased performance in the CCTV condition was unlikely to be merely a result of the increased screen size of the TV monitor.

In the version of the search task used in this research, children had to listen to the person on the screen tell them where to find Piglet. No matter what screen size, the information was conveyed verbally. Therefore, the size of the screen only mattered for children's interpretation of the status of the person on the screen, whether they were someone to be listened to. Children did not see a hiding event on the video screen, which would have had a clearer perceptual reason to affect search performance.

While the bigger perceptual images on the TV was not the cause of children's better performance in this condition, it was possible that the good modeling behaviors of parents during the five-minute contingent video experience had an effect. Parents' good modeling behaviors could result from parents' increased interests and attention due to the novelty of using TV to video chat. Parents in this condition displayed similar patterns of scaffolding and modeling behaviors during the contingent video experience as in the Troseth et al. (2006) study. By responding to the on-screen researcher with social cues and accurate information about their child, parents showed their child the relevance of the live video to the real life (Strouse & Troseth, 2014). In fact, in later research, Strouse and Troseth (2014) found that the social cues from a live co-viewer, in this case the parents watching the video chat with the child, could be more influential than the social contingency of the on-screen person on children's learning from the live video. Even indirect supports from parents, such as responding to the on-screen researcher, could lead to more children's responsiveness to the live video chat interactions (Myers, Crawford, Murphy, Aka-Ezoua, & Felix, 2018). As a result, this follow-up CCTV condition supported the hypothesis that children's low performance in the video condition from the main study were likely the result of the lower attention and less support of the co-viewing parents.

General Discussion

The results of the study showed another instance of the video deficit, as in the main study, children in the video condition performed significantly worse than those in the control condition on the first day. While children in the follow-up CCTV condition had about the

same average performance as those in the control condition, more data would be needed to draw constructive conclusions. A practice effect was also founded in this study. Although children in the direct condition never received any additional training on symbolic media in the lab, their performance on the symbolic object retrieval task using pictures as a source of information were better than those recorded in the former studies (Russo Johnson et al., 2018). This suggested that children's experience on the first day with the object retrieval task had positive effects on their performance on the second day, as they better understood the rules and the social pragmatics of the task.

Compared to 2006 when Troseth and the colleague conducted their study, video chat became a much more common experience for both children and parents. In the current study, 97% of the children was exposed to video chat prior coming into the lab. However, such an increase in children's prior experience with representational media did not help their performance in the object retrieval task. In fact, as video chatting became a less novel experience for both parents and their child, parents might misconstrue that their child could understand video chat without their help. For example, in the main study where either a phone or a tablet was used by parents to achieve video chat, parents might provide enough support to their child when co-viewing the 5-minute contingent video chat. However, in the follow-up CCTV condition, video chatting became a novel experience again for both parents and their child, as it was rare for families to use TV as a device to video chat. In this case, similar levels of parental support as in the Troseth et al. (2006) study were observed during five-minute interactive video chat, and children's performance using the person on video's

cue for information on the first day in this condition was also comparable to children's performance in the interactive video condition in the Troseth et al. (2006) study. This indicates the importance of the parent's role when co-viewing videos with the child for children to better develop representational insight—in other words, for children to understand a video image as a representation of something in the real world.

References

- American Academy of Pediatrics: Council on Communications and Media. (2016). Media and young minds. *Pediatrics*, *138*(5), 1-6.
- Anderson, D. R., & Pempek, T. A. (2005). Television and very young children. *American Behavioral Scientist*, *48*(5), 505–522.
- Barr, R. (2010). Transfer of learning between 2D and 3D sources during infancy: Informing theory and practice. *Developmental Review*, *30*(2), 128-154.
- Barr, R., & Hayne, H. (1999). Developmental changes in imitation from television during infancy. *Child Development*, *70*, 1067–1081.
- Carver, L. J., Meltzoff, A. N., & Dawson, G. (2006). Event-related potential (ERP) indices of infants' recognition of familiar and unfamiliar objects in two and three dimensions. *Developmental Science*, *9*, 51-62.
- Choi, K., Kirkorian, H. L., Pempek, T., & Schroeder, E. (2016). Toddlers object retrieval using touchscreens: The role of working memory skills. *Psychological Science*, *27*(5), 726-736.
- Dale, P. S. (1991). The validity of a parent report measure of vocabulary and syntax at 24 months. *Journal of Speech and Hearing Research*, *34*, 565-571.
- DeLoache, J. S. (1987). Rapid change in the symbolic functioning of very young children. *Science*, *238*, 1556–1557.
- DeLoache, J. S. (1995a). Early symbolic understanding and use. In D. Medin (Ed.), *The Psychology of learning and motivation* (Vol. 33, pp. 65–114). New York, NY: Academic Press.

- DeLoache, J. S. (1995b). Early understanding and use of symbols. *Current Directions in Psychological Science*, 4, 109–113.
- 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: Erlbaum.
- DeLoache, J. S. (2002). The symbol-mindedness of young children. In W. W. Hartup & R. A. Weinberg (Eds.), *Child psychology in retrospect and prospect: The Minnesota Symposia on child psychology* (Vol. 32, pp. 73-101). Los Gatos, CA: LEA Publishing.
- DeLoache, J. S. & Burns, N. M. (1994). Early understanding of the representational function of pictures. *Cognition*, 52, 83-110.
- DeLoache, J. S., Strauss, M. S., & Maynard, J. (1979). Picture perception in infancy. *Infant Behavior & Development*, 2, 77–89.
- Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E. (2007). *MacArthur-Bates Child Development Inventory, Infant Short Form* (2nd ed.). Baltimore: Brooks.
- Hayne, H., Herbert, J., & Simcock, G. (2003). Imitation from television by 24- and 30-month-olds. *Developmental Science*, 6, 254–261.
- Myers, L. J., Crawford, E., Murphy, C., Aka-Ezoua, E., & Felix, C. (2018). Eyes in the room trump eyes on the screen: Effects of a responsive co-viewer on toddlers' responses to and learning from video chat. *Journal of Children and Media*, 12(3), 275-294.

- Nielsen, M., Simcock, G., & Jenkins, L. (2008). The effect of social engagement on 24-month-olds' imitation from live and televised models. *Developmental Science, 11*(2008), 722-731.
- Preissler, M. A., & Carey, S. (2004). Do both pictures and words function as symbols for 18- and 24-month-old children? *Journal of Cognition and Development, 5*, 185-212.
- Russo Johnson, C., Troseth, G. L., & Flores, I. (2018). Do young children in the “selfie generation” understand digital photos as representations? Manuscript submitted for publication.
- 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), 2368–2382.
- Strouse, G. A., Troseth, G. L., O' Doherty, K. D., & Saylor, M. M. (2017). Co-viewing supports toddlers' word learning from contingent and noncontingent video. *Journal of Experimental Child Psychology, 166*, 310-326.
- Strouse, G. A., & Troseth, G. L. (2014). Supporting toddlers' transfer of word learning from video. *Cognitive Development, 30*, 47–64.
- Troseth, G. L. (2003). TV guide: Two-year-old children learn to use video as a source of information. *Developmental Psychology, 39*, 140–150.
- Troseth, G. L. (2010). Is it life or is it Memorex? Video as a representation of reality. *Developmental Review, 30*, 155-175.

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.

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., Flores, I., & Stuckelman, Z. D. (2019). *When representation becomes reality: Interactive digital media and symbolic development*. Unpublished manuscript, Department of Psychology and Human Development, Vanderbilt University, Nashville, USA.

Troseth, G. L., Saylor, M., & Archer, A. H. (2006). Young children's use of video as a source of socially relevant information. *Child Development, 77*, 786–99.